



HAL
open science

Contribution à l'étude des déterminants génétiques impliqués dans le processus infectieux de *Melampsora larici-populina*, l'agent de la rouille foliaire du peuplier

Stéphane Hacquard

► To cite this version:

Stéphane Hacquard. Contribution à l'étude des déterminants génétiques impliqués dans le processus infectieux de *Melampsora larici-populina*, l'agent de la rouille foliaire du peuplier. Sylviculture, foresterie. Université Henri Poincaré - Nancy 1, 2010. Français. NNT: 2010NAN10112. tel-01746338

HAL Id: tel-01746338

<https://hal.univ-lorraine.fr/tel-01746338>

Submitted on 29 Mar 2018

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Faculté des Sciences et Technologies, secteur biologie
Ecole Doctorale Ressources Procédés Produits Environnement (RP2E)

THESE

Présentée pour l'obtention du titre de
Docteur de l'Université Henri Poincaré, Nancy I
En Biologie Végétale et Forestière

Par **Stéphane HACQUARD**

Contribution à l'étude des déterminants génétiques impliqués dans le processus infectieux de *Melampsora larici-populina*, l'agent de la rouille foliaire du peuplier

Approches génomique et transcriptomique

Soutenance publique le 18 novembre 2010

Membres du Jury :

Président :

Pr. Jean-Pierre Jacquot Professeur, Université Henri Poincaré, Nancy

Rapporteurs :

Dr. Marc-Henri Lebrun Directeur de Recherches, INRA de Versailles
Dr. Diana Fernandez Directeur de Recherches, IRD de Montpellier

Examineurs :

Dr. Elodie Gaulin Maître de conférences, Université Paul Sabatier, Toulouse
Dr. Francis Martin Directeur de Recherches, INRA de Nancy (Directeur de thèse)
Dr. Sébastien Duplessis Chargé de Recherches, INRA de Nancy (Directeur de thèse)

Remerciements

Francis, Merci de m'avoir fait confiance pour cette thèse. Je suis ravi d'avoir réalisé ce travail sous ta direction et j'espère avoir répondu à tes attentes. Je sais que j'ai été dans les meilleures conditions pour réaliser cette thèse et c'est en grande partie grâce à toi. Je tenais également à te remercier pour ton soutien durant cette période difficile de fin de thèse, que ce soit pour la recherche de Post docs ou pour le renouvellement de mon contrat.

Sébastien, je ne sais pas trop par où commencer. Tout d'abord merci pour ces trois magnifiques années de thèse passées sous ta direction. Cela a été un honneur pour moi de travailler à tes côtés et d'évoluer sous ton impulsion. Tu m'as transmis ta rigueur scientifique ainsi que ta méthodologie qui me suivront durant toute ma carrière professionnelle. Tu as également été tout le temps disponible pour répondre à mes questions et tu as été d'une grande efficacité dans la correction de mes divers rapports. Hormis l'aspect scientifique, nous partageons les mêmes goûts musicaux, certes douteux, et je te remercie de m'avoir fait connaître des pépites qui trônent maintenant dans la top liste de ma discographie. Comme on dit chez moi, à bientôt gros.

Merci à Marc-Henri Lebrun, Diana Fernandez, Elodie Gaulin et Jean-Pierre Jacquot d'avoir accepté de lire ce manuscrit et de faire partie de mon jury de thèse.

Pascal, en plus d'une bouille incroyable, tu as une 'tchatche' mémorable. Merci pour tout ce que tu m'as appris sur les rouilles en cage.

Valérie, c'est toi qui m'as fait mettre un pied chez les microbes, qui m'as fait découvrir le monde de la recherche et qui m'as tout appris sur la microscopie alors merci mille fois pour tout.

Manue, tu venais de Mirecourt, je venais de St Dié. De ton côté le p'tit Johnny, de mon côté le p'tit Gregory. On était mal partis... mais on a quand même réussi. A bientôt dans World of Warcraft...

Ben, c'était un vrai plaisir de te côtoyer tous les jours au bureau, de partager des vannes bien souvent douteuses avec toi et de déconner sur à peu près tout et n'importe quel sujet. Toutes nos soirées sportives, 'houblonisées', ou culturelles resteront également marquées dans ma mémoire en attendant les prochaines j'en suis sûr...

JB, j'ai survécu sans le savoir à tous tes tests allergiques ; que ce soit la galette, la tarte aux poires piégée à la poudre d'amande ou encore le 'traquenardesque' rhum arrangé noix-amandes invisibles alors je ne sais pas si je dois te remercier pour ça mais en tout cas un grand merci pour ton accueil et ta cuisine délicieuse.

PJ, gamin de la région Nord, dans l'enfer du terrib, vers la Lorraine tu te dérobes, t'as vraiment pas la vie facile. En tout cas merci d'avoir atterri ici et merci pour toutes ces soirées bières-pizzas qui vont me manquer. A plus cht'i gars.

Jaime, j'ai pu voir le dieu Lemmy avec toi et cela restera un super moment de Rock'n roll. Je suis sûr que cette douce musique nous rassemblera de nouveau un jour.

Angela, tou es folie, tou es chaleul, tou es bonheur, tou as la bonne humeur et même si tou chantes Toto Cutugno à tou tête dans le labo, c'était un plaisir de te rencontrer. On gardé le contact...

A2line, je ne sais toujours pas si tu as trouvé le « putain con qu'a cagé » mais en tout cas c'est l'une de tes expressions qui m'a le plus scotché, même le vosgien que je suis ne peut rivaliser.

Yann, c'était aussi un plaisir de te rencontrer. Reste l'homme aux dreads, joueur de djeridoo et amateur de dub music et de cinéma que tu es. Bon courage pour la suite.

François, mon p'tit bouchon, je n'oublierai pas tous ces moments passés à chanter l'helicon et bobo Léon dans le labo.

André, l'image d'un brésilien sur les pistes de ski des Vosges restera pour moi un grand moment de bonheur. La prochaine fois c'est moi qui viens chez toi ...

Marlieseeee, un grand merci à toi aussi pour toutes ces années passées dans le même bureau, à partager des capuccino et à se tenir compagnie jusque tard dans la soirée.

Christine, ça a été également un plaisir de travailler avec toi au quotidien et je tenais vraiment à te remercier pour tout le travail que tu as fourni pour m'épauler. Je te souhaite bon courage ainsi qu'à tes filles pour la suite.

Béatrice ... et je sais à quel point tu aimes ça que je t'appelle Béatrice ... Merci pour ton franc parler, pour ta bonne humeur et ta joie de vivre. Que l'imberbe et douce peau des Viets te fasse toujours fantasmer et que le Shiatsu, Chi kông, Tai-chi, yoga et autre sophro je ne sais quoi t'apportent la paix intérieure.

Je tenais également à adresser un remerciement général à Cendrella, Yohann, Carine, Aude, Emilie, Aurore, PE, Abdala, Veronica, Cécile, Aurélie, Julien, Jessy, Judith, Sara, Jonathan, Panos, Nicolas, Dominique, Agnès, Nicole, Mauricio, Annegret, Marc, Stéphane, Claire, Annick, Cyrille, Claude, Pascale, François et Jean ainsi qu'à tous les autres microbes que j'ai côtoyés durant ces quatre magnifiques années.

Enfin, un grand merci à mon frangin pour ces week ends de glisse dans les Alpes qui m'ont permis de me déconnecter du labo et à mes parents qui m'ont toujours soutenu que ce soit dans mes choix d'orientation ou même financièrement et sans qui je ne serais peut-être pas en train d'écrire ce pavé.

Mes derniers mots iront pour toi Delphine qui me supportes depuis plus de 6 ans déjà. Même si ces dernières semaines j'ai passé plus de temps avec mon ordinateur qu'avec toi, je sais que tu ne m'en tiendras pas rigueur puisque les rôles s'inverseront bientôt ... Alors merci mille fois de m'avoir soutenu, épaulé et encouragé au cours de la rédaction de ce manuscrit mais également pour tout ce que tu m'apportes dans la vie de tous les jours en général.

Sommaire

Chapitre I : Analyse bibliographique	1
I - La reconnaissance entre plantes et agents pathogènes	3
1 - Des anciens modèles aux modèles actuels de reconnaissance plantes-pathogènes	3
2 - Implication des effecteurs dans les interactions plantes/pathogènes	7
2.1- Implication des effecteurs dans la <u>virulence</u> du pathogène (ETS).....	7
2.2- Implication des effecteurs (Avr) dans l' <u>avirulence</u> du pathogène (ETI)	13
3 - L'export des effecteurs dans la cellule hôte	18
3.1 - Le système de sécrétion de type III : une nanomachine pour délivrer les effecteurs bactériens.	19
3.2 - RxLR : Un motif signal conservé chez les effecteurs d'oomycètes pour l'entrée dans les cellules de l'hôte.....	19
3.3 - l'export des effecteurs de champignons et motifs RXLR-like ?.....	21
4 - La reconnaissance des effecteurs (Avr) par les protéines R de la plante.....	22
4.1 - Les gènes de résistance R	23
4.2 - Perception de l'effecteur par les protéines R.....	25
4.3 - Coévolution entre l'effecteur et la protéine de résistance.....	27
5 - Identification de nouveaux effecteurs chez les oomycètes et les champignons.....	28
5.1 - Taille des effecteurs et signal de sécrétion	28
5.2 - Richesse en cystéine des effecteurs	29
5.3 - Présence de motifs conservés chez les effecteurs.....	29
5.4 - Variabilité nucléotidique des effecteurs.....	31
5.5 - Expression des gènes codant pour ces effecteurs.....	32
II - Les Pucciniales	33
1 - Généralités	33
2 - Phylogénie et taxonomie	35
3 - Cycle de vie	37
4 - Processus d'infection au cours de la phase urédienne	39
5 - L'haustorium.....	41
5.1 - Implication de l'haustorium dans la biosynthèse de métabolites.....	41
5.2 - Implication de l'haustorium dans le transfert des nutriments.....	43
5.3 - Implication de l'haustorium dans la sécrétion de molécules signal.....	45
III - L'interaction peuplier-<i>Melampsora larici-populina</i>	47
1 - Le peuplier	47
1.1 - La populiculture.....	47
1.2 - Le peuplier: un arbre modèle dont le génome a été séquencé.....	49
1.3 - Les maladies des peupliers	49
2 - <i>Melampsora larici-populina</i>	51
2.1 - Cycle biologique de <i>Melampsora larici-populina</i>	53
2.2 - Virulences connues chez <i>Melampsora larici-populina</i>	55
3 - Bases moléculaires de l'interaction Peuplier/ <i>Melampsora</i>	57
IV - La génomique	61
1 - Etat des lieux sur le séquençage des génomes de champignons et oomycètes phytopathogènes ..	61
2 - L'annotation des Génomes.....	65
3 - La transcriptomique	67
V - Objectifs de la thèse	71
1 - Contexte	71
2 - Projet de thèse.....	71

Chapitre II : Analyse globale du génome et du transcriptome de <i>Melampsora larici-populina</i>	77
Chapitre III : Analyse globale du secrétome de <i>Melampsora larici-populina</i>	131
Chapitre IV : La microdissection à capture laser, un outil puissant pour l'analyse du transcriptome de tissus complexes: le cas du stade urédien chez <i>Melampsora larici-populina</i>	183
Chapitre V : Discussion générale	201
I - Analyse du développement de <i>Melampsora larici-populina in planta</i>	203
II- Analyse du génome de <i>Melampsora larici-populina</i>	206
III - Analyse du transcriptome de <i>Melampsora larici-populina</i>	207
1 - La biotrophie, principale limitation à l'étude de l'expression des gènes de <i>Melampsora larici-populina</i>	207
2 - Identification de gènes induits durant le processus infectieux	209
IV - Analyse du secretome de <i>Melampsora larici-populina</i>	214
1 - Identification de petites protéines secrétées présentant des homologies avec des effecteurs précédemment caractérisés.	215
2 - Identification d'effecteurs candidats spécifiques à <i>Melampsora larici-populina</i> riches en cystéine et sous pression de sélection positive.....	218
3 - Identification d'effecteurs candidats induits <i>in planta</i>	221
Conclusions et Perspectives	227
Références bibliographiques	233

Chapitre I : Analyse bibliographique

Chapitre I : Analyse bibliographique

Dans la nature, les plantes sont en contact permanent avec des microorganismes présents dans leur environnement et les interactions biologiques en découlant peuvent être bénéfiques, neutres ou néfastes pour la plante. Alors que certains microorganismes jouent un rôle déterminant dans la vigueur et l'état sanitaire des plantes, qu'elles soient annuelles ou pérennes, d'autres peuvent engendrer des maladies qui impactent leur santé. Ces organismes phytopathogènes, qu'ils soient des champignons, des oomycètes, des bactéries ou des virus, peuvent entraîner des pertes économiques considérables. Au cours de l'histoire par exemple, certains microorganismes pathogènes ont dévasté des cultures entières entraînant dans les cas les plus graves des épisodes de famines importants. Le plus célèbre, causé par la maladie du mildiou de la pomme de terre, a eu lieu en Irlande entre 1845 et 1849 et a entraîné des pertes humaines conséquentes ainsi qu'un exode massif d'irlandais vers le nouveau monde (Haas et al. 2009).

I - La reconnaissance entre plantes et agents pathogènes

Du fait de leur statut figé, les plantes sont soumises en permanence à des attaques d'agents pathogènes. Néanmoins, la maladie reste, dans la majorité des cas une exception car les plantes ont mis en place, au cours de leur évolution, des systèmes de défense reposant sur de nombreux protagonistes moléculaires afin de lutter contre ces agressions extérieures. Ces réactions de défense sont gouvernées par la reconnaissance de l'agent pathogène par la plante hôte.

1 - Des anciens modèles aux modèles actuels de reconnaissance plantes-pathogènes

Le premier modèle sur le mode de reconnaissance entre plantes et organismes phytopathogènes est issu des travaux de Flor en 1947 qui ont été conduits sur l'agent de la rouille foliaire du lin. Suite à ses travaux, Flor avait observé que la résistance était gouvernée par un locus dominant chez l'hôte, que la virulence était récessive chez l'agent pathogène et que cette virulence était gouvernée par un nombre de gènes qui est égal au nombre de gènes de résistance présents chez l'hôte. En 1956, Flor énonce le concept de relation gène-pour-gène selon lequel: 'A chaque gène qui conditionne la résistance de la plante hôte correspond chez l'agent pathogène un gène qui conditionne la virulence'. La reconnaissance de la plante est ainsi gouvernée par l'interaction entre un produit de l'agent pathogène codé par les gènes d'avirulence *Avr*, et un

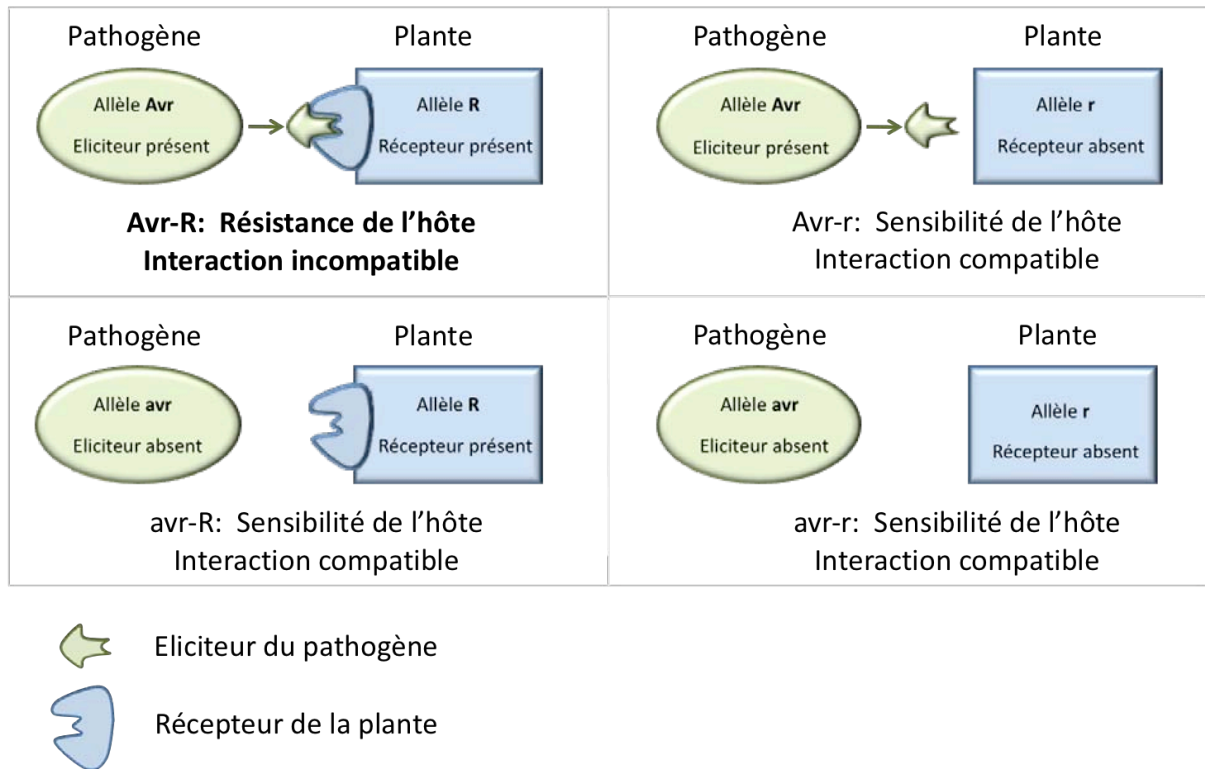


Figure 1: Représentation schématique de la relation gène-pour-gène. (Adapté de Mc Donald 2004). L'allèle de résistance de la plante code pour un récepteur qui reconnaît un éliciteur produit par le pathogène. La reconnaissance de l'éliciteur du pathogène par le récepteur de la plante induit la mise en place des mécanismes de défense conduisant à la résistance (R-Avr). On parle alors d'interaction incompatible. Si la reconnaissance n'a pas lieu (éliciteur et/ou récepteur absent), la plante est sensible (r-Avr, r-avr, R-avr). On parle alors d'interaction compatible puisque le pathogène se développe.

produit de la plante codé par les gènes de résistance *R* (**Figure 1**). Cette interaction R-Avr déclenche une réponse immédiate de la plante appelée réponse d'hypersensibilité (HR, Hypersensitive Response) qui consiste en une mort cellulaire localisée aux points d'infection et qui a pour objectif de limiter la croissance de l'agent pathogène. Si cette reconnaissance a lieu, le pathogène est dit avirulent et l'interaction avec la plante est incompatible puisque celui-ci ne parvient pas à se développer dans les tissus foliaires. Dans le cadre de cette interaction, l'hôte exprime une résistance totale (qualitative). A l'inverse, si cette reconnaissance n'a pas lieu, le pathogène est dit virulent et l'interaction avec la plante est compatible puisque celui-ci se développe sur son hôte. Dans le cadre de cette interaction, l'hôte est sensible et exprime une résistance basale. Cette relation gène-pour-gène a été vérifiée pour de nombreux pathosystèmes (Thompson et Burdon 1992; Crute et al. 1997) et représente toujours l'un des fondements de la reconnaissance plante pathogène.

Il apparaît aujourd'hui clairement que la reconnaissance d'agents pathogènes par les plantes ne repose pas sur ce seul niveau de reconnaissance. Un modèle récent, dit en ZigZag et selon lequel les plantes répondent à l'infection par un 'système immunitaire' à deux niveaux de reconnaissance est maintenant admis (Jones et Dangl 2006). Alors que le premier niveau (réponse immunitaire primaire) reconnaît et répond aux molécules communes à de nombreuses classes de microbes, le second (réponse immunitaire secondaire) répond aux facteurs de virulence du pathogène suivant le modèle gène-pour-gène de Flor. Au cours de la co-évolution entre une plante et un agent pathogène, quatre phases peuvent ainsi être distinguées (Figure 2):

- Lors de la première phase, des éliciteurs généraux (MAMPs/PAMPs, Microbe/Pathogen Associated Molecular Patterns), qu'ils soient de nature polysaccharidiques, glycoprotéiques ou lipidiques (Boller et Felix 2009) sont libérés par le pathogène et reconnus par des récepteurs de la plante (PRR, Pattern Recognition Receptors) déclenchant ainsi les mécanismes de défense de l'hôte (PTI, PAMP-Triggered Immunity). Ce premier niveau de reconnaissance correspond à la réponse immunitaire primaire/basale.

- Un pathogène pouvant contourner ou passer outre cette barrière de défense délivre des effecteurs (§ I-2.1) qui vont interagir avec la cellule hôte et interférer avec cette première barrière de l'immunité rendant ainsi la plante sensible. (ETS, Effector-Triggered Susceptibility)

- La reconnaissance de l'un des effecteurs délivrés par le pathogène par un gène de résistance de l'hôte (§ I-2.2) déclenche les mécanismes de défense de la plante (ETI, Effector-Triggered Immunity), définissant un second niveau de reconnaissance. Celui-ci, basé sur la reconnaissance R-Avr (relation gène-pour-gène), correspond à la réponse immunitaire secondaire de la plante.

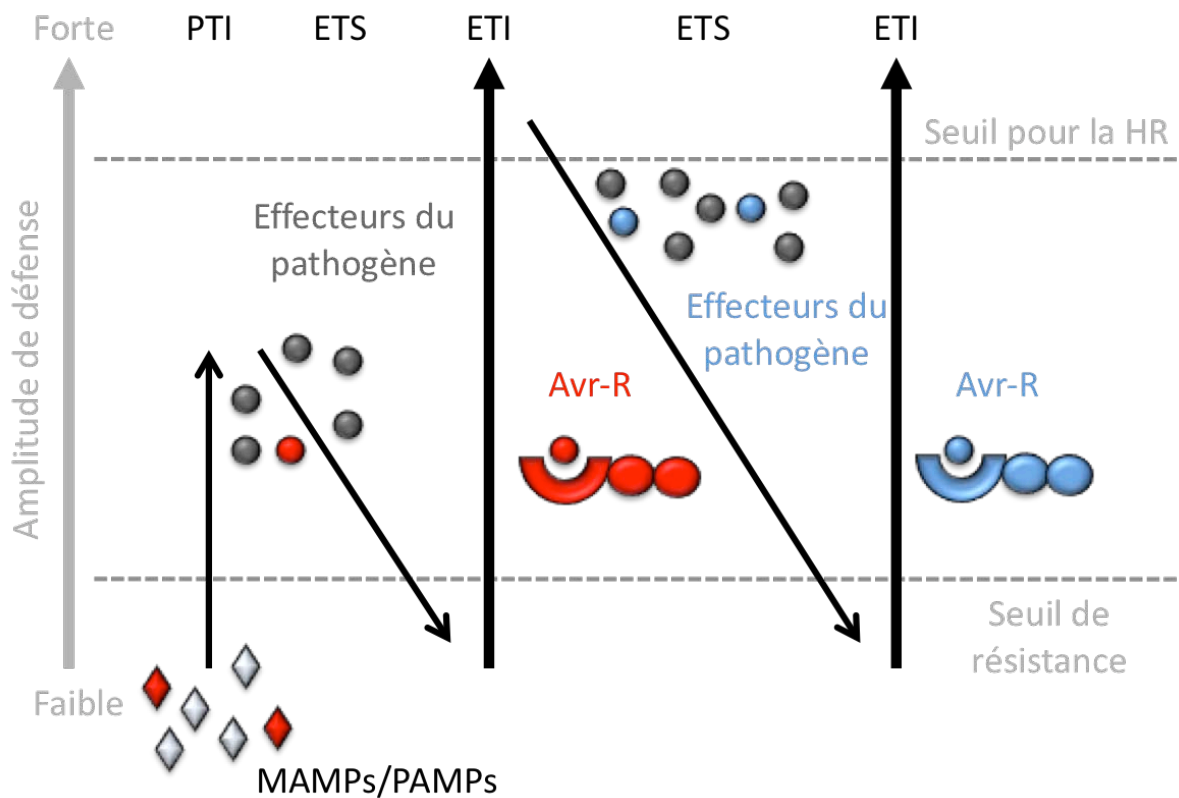


Figure 2: Représentation schématique du modèle en ZigZag (Jones et Dangl 2006). **Phase 1:** La plante détecte les éliciteurs généraux (en rouge) du pathogène (microbial/pathogen-associated molecular patterns MAMPs/PAMPs) via ses récepteurs PRRs (Pathogen Recognition receptors) pour déclencher la première barrière de l'immunité (PAMP-Triggered Immunity, PTI). **Phase 2:** Le pathogène virulent délivre des effecteurs qui interfèrent avec la première barrière de l'immunité (PTI) entraînant la sensibilité de la plante (Effector-Triggered Sensibility, ETS). **Phase 3:** Un effecteur du pathogène (en rouge) est reconnu par une protéine de résistance de la plante R (R-Avr), activant la deuxième barrière de l'immunité (Effector-Triggered Immunity, ETI) qui dépasse le seuil d'induction de la réponse hypersensible (Hypersensitive Response, HR) et induit la mort cellulaire. **Phase 4:** Des isolats du pathogène ont perdu l'effecteur rouge et/ou accumulé de nouveaux effecteurs (en bleu). Ces nouveaux effecteurs peuvent une nouvelle fois interférer avec le système immunitaire de l'hôte (ETS). Au cours de l'évolution, la sélection favorise l'apparition de nouveaux allèles R chez la plante qui reconnaissent les effecteurs du pathogène (ETI).

- Les isolats du pathogène qui ont, au cours de leur évolution perdu cet effecteur et/ou qui ont accumulé de nouveaux effecteurs (transfert horizontal, duplication de gènes, mutations) peuvent supprimer de nouveau cette immunité de la plante.

2 - Implication des effecteurs dans les interactions plantes/pathogènes

Le terme effecteur englobe l'ensemble des molécules produites par des bio-agresseurs afin de manipuler la structure et le fonctionnement de la cellule hôte tels que les facteurs de virulence et les toxines (ETS). Il inclut également les molécules qui sont reconnues par la plante et qui déclenchent des réactions de défense tels que les facteurs d'avirulence et les éliciteurs (ETI). Cette dualité d'activité des effecteurs impliqués dans la virulence et/ou l'avirulence du pathogène a été largement rapportée dans de nombreuses interactions entre plantes et microorganismes et présente actuellement un intérêt particulier dans la compréhension des interactions plantes-pathogènes.

2.1- Implication des effecteurs dans la virulence du pathogène (ETS)

Afin de contourner ou supprimer les mécanismes de défense activés par le système immunitaire de l'hôte, les bactéries, oomycètes ou encore champignons pathogènes sont capables de sécréter des protéines effectrices (Göhre et Robatzek 2008) afin de manipuler le fonctionnement de la plante hôte pour faciliter l'infection.

2.1.1 - Chez les bactéries

Actuellement, une large part des connaissances acquises sur la virulence des organismes phytopathogènes et le contournement des mécanismes de défense de la plante par les effecteurs des microorganismes proviennent d'études menées chez les bactéries (Göhre et Robatzek 2008; Génin et al. 2010). La plupart des effecteurs bactériens ont été identifiés sur la base de leur activité d'avirulence et ont ainsi été nommés gènes d'avirulence *Avr*, leur interaction avec les gènes de résistance correspondants conduisant à la mise en place des mécanismes de défense chez l'hôte. Néanmoins, chez les plantes sensibles, la plupart de ces gènes codent des fonctions impliquées dans la virulence de l'agent pathogène et interfèrent avec le système immunitaire de la plante à de multiples niveaux, notamment en surmontant les barrières physiques de l'hôte, en inhibant l'activation des récepteurs de la plante (PRRs), en modifiant le transcriptome de la plante, en dérégulant les voies de signalisation ou encore en exploitant la machinerie de dégradation de l'hôte (Göhre et Robatzek 2008) (**Figure 3**).

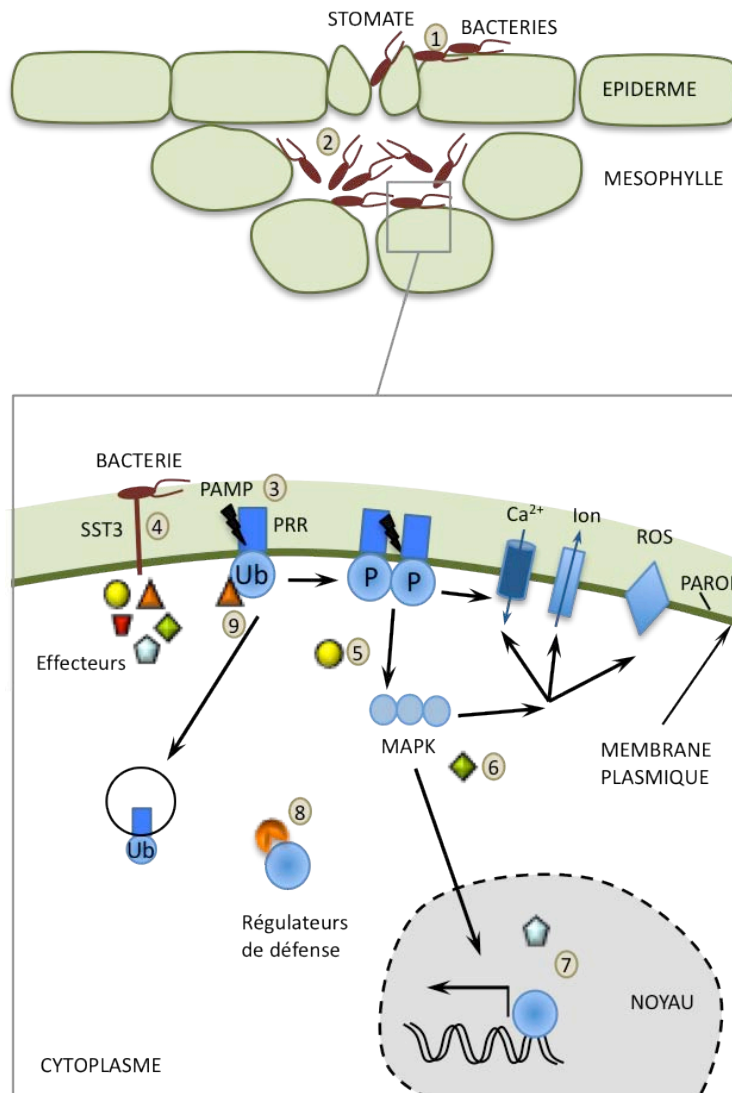


Figure 3: Modifications des mécanismes de défense induits par les effecteurs bactériens (Adapté de Göhre et Robotzek 2008). **1.** Les bactéries se développent à la surface de la feuille et sont stoppées par la cuticule de la plante. **2.** Les bactéries pénètrent au sein des tissus foliaires par les stomates qui se referment lors de la reconnaissance des PAMPs. Certains effecteurs (Coronatine) conduisent à leur réouverture. **3.** La reconnaissance des PAMPs par les récepteurs PRR conduit à l'activation des mécanismes de défense (Flux d'ions, production de ROS ...) via une cascade de signalisation (MAPK). **4.** Les bactéries possédant un système de sécrétion de type III libèrent des effecteurs dans la cellule hôte. **5.** Certains effecteurs tels que AvrPto et AvrPtoB interfèrent avec l'activation des récepteurs PRR ou les étapes précoces de la signalisation. **6.** D'autres tel que HopAI-1 bloquent les voies de signalisation. **7.** Certains induisent des modifications de l'expression des gènes tels que les membres de la famille AvrBs3 (activateurs transcriptionnels). **8.** Afin de dégrader les protéines de l'hôte, des protéases (protéases à cystéines ou serines) sont libérées par les bactéries pathogènes. **9.** Enfin, Certains effecteurs agissent sur le système d'ubiquitination de la plante tel que AvrPtoB.

Lors d'un contact avec un pathogène, les PAMPs tels que la flagelline ou les lipopolisaccharides bactériens sont perçus par des récepteurs de la plante (PRR) conduisant à la fermeture immédiate des stomates. Certaines bactéries telle que *Pseudomonas syringae*, une bactérie Gram négative infectant une large gamme d'hôtes, sont ainsi capables de sécréter de la coronatine, une phytotoxine proche de l'acide jasmonique qui induit la réouverture des stomates (Melotto et al. 2006). Une fois dans l'apoplaste de la plante, celles-ci ont également la capacité d'injecter de multiples effecteurs directement dans la cellule hôte à travers la paroi grâce à un système de sécrétion de type III (§ I-3.1) (Galan et Wolf-Wats 2006). Par ce système sophistiqué, des molécules effectrices sont secrétées pour supprimer les mécanismes de défense mis en place par l'hôte en réponse au pathogène (Jha et al. 2007). Lors de la formation de ce complexe de sécrétion, de nombreuses enzymes lytiques sont produites par la bactérie pour dégrader la structure de la paroi cellulaire de la plante. Cette dégradation peut être accompagnée par la libération de PAMPs directement reconnus par les récepteurs de la plante (PRRs). Cette reconnaissance conduit à l'activation immédiate de ces récepteurs et la mise en place des mécanismes de défense de l'hôte (Zipfel 2009). Chez *P. syringae*, l'effecteur AvrPto cible ces récepteurs PRRs de la plante bloquant ainsi les réponses de défense en découlant telles que la production d'espèces oxygénées réactives, l'activation des MAPK (pour Mitogen-Activated Protein Kinase), l'induction de gènes de défense ou encore la déposition de callose (Xiang et al. 2008).

La perception des PAMPs entraîne chez la plante l'activation de voies de signalisation faisant notamment intervenir des protéines kinases (MAPK). Les MAPKs sont activés en cascade par des processus de phosphorylation/déphosphorylation et conduisent à la mise en place des mécanismes de défense de l'hôte (Pedley et Martin 2005). Les effecteurs bactériens peuvent avoir une action sur cette cascade de transduction du signal chez les plantes comme chez les animaux (Shan et al. 2007). Certains effecteurs, identifiés chez *P. syringae* tels que HopAI-1 et AopA01, sont capables de modifier les composants de la voie de signalisation. Alors que le premier altère les résidus thréonine des MAPKs (Zhang et al. 2007), le second (HopA01) possède une activité phosphatase conduisant à une déphosphorylation des MAP kinases et une suppression des signaux de défense de la plante (Underwood et al. 2007).

Les effecteurs des organismes phytopathogènes peuvent également directement modifier le transcriptome de la plante. Par exemple, les protéines se liant à l'ARN sont ciblées pour altérer le turnover des ARN. C'est notamment le cas de l'effecteur bactérien HopU1 de *P. syringae*, une mono-ADP-ribosyltransférase qui agit sur des protéines riches en glycine se liant à l'ARN telles que les protéines chaperonnes. Cette interaction influence le transcriptome de la plante et régule

négativement les réponses de défense (Fu et al. 2007). De manière différente, les effecteurs bactériens de la famille AvrBs3, identifiés chez la protéobactérie pathogène du poivron et de la tomate *Xanthomonas campestris* pv. *vesicatoria*, agissent directement comme activateur transcriptionnel au niveau du noyau de la plante en se liant à l'ADN (Bonas et al. 1989, Boch et al. 2009). Ces effecteurs possèdent un signal de localisation nucléaire, un domaine leucine zipper pouvant permettre la ligation à l'ADN et un domaine acide d'activation, induisant l'expression des gènes de l'hôte (Kay et al. 2007). Pour certains d'entre eux, le ciblage au niveau du noyau de la plante a été confirmé expérimentalement (Szurek et al. 2002).

Afin de dégrader les molécules de la plante, de nombreux effecteurs identifiés chez les bactéries portent une activité protéase de type cystéine protéase (Hotson et al. 2004) ou encore ciblent la machinerie du protéasome de leur hôte. Par exemple, l'effecteur AvrPtoB de *P. syringae* pv. *tomato* est une protéine modulaire possédant à son extrémité une activité E3 ligase (Janjusevic et al. 2006). Cette activité E3 ligase permet l'ubiquitination d'une protéine kinase de la plante jouant un rôle clé dans l'activation du système immunitaire. La dégradation de cette kinase via le protéasome de la plante induit la suppression de la mort cellulaire programmée de l'hôte et est essentielle pour la virulence de la bactérie (Rosebrock et al. 2007). De la même manière les effecteurs GALA de *Ralstonia solanacearum* possèdent un domaine F-box ainsi que des répétitions riches en leucine. Ce domaine F-box interagit avec les sous-unités Skp1 et Cullin1 de la plante pour former une ubiquitine E3 ligase de type SCF active. Ce domaine F-box étant essentiel à la virulence de la bactérie, les auteurs ont avancé l'hypothèse selon laquelle ces effecteurs pourraient détourner cette ubiquitine E3 ligase de l'hôte pour interférer avec le protéasome et faciliter l'infection (Angot et al. 2006). Néanmoins, les cibles de ces effecteurs restent inconnues.

2.1.2 - Chez les champignons extracellulaires

Bien qu'ils pénètrent dans les tissus de l'hôte, les champignons extracellulaires ne différencient pas de structures d'infections à l'intérieur des cellules de la plante tels que les haustoria (§ II-5). L'espace apoplastique dans lequel ils se développent est un environnement riche en enzymes de dégradation telles que des protéases, des enzymes de dégradation des parois ou des hydrolases, qui, bien qu'inoffensives pour la plante, représentent une menace pour les composants structuraux du pathogène.

Le champignon biotrophe *Cladosporium fulvum* (ascomycète) prolifère et sécrète des effecteurs dans l'espace apoplastique de la tomate. Parmi eux, certains ont la capacité d'inhiber la

dégradation enzymatique de l'hôte. Par exemple l'effecteur Avr2 se lie et inhibe la protéase à cystéine de plante Rcr3 (Rooney et al 2005), alors que l'effecteur Avr4 protège le champignon contre les effets délétères des chitinases de la plante (Van den Burg et al. 2006; Van Esse et al. 2007). En se protégeant contre cette activité de dégradation, les organismes phytopathogènes n'assurent pas seulement leur persistance dans l'apoplaste, mais empêchent également leur reconnaissance par la plante via les PAMPs (PTI) que libèreraient de telles enzymes de dégradation (Libault et al. 2007). Récemment, de Jong et collaborateurs ont montré que le domaine LysM présent au sein de la séquence protéique d'un autre effecteur de *C. fulvum* (Ecp6) est essentiel à la virulence du champignon. Ce domaine permet la séquestration des oligosaccharides de chitine qui sont relargués de la paroi cellulaire des hyphes d'infection durant l'infection, empêchant ainsi la reconnaissance de la chitine par le système immunitaire de l'hôte (de Jonge et al. 2010).

Chez *Fusarium oxysporum* f.sp. *lycopersici*, champignon ascomycète biotrophe qui colonise les vaisseaux du xylème de la tomate, un rôle direct dans la virulence a également été démontré pour les effecteurs Avr2 (Six3) et Avr3 (Six1). Les mutants de délétion pour ces deux gènes rendent le pathogène moins virulent, notamment sur les plants de tomates matures (Rep et al. 2004; Houterman et al. 2009). Néanmoins, les cibles de ces effecteurs chez la plante restent inconnues. Un troisième effecteur nommé Avr1, également identifié chez ce pathogène, a un effet encore plus marqué en supprimant les réponses de défense induites par la reconnaissance R-Avr (I2-Avr2 et I3-Avr3) mises en place par la plante. Cette suppression pourrait provenir d'une action directe sur la reconnaissance R-Avr ou de l'inhibition des voies de signalisation induites par celle-ci (Houterman et al. 2008).

Ustilago maydis est un basidiomycète biotrophe responsable de la maladie du charbon du maïs qui prolifère rapidement au sein des tissus foliaires de son hôte et induit la formation de tumeurs chez la plante. Un exemple élégant d'implication d'effecteur dans la virulence de ce champignon a été démontré pour la protéine Pep1 (Doehlmann et al. 2009). Pep1 est un gène orphelin spécifiquement exprimé durant le processus infectieux et codant une protéine de 178 acides aminés. Pep1 est essentiel pour la virulence du champignon puisque le mutant de délétion est incapable d'entrer dans la cellule hôte et élicite de fortes réponses de la plante via l'expression de tout un panel de gènes de défense. Cet effecteur, indispensable à la virulence du champignon pourrait modifier/inhiber la fonction des protéines de l'hôte ou interagir avec les molécules d'ADN affectant ainsi l'expression des gènes de défense de la plante (Doehlmann et al. 2009).

2.1.3 - Chez les champignons et oomycètes formant des haustoria

Certains champignons ascomycètes, basidiomycètes et oomycètes pathogènes biotrophes se caractérisent par la formation d'une structure d'infection particulière mise en place lors de l'infection de leur hôte appelée haustorium (§ II-5). Cette structure, invaginée dans le plasmalemme de la cellule hôte, est impliquée dans le détournement des métabolites de la plante (Voegelé et al. 2003) ainsi que la sécrétion de molécules signal lors de l'interaction avec l'hôte (Kamoun 2006; Dodds et al. 2009). Chez les champignons et oomycètes formant des haustoria, un large nombre d'effecteurs ont été décrits (Catanzariti et al. 2007) mais peu ont été identifiés comme essentiels dans la virulence du pathogène.

Chez le champignon ascomycète biotrophe obligatoire *Blumeria graminis* responsable de l'oïdium de l'orge, l'expression transitoire de Avr-a10 et Avr-k1 dans les cellules de l'hôte augmente la capacité d'infection des plants sensibles par le champignon (Ridout et al. 2006). Chez *Phytophthora sojae*, un oomycète pathogène du soja, la surexpression de l'effecteur Avr1b chez des souches transgéniques entraîne une suppression de la mort cellulaire chez la plante permettant ainsi au pathogène de s'établir (Dou et al. 2008). Une augmentation de la taille des lésions ainsi qu'une mort plus rapide des plantules de soja sensibles ont également été observées. Néanmoins, l'inactivation du gène n'affecte pas la virulence vraisemblablement du fait de la redondance de l'effecteur au sein du génome de *Phytophthora* spp. (Jiang et al. 2008). Chez *Phytophthora infestans*, l'oomycète responsable de la maladie du mildiou de la pomme de terre, l'effecteur Avr3a est également essentiel à la virulence du pathogène (Bos et al. 2006; Bos et al. 2010). Cet effecteur manipule l'immunité de la plante en stabilisant l'ubiquitine E3 ligase CMPG1 de l'hôte conduisant ainsi à la suppression de la mort cellulaire programmée mise en place par la plante (Bos et al. 2010). Chez les oomycètes, un autre exemple d'implication d'effecteurs dans la virulence du pathogène a été mis en évidence chez *Hyaloperonospora arabidopsidis*, un agent biotrophe obligatoire attaquant les crucifères tels que le chou ou encore la plante modèle *Arabidopsis thaliana*. Les auteurs ont utilisé le système de sécrétion de type III de *P. syringae* via une fusion entre les effecteurs ATR1 et ATR13 de *H. arabidopsis* et la région N-terminale de la protéine effectrice bactérienne AvrPS4 (Sohn et al. 2007). Alors que chez les plants résistants d'*Arabidopsis*, la sécrétion de ces effecteurs déclenche une HR, une virulence exacerbée de la bactérie a été observée chez les plants sensibles. Ces effecteurs supprimeraient les réactions de défense de la plante induite en réponse aux PAMPs de la bactérie telles que la production d'espèces oxygénées réactives et la déposition de callose (Sohn et al. 2007).

2.2- Implication des effecteurs (Avr) dans l'avirulence du pathogène (ETI)

Parmi les effecteurs qui sont sécrétés par les organismes phytopathogènes dans la plante hôte pour y accomplir leurs fonctions (§ I-2.1), certains sont reconnus par la plante et déclenchent des réponses de défense. Le produit de ces gènes dit d'avirulence, est directement ou indirectement reconnu par le produit des gènes de résistance de la plante (§ I-4.2). Cette stratégie de défense, correspondant à la réponse immunitaire secondaire de la plante (ETI), conduit à une mort cellulaire localisée aux points d'infection qui permet de limiter la progression de l'agent pathogène (HR). Comme il l'a été mentionné dans le paragraphe précédent, il est maintenant établi que ces gènes d'avirulence peuvent initialement correspondre à des facteurs de virulence qui sont nécessaires à l'expression de la pathogénicité de l'agent infectieux (Stergiopoulos et de Wit 2009). Les gènes d'avirulence ont été mis à jour aussi bien chez les bactéries, les virus ou les champignons que chez les insectes et les nématodes (White et al. 2000) et, contrairement aux gènes de résistance de la plante (§ I-4.1), ils ne possèdent pas de similarité de structure. Le premier gène d'avirulence bactérien a été cloné en 1984 (Staskawicz et al. 1984), le premier gène d'avirulence fongique en 1991 (van Kan et al. 1991), suivi en 2004 par le premier gène d'avirulence d'un oomycète (Tian et al. 2004; Kamoun et al. 2006). Durant ces 20 dernières années, de nombreux gènes d'avirulence mais aussi les gènes de résistance correspondants ont été identifiés, permettant ainsi une meilleure compréhension de la relation gène-pour-gène. Nous nous focaliserons ici uniquement sur les facteurs d'avirulence identifiés chez les champignons et les oomycètes phytopathogènes.

2.2.1 - Gènes d'avirulence des champignons extracellulaires

De nombreux gènes d'avirulence ont été identifiés chez les champignons extracellulaires (**Tableau 1**) dont le développement se limite à l'espace apoplastique (Stergiopoulos et de Wit 2009). Ceux-ci sécrètent des protéines d'avirulence qui sont reconnues ou exécutent leurs fonctions dans l'apoplaste des cellules infectées.

Chez *C. fulvum*, 4 gènes d'avirulence ont été clonés et codent tous pour de petites protéines riches en cystéine qui sont sécrétées durant l'infection. Il s'agit des protéines effectrices Avr2, Avr4, Avr4E et Avr9 pour lesquelles la reconnaissance avec la plante est basée sur l'interaction avec les protéines de résistance Cf-2, Cf-4, Cf-4E et Cf-9 correspondantes (Joosten et De Wit 1999; Thomma et al. 2005). Pour deux d'entre eux (Avr2 et Avr4) une fonction liée à la virulence du champignon a été démontrée (§ I-2.1.2). Cinq protéines extracellulaires additionnelles

Tableau 1: Gène Avr identifiés chez les champignons extracellulaires (D'après Stergiopoulos et De Wit 2009).

Effecteur (Avr)	Taille (aa)	Cystéines	Secretion	Fonction et/ou homologie	Localisation (plante)	Expression	Rôle dans la virulence	Gène-R (type)
Cladosporium Fulvum								
Avr2	78	8	Oui	Inhibiteur de protéases	Apoplaste	In planta	Inhibe Rcr3 et d'autres protéases	Cf-2 (eLRR-TM)
Avr4	135	8	Oui	Fixation à la chitine	Apoplaste	In planta	Protège contre les chitinases	Cf-4 (eLRR-TM)
Avr4E	121	6	Oui	-	Apoplaste	In planta	-	Hcr9-4E (eLRR-TM)
Avr9	63	6	Oui	Inhibiteur de carboxypeptidase	Apoplaste	In planta	-	Cf-9 (eLRR-TM)
Ecp1	96	8	Oui	-	Apoplaste	In planta	Essentiel pour la complète virulence	Cf-Ecp1 (non cloné)
Ecp2	165	4	Oui	-	Apoplaste	In planta	Essentiel pour la complète virulence	Cf-Ecp2 (non cloné)
Ecp4	119	6	Oui	-	Apoplaste	In planta	-	Cf-Ecp4 (non cloné)
Ecp5	115	6	Oui	-	Apoplaste	In planta	-	Cf-Ecp5 (non cloné)
Ecp6	222	8	Oui	Domaine LysM	Apoplaste	In planta	Essentiel pour la complète virulence	-
Ecp7	-	6	-	-	Apoplaste	In planta	-	-
Leptosphaeria maculans								
AvrLm1	205	1	Oui	-	-	In planta et In vitro	-	Rim1 (non cloné)
AvrLm6	144	6	Oui	-	-	In planta et In vitro	-	Rim6 (non cloné)
AvrLm4-7	143	8	Oui	-	-	In planta et In vitro	-	Rim4 et/ou Rim7
Fusarium oxysporum f. sp. lycopersici								
Avr3 (Six1)	284	8	Oui	-	Xyleme	In planta	Essentiel pour la complète virulence	I-3 (non cloné)
Six2	232	8	Oui	-	Xyleme	In planta	-	-
Six3	163	3	Oui	-	Xyleme	In planta	Essentiel pour la complète virulence	I-2 (CC-NBS-LRR)
Avr1 (Six4)	242	6	Oui	-	Xyleme	In planta	Suppression résistance induite par I-2 et I-3	I ou I-1 (non cloné)
Rhynchosporium secalis								
Nip1	82	10	Oui	Toxine non spécifique/induit nécrose et pompe H ⁺ ATPase	-	In planta et In vitro	Non essentiel pour la virulence	Rrs-1 (non cloné)
Nip2	109	7	Oui	Toxine non spécifique/induit nécrose	-	In planta et In vitro	Non essentiel pour la complète virulence	-
Nip3	115	9	Oui	Toxine non spécifique/induit nécrose	-	In planta et In vitro	Non essentiel pour la complète virulence	-
Magnaporthe oryzae								
Avr-Pita	224	8	Oui	Metalloprotéases	Cytoplasme	In planta	Non essentiel pour la virulence	PI-ta (CC-NBS-LRR)
Avr-Pita2	224	8	Oui	Metalloprotéases	-	-	-	PI-ta (CC-NBS-LRR)
Avr-Pita3	226	8	Oui	Metalloprotéases	-	-	-	-
Pw1	147	2	Oui	Protéine hydrophobe riche en glycine	-	-	-	-
Pw2	145	1	Oui	Protéine hydrophobe riche en glycine	Cytoplasme	-	-	-
Pw3	137	0	Oui	Protéine hydrophobe riche en glycine	-	-	non fonctionnel	-
Pw4	138	0	Oui	Protéine hydrophobe riche en glycine	-	-	non fonctionnel	-
Ace1	4035	43	Non	Polyketide synthase et peptidase synthase nonribosomale	non secrété	appressoria	-	PI33
Avr1-CO39	-	-	-	-	-	-	-	PI-CO39

appelées Ecp1, Ecp2, Ecp4 et Ecp5 ont également été caractérisées et conduisent à une HR chez les lignées de tomates portant les gènes *Cf-Ecp* correspondants (Laugé et al. 2000).

Chez *F. oxysporum* f.sp. *lycopersici*, 4 petites protéines appelées Six1 à Six4 (Secreted in xylem) ont été identifiées du fait de leur abondance dans la sève du xylème (Rep et al. 2004; Houterman et al. 2007). Certaines ont un rôle dans la virulence du champignon (§ I-2.1.2) et bien qu'il ait été suggéré que leur action soit extracellulaire, la protéine Six3 (Avr2) peut être reconnue de manière intracellulaire par le gène de résistance I-2, impliquant un transfert de l'effecteur dans la cellule hôte (Houterman et al. 2009).

Au moins 9 gènes *Avr*, nommés *AvrLm1* à *AvrLm9*, ont été génétiquement identifiés chez le champignon ascomycète *Leptosphaeria maculans* responsable de la « nécrose du collet » des crucifères. De manière intéressante les gènes *AvrLm1*, *AvrLm6* et *AvrLm4-7* sont présents dans des régions génomiques riches en transposons et partagent un contenu faible en GC comparé aux autres gènes *Avr* identifiés jusqu'à maintenant (Gout et al. 2006; Fudal et al. 2007; Parlange et al. 2009). Ceci est dû à la particularité du génome de *L. maculans*, composé d'isochores équilibrés en GC entremêlés d'isochores riches en AT au sein desquels ont été localisés ces gènes d'avirulence. Les produits de ces gènes d'avirulence sont reconnus, vraisemblablement dans l'apoplaste, par les produits des gènes de résistance de l'hôte *Rlm1*, *Rlm6* et *Rlm4/Rlm7* entraînant une mort cellulaire programmée des cellules de la plante aux points d'infection.

Rhynchosporium secalis, l'agent necrotrophe responsable de la « tache pâle » sur orge sécrète dans l'apoplaste des peptides de faible poids moléculaire (Nip1, Nip2 et Nip3). Parmi eux, Nip1 et Nip2 sont produits *in vitro* et leur présence *in planta* est corrélée avec la formation de lésions nécrotiques qui sont causées par une stimulation indirecte de l'activité d'une pompe H⁺ ATPase dans la membrane plasmique de la plante (Wevelsiep et al. 1991; Wevelsiep et al. 1993). De plus, Nip1 possède un grand nombre de résidus cystéines essentiel à la structuration de la protéine et déclenche une réponse spécifique de la plante (non-HR) chez les cultivars d'orge possédant le gène de résistance *Rrs1* (Hahn et al. 1993; Van't Slot et al. 2007).

De nombreux gènes d'avirulence nommés *Avr-Pita* (Valent et al. 1991), *Avr1-CO39* (Farman et al. 1998), *Ace1* (Böhnert et al. 2004) et *Pwl1*, *Pwl2*, *Pwl3* et *Pwl4* (Kang et al. 1995; Sweigard et al. 1995) ont été identifiés chez *Magnaporthe oryzae*, l'agent ascomycète hémibiotrophe responsable de la pyriculariose du riz. Ce pathogène pénètre dans les tissus de l'hôte grâce à une structure spécialisée appelée appressorium puis différencie des hyphes d'infection intracellulaire. Tous les gènes d'avirulence identifiés chez *M. oryzae* codent de petites protéines

Tableau 2: Gène Avr identifiés chez les champignons formant des haustoria (D'après Stergiopoulos et De Wit 2009).

Effecteur (Avr)	Taille (aa)	Cystéines	Secretion	Fonction et/ou homologie	Localisation (plante)	Expression	Rôle dans la virulence	Gène-R (type)
<i>Melampsora lini</i>								
AvrL567	150	1	Oui	-	Cytoplasme	haustoria	-	L5, L6 et L7 (TIR-NBS-LRR)
AvrM	314	1	Oui	-	Cytoplasme	haustoria et in vitro	-	M (TIR-NBS-LRR)
AvrP123	117	11	Oui	Inhibiteur de protéase Kazal Ser	Cytoplasme	haustoria	-	P, P1, P2 et/ou P3 (TIR-NBS-LRR)
AvrP4	95	7	Oui	-	Cytoplasme	haustoria	-	P4 (TIR-NBS-LRR)
<i>Blumeria graminis</i> f. sp. hordei								
Avr10	286	4	Non	-	-	-	Augmente la virulence	Mia10 (CC-NBS-LRR)
AvrK1	177	3	Non	-	-	-	Augmente la virulence	MiK1 (CC-NBS-LRR)

secrétées à l'exception de *Ace1* qui code un polypeptide fusion de 4035 acides aminés impliqué dans la production de métabolites secondaires (Böhnert et al. 2004). *Ace1* activerait de façon indirecte le gène de résistance *Pi33* du riz par la biosynthèse d'un métabolite secondaire, aboutissant ainsi à la mise en place des mécanismes de défense de la plante. *Avr1-CO39* entraîne l'avirulence du champignon sur les plants de riz portant le gène de résistance *Pi-39* correspondant (Farman et al. 1998). *Avr-Pita* code quant à lui une petite protéine sécrétée de 224 acides aminés présentant une homologie avec des métalloprotéases zinc dépendantes. L'interaction directe de la protéine mature (176 acides aminés) avec la protéine de résistance *Pita* de l'hôte a été démontrée par des expériences de double hybride et d'accrochage *in vitro* (Jia et al. 2000).

2.2.2 - Gènes d'avirulence des champignons et oomycètes formant des haustoria

Certains gènes d'avirulence ont également été identifiés chez les champignons et les oomycètes formant des haustoria. Dans la majorité des cas, ces protéines sont reconnues ou exécutent leurs fonctions à l'intérieur de la cellule hôte (**Tableau 2**).

Chez *Melampsora lini*, l'agent pathogène basidiomycète responsable de la maladie de la rouille foliaire du lin, les facteurs d'avirulence ont été clonés à partir de 4 loci (Lawrence et al. 2007). Deux de ces loci contiennent des gènes d'avirulence qui interagissent avec différents gènes de résistance, nommés *L5*, *L6*, *L7* et *P*, *P1*, *P2*, *P3*. Les deux autres loci contiennent les gènes d'avirulence qui interagissent avec un seul gène de résistance nommés *M* et *P4*. L'approche utilisée pour identifier ces effecteurs a été dans un premier temps de présélectionner un lot de gènes exprimés durant l'infection ou au niveau de l'haustorium pouvant contenir un ou plusieurs gènes d'avirulence (Dodds et al. 2004; Catanzariti et al. 2006). Chaque gène de ce lot sert ensuite de sonde ADN pour détecter des polymorphismes de taille des fragments de restriction (RFLPs) qui sont ensuite cartographiés parmi une famille F2 au sein de laquelle ségrége 16 gènes d'avirulence (Lawrence et al. 1981). Les sondes pour lesquelles un polymorphisme de taille des fragments de restriction co-ségrége avec un gène d'avirulence dans la famille F2 devient un gène d'avirulence candidat (Lawrence et al. 2007). La fonction d'avirulence de ces candidats est ensuite testée par expression transitoire dans différentes lignées de lin présentant des résistances différentes au pathogène. Ces expériences d'expression transitoire impliquent l'infiltration de feuilles de lin avec une souche d'*Agrobacterium tumefaciens* capable de délivrer dans les cellules de l'hôte le gène candidat exprimé avec un promoteur de plante (Dodds et al. 2004). La fonction d'avirulence est démontrée si une HR est observée chez une seule lignée ou chez des lignées possédant le(s) gène(s) de résistance

correspondant aux spécificités d'avirulence du gène candidat. Les quatre facteurs d'avirulence identifiés chez *M. lini* par cette approche ont été nommés *AvrL567*, *AvrM*, *AvrP123* et *AvrP4* en fonction de l'interaction avec leurs gènes de résistance correspondants chez l'hôte (Dodds et al. 2004; Catanzariti et al. 2006; Catanzariti et al. 2010). Tous codent de petites protéines sécrétées exprimées au niveau de l'haustorium et induisent une HR chez les plants possédant le(s) gène(s) de résistance correspondant

Chez *B. graminis* pv. *hordei*, deux gènes d'avirulence *Avr-a10* et *Avr-k1*, essentiels pour la virulence du champignon chez les plants d'orge sensibles (§ I-2.1.3) sont respectivement reconnus par les gènes de résistance *Mla10* et *MLK* chez des individus résistants (Ridout et al. 2006). A l'inverse d'autres protéines *Avr*, ces protéines d'avirulence ne possèdent pas de peptide signal pour la sécrétion mais sont néanmoins reconnues intracellulairement par les protéines de résistance correspondantes lorsqu'exprimées transitoirement *in planta*. Ceci suggère qu'elles pourraient être sécrétées par une voie de sécrétion alternative. Ces deux gènes partagent de fortes similarités et appartiennent à une famille diversifiée plus large contenant au moins 30 membres.

Chez les oomycètes phytopathogènes, 4 gènes d'avirulence ont été à l'origine isolés par cartographie génétique et sur la base de leurs phénotypes d'avirulence: *Avr1b-1* de *P. sojae*, *Avr3a* de *P. infestans*, et *ATR13* et *ATR1* de *H. arabidopsidis* (Allen et al. 2004; Shan et al. 2004; Armstrong et al. 2005; Rehmany et al. 2005). Comme pour de nombreux autres gènes *Avr*, ces gènes codent de petites protéines sécrétées qui sont reconnues dans le cytoplasme de l'hôte par les protéines R, conduisant ainsi à la mise en place de la HR. L'identification d'un motif RxLR (§ I-3.2) conservé au sein de ces protéines *Avr* a conduit à l'identification d'autres facteurs d'avirulence: *Avr1a*, *Avr3a* et *Avr4/6* de *P. sojae* (Dou et al. 2008a; Qutob et al. 2009), ainsi qu'*Avr4* et *AvrBlb1* chez *P. infestans*. (van Poppel et al. 2008; Vleeshouwers et al. 2008).

3 - L'export des effecteurs dans la cellule hôte

Certains effecteurs ciblent la cellule de l'hôte et doivent donc traverser la membrane plasmique de la plante pour être libérés dans le cytoplasme de la plante. Les mécanismes d'export des effecteurs ont été largement étudiés chez les oomycètes et les bactéries phytopathogènes et des travaux plus récents ont pu établir des points de convergence entre oomycètes et champignons pour ces mécanismes.

3.1 - Le système de sécrétion de type III : une nanomachine pour délivrer les effecteurs bactériens.

Les bactéries pathogènes des plantes et des animaux possèdent 6 systèmes de sécrétion. Parmi eux, le système de sécrétion de type trois (Type Three Secretion System, TTSS) est indispensable pour la virulence de certaines bactéries et est requis pour la translocation des effecteurs à l'intérieur de la cellule hôte (Alfano et al. 2004; Espinosa et al. 2004; Abramovitch et al. 2006). Ce système de sécrétion est certainement le plus sophistiqué des appareils d'export de protéines décrit jusqu'à présent (**Figure 4**). De nombreux effecteurs injectés par ce système interfèrent avec le système immunitaire de la plante (ETS). Pour certains d'entre eux, les mécanismes de suppression des défenses de l'hôte ont été élucidés (§ I-2.1.1). Contrairement à la variabilité des effecteurs injectés par ce système chez différents genres bactériens, la machinerie de sécrétion en elle-même reste relativement bien conservée. Ce système n'a pas d'équivalent chez les champignons ou les oomycètes chez qui d'autres mécanismes sont mobilisés pour délivrer les effecteurs dans la cellule hôte.

3.2 - RxLR : Un motif signal conservé chez les effecteurs d'oomycètes pour l'entrée dans les cellules de l'hôte.

Tous les facteurs d'avirulence identifiés jusqu'à présent chez les oomycètes (*Phytophthora* spp. et *H. parasitica*) possèdent un motif en aval de leur signal de sécrétion (Birch et al. 2008; Tyler 2009) qui est requis pour leur internalisation dans la cellule de l'hôte. Ce motif consiste en une séquence consensus d'acides aminés RxLR (Arginine-x-Leucine-Arginine, x représentant n'importe quel acide aminé), pouvant être suivie par une région acide dEER (Aspartate-Glutamate-Glutamate-Arginine) chez *Phytophthora* spp. Des acides aminés additionnels de part et d'autre de ces résidus conservés sont également indispensables pour la translocation dans la cellule hôte suggérant l'importance de la structure protéique dans la reconnaissance du motif (Bhattacharjee et al. 2006; Dou et al. 2008b). De manière intéressante, Whisson et collaborateurs (2007) ont montré que ce motif n'est pas requis pour la sécrétion de Avr3a chez *P. infestans* mais nécessaire pour le transfert de l'effecteur de l'haustorium vers la cellule hôte. Le motif RxLR rappelle le signal de translocation d'effecteurs identifiés chez *Plasmodium falciparum*, parasite de l'homme transmettant la malaria. Durant l'infection, les protéines sécrétées par *P. falciparum* sont exportées à l'intérieur des hématies grâce au motif conservé RxLxE/D/Q (Arginine-x-Leucine-Glutamate/Aspartate/Glutamine) (Hiller et al. 2004; Marti et al. 2004). L'interversion des motifs de translocation entre les effecteurs de *P. infestans* de *P. falciparum* (c'est à dire le motif RxLR sur l'effecteur de *P. falciparum* et le motif RxLxE/D/Q sur

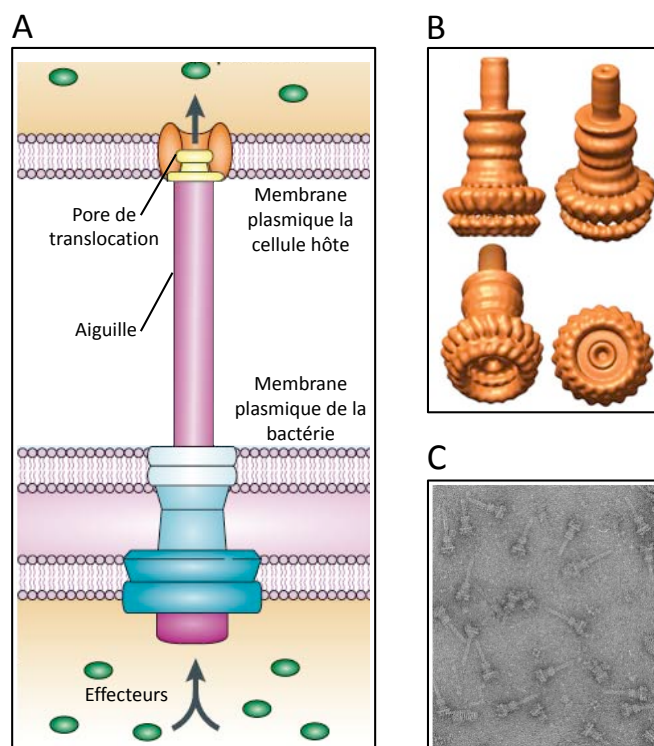


Figure 4: Structure de l'aiguille du système de sécrétion de type III bactérien. **A**, Représentation schématique de l'injectisome (adapté de Cornelis, 2006). La membrane bactérienne et celle de la cellule cible sont séparées par un espace qui est comblé par l'aiguille de l'injectisome permettant l'export des effecteurs bactériens du cytosol de la bactérie vers le cytosol de la cellule cible. **B**, Différentes vues de la modélisation de la structure du complexe de l'aiguille (Galan et Wolf-Wats, 2006). **C**, Complexe de l'aiguille de *Salmonella tyhimurium* visualisé par microscopie électronique à transmission (Galan et Wolf-Wats, 2006)

l'effecteur de *P. infestans*) n'affecte pas leur entrée dans la cellule hôte suggérant un mécanisme d'export commun entre les deux parasites (Bhattacharjee et al. 2006; Dou et al. 2008b). Chez *P. falciparum*, deux études ont montré que le clivage de ce motif entre les acides aminés RxL et xE/D/Q par une aspartyl protéase, était indispensable à l'export de la protéine mature dans l'érythrocyte (Russo et al. 2010; Boddey et al. 2010). Bien que ces motifs soient essentiels à l'export de l'effecteur du pathogène dans la cellule de l'hôte, le mécanisme de transport restait jusqu'à peu méconnu. Une étude récente a montré que le motif RxLR identifié chez *P. infestans* ainsi que d'autres motifs présentant des structures dégénérées proches permettent la fixation des effecteurs aux phospholipides PI3P (Phosphatidylinositol-3-phosphate) et PI4P (Phosphatidylinositol-4-phosphate) (Kale et al. 2010). Cette étude a également révélée l'abondance de ces phospholipides à la surface externe de la membrane plasmique de cellules de plante mais également à la surface de certaines cellules animales suggérant que ce mécanisme pourrait être utilisé par un large panel d'organismes pathogènes eucaryotes pour permettre l'entrée des effecteurs dans la cellule de l'hôte. Ces phospholipides pourraient être accumulés au niveau des radeaux lipidiques de la membrane plasmique de l'hôte conduisant ainsi à l'entrée de l'effecteur dans la cellule de l'hôte par endocytose (Kale et al. 2010). D'autres motifs ont également été identifiés chez les effecteurs CRN de *P. infestans* (LxLFLAK, Haas et al. 2007) et CRN-like d' *Aphanomyces euteiches* (F/LxLYLALK, Gaulin et al. 2008), un oomycète responsable de pourritures racinaires chez diverses légumineuses. Très récemment, une étude a révélé que ces motifs sont également essentiels à l'export de ces effecteurs dans la cellule hôte et que plusieurs de ces effecteurs ciblent le noyau de la plante (Shornack et al. 2010).

3.3 – l'export des effecteurs de champignons et motifs RXLR-like ?

Certains effecteurs de champignons pathogènes des plantes seraient également exportés à travers la membrane de l'hôte et accumulés dans le cytoplasme de la plante. L'interaction de certains effecteurs, ayant une activité d'avirulence, avec des protéines R cytoplasmiques ou nucléaires de la plante est une preuve indirecte de cette translocation (Jia et al. 2000; Dodds et al. 2004; Rafiqi et al. 2010). L'immunolocalisation d'effecteurs de certains champignons pathogènes par microscopie électronique à balayage a apporté la preuve directe de leur translocation dans la cellule hôte. Chez *Uromyces fabae*, l'agent responsable de la rouille foliaire de la féverole, l'accumulation de la protéine sécrétée RTP1 (Rust Transferred Protein 1) a été observée dans le cytoplasme ainsi que dans le noyau de l'hôte confirmant l'export de cette protéine de fonction non connue (Kemen et al. 2005). De la même manière, Le produit du gène d'avirulence AvrM de *M. lini* a pu être détecté au niveau de l'haustorium ainsi que dans le cytoplasme des cellules de lin infectées (Rafiqi et al. 2010) confirmant l'implication de l'interface

haustoriale dans la sécrétion de ces effecteurs chez les agents de rouilles. Dans le cadre de l'interaction *M. oryzae*/riz, l'interface plante champignon a également été analysée en détail afin de comprendre le transfert des effecteurs depuis les hyphes d'infection du champignon vers le cytoplasme de l'hôte (Kankanala et al. 2007; Khang et al. 2010). Kankanala et collaborateurs (2007) ont ainsi montré que les hyphes d'infection sont séparés du cytoplasme de la plante par une membrane dérivée de celle de la plante appelée membrane de l'hyphe extrainvasive (EIHM, Extrainvasive Hyphal Membrane). De manière intéressante, cette membrane renferme une structure renflée caractéristique se formant au niveau de la première cellule de l'hyphe d'infection entrant dans la cellule hôte. Cette structure, récemment nommée BIC pour « Biotrophic Interfacial Complex » semble être l'interface privilégiée pour la sécrétion de molécules signal (Khang et al. 2010). L'accumulation de l'effecteur PWL2 ainsi que de la protéine sécrétée BAS1 de *M. oryzae* au niveau de cette interface ainsi que dans le cytoplasme des cellules de riz infectées suggère l'importance du BIC dans le relargage de ces protéines dans la cellule de l'hôte (Khang et al. 2010).

Les effecteurs de champignons identifiés jusqu'à maintenant ne semblent pas présenter de motifs fortement conservés et facilement reconnaissables tels que le motif RxLxE/Q/D des effecteurs de *P. falciparum* ou le motif RxLR-dEER présent chez *P. infestans*. Néanmoins, des motifs présentant des structures dégénérées proches ont été récemment mis à jour tels que les motifs RFYR et KFLK identifiés respectivement chez AvrL567 et AvrM de *M. lini*, le motif RYWT présent chez AvrLm6 de *L. maculans* ou encore la région RIYER identifiée chez Avr2 de *F. oxysporum* f.sp. *lycopersici* (Rafiqi et al. 2010, Kale et al. 2010). La fusion de la région N-terminale de ces effecteurs avec la GFP (Green Fluorescent Protein) suivie de leur incubation avec des racines de soja a révélé l'accumulation de ces protéines dans les cellules racinaires de soja reflétant également l'importance de ces motifs dans l'export des effecteurs chez les champignons (Kale et al. 2010). Cette étude a révélé que ces motifs permettraient également l'accrochage des effecteurs aux phospholipides PI3P et PI4P présents au sein de la membrane plasmique de la plante suggérant un mode de translocation commun entre les effecteurs d'oomycètes et de champignons (Kale et al. 2010).

4 - La reconnaissance des effecteurs (Avr) par les protéines R de la plante

Une fois exportés dans la cellule de l'hôte ou dans l'espace apoplastique, les effecteurs du pathogène s'exposent à la seconde barrière du système immunitaire de la plante (ETI). La reconnaissance de certains d'entre eux par la plante est gouvernée par l'interaction avec une protéine de résistance R correspondante de l'hôte (R-Avr).

4.1 - Les gènes de résistance R

Bien qu'ils confèrent la résistance à divers agents biotiques et abiotiques (virus, champignons, bactéries, nématodes, blessures ...), les produits des gènes R possèdent de fortes similarités de structure (Dangl et Jones 2001) ce qui permet de les distinguer en sept classes différentes (**Figure 5**).

La première classe de protéines de résistance R chez les plantes ne comprend qu'un seul membre, Pto, qui possède un domaine sérine thréonine kinase intracellulaire. Cette protéine est impliquée dans l'interaction entre la bactérie *P. syringae* et la tomate (Martin et al. 1993). Trois autres classes de protéines de résistance R chez les plantes correspondent à des protéines intracellulaires NB-LRR (aussi appelées Nimbler) possédant un domaine NB (Nucléotide Binding) central et un domaine C-terminal riche en leucines répétées (LRR). Les deux premières classes se distinguent par la présence d'un domaine TIR (Toll and human interleukin 1 receptor) et d'un motif CC (Coiled-Coil) en N-terminal. Les gènes de résistance possédant un domaine TIR constituent la classe des TIR-NB-LRR (TNL) et les gènes de résistance possédant un motif Coiled-Coil constituent la classe des CC-NB-LRR (CNL). Ces classes sont les plus représentées au sein des génomes de plantes et également les plus documentées dans la littérature (Collier et Moffett 2009; Eitas et Dangl 2010). Ces protéines de résistance sont impliquées dans de nombreuses interactions plantes-microorganismes faisant intervenir aussi bien des champignons (cas du gène *L* du lin; Lawrence et al. 1995), des bactéries (cas du gène *RPM1* d'*A. thaliana*; Grant et al. 1995) que des virus (cas du gène *N* du tabac; Whitham et al. 1994). Chez les plantes pérennes, l'analyse des familles de gènes codant des NB-LRR a révélé l'expansion d'un nouveau type de NB-LRR possédant un domaine de type doigt de zinc (BED finger) en N-terminal constituant la classe des BED-finger-NB-LRR (BNL) (Kohler et al. 2008; Yang et al. 2009; Duplessis et al. 2009). La cinquième classe de gènes de résistance chez les plantes correspond à des protéines ancrées dans la membrane et possédant un domaine LRR extracellulaire. Les représentants les plus connus de cette classe sont les protéines Cf-2, Cf-4 et Cf-9 de la tomate impliquées dans l'interaction avec les effecteurs Avr2, Avr4 et Avr9 du champignon phytopatogène *C. fulvum* (§ I-2.2.1) (Rivas et Thomas 2005; Stergiopoulos et de Wit 2009). La sixième classe correspond à des protéines ancrées dans la membrane possédant à la fois un domaine LRR extracellulaire et un domaine kinase intracellulaire. Le seul représentant caractérisé correspond à la protéine Xa21, impliquée dans l'interaction entre le riz et la bactérie *X. oryzae* (Song et al. 1997). Une dernière classe regroupe des gènes de résistance sans domaine conservé. Parmi ces gènes, nous retrouvons les gènes codant les protéines Ve1 et Ve2 chez la tomate, impliquées dans l'interaction avec le champignon *Verticillium alboatrum*, ou encore le

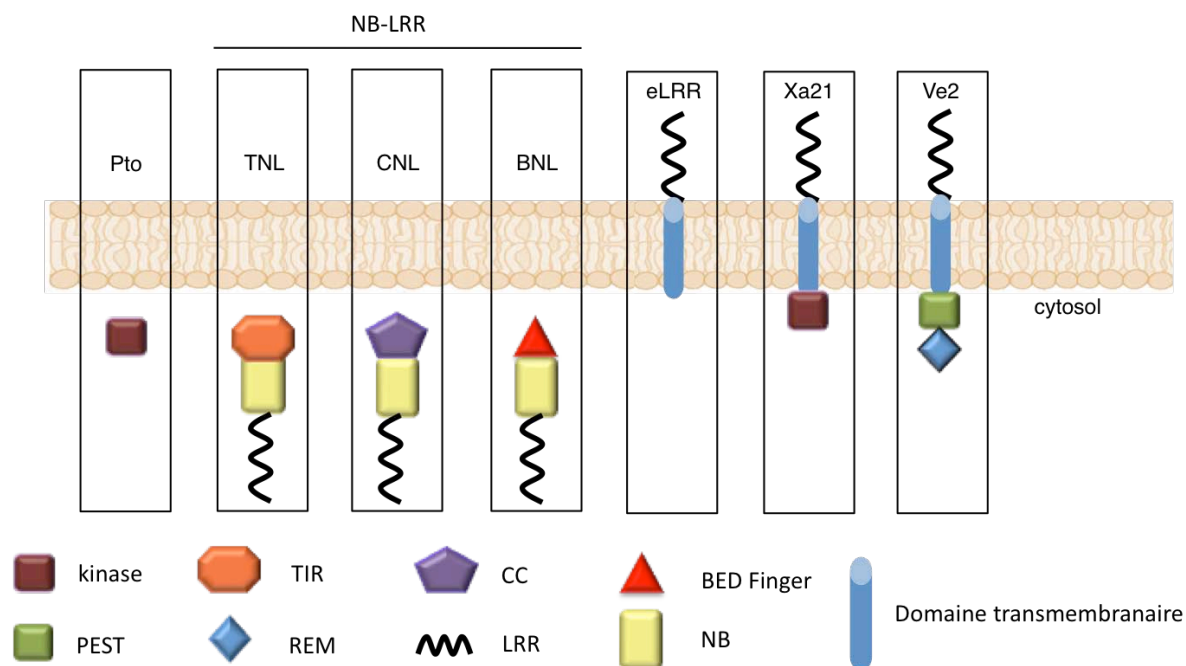


Figure 5: Structure des différentes classes des gènes de résistance R chez les plantes (D'après Dangl et Jones 2001; McDowell et Woffenden 2003; Kohler et al. 2009; Eitas et Dangl 2010). La majorité des protéines de résistance R des plantes possèdent un domaine riche en leucine (LRR) participant à la reconnaissance de l'agent pathogène ainsi qu'un domaine Nucleotide Binding (NB) impliqué dans l'activation de la protéine R. En plus de ces deux domaines intracellulaires, la classe 2 possède un domaine Toll like and Interleukin-1-like Receptor (TIR) impliqué dans la signalisation cellulaire et la reconnaissance, et la classe 3 possède un domaine Coiled-Coil (CC) impliqué dans la signalisation. La classe 4 consiste en des protéines de résistance ancrées dans la membrane et possédant un domaine extracellulaire riche en leucine. Les classes 1 et 5 correspondent à des protéines uniques. La classe 1 correspond à la protéine Pto de la tomate constituée d'un domaine sérine-thréonine kinase intracellulaire, et la classe 5 correspond à la protéine du riz Xa21 constituée d'un domaine riche en leucine extracellulaire et d'un domaine sérine-thréonine kinase intracellulaire reliés entre eux par un domaine transmembranaire. La classe 6 correspond à des protéines de résistance qui possèdent des structures uniques tels que les domaines Pro-Glu-Ser-Tyr (PEST) et Receptor-Mediated Endocytosis (REM) de la protéine Ve2 de la tomate. Le domaine de dégradation de protéine PEST impliqué dans la dégradation des protéines et le domaine REM est impliqué dans l'endocytose couplée à un récepteur.

gène de betterave codant HS1pro-1, une protéine impliquée dans l'interaction avec le nématode *Heterodera schachtii* (Cai et al. 1997; Kawchuck et al. 2001).

4.2 - Perception de l'effecteur par les protéines R

La caractérisation croissante de gènes de résistance et de leurs effecteurs apparentés a permis une meilleure compréhension des mécanismes moléculaires associés à la perception directe et indirecte de l'effecteur par les protéines R de l'hôte conduisant à l'activation de ces protéines de résistance et la mise en place des réactions de défense de la plante (Bent et Mackey 2007).

4.2.1 - Perception indirecte de l'effecteur

La plupart des plantes ont développé des protéines R qui répondent, non pas à l'effecteur en lui-même, mais « gardent » les cibles cellulaires de ces effecteurs et répondent aux altérations induites par ceux-ci (Van der Hoorn 2008). Contrairement au modèle récepteur ligand (§ I-4.2.3), qui permet la reconnaissance d'un panel limité d'effecteurs structurellement proches, ce modèle dit de protéine de garde, permet la détection de multiples effecteurs, de nature diverse, qui interagissent avec la même cible de la plante gardée par une seule protéine R (**Figure 6**). La première démonstration de ce modèle a été mise en évidence entre les effecteurs AvrRpm1 et AvrB de *P. syringae* et la protéine de résistance d'*A. thaliana* RPM1 (Leister et al. 1996, Mackey et al. 2002). Ces effecteurs ne se fixent pas directement à cette protéine de résistance, mais ciblent une seconde protéine associée à la membrane appelée RIN4 (RPM1-interacting protein) qui colocalise physiquement avec RPM1 (Mackey et al. 2002). La protéine de résistance RPM1 est ainsi activée de manière indirecte en réponse aux modifications induites par ces effecteurs sur la protéine RIN4. Chez les champignons, un exemple classique de ce modèle est représenté par l'interaction indirecte entre l'effecteur Avr2 et la protéine Cf-2 de *C. fulvum*, médié par la protéase à cystéine Rcr3 et déclenchant l'induction des mécanismes de défense de la plante (Rooney et al. 2005). Dans de nombreux autres pathosystèmes, des interactions indirectes entre une protéine R et l'effecteur associé ont été rapportées suggérant qu'une majorité d'interactions se dérouleraient par l'intermédiaire de ce modèle de garde (Jones et Dangl 2006).

4.2.2 - Perception directe de l'effecteur

A l'inverse du modèle de garde présenté dans le paragraphe précédent, un second modèle de perception, dit récepteur-ligand, est basé sur la reconnaissance directe entre la protéine de résistance R et l'effecteur correspondant. Chez les champignons, quelques exemples

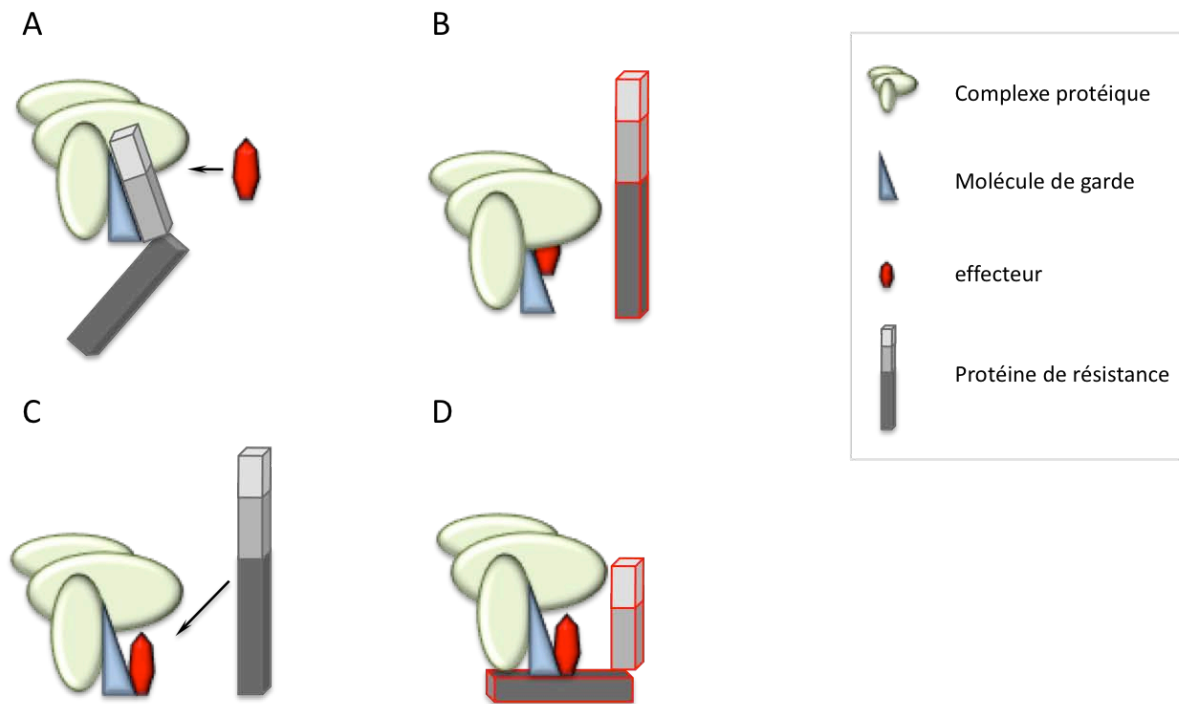


Figure 6: Représentation schématique du modèle de garde (Dangl et Jones 2001). **A**, Un complexe cellulaire protéique (vert) contenant la molécule de garde (bleu) et la protéine de résistance NB-LRR (Du gris clair au gris foncé: domaines N terminal, NB et LRR) est la cible d'un effecteur (rouge). **B**, La fixation de l'effecteur sur sa cible induit une dissociation du complexe, l'activation de la protéine de résistance et la mise en place des mécanismes de défense. **C**, De manière alternative, la protéine de résistance ne pourrait faire partie du complexe qu'après la fixation de l'effecteur. **D**, La fixation de l'effecteur sur le complexe activerait ensuite la protéine de résistance.

d'interactions directes entre un effecteur fongique et une protéine R ont été décrits. Le système de levure double hybride ainsi que des expériences d'accrochage *in vitro* ont ainsi permis de montrer l'interaction physique entre Avr-Pita de *M. oryzae* et la protéine de résistance Pi-ta associée du riz (Jia et al. 2000). Pi-ta code une protéine NB-LRR intracellulaire pour laquelle une seule substitution d'acide aminé au niveau du domaine LRR perturbe l'interaction physique Avr-Pita/Pi-ta et abolit les réponses de défense (Jia et al. 2000). De la même manière, l'interaction physique entre un effecteur de *M. lini* et les protéines de résistance associées chez le lin a également été rapporté (Dodds et al. 2006; Ellis et al. 2007; Wang et al. 2007). Le locus polymorphique d'AvrL567 de *M. lini* consiste en au moins trois gènes (*AvrL567A*, *AvrL567B*, *AvrL567C*) dont les produits sont reconnus par les protéines de résistance (NB-LRR) du lin L5, L6, L7 associées. Des expériences en système double hybride chez la levure ont permis de démontrer l'interaction physique entre AvrL567 et les protéines L5/L6 associées, alors qu'une mutation au niveau du domaine LRR de la protéine L6 abolit l'interaction, suggérant un accrochage de l'effecteur au niveau du domaine LRR (Dodds et al. 2006; Wang et al. 2007).

4.3 - Coévolution entre l'effecteur et la protéine de résistance

Dans un système hôte parasite, les deux espèces exercent l'une sur l'autre des pressions sélectives. Chaque fois que l'hôte s'adapte à un nouveau mécanisme d'attaque en développant une résistance, le parasite est contraint de s'adapter à son tour afin de contourner les systèmes de reconnaissance de son hôte et assurer ainsi son développement et sa reproduction (Burdon et Thrall 2009). Ceci explique le fait que dans la plupart des agrosystèmes dans lesquels des résistances sont développées chez l'hôte, on assiste à un contournement de ces différentes résistances et à l'émergence de nouvelles virulences chez le pathogène. Les effecteurs fongiques ainsi que les protéines de résistance de l'hôte sont en première ligne de cette « course aux armements » (arm race en anglais) puisque de nombreuses études ont rapporté l'évolution rapide des effecteurs ainsi que la forte diversité allélique des loci *R* correspondant de la plante (Allen et al. 2004; Dodds et al. 2006; Ma et Guttman 2008; Soanes et al. 2008). Cette évolution rapide des effecteurs, résultat de mutation, sélection et reproduction, a causé une intense diversification de leurs séquences expliquant ainsi l'absence d'homologie parmi les différentes protéines effectrices identifiées à ce jour. De nombreux gènes codant des effecteurs ont été localisés dans les régions dynamiques des génomes, telles que les régions télomériques ou les régions riches en éléments transposables permettant une grande flexibilité génétique, indispensable au contournement rapide des mécanismes de résistance de la plante induits par la reconnaissance avec les protéines R. Un exemple récent de coévolution entre une famille d'effecteurs (*Avrk1*) et une classe particulière de rétrotransposons LINE-1 (nommée TE1a) a été

mis à jour à travers l'analyse de la séquence génomique de *B. graminis* (Sacristan et al. 2009). Cette coévolution suggère un bénéfice mutuel permettant au champignon d'amplifier et de diversifier ces effecteurs pour échapper à la reconnaissance de l'hôte tout en maintenant la présence de ces transposons au sein du génome (Sacristan et al. 2009). Dans de nombreux cas, le gain de virulence dû à l'insertion de ces transposons, aux délétions de gène ou à d'autres réarrangements chromosomiques a été fréquemment observé (Orbach et al. 2000; Farmann et al. 2002; Fudal et al. 2007; Gout et al. 2007; Khang et al. 2008). Très récemment, l'analyse comparative de trois espèces phénotypiquement différentes de *Fusarium* a révélé la présence de chromosomes spécifiques à *F. oxysporum* f. sp. *lycopersici* riches en éléments transposables. De manière intéressante, le transfert de l'un de ces chromosomes au sein d'une souche avirulente de *F. oxysporum* a conduit à l'émergence d'une souche virulente (Ma et al. 2010).

5 - Identification de nouveaux effecteurs chez les oomycètes et les champignons

La grande majorité des effecteurs (gènes Avr) d'organismes phytopathogènes mis à jour au cours des 2 dernières décennies ont été identifiés par cartographie génétique, basée sur leur phénotype d'avirulence (Ellis et al. 2007; Gout et al. 2006; Fudal et al. 2007; van Poppel et al. 2008). Ce n'est que récemment avec la disponibilité croissante de séquences de génomes de champignons et d'oomycètes, associée à leurs transcriptomes, qu'il est devenu possible de prédire les gènes codant des effecteurs candidats en ciblant le sécrétome (ensemble des protéines sécrétées) des pathogènes (Kämper et al. 2006; Martin et al. 2008; Mosquera et al. 2009). L'identification à une plus large échelle des effecteurs parmi ces protéines sécrétées va ainsi reposer sur des caractéristiques de séquence ou de fonctions retrouvées chez les effecteurs identifiés jusqu'à présent (Rep et al. 2005; Ellis et al. 2009; Rouxel et al. 2009).

5.1 - Taille des effecteurs et signal de sécrétion

Les effecteurs décrits jusqu'à présent chez les oomycètes ainsi que chez les champignons sont de petite taille (< 300 acides aminés, voir **Tableau 1**) et possèdent un signal de sécrétion à leur extrémité N-terminal permettant leur adressage au système membranaire endocellulaire. Néanmoins il existe des exceptions puisque certains effecteurs ne possèdent pas ce signal de sécrétion tels que les effecteurs Avr-a10 et Avr-k1 de *B. graminis* suggérant un mécanisme alternatif de sécrétion (Ridout et al. 2006). Du fait de la richesse en acides aminés hydrophobes (Phe, Leu, Ile, Met, Val) de ce signal, permettant son insertion dans la membrane du reticulum endoplasmique, de nombreux programmes bioinformatiques sont capables de prédire leur présence et de définir le site de clivage au niveau duquel la protéine sera libérée dans l'espace

intracellulaire (Bendtsen et al. 2004). Afin d'améliorer et d'affiner la prédiction bioinformatique du sécrétome d'un organisme, la combinaison de l'ensemble de ces programmes est généralement mise en œuvre (Choi et al. 2010).

5.2 - Richesse en cystéine des effecteurs

De nombreux effecteurs d'oomycètes et de champignons présentent un nombre important de cystéines au sein de leurs séquences protéiques. Ces cystéines pourraient être impliquées dans la formation de ponts disulfures et permettraient de stabiliser la structure tertiaire de la protéine (Van't Slot et al. 2005; Van den Burg et al. 2003; Vervoort et al. 1997; Van den Hooven et al. 2001). Par exemple, la sensibilité aux protéases des mutant Avr4 de *C. fulvum* mutés au niveau de deux cystéines (Van den Burg et al. 2003) ou encore la présence indispensable de deux cystéines à l'activité biologique de ToxA, une toxine produite par le champignon *Pyrenophora tritici-repentis* (Tuori et al. 2000), confirment cette hypothèse. Néanmoins, cette caractéristique n'est pas une spécificité des effecteurs puisque certains possèdent très peu de cystéines (< à 2) au sein de leurs séquences tels que AvrL567 et AvrM de *M. lini*, Pwl1, Pwl2 et Pwl3 de *M. grisea* ou encore AvrLm1 de *L. maculans* (Stergiopoulos et de Witt 2009).

5.3 - Présence de motifs conservés chez les effecteurs

La grande majorité des effecteurs découverts jusqu'à présent possèdent leurs propres caractéristiques de séquence hormis la présence d'un signal de sécrétion dans leur partie N-terminal (Ellis et al. 2009). Néanmoins certains possèdent des motifs caractéristiques qui permettent de les identifier tel que le motif RxLR, identifié chez les oomycètes et impliqué dans l'export de l'effecteur dans la cellule hôte (§ I-3.2). La présence de ce motif conservé a notamment permis de dresser le catalogue d'effecteurs candidats au sein des séquences génomiques de plusieurs oomycètes pathogènes. (Tyler et al. 2006; Win et al. 2007; Gaulin et al. 2008; Jiang et al. 2008; Haas et al. 2009). Ainsi approximativement 700 effecteurs candidats portant ce motif ont été identifiés chez deux espèces de *Phytophthora* (Jiang et al. 2008) et 150 ont été mis à jour chez *H. arabidopsis* (Win et al. 2007) suggérant l'abondance de ce motif et la grande diversité de ces effecteurs candidats chez les oomycètes. Comme mentionné précédemment, un motif consensus RxLR n'a pas été formellement identifié chez les effecteurs de champignons, malgré l'identification de motifs présentant des structures dégénérées proches (Rafiqi et al. 2010; Kale et al. 2010). Ces études suggèrent que de nombreux motifs proches du motif RxLR sont présents chez les champignons et permettraient également l'entrée de

Tableau 3: Motifs identifiés chez les effecteurs d'oomycètes et de champignons (Oliva et al. 2010; Rafiqi et al. 2010; Godfrey et al. 2010; Kale et al. 2010).

Motif	Protéines	Pathogènes	Fonction	Références
RxLR (quelques fois RxLR-dEER)	Effecteurs RxLR	Phytophthora spp. et Hyaloperonospora arabidopsidis	Translocation dans la cellule hôte	Whisson et al. 2007; Dou et al. 2008
KFLK	AvrM	Melampsora lini	Translocation dans la cellule hôte	Rafiqi et al. 2010
REYR	AvrL567	Melampsora lini	Translocation dans la cellule hôte	Rafiqi et al. 2010
RIYER	Avr2	Fusarium oxysporum f. sp. lycopersici	Translocation dans la cellule hôte	Kale et al. 2010
RWWT	AvrLm6	Leptosphaeria maculans	Translocation dans la cellule hôte	Kale et al. 2010
[R/H/K]X[L/M/I/F/Y/W]	Effecteurs champignons (RxLR-like)	Champignons	Translocation dans la cellule hôte	Kale et al. 2010
LxFLAK	Effecteurs Crinkler (CRN); CRN1, CRN2, CRN8 ...	Phytophthora spp.	Translocation dans la cellule hôte	Win et al. 2007; Haas et al. 2009; Shornack et al. 2010
HVLVxxP	Effecteurs Crinkler (CRN); CRN1, CRN2, CRN8 ...	Phytophthora spp.	Inconnue	Haas et al. 2009
[F/L]xLxLxLxK	Protéines CRN-like: AeCRN5, AeCRN13	Aphanomyces euteiches	Translocation dans la cellule hôte	Gaulin et al. 2008; Shornack et al. 2010
LxLAR/K	Protéines CRN-like: PuCRN3, PuCRN10	Pythium ultimum	Inconnue	Cheung et al. 2008; Lévesque et al. 2010
W-Box, L-Box, Y-Box	Superfamille des effecteurs RxLR: IPO1, AVR4 ...	Phytophthora spp.	Requis pour l'activité d' avirulence	Jiang et al. 2008; van Poppei et al. 2008
KECD	Famille de toxines PGF	Phytophthora spp.	Inconnue	Nicastro et al. 2009
[V/F/Y]XC	Effecteurs putatifs exprimés au niveau de l'haustorium	Didium et rouilles	Inconnue	Godfrey et al. 2010
[R/K]WYL/IJR	Famille AvrK1	Blumeria graminis f.sp. Hordei	Inconnue	Ridout et al. 2006
GI[F/W]IA[LS/7]R	AvrP4, AvrP123, AvrL567, AvrM	Melampsora lini	Pas d'effet sur la Translocation ou l'activité d'avirulence	Catanzariti et al. 2006
ROHKKRx _n HRHK	RIP1p	Uromyces viciae-fabae	Localisation nucléaire	Kemen et al. 2005

l'effecteur dans la cellule de l'hôte. Le motif consensus R/H/KxL/M/I/F/Y/W a ainsi pu être établi à partir des motifs les plus redondants identifiés chez les champignons et ouvre ainsi de nouvelles perspectives dans l'identification à grande échelle de nouveaux effecteurs chez les champignons (Kale et al. 2010). D'autres motifs ont également été identifiés chez les effecteurs de pathogènes filamenteux (Oliva et al. 2010; **Tableau 3**). Parmi les mieux caractérisés on retrouve les motifs LxLFLAK et HVLVxxP, présents chez les effecteurs CRN (Crinkler effectors) de *Phytophthora* spp. (Win et al. 2007; Haas et al. 2009) ainsi que les motifs similaires F/LxLYLALK et LxLYLAR/K, retrouvés respectivement chez *A. euteiches* (Gaulin et al. 2008) et chez *Pythium ultimum*, un autre oomycète nécrotrophe s'attaquant à un large spectre d'hôte (Lévesque et al. 2010). Tout comme les motifs RxLR, ces motifs sont essentiels à l'export des effecteurs CRN à l'intérieur la cellule hôte (Schornack et al. 2010). Récemment, chez *B. graminis* f.sp. *hordei*, un autre motif Y/F/WxC a été identifié dans la partie N terminal de la séquence mature d'effecteurs candidats. Ce motif a également été identifié chez certaines protéines sécrétées de *Puccinia graminis*, l'agent responsable de la rouille du blé formant également des haustoria, suggérant le rôle potentiel de ce motif au niveau de l'haustorium (Godfrey et al. 2010). D'autres motifs ont également été identifiés chez les pathogènes filamenteux et notamment chez les agents de rouilles tel que le domaine d'adressage nucléaire uniquement identifié chez la protéine RTP1 d'*U. fabae* (Kemen et al. 2005)

5.4 - Variabilité nucléotidique des effecteurs

En accord avec la théorie de la « course aux armements » entre les acteurs du système immunitaire du côté de la plante et les effecteurs du côté du pathogène, les séquences nucléotidiques codant pour ces effecteurs présentent, dans de nombreux cas, des évidences de niveaux d'évolution accélérés (Dodds et al. 2006; Guttman et al. 2006; Win et al. 2007; Brunner et al. 2009). Cette pression de sélection positive (aussi appelée sélection diversifiante) semble, dans de nombreux cas, essentielle au contournement des mécanismes de défense de l'hôte et pourrait être un indicateur de l'importance fonctionnelle de certains loci qui contribuent à la survie de l'organisme. De nombreux effecteurs, présentant des niveaux extrêmes de polymorphisme au niveau de leur séquence d'acides aminés sont identifiés comme étant sous sélection positive (Stergiopoulos et de Witt 2009). En effet, certaines protéines effectrices d'oomycètes telles que ATR13 et scr74 (Allen et al. 2004; Liu et al. 2005) ainsi que celles de champignons telles que AvrM, AvrP123, AvrP4 et AvrL567 identifiées chez *M. lini* par exemple sont sous pression de sélection positive (Dodds et al. 2006; Catanzariti et al. 2006; Barrett et al. 2009; Van der Merwe et al. 2009). Du fait de la disponibilité croissante des séquences génomiques de nouvelles espèces de pathogènes, des analyses plus globales de génomique

comparative, basées sur l'identification des petites protéines sécrétées sous pression de sélection positive vont accélérer l'identification de nouveaux effecteurs candidats (Liu et al. 2005; Kamoun et al. 2006).

5.5 - Expression des gènes codant pour ces effecteurs

L'une des autres particularités des effecteurs identifiés à ce jour concerne le niveau d'expression de leurs transcrits lors du processus infectieux. En effet, la majorité des effecteurs présentent des niveaux de transcrits importants au cours de l'interaction avec l'hôte et faible voire nul dans le mycélium sur milieu synthétique ou dans les spores (Ellis et al. 2009). Ainsi, l'analyse détaillée des transcriptomes d'organismes phytopathogènes a permis de dresser la liste des transcrits spécifiquement exprimés durant le processus infectieux et pouvant coder des effecteurs potentiels. Chez *U. maydis*, un large répertoire de 554 protéines sécrétées, la plupart regroupées en clusters de gènes au sein du génome, a été prédit bioinformatiquement à partir de la séquence du génome du champignon. Des analyses d'expression ont montré que la plupart des gènes présents dans ces clusters sont induits dans les tissus infectés et la délétion de ces clusters de gènes a indiqué que la plupart étaient essentiels pour la virulence du champignon et sa prolifération dans les tissus foliaires (Kämper et al. 2006). Récemment, l'analyse du transcriptome de champignons pathogènes et mycorrhiziens a conduit à l'identification de transcrits fortement accumulés lors de l'interaction avec leurs hôtes correspondant à des effecteurs potentiels. Ainsi, *M. oryzae* présente 59 gènes fortement induits (>10 fois) durant l'infection qui codent des protéines sécrétées (Mosquera et al. 2009). De manière intéressante, cette étude a révélé la présence de facteurs d'avirulence parmi ces gènes et a permis l'identification de nouveaux gènes candidats spécifiquement associés à la biotrophie (BAS, Biotrophy-Associated Secreted Proteins) et dont les produits sont préférentiellement accumulés au niveau du BIC (Mosquera et al. 2009; Khang et al. 2010). Dans le cadre de l'interaction entre *Populus trichocarpa* et *Laccaria bicolor*, un champignon symbiotique ectomycorhizien, l'analyse du transcriptome du champignon a conduit à l'identification de nombreux gènes induits durant la formation de l'ectomycorhize codant de petites protéines sécrétées (MISSPs, Mycorrhiza-Induced Small Secreted Proteins) (Martin et al. 2008). Le plus induit d'entre eux code une petite protéine de 7 kDa nommée MISSP7 qui est fortement accumulée au niveau de l'interface entre la plante et le champignon et qui ciblerait le noyau de la cellule hôte. L'absence de mycorrhization observée chez les souches de champignon mutant suggère le rôle clé de cet effecteur dans l'établissement de la symbiose ectomycorhizienne (Plett et al., non publié).

II - Les *Pucciniales*

1 - Généralités

Les champignons appartenant à l'ordre des *Pucciniales* (anciennement *Urédinales*) sont couramment appelés agents de rouilles. Tous les membres appartenant à cet ordre sont des parasites des plantes et causent des pertes économiques importantes, notamment dans le domaine de la céréaliculture (Alexopoulos et al. 1996). Avec les agents d'oïdiums et certains oomycètes, les agents de rouilles sont des parasites biotrophes obligatoires dépendant de leur hôte pour se développer et compléter leur cycle de vie. D'après Voegelé et collaborateurs (2009), ces agents pathogènes biotrophes obligatoires se caractérisent par:

- l'impossibilité de les cultiver *in vitro*
- la formation de structures d'infection très différenciées
- leur activité de sécrétion limitée
- l'établissement d'une zone de contact étroite séparant la membrane du champignon de celle de la plante
- l'induction d'une suppression des défenses de l'hôte à long terme
- la formation de structures d'infection intracellulaires, les haustoria (§ II-5)

La récente émergence de la nouvelle race UG99 de rouille noire du blé (causée par *P. graminis* f. sp. *tritici*) en Afrique ainsi que la rapide dispersion de la rouille asiatique du soja (causée par *Phakopsora pachyrhizi*) dans le continent américain illustrent le danger de ces agents de rouilles pour l'agriculture. En effet, la plupart des cultivars de blé plantés actuellement sont sensibles à la race UG99 de *P. graminis* qui a été détectée en 1999 en Ouganda (Pretorius et al. 2000). Cette nouvelle race se disperse rapidement dans l'Est du continent africain et au Moyen-Orient et constitue une nouvelle menace sur la production de blé partout où le gène de résistance *Sr31* a été introduit (Wanyera et al. 2006). Néanmoins, certaines variétés de blé résistantes sont actuellement développées afin de remplacer les variétés sensibles dans l'est de l'Afrique ainsi qu'en Asie. De manière similaire, la rapide dispersion de la rouille asiatique du soja à travers le continent américain est l'exemple typique de l'introduction d'un « vieil ennemi » dans une région jusqu'alors exempte du pathogène. Ces dernières années, la maladie a rapidement progressé depuis l'Asie vers Hawaï où elle a été détectée en 1994 puis a gagné l'Afrique en 1996, l'Amérique du Sud en 2001 pour ensuite se disperser sur le continent américain (Pivona et Yang 2004). Actuellement, seules les conditions climatiques peuvent ralentir cette épidémie de rouille dans les plantations de soja aux Etats-Unis. De nombreux autres agents de rouilles font

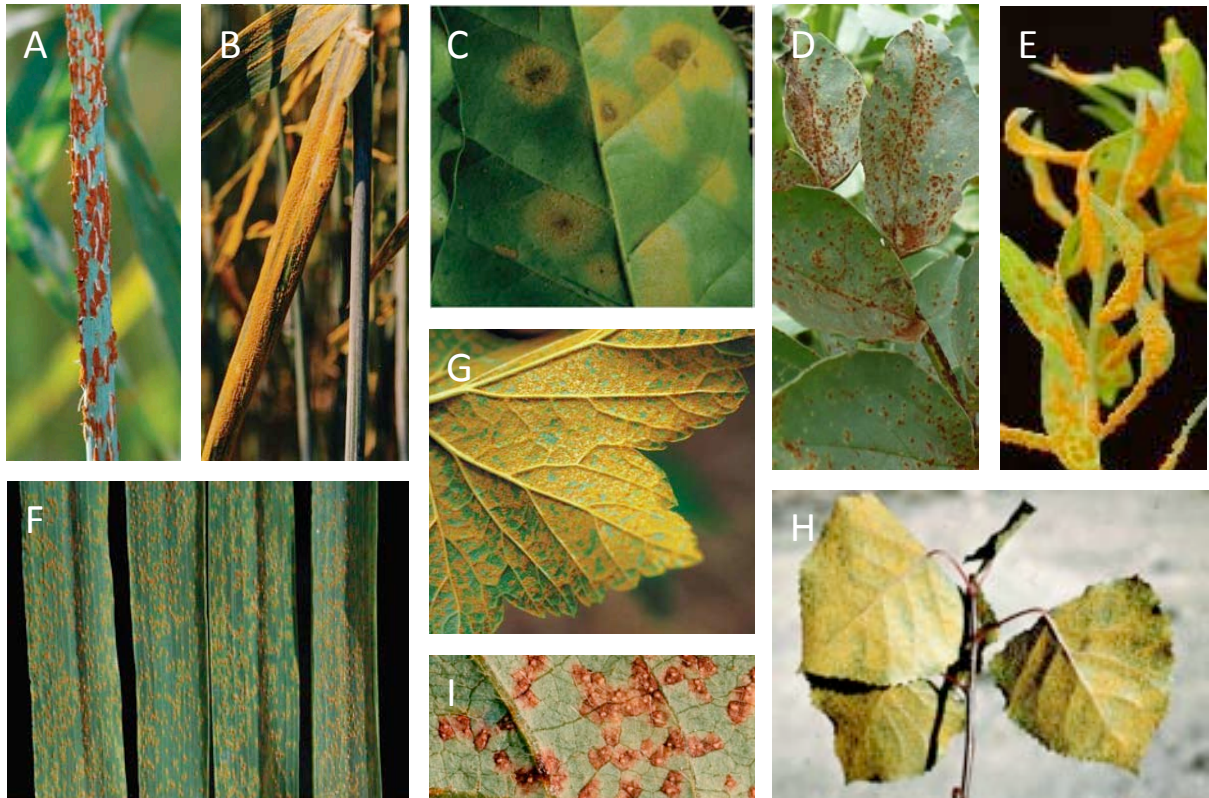


Figure 7: Exemples de symptômes causés par différentes rouilles. **A**, Rouille noire du blé causée par *Puccinia graminis* f.sp. *tritici* (photographie de Yue Jin). **B**, Rouille brune du blé causée par *Puccinia triticina* (photographie de James Kolmer). **C**, Rouille orangée du caféier causée par *Hemileia vastatrix* (photographie de Howard Schwartz). **D**, Rouille causée par *Uromyces viciaefabae* sur haricot (photographie DRAF SRPV Lorraine). **E**, Rouille du lin causée par *Melampsora lini* (photographie du CSIRO, Australie). **F**, Rouille jaune du blé causée par *Puccinia striiformis* f. sp. *tritici* (photographie de James Kolmer). **G**, Rouille vésiculeuse causée par *Cronartium ribicola* sur groseillier (photographie de Robert Anderson). **H**, Rouille foliaire du peuplier causée par *Melampsora larici-populina* (photographie Jean Pinon). **I**, Rouille asiatique du soja causée par *Phakopsora pachyrhizi* (photographie de l'Agricultural Research Service USDA)

actuellement l'objet de recherches approfondies afin de mettre en place des moyens efficaces pour lutter contre ces maladies. On retrouve notamment les agents de la rouille brune (*Puccinia triticina*) et la rouille jaune (*P. striiformis* f. sp. *tritici*) du blé, de la rouille du lin (*M. lini*), de la rouille de la fève (*U. fabae*) et du haricot (*U. appendiculatus*). Chez les plantes pérennes, d'autres agents de rouilles font également l'objet de recherche tels que les agents responsables de la rouille du peuplier (*M. larici-populina*), du chêne (*Cronartium quercuum*), du pin (*Cronartium ribicola*) ou encore du caféier (*Hemileia vastatrix*). Les principaux symptômes sur plant de ces rouilles sont illustrés dans la **figure 7**.

2 - Phylogénie et taxonomie

Les agents de rouilles appartiennent au phylum des basidiomycètes puisqu'ils produisent des basidiospores haploïdes à partir de leurs basides. Traditionnellement, les agents de rouilles ont été regroupés avec les agents du charbon dans la classe des *Teliomycètes* (Jülich 1981). Cependant, de récentes données moléculaires, basées sur le séquençage de la large sous unité de l'ADN ribosomique, ont permis de distinguer 3 classes au sein du phylum des basidiomycètes: les *Pucciniomycotina* (incluant les agents de rouilles), les *Ustilaginomycotina* et les *Agaricomycotina* (Bauer et al. 2006; Hibbett et al. 2007). Au sein des *Pucciniomycotina*, 8 classes peuvent être distinguées parmi lesquelles on retrouve la classe des *Pucciniomycètes* (**Figure 8**). Au sein de cette classe, l'ordre des *Pucciniales* comporte plus de 100 genres et environ 7000 espèces (Maier et al. 2003), représentant plus de 95% des espèces de la classe des *Pucciniomycètes*.

La classification des agents de rouilles par famille a été traditionnellement basée sur la morphologie des télies et des télisporos (Voir § II-3) et a conduit à l'identification de deux grandes familles définies par Dietel (Dietel 1928; cité par Aime 2006): Les *Pucciniaceae* et les *Melamporaceae*. Néanmoins, certains caractères, basés sur d'autres critères que la morphologie des télies/télisporos ont été identifiés comme importants dans l'interprétation phylogénétique des familles tels que l'arrangement des pores des urédosporos, la morphologie de l'urédie, des écidies ou des spermogonies (Hiratsuka and Cummins 1963; Hiratsuka and Hiratsuka 1980). Avec ces nouveaux critères morphologiques, Cummins et Hiratsuka (2003) ont ainsi proposé la division de l'ordre en 13 familles. Plus récemment, Aime (2006) a examiné des espèces représentatives des 13 familles proposées par Cummins et Hiratsuka (2003) par séquençage des régions 18S et 28S de l'ADN ribosomique. Les 13 familles ont ainsi été placées dans 3 sous ordres : le sous ordre des *Uredinineae*, celui des *Melamporineae* and celui des *Mikronegeriineae*.

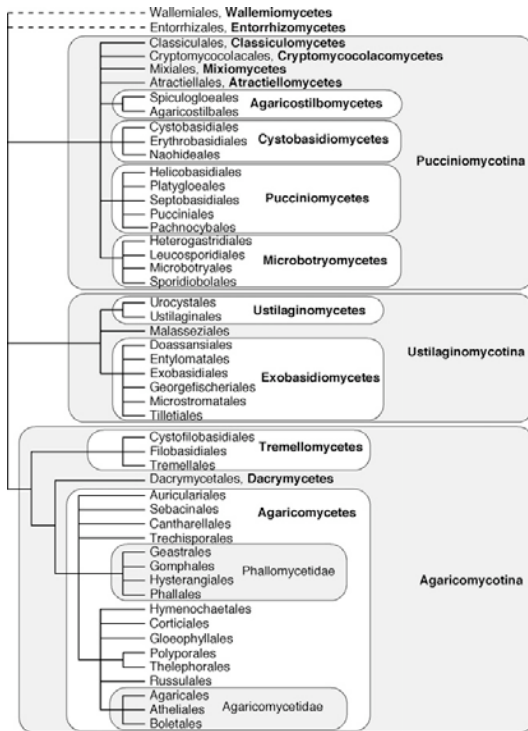


Figure 8: Phylogénie et classification des Basidiomycètes (Hibbett et al. 2007). Au sein de ce phylum, 3 sous phyla se distinguent; les *Pucciniomycotina*, les *Ustilaginomycotina* et les *Agaricomycotina*. Les *Pucciniomycetes* correspondent à l'une des huit classes appartenant au sous phylum *Pucciniomycotina*. Au sein de cette classe, l'ordre des *Pucciniales* représente plus de 95% des espèces.

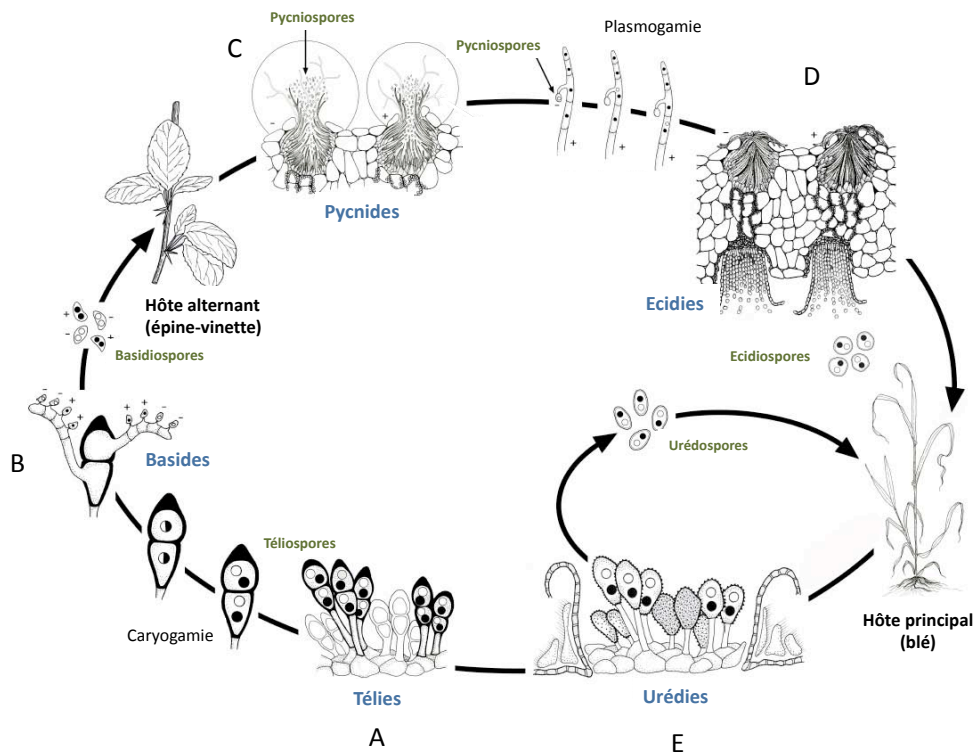


Figure 9: Cycle de vie d'une rouille macrocyclique hétéroïque (*Puccinia graminis* f. sp. *tritici*) dont l'hôte principal est le blé et l'hôte alternant l'épine vinette. A, Télios et télisporangies. B, Basides et basidiosporangies. C, Pycnides et pycniosporangies (noms synonymes: spermogonies et spermaties). D, Ecidies et ecidiosporangies. E, Urédies et urédiosporangies (Adapté de Leonard et Szabo 2005)

Cette étude a ainsi pu montrer que les caractéristiques morphologiques ne sont pas de bons indicateurs pour la discrimination des familles au sein de l'ordre des *Pucciniales*.

3 - Cycle de vie

L'une des caractéristiques observées chez la plupart des agents de rouilles est la complexité de leur cycle biologique, faisant intervenir jusqu'à cinq types de spores distincts et parfois deux hôtes différents. Le cycle de vie est dit macrocyclique lorsqu'il comporte cinq types de spores distincts, et hétéroïque lorsque deux hôtes sont requis pour compléter le cycle biologique du pathogène. C'est par exemple le cas des agents responsables de la rouille du blé *P. graminis* et de la rouille du peuplier *M. larici-populina*. Ce cycle de vie complexe, présenté dans la **figure 9** (Leonard et Szabo 2005), débute au printemps par la germination des téliospores (**Figure 9a**). Celles-ci ont été produites la saison précédente et ont hiverné sur les tissus morts de l'hôte au sein de pustules noirâtres appelées télies. Les téliospores diploïdes germent sur place et subissent la méiose pour produire au moins quatre basidiospores haploïdes qui sont disséminés dans l'air (**Figure 9b**). Chez les agents de rouilles hétéroïques, les basidiospores doivent rencontrer le second hôte (hôte alternant ou hôte écidien) afin de boucler le cycle biologique du pathogène. Ces basidiospores germent et envahissent les tissus de l'hôte pour produire des colonies haploïdes appelées pycnides. Ces pycnides peuvent être de deux types sexuels et produisent des pycniospores qui vont fusionner par plasmogamie pour former une cellule contenant 2 noyaux haploïdes distincts (**Figure 9c**). Une fois cette fusion réalisée, le champignon différencie de nouvelles structures appelées écidies (**Figure 9d**). Les écidiospores libérées à partir de cette structure sont produites en grande quantité et peuvent parcourir de longues distances afin d'atteindre l'hôte principal (ou hôte télieu) (Nagarajan et Singh 1990). Une fois en contact avec les tissus de l'hôte, ces écidiospores germent rapidement et envahissent les tissus foliaires par l'intermédiaire des stomates. Il en résulte la formation d'urédies à partir desquelles des quantités massives d'urédospores sont libérées (**Figure 9e**). Ce stade de reproduction asexuée, caractérisé par plusieurs cycles successifs d'infection sur le même hôte durant l'été, est le plus étudié par les phytopathologistes car il peut conduire à de sévères épidémies (Voegelé et al. 2009). Néanmoins, de nombreux symptômes peuvent également être observés au cours de la reproduction sexuée, particulièrement lors du stade écidien. Ceux-ci sont particulièrement dommageables dans le cadre de l'interaction entre *C. ribicola* et le pin par exemple, causant des chancres à l'origine du dépérissement de nombreux pins blancs.

Bien que ce cycle macrocyclique hétéroïque soit commun, de nombreux agents de rouilles ont un cycle de vie moins complexe, la plupart du temps mis en place du fait d'une haute latitude, d'une

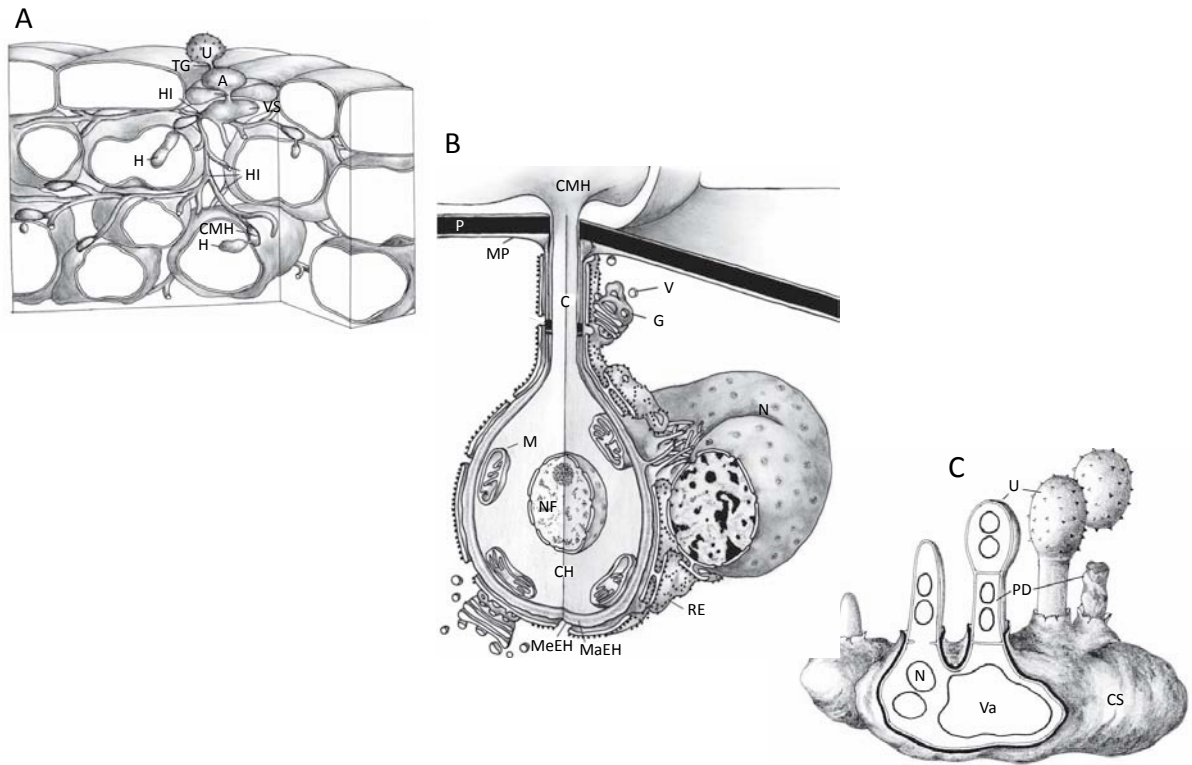


Figure 10: Illustration des structures d'infection d'une rouille de céréales (Kolmer et al. 2009). **A**, Structures d'infection mises en place au cours du stade urédien à 60 heures post inoculation (hpi). U, urediniospore; TG, tube germinatif; A, appressorium, VS, vésicule sous-stomatique; HI, hyphes d'infection; CMH, cellule mère haustoriale; H, haustorium. **B**, Section d'une cellule hôte infectée au niveau du site de pénétration montrant la structure tridimensionnelle d'un haustorium mature. CMH, cellule mère haustoriale; C, « col » de l'haustorium; CH, corps de l'haustorium; NF, noyau fongique; M, mitochondrie; MaEH, matrice extra-haustoriale; MeEH, membrane extra-haustoriale; RE, reticulum endoplasmique de l'hôte; G, appareil de Golgi de l'hôte; N, noyau de l'hôte; V, vésicule. **C**, formation des urédospores en fin du stade urédien. N, noyau; Va, vacuole; CS, cellule sporifère; PD, pédicelle; U, urédospores.

haute altitude ou de conditions désertiques (Savile 1953; cité par Kolmer et al. 2009). Certains agents de rouilles, dits autoïques, sont capables de compléter leur cycle de vie sur un seul hôte. Parmi les rouilles macrocycliques autoïques, on retrouve notamment l'agent de la rouille du lin *M. lini* ou encore celui de rouille de la féverole *U. fabae*. Chez les agents de rouilles appelés hémicycliques, les stades urédien et tellien sont présents alors que chez les agents des rouilles microcycliques, seuls les stades pynidien et telien sont présents. Le stade écidien étant absent chez ces derniers, tous les agents de rouilles microcycliques sont nécessairement autoïques. De manière contrastée, les agents de rouilles hémicycliques et macrocycliques peuvent être autoïques ou hétéroïques (Shattock et Preece 2000; Voegelé et al. 2009).

4 - Processus d'infection au cours de la phase urédienne

Au cours de la phase urédienne (§ II-3), le processus d'infection a particulièrement bien été étudié chez *P. graminis* (Staples et Macko 1984), *U. fabae* (Mendgen et Hahn 2002) et *M. laricipopulina* (Laurans et Pilate 1999; Rinaldi et al. 2007) (**Figure 10**). Une fois en contact avec la fine couche d'eau présente sur la face inférieure des feuilles, les urédospores se gonflent et produisent un tube germinatif dont l'orientation est contrôlée par la topographie de la surface de la feuille (Wynn et Staples 1981). Une fois en contact avec un stomate, la croissance du tube germinatif s'arrête et celui-ci se différencie la plupart du temps en une structure appelée appressorium, permettant la pénétration de l'agent pathogène au sein des tissus foliaires. En quelques heures, le champignon développe une vésicule sous-stomatique dans le parenchyme lacuneux sous l'épiderme, à partir de laquelle se différencie un hyphe d'infection primaire. Celui-ci se propage entre les cellules du mésophylle et au contact de celles-ci, l'extrémité de l'hyphe se cloisonne par un septum, aboutissant à la formation d'une cellule-mère haustoriale. Un hyphe de pénétration est formé à partir de cette cellule mère haustoriale, permettant la formation d'une structure d'infection intracellulaire: l'haustorium (§ II-5). Cette structure, invaginée dans le plasmalemme de la cellule végétale facilite le détournement des métabolites de la plante et la sécrétion de molécules signal (Staples 2001; Mims 2002; Voegelé et Mendgen 2003; Catanzariti et al. 2007). Ces échanges sont facilités par la présence d'une membrane, dérivée du plasmalemme de la cellule hôte appelée membrane extrahaustoriale (MeEH) qui est séparée de l'haustorium par la matrice extrahaustoriale (MaEH) (Kolmer et al. 2009). La différenciation croissante des haustoria dans les premières heures ou premiers jours après l'infection va permettre une croissance rapide du champignon, envahissant ainsi les espaces intercellulaires des parenchymes lacuneux et palissadiques. Après une semaine ou plus suivant les espèces considérées, le mycélium forme des cellules sporifères au sein de l'espace intercellulaire situé sous l'épiderme inférieur de la feuille à partir desquelles vont se différencier les urédospores. La

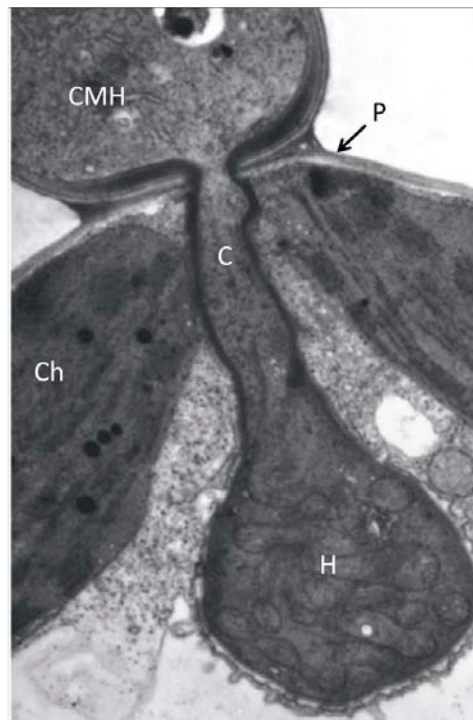


Figure 11: Visualisation d'un haustorium au cours de l'interaction compatible entre *Melampsora lini* et le lin par microscopie électronique à transmission (Ellis et al. 2007). CMH: cellule mère haustoriale, H: haustorium; C: « col » de l'haustorium; Ch: chloroplaste; P: paroi cellulaire de la plante.

pression exercée par ces spores nouvellement formées et par les paraphyses (filaments stériles) sur l'épiderme de la feuille aboutit à son déchirement et à la libération des spores qui peuvent ensuite être disséminées par le vent sur de larges distances afin d'infecter de nouveau leur hôte (Nagarajan et Singh 1990; Barrès et al. 2008).

5 - L'haustorium

L'haustorium représente la structure d'infection caractéristique des organismes biotrophes obligatoires (**Figure 11**). La plupart des approches moléculaires conduites actuellement chez les rouilles se focalisent sur cette structure caractéristique afin de comprendre sa fonction au cours de l'interaction entre le champignon pathogène et sa plante hôte. Cependant ces approches sont difficiles à développer du fait du caractère biotrophe des rouilles et de la formation spécifique des haustoria dans les cellules vivantes de l'hôte. Durant les années 90, des méthodes ont été mises en place afin d'isoler des haustoria à partir des tissus de plante infectée (Hahn et Mendgen 1992; Tiburzy et al. 1992; Cantrill et Deverall 1993). La méthode qui a été la plus utilisée pour cibler cette structure fongique a été développée par Hahn et Mendgen en 1992. Elle est basée sur l'accrochage des oligosaccharides présents à la surface de la paroi de l'haustorium à la lectine concanavaline A, elle-même immobilisée sur une colonne de sepharose 6B. Cette méthode permet d'isoler un nombre suffisant d'haustoria pour réaliser l'extraction des acides nucléiques et ainsi identifier des gènes préférentiellement ou spécifiquement exprimés au niveau de cette structure. Cette méthode a ainsi été appliquée avec succès chez de nombreux agents de rouilles tels que *U. fabae* (Hahn et Mendgen 1997a), *M. lin* (Catanzariti et al. 2006), *M. larici-populina* (Feau et al. 2007; Joly et al. 2010), *U. appendiculatus* (Puthoff et al. 2008) ou encore *P. striiformis* f. sp. *tritici* (Yin et al. 2009).

5.1 - Implication de l'haustorium dans la biosynthèse de métabolites

Chez *U. fabae*, cette méthode de purification des haustoria a permis le séquençage d'étiquettes de séquences exprimées (ESTs) à partir d'une banque d'ADNc de ces structures et a conduit à l'identification de gènes fortement exprimés *in planta*, nommés *PIGs* (*in Planta* Induced Genes) (Hahn et Mendgen 1997a). Parmi les transcrits les plus fortement représentés dans la banque d'ADNc, *THI1* et *THI2* codent des enzymes impliquées dans la synthèse de la vitamine B1 (Hahn et Mendgen 1997a). Les transcrits de ces deux gènes représentent 5% du nombre total de transcrits détectés dans la banque suggérant un rôle déterminant dans le métabolisme de l'haustorium. L'accumulation des transcrits de *THI1* et *THI2* dans les haustoria et les cellules infectées a pu être confirmée par Northern blot et l'immunolocalisation de la protéine *THI1* a

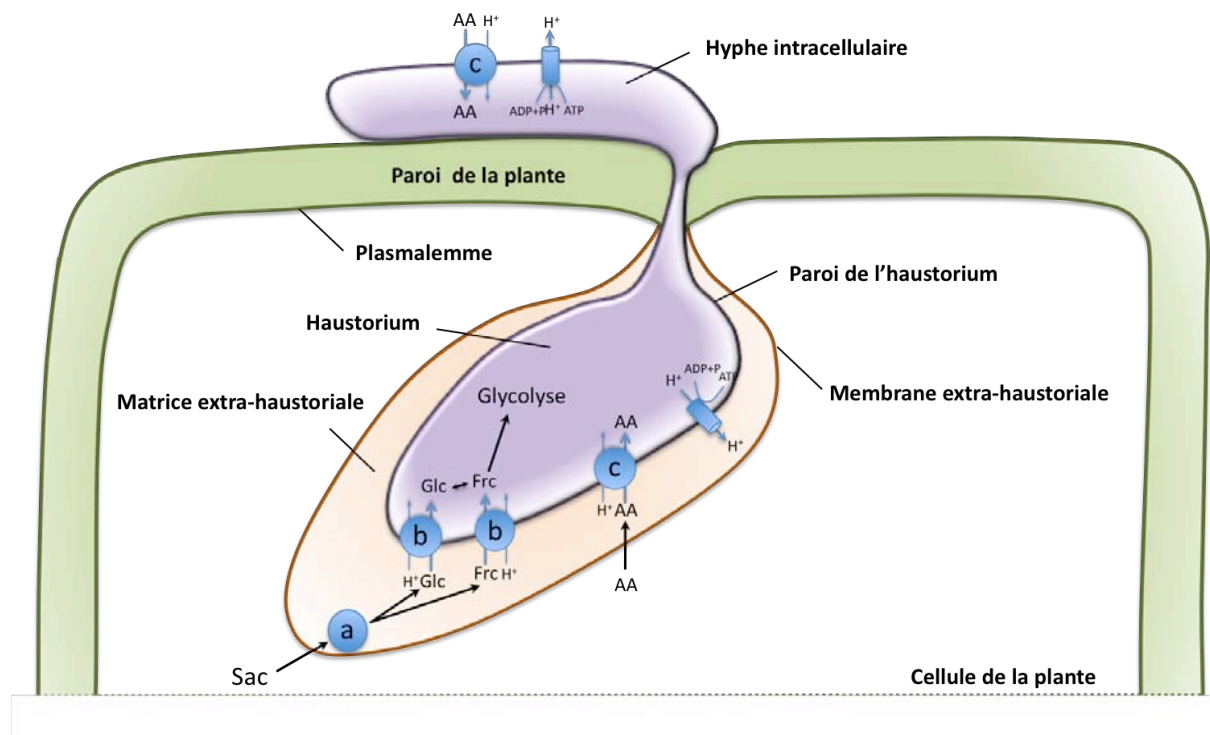


Figure 12: Modèle de transport des acides aminés et des hexoses chez les rouilles (D'après Voegelé et Mendgen 2003). Le saccharose (Sac) de la plante présent dans la matrice extracellulaire est clivé en glucose (Glc) et fructose (Frc) grâce à l'invertase fongique INV1p (a). HXT1 (b), présent spécifiquement au niveau de l'haustorium, transporte ces monosaccharides de la matrice extra-haustoriale vers l'haustorium. AAT1p et AAT2p (c) transportent les acides aminés (AA) dans l'haustorium mais également dans les hyphes intracellulaires. Une augmentation de l'activité des pompes à protons H^+ ATPases permet le transport intense.

révélé une forte accumulation au niveau de l'haustorium et de la cellule mère haustoriale (Sohn et al. 2000). Ces résultats suggèrent l'importance de l'haustorium pour la biosynthèse *de novo* de métabolites telle que la vitamine B1, cofacteur essentiel pour l'activité de nombreuses enzymes impliquées dans le métabolisme carboné (Sohn et al. 2000).

5.2 - Implication de l'haustorium dans le transfert des nutriments

De par son appellation (haustor : Celui qui puise), de Barry (1863) indique une des fonctions de l'haustorium, à savoir le transfert des nutriments de la plante hôte vers le champignon. Ce n'est que bien plus tard, avec l'avènement de la biologie moléculaire, que le rôle de l'haustorium dans le transfert de nutriments de la plante hôte a pu être établi. Parmi les *PIGs* d'*U. fabae* préférentiellement exprimés dans l'haustorium, certains codent des transporteurs de sucres et d'acides aminés (Hahn et Mendgen 1997a; Hahn et al. 1997b) soulignant l'implication de l'haustorium dans le transfert des nutriments de la plante (**Figure 12**).

5.2.1 - Transporteurs d'acides aminés

Au cours des années 90, des études menées par le groupe de Mendgen et Voegelé sur *U. fabae* ont révélé une augmentation de l'activité H⁺ATPase au niveau de la membrane de haustorium comparée aux membranes isolées de spores et spores germées (Struck et al. 1996; Struck et al. 1998). Il a été suggéré que le gradient de proton généré par cette activité ATPase pourrait piloter le système de transport actif secondaire mis en place par le champignon (Hahn et al. 1997b). Parmi les transporteurs d'acides aminés identifiés chez *U. fabae*, AAT2p (PIG2) est spécifiquement localisé au niveau de l'haustorium mais aucune activité spécifique de transport pour un acide aminé n'a pu être mise en évidence (Mendgen et al. 2000). AAT1p, un proche homologue d'AAT2p est quant à lui impliqué dans le transport secondaire d'un large panel d'acides aminés avec une spécificité plus importante pour les acides aminés L-histidine et L-lysine (Struck et al. 2002). Des expériences de Northern blot indiquent que, contrairement à son homologue *AAT2p*, l'expression de *AAT1p* n'est pas restreinte à l'haustorium. C'est également le cas pour *AAT3p* qui, bien que fortement exprimé au niveau de l'haustorium, présente également des niveaux de transcrits importants dans les structures d'infection précoces formées *in vitro* (Struck et al. 2004). De manière contrastée avec AAT1p, la caractérisation d'AAT3p a révélé une spécificité pour la L-leucine ainsi que pour les acides aminés soufrés L-cystéine et L-méthionine (Struck et al. 2004). Ces données suggèrent que la prise en charge des acides aminés chez *U. fabae* s'effectue via l'haustorium mais également via d'autres structures fongiques telles que les hyphes intercellulaires par exemple (**Figure 12**).

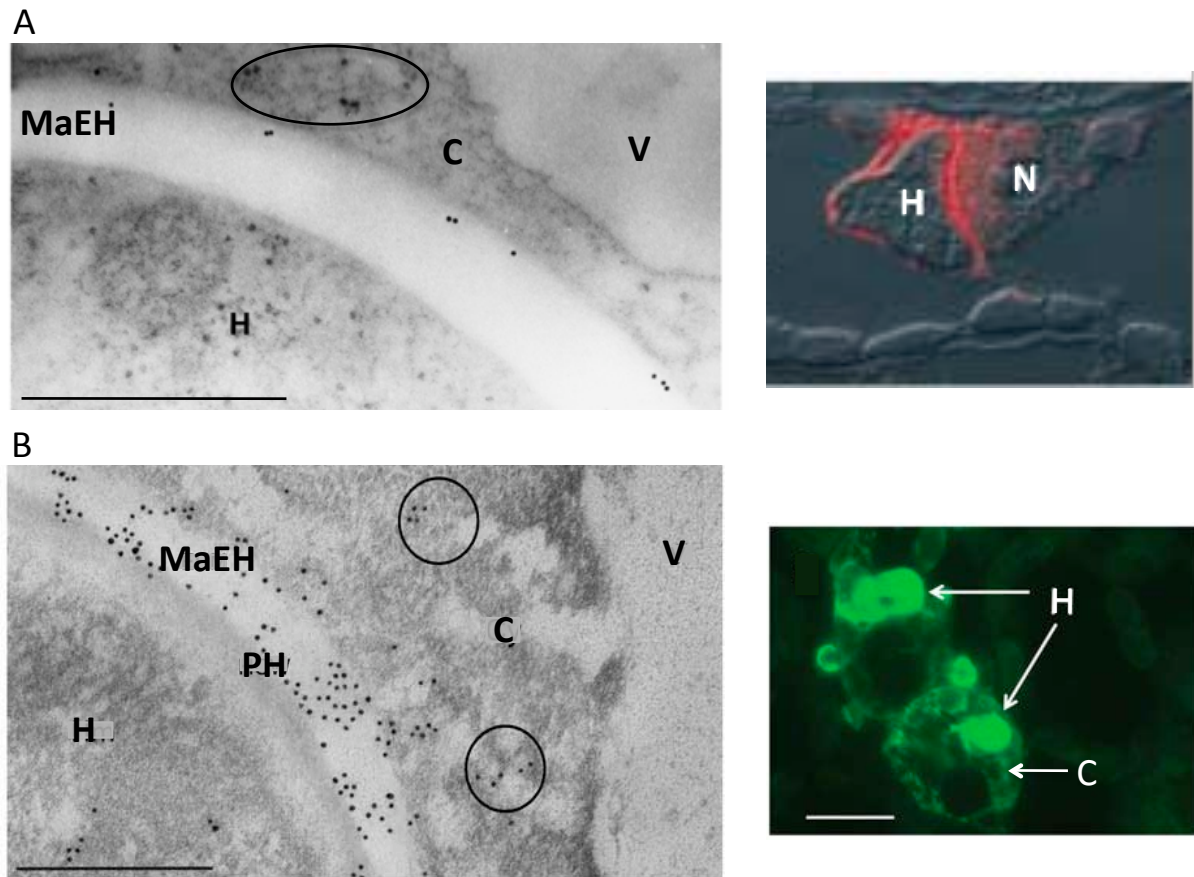


Figure 13: Immunolocalisation de *Uf*-RTP1 (*Uromyces fabae*) et de *Ml*-AvrM (*Melampsora lini*) au niveau de l'interface plante-haustorium par la technique d'immunogold (à gauche) et d'immunofluorescence (à droite) (Kemen et al. 2005; Rafiqi et al. 2010). **A**, Localisation de la rust transferred protein RTP1 dans une cellule de haricot infectée par *U. fabae* (Kemen et al. 2005) et **B**, de la protéine d'avirulence AvrM de *M. lini* dans une cellule de lin (Rafiqi et al. 2010). Le marquage des deux protéines, révélé par la présence de billes d'or, a été détecté au niveau de l'haustorium (H), de la Matrice extra-haustoriale (MaEH) mais également dans le cytoplasme (C) de la cellule hôte (cercles) par microscopie électronique à transmission (Bar= 500nm). L'accumulation de AvrM au niveau de l'haustorium et du cytoplasme et de RTP1 au niveau de l'haustorium, du cytoplasme et du noyau de la cellule hôte a également été visualisée par immunofluorescence (Bar= 50µm). H: haustorium; PH, paroi de l'haustorium, MaEH: membrane extra-haustoriale; C: cytoplasme de l'hôte; V: vacuole de l'hôte, CH: chloroplastes, N: noyau de l'hôte.

5.2.2 - Transporteurs de sucres

Contrairement au transport des acides aminés qui n'est pas restreint à l'haustorium, le transport des sucres mis en évidence chez *U. fabae* semble s'effectuer spécifiquement au niveau de cette structure d'infection (Voegelé et al. 2001) (**Figure 12**). Le transporteur d'hexose HXT1 a été localisé par immunolocalisation à la périphérie du corps des haustoria dicaryotiques (Voegelé et al. 2001) et à l'extrémité des haustoria monocaryotiques de *U. fabae* (Voegelé et Mendgen 2003). Aucun marquage n'a été détecté dans les hyphes intercellulaires ou dans les spores. Ni la « nested » PCR, ni les analyses de Southern blot ont permis de mettre en évidence d'autres transporteurs d'hexoses présents chez *U. fabae* dans les différents stades de développement testés (Voegelé et al. 2001). L'expression hétérologue dans la levure et dans des oocytes de Xénope a révélé que HXT1p est un transporteur de monosaccharide proton dépendant. Ce transporteur a une forte affinité pour le D-glucose, le D-fructose et le D-mannose avec des valeurs de K_m croissantes dans cet ordre. De manière intéressante, la présence d'une invertase fongique INV1p, localisée au niveau de la membrane extrahaustoriale pourrait permettre de cliver le saccharose de la plante en glucose et fructose directement utilisables par le transporteur de monosaccharide HXT1 (Voegelé et al. 2006).

5.3 - Implication de l'haustorium dans la sécrétion de molécules signal

Au delà de l'aspect trophique considéré pour l'haustorium, celui-ci semble également être l'interface privilégiée pour la sécrétion de molécules effectrices (Catanzariti et al. 2007). Chez les agents de rouilles, de nombreuses études sont ou ont été récemment menées afin d'identifier ces facteurs de virulence/avirulence sécrétés par le champignon afin de manipuler la structure et le fonctionnement de la cellule hôte.

Chez *U. fabae*, des gènes codant des protéines sécrétées ont également pu être identifiés parmi les *PIGs* préférentiellement exprimés au niveau de l'haustorium (Hahn et Mendgen 1997a). Par exemple, la petite protéine sécrétée RTP1 (Rust Transferred Protein) de 24 kDa a été détectée par deux techniques d'immunolocalisation, l'immunogold et l'immunofluorescence, au niveau de la matrice extrahaustoriale mais également dans les cellules de la plante hôte suggérant son transfert de l'haustorium vers le cytoplasme (§ I-3.3) (**Figure 13**; Kemen et al. 2005). De manière intéressante, cette protéine présente un signal d'adressage nucléaire (Nuclear localization signal, NLS) au niveau de sa séquence protéique et l'immunolocalisation a confirmé l'accumulation de la protéine dans le noyau de la cellule hôte (Kemen et al. 2005). Néanmoins, le mécanisme de transport par lequel la protéine traverse la membrane extrahaustoriale et

s'accumule dans le cytoplasme de l'hôte reste inconnu. Cette localisation atypique de RTP1 suggère le rôle important de l'haustorium pour la sécrétion de protéines chez l'hôte.

Chez *M. lini*, une banque d'ADNc d'haustoria isolés de feuilles de lin infectées a permis l'identification de 21 gènes exprimés au niveau de l'haustorium et codant des protéines sécrétées (Haustorially Expressed Secreted Proteins, HESPs). Parmi ces 21 HESPs, quatre correspondent aux facteurs d'avrulence AvrL567, AvrP4, AvrP123 et AvrM suggérant le rôle de ces HESPs dans la pathogénie de l'agent responsable de la rouille du lin (Catanzariti et al. 2006). La reconnaissance intracellulaire de ces facteurs d'avrulence par les protéines de résistance intracellulaires du lin indique une sécrétion au-delà de la membrane de l'hôte (Dodds et al. 2004; Catanzariti et al. 2006). Plus récemment, la protéine AvrM de *M. lini* a pu également être observée par les techniques d'immunogold et d'immunofluorescence au niveau de la matrice extrahaustoriale ainsi que dans le cytoplasme de la cellule hôte infectée renforçant le rôle essentiel de l'haustorium pour la sécrétion de protéines effectrices dans le cytoplasme de l'hôte (**Figure 13**; Rafiqi et al. 2010).

Certains gènes codant des petites protéines sécrétées ont également été mis à jour parmi les ESTs d'hyphes d'infection et d'haustoria de feuilles de haricot infectées par *U. appendiculatus* (Puthoff et al. 2008). La comparaison de ce set avec d'autres protéines sécrétées d'agents de rouilles a notamment permis l'identification d'une petite famille de protéines sécrétées RTP1-like également présente chez *U. appendiculatus* et *P. graminis*. Très récemment, le séquençage d'ESTs d'haustoria et de spores issu de 4 espèces de *Melampsora* affectant le peuplier a permis d'identifier un large catalogue de protéines sécrétées (Joly et al. 2010). Bien que certains homologues d'effecteurs de *M. lini* ou *U. fabae* aient été identifiés, la majorité des gènes mis à jour dans ce catalogue semblent spécifiques des *Melampsora* spp. affectant le peuplier. Une approche de génomique comparative, basée sur l'identification de gènes homologues entre ces 4 espèces a permis l'identification de nombreux gènes sous pression de sélection positive au sein de ce catalogue de protéines sécrétées (Joly et al. 2010).

III - L'interaction peuplier-*Melampsora larici-populina*

1 - Le peuplier

Le peuplier appartient avec le genre *Salix* à la famille des Salicacées. Les différentes espèces de peupliers se répartissent en six sections selon leurs caractéristiques phénotypiques et biologiques (Eckenwalder 1996; **Figure 14**). Les sections *Aigeiros*, *Tacamahaca* et *Populus* sont les plus importantes. La section *Aigeiros* comprend notamment *Populus deltoides* et *Populus nigra*, la section *Tacamahaca*, *Populus trichocarpa* et la section *Populus*, *Populus alba* et *Populus tremula*. Les espèces de peupliers issues de la section *Aigeiros* poussent principalement en Amérique du Nord (*P. deltoides*) et en Eurasie (*P. nigra*). Les espèces *P. deltoides* et *P. trichocarpa* sont toutes deux originaires d'Amérique du Nord alors que l'espèce *P. nigra* est autochtone d'Eurasie. Ces trois espèces ainsi que leurs hybrides sont les plus communément utilisées en populicultures dans le monde entier.

1.1 - La populiculture

La surface des peupleraies cultivées en France est passée de 100 000 ha après la seconde guerre mondiale à environ 244 000 ha en 2006 représentant ainsi 1,6% de la surface boisée. La peupleraie est caractérisée par une courte révolution liée à la croissance rapide des cultivars sélectionnés. Les cultivars issus de sélections récentes à croissance rapide et plantés dans des conditions favorables peuvent ainsi être exploités après 15 à 20 ans seulement (Frey et Pinon 2004). La répartition géographique des peupleraies en France est inégale selon les régions avec une prédominance pour les régions Nord Pas-de-Calais, Picardie, Champagne-Ardenne, Pays de Loire et Aquitaine (**Figure 15**). En France, le peuplier est l'essence feuillue la plus récoltée après le chêne et le hêtre avec environ 1,3 millions de m³ de bois récolté par an (AFOCEL 2006). Son utilisation comme bois d'industrie (trituration pour la papétrie et les panneaux de particules) est moins importante que son utilisation en tant que bois d'œuvre, avec le sciage (caisserie, ameublement et bâtiment) et le déroulage (contreplaqué et emballages légers) (Viart 1999). Plus récemment, le peuplier a présenté un regain d'intérêt pour différents programmes de production de bois pour des finalités bioénergétiques (bois-énergie et transformation en bio-carburant). Ainsi, aux Etats-Unis et en Europe de vastes programmes de recherche ont été mis en place afin d'utiliser le peuplier pour la production de biocarburant de nouvelle génération (Rubin et al. 2009 et Energy Poplar; <http://www.energypoplar.eu>).

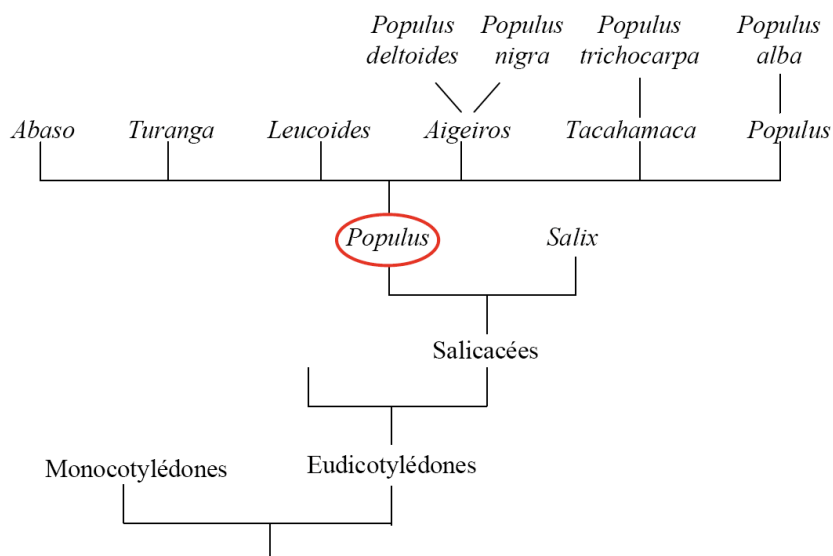


Figure 14: Les différentes espèces de peuplier. Le peuplier appartient avec le genre *Salix* à la famille des Salicacées. Les différentes espèces de peuplier se répartissent en 6 sections selon leurs caractéristiques phénotypiques et biologiques (Eckenwalder 1996). Les 4 espèces de peuplier les plus communes en France sont indiquées au sein des sections auxquelles elles appartiennent.

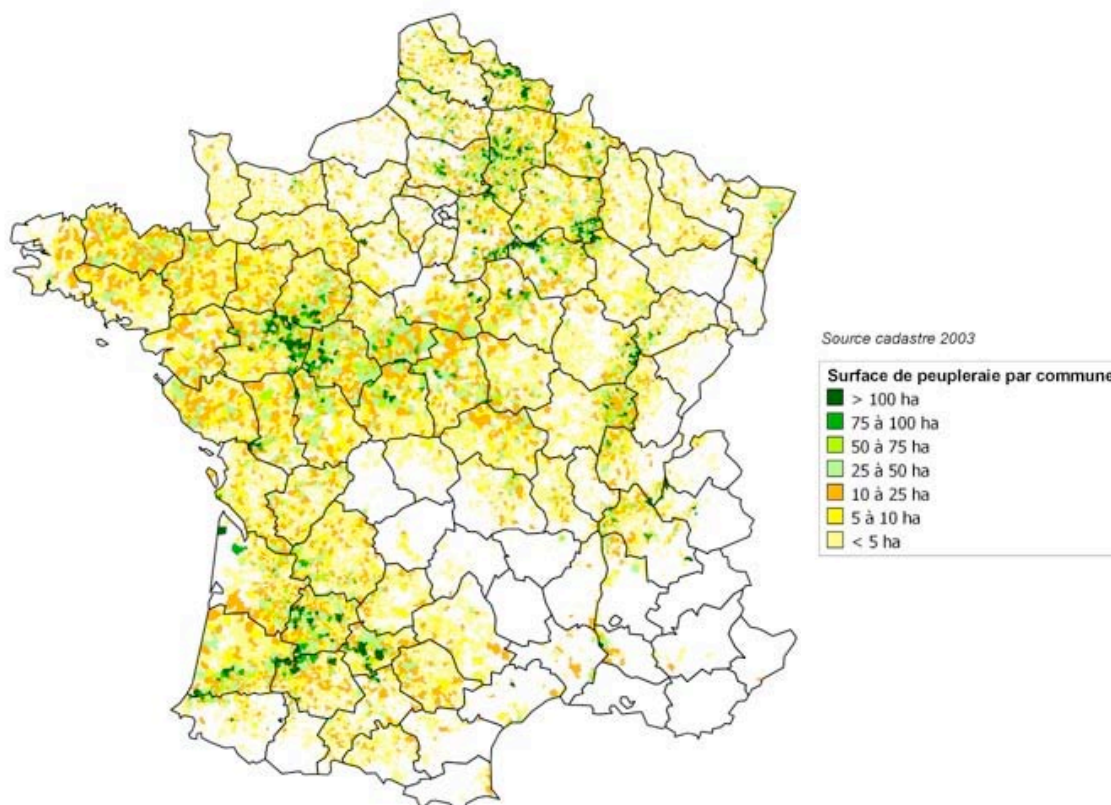


Figure 15: Carte de répartition et superficie des peupleraies par commune française (hectare, ha). (Source cadastre 2003 et site internet <http://www.foretpriveefrancaise.com/repartition-geographique>)

Les cultivars euraméricains (*P. deltoides* x *P. nigra* = *P. x euramericana*) et interaméricains (*P. deltoides* x *P. trichocarpa* = *P. x interamericana*) sont actuellement les cultivars prédominants dans la populiculture française (Berthelot et Chantre 2004). Ces hybrides présentent une croissance rapide, une très grande vigueur, et sont plus ou moins résistants aux stress biotiques et abiotiques. Jusqu'en 2000, le cultivar interaméricain 'Beaupré' était le cultivar le plus vendu par les pépiniéristes français (Berthelot et Chantre 2004). Ce cultivar, du fait de ses qualités exceptionnelles de croissance et de qualité de bois et de sa résistance complète (qualitative) à la rouille à *Melampsora*, a connu un succès extraordinaire dans la moitié nord de la France représentant jusqu'à 80% des peupliers plantés en Picardie dans les années 1990. Néanmoins, les sévères épidémies de rouilles à *M. larici-populina* sur ce cultivar suite au contournement de sa résistance ont conduit à son abandon progressif.

1.2 - Le peuplier: un arbre modèle dont le génome a été séquencé

Outre son intérêt économique lié à sa croissance rapide et son intérêt écologique (séquestration du CO₂, biocarburants), le peuplier est également facilement manipulable par génie génétique et de nombreuses cartes génétiques sont disponibles pour différents pedigrees faisant de lui un excellent modèle de laboratoire (Taylor et al. 2007; Jansson et al. 2009). En 2001, le Joint Genome Institute (JGI, Département de l'énergie américain; <http://www.jgi.doe.gov/>), avec le concours de l'International Poplar Genome Consortium (IPGC, <http://www.ornl.gov/sci/ipgc/>) a entrepris le séquençage du génome de *P. trichocarpa* faisant ainsi du peuplier le premier arbre dont la séquence génomique a été séquencée (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html) (Tuskan et al. 2006). Le séquençage des 485 Mb ± 10 Mb a permis la prédiction de plus de 45000 gènes putatifs. La mise à disposition de ce génome à la communauté scientifique a ouvert de nouvelles perspectives afin d'identifier les déterminants génétiques qui contrôlent le développement des arbres ainsi que les mécanismes mis en place en réponse aux stress environnementaux, qu'ils soient biotiques ou abiotiques. (Jansson et al. 2009; Duplessis et al. 2009).

1.3 - Les maladies des peupliers

Les peupliers sont sensibles à de nombreux agresseurs dont des insectes (Saperdes, Chrysomelles...), des bactéries, des virus et de nombreux champignons pathogènes. En Europe, la résistance des clones de peuplier à *Venturia populina*, agent responsable de la tavelure, à *Marssonina brunnea*, agent de la brunissure des feuilles, et surtout à *M. larici-populina*, responsable de la rouille foliaire du peuplier, sont des critères de sélection primordiaux. La

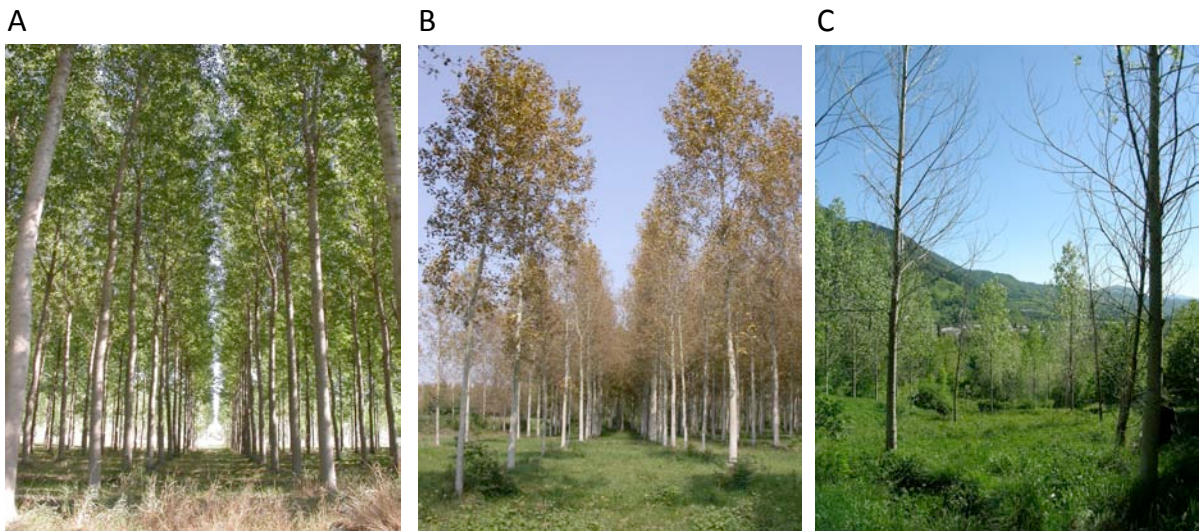


Figure 16: Symptômes observés en plantation de ‘Beaupré’ lors d’attaques par l’agent de la rouille foliaire *Melampsora larici-populina*. **A**, plantation de Beaupré de 12 ans très peu infectée (Manosque (04), Septembre 2004). **B**, Brunissement des feuilles de peupliers observé sur une plantation de ‘Beaupré’ d’environ 8 ans fortement infectée par la rouille (Aoste (38), septembre 2004). **C**, plantation de ‘Beaupré’ d’environ 15 ans ayant subi des attaques répétées de rouille depuis 1999, entraînant la mortalité de nombreux arbres (Pont-du-Fossé (05), mai 2006). Photos de Pascal Frey, INRA Nancy.

résistance au chancre bactérien, causée par la bactérie *Xanthomonas populi* pv. *populi*, est également prise en compte chez les cultivars plantés en France au nord de la Loire. Actuellement, le parasite fongique primaire occasionnant les dégâts les plus importants sur les peupliers cultivés est *M. larici-populina*, agent de la rouille foliaire (**Figure 16**). La description de ce parasite et les dégâts qu'il occasionne sont développés au cours des prochains paragraphes.

La tavelure, provoquée par différentes espèces de *Venturia* spp. est observable au printemps sur les peupliers des sections *Populus* et *Aigeiros*. Cette maladie, causant des défeuillaisons printanières, provoqua des pertes de production de l'ordre de 30% au début du XX^{ème} siècle en Italie principalement (Ridé 1995). Suite à ce constat, la résistance partielle à ce parasite fut intégrée dans les programmes italiens de sélection. Ce champignon, toujours commun en Europe et en Amérique du Nord, ne provoque plus de dégâts considérables. Dans les années 60 dans la vallée du Pô, *M. brunnea*, entraîna une baisse de production annuelle de 16% (Lanier 1976) chez les cultivars euraméricains. De nombreux dégâts ont également été observés ailleurs en Europe et de nouveaux cultivars ayant comme parent *P. deltoides*, furent sélectionnés pour leur résistance partielle à *M. brunnea*. Au cours des années 60 dans le nord de la France, *X. populi* pv *populi* provoqua des dégâts importants sur les cultivars euraméricains (FAO 1981). Enfin, le virus de la mosaïque du peuplier ('PopMV'), causé par un carlavirus du groupe S de la pomme de terre peut entraîner dès le mois de juin des symptômes foliaires correspondant à des plages vert clair situées à l'embranchement de nervures et qui s'étendent pour former des taches rondes appelées 'taches d'huile' (FAO 1981). De nombreux auteurs ont constaté un effet dépressif considérable du virus sur la taille, le diamètre des plants et l'aptitude à l'enracinement (Cooper 1979).

2 - *Melampsora larici-populina*

Les *Melampsora* spp. appartiennent à la classe des *Pucciniomycètes* et à l'ordre des *Pucciniales* (§ II-2). Trois espèces sévissent sur les peupliers hybrides cultivés en Europe: *M. larici-populina* et *Melampsora allii-populina*, toutes deux d'origine eurasiatique et *Melampsora medusae* f. sp. *deltoidae* d'origine nord-américaine. Seules les deux premières espèces causent des dégâts importants sur les peupliers cultivés en France (Frey et al. 2005; Pinon 1995). En outre, il existe cinq autres espèces de *Melampsora* qui infectent les peupliers trembles telles que *Melampsora larici-tremulae*, *Melampsora magnusiana*, *Melampsora pinitorqua*, *Melampsora pulcherrima* et *Melampsora rostrupii*. Les feuilles infectées par *Melampsora* spp. voient leur activité photosynthétique nettement réduite, entraînant une chute prématurée des feuilles et perturbant fortement la mise en réserve des substances carbonées et azotées qui a lieu normalement à

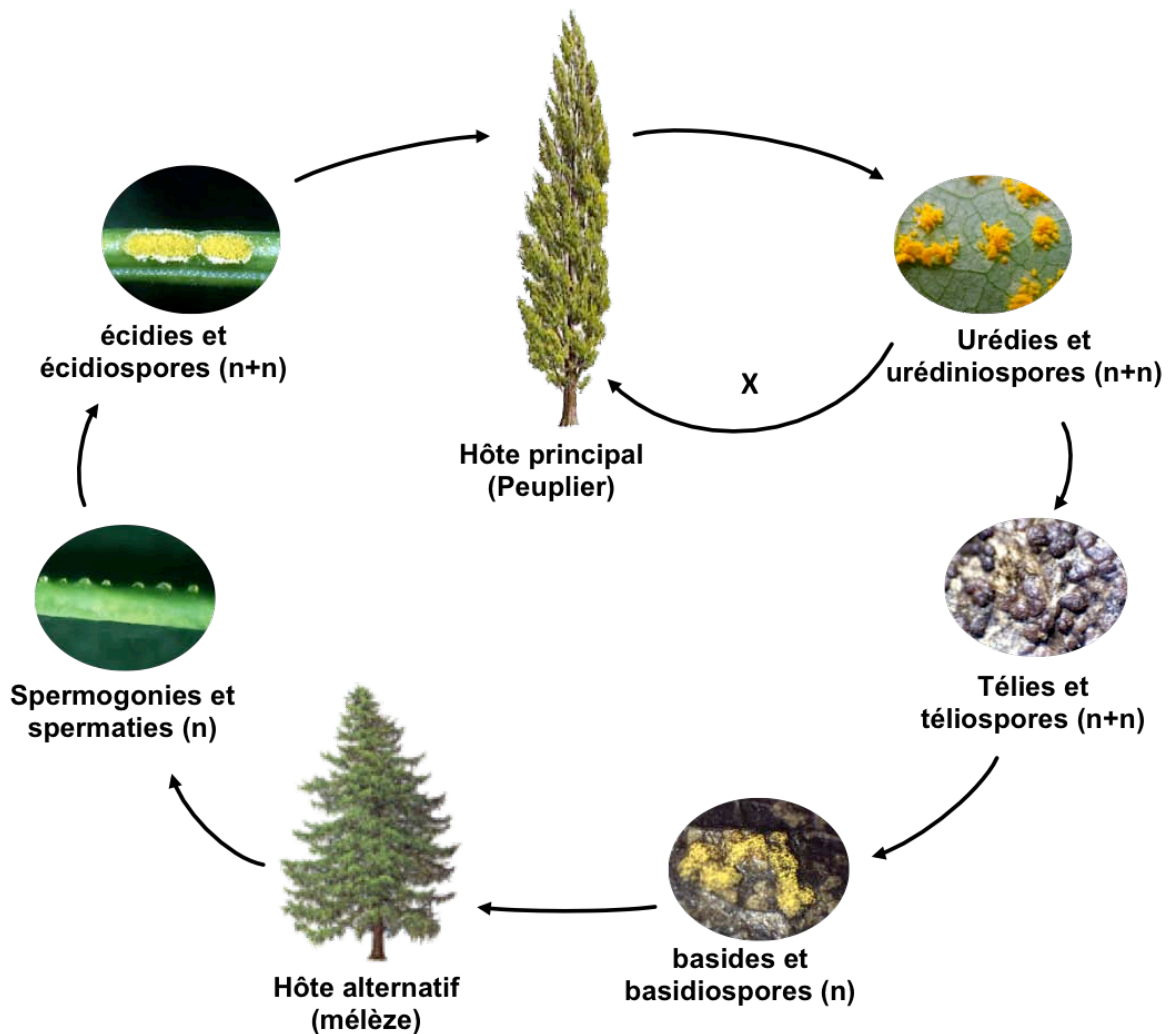


Figure 17: Cycle biologique de *Melampsora larici-populina*. Le cycle est hétéroïque et macrocyclique car il fait intervenir respectivement deux hôtes et cinq types de spores différents. La reproduction sexuée a lieu sur le mélèze (hôte alternatif ou hôte écidien) à partir duquel des écidiospores vont être produites avant d’infecter des feuilles de peuplier (hôte principal ou hôte télieu). La reproduction asexuée a lieu sur le peuplier avec la production répétée d’urédies à partir desquelles des urédospories vont être relarguées durant tout l’été conduisant à de sévères épidémies. A l’automne, le champignon forme des télispories qui vont hiverner sur les feuilles mortes tombées au sol. Au printemps, la formation de basides à partir de ces télispories va permettre la production de basidiospories infectant le mélèze

l'automne. Des épidémies répétées de rouille peuvent conduire à des pertes de croissance pouvant aller jusqu'à 60% (Stépien 1992; Miot et al. 1999). Par ailleurs, ces attaques à répétition peuvent entraîner un affaiblissement général des plants, les prédisposant à certains parasites secondaires tels que *Cytospora chrysoperma* et *Discosporium populeum*. A terme, ces épidémies chroniques peuvent conduire à la mort des arbres (**Figure 16**).

2.1 - Cycle biologique de *Melampsora larici-populina*

2.1.1 - Un cycle biologique complexe

La plupart des *Pucciniales* se distinguent des autres champignons par leur cycle biologique complexe (§ II-3) C'est notamment le cas de *M. larici-populina* qui présente un cycle hétéroïque macrocyclique faisant référence à ses deux hôtes ainsi qu'aux cinq types de spores produits durant le cycle (**Figure 17**) (Frey et Pinon 2004). Alors que la reproduction sexuée se déroule sur les aiguilles de mélèze (hôte alternant ou hôte écidien), la reproduction asexuée a lieu sur les feuilles de peuplier (hôte principal ou hôte télén). La phase importante dans la maladie est le stade urédien sur peuplier pendant laquelle le champignon est sous forme dicaryotique ($n + n$ chromosomes). La cinétique d'infection des urédospores aboutit à la formation d'urédies (§ II-2.1.2), pustules orangées caractéristiques apparaissant à la face inférieure des feuilles. Ces urédies produisent des quantités massives d'urédospores (2500 urédospores par jour et par urédie; Dowkiw et al. 2003) qui sont disséminées par le vent sur de très longues distances permettant ainsi l'infection de nouvelles feuilles de peuplier (Barrès et al 2008). Lorsque les conditions climatiques sont favorables, une dizaine de cycles peuvent ainsi se succéder pendant l'été. A l'automne, le champignon forme des télies contenant des téliospores ($n + n$) qui apparaissent sous forme de croûtes noirâtres sur la surface des feuilles de peuplier tombées au sol. Le champignon hiverne ainsi dans les feuilles mortes jusqu'au printemps où la caryogamie se produit dans les téliospores donnant naissance à des basidiospores haploïdes. Ces basidiospores haploïdes (n), de polarité + ou - sont disséminées par le vent jusqu'aux aiguilles de mélèze où elles vont germer et former des fructifications appelées spermogonies, également de polarité + ou -. Les spermogonies produisent à la surface des aiguilles du mélèze des spermaties haploïdes (n), qui fécondent les spermogonies de polarité opposée (plasmogamie). Ceci aboutit à la formation d'un mycélium dicaryotique, à partir duquel une nouvelle fructification, l'écidie, est formée sur l'aiguille du mélèze. A partir de cette structure, des écidiospores sont libérées et disséminées par le vent et vont infecter de nouveau l'hôte principal durant tout l'été.

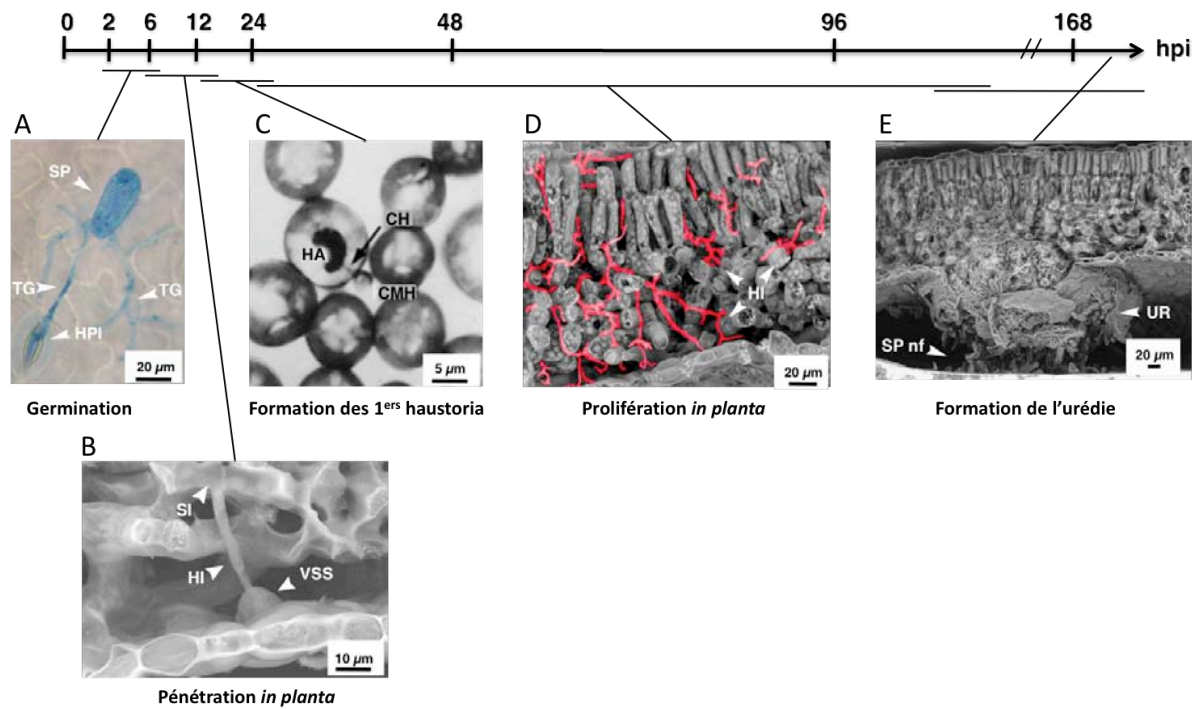


Figure 18: Cinétique d'infection de feuilles de peuplier par *Melampsora larici-populina* lors d'une interaction compatible (D'après Laurans et Pilate 1999 et Rinaldi et al. 2007). Développement des structures d'infection de *M. larici-populina* durant la cinétique d'infection de feuilles de *Populus trichocarpa* x *Populus deltoides* 'Beaupré' (A, B, D, E) et de *Populus deltoides* x *Populus nigra* 'Ogy' (C). **A**, Coloration au bleu d'aniline révélant la formation de tubes germinatifs (TG) et la pénétration par un stomate 2 heures post inoculation (hpi). **B**, Formation d'une vésicule sous-stomatique et différenciation des premiers hyphes d'infection à 6hpi. **C**, Formation des premiers haustoria observés à 17 hpi. **D**, Suite à la mise en place des haustoria, prolifération du champignon au sein du parenchyme lacuneux et palissadique entre 24 et 96 hpi. **E**, Urédie mature libérant de grandes quantités d'urédospores à 168 hpi. SP, Urédospores; HI: hyphe d'infection; HPI: hyphe primaire d'infection; HA: haustorium; CMH: cellule mère haustoriale; CH: 'cou' de l'haustorium; SP nf: urédospores nouvellement formées; UR: urédie; VSS: vésicule sous-stomatique; SI: site d'infection.

2.1.2 - Cinétique d'infection au cours du stade urédien

La cinétique d'infection de *M. larici-populina* au cours du stade urédien est en tout point comparable à celle décrite pour la majorité des agents de rouilles (§ II-4). Les principales structures d'infection formées durant le processus infectieux ont été décrites sur différents cultivars de peuplier (Laurans et Pilate 1999; Rinaldi et al. 2007) (**Figure 18**). Ainsi, ces deux études ont montré qu'après seulement 6 hpi l'urédospore a germé, formé un tube germinatif et pénétré au sein des tissus foliaires via un stomate. Une structure appressoriale est généralement formée avant la pénétration par le stomate, mais pas de manière stricte suggérant que cette structure ne serait pas nécessaire aux étapes ultérieures d'infection de la plante (Rinaldi et al. 2007). Les premières vésicules sous-stomatiques ont été observées entre 6 et 12 hpi alors que les premiers haustoria ont été observés dans le cadre de l'interaction entre *M. larici-populina* et le cultivar 'Ogy' dès 17 hpi (Laurans et Pilate 1999). Par la suite, le pourcentage de cellules infectées à 2 jours post inoculation (jpi) est passé à 2%, puis à 10% après 4 jpi et finalement à 11% après 6 jpi (Laurans et Pilate 1999). Dès la formation des premiers haustoria, une forte colonisation est observée au sein des tissus foliaires dans le cadre d'une interaction compatible entre le champignon et son hôte aboutissant après 168 hpi à la formation de l'urédie, structure à partir de laquelle des quantités importantes d'urédospores sont libérées. A l'inverse, dans le cadre d'une interaction incompatible, la reconnaissance de molécules du champignon par les récepteurs de la plante déclenche rapidement les mécanismes de défense de l'hôte. Chez le peuplier, cette réponse consiste en une réaction d'hypersensibilité très localisée. Les spots nécrotiques étant très localisés, aucune différence n'est observable macroscopiquement entre une feuille infectée par un isolat avirulent et une feuille non infectée (Laurans et Pilate 1999; Rinaldi et al. 2007).

2.2 - Virulences connues chez *Melampsora larici-populina*

La mise en évidence des réactions de compatibilité et d'incompatibilité en fonction des combinaisons d'individus hôtes et pathogènes indique une relation gène-pour-gène (Flor 1971) pour le pathosystème *M. larici-populina*/peuplier bien que l'interaction R-Avr n'ait pas encore été démontrée expérimentalement. Ces réactions de compatibilité/incompatibilité permettent d'établir les caractères de résistance des clones de peuplier ainsi que les virulences des isolats de *M. larici-populina*.

Au cours des 30 dernières années, des programmes de sélection variétale pour la résistance à la rouille foliaire ont conduit à la sélection de nombreux cultivars de peuplier présentant des

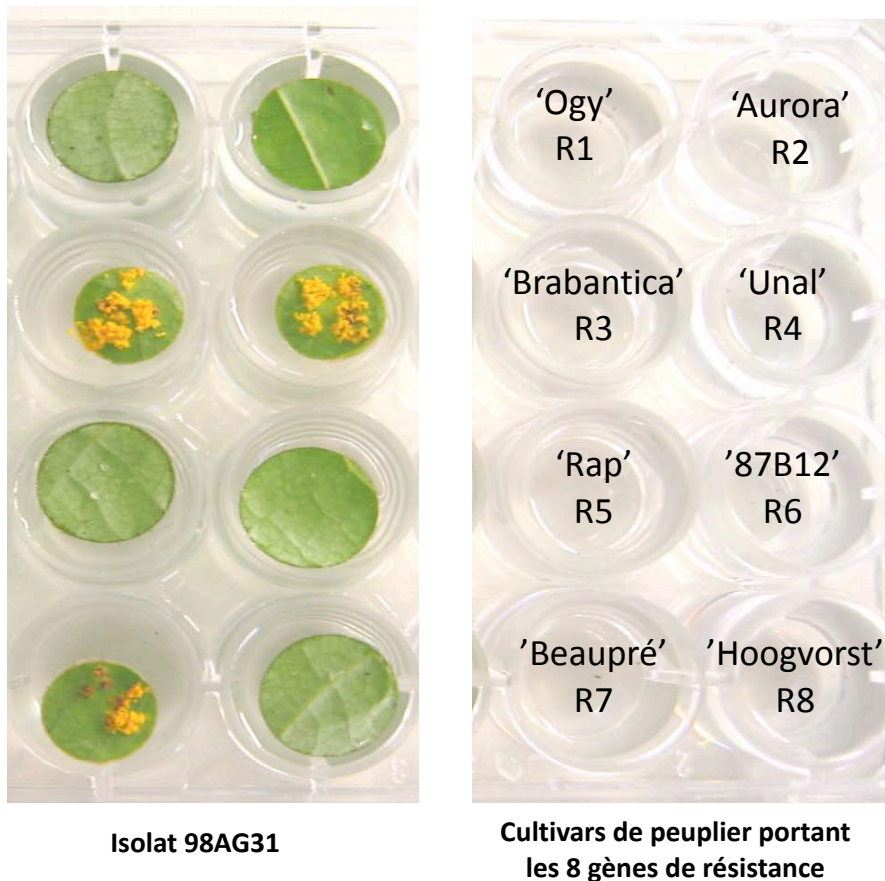


Figure 19: Exemple de pathotypage de l'isolat 98AG31 de *Melampsora larici populina*. Les virulences sont discriminées sur une gamme différentielle de 8 cultivars de peupliers possédant différentes résistances complètes à *M. larici-populina* allant de R1 à R8. Dans cet exemple, l'isolat 98AG31, virulent sur les cultivars 'Brabantica', 'Unal' et 'Beaupré', aura donc le pathotype 3-4-7. Photographie de Cécile Rinaldi.

résistances complètes à *M. larici-populina*. Néanmoins, après plusieurs années de culture, ces résistances complètes ont rapidement été contournées par le champignon du fait de l'acquisition de nouvelles virulences, rendant ces cultivars sensibles à la maladie (Frey et Pinon 2004). A ce jour, 8 virulences ont été décrites pour *M. larici-populina* (« Vir1 » à « Vir8 ») et chacune d'elles confère au champignon la capacité de contourner un gène de résistance du peuplier. Le contournement du gène de résistance R7 porté par le cultivar 'Beaupré' par *M. larici-populina* (Pinon 1995) a entraîné des épidémies de rouille sans précédent et causé des pertes conséquentes pour la populiculture française entre 1995 et 1999 (Frey et Pinon 2004). Afin de caractériser les virulences des isolats de *M. larici-populina* collectées dans la nature, une gamme différentielle de cultivars, présentant des résistances contrastées vis-à-vis de la maladie a été établie à l'INRA de Nancy dans l'équipe « Ecologie des champignons pathogènes forestiers » par le groupe de Jean Pinon et Pascal Frey (Gerard et al. 2006). Les isolats peuvent ainsi être caractérisés par leur aptitude à infecter un ou plusieurs cultivars de cette gamme et la combinaison de virulences qui caractérise un isolat est appelé race physiologique ou pathotype. En théorie, avec 8 virulences, on peut définir 2^8 soit 256 pathotypes différents. Néanmoins à ce jour, seuls 57 pathotypes ont été identifiés dans la nature (Pinon et Frey 2005). Un exemple de pathotypage, présenté dans la **figure 19**, montre que l'isolat 98AG31 de *M. larici-populina* est virulent sur les hôtes possédant les résistances 3, 4 et 7. Cet isolat appartient donc au pathotype 3-4-7.

3 - Bases moléculaires de l'interaction Peuplier/*Melampsora*

Alors que l'épidémiologie de la maladie de la rouille foliaire a particulièrement bien été étudiée au cours des 20 dernières années (Frey et al. 2005), les mécanismes moléculaires mis en jeu lors de l'interaction entre le peuplier et ce pathogène commencent seulement à être élucidés. Les réponses de défense mises en place par le peuplier durant le processus infectieux de différentes *Melampsora* spp. ont plus particulièrement été étudiées (Duplessis et al. 2009; **Figure 20**).

Les déterminants génétiques de la résistance à *M. larici-populina* ont tout d'abord été identifiés par plusieurs approches de génomique et de transcriptomique, notamment par l'analyse comparative entre les interactions compatible et incompatible entre *M. larici-populina* et *P. trichocarpa* x *P. deltoides* 'Beaupré' (Rinaldi et al. 2007). Un pathosystème a tout d'abord été établi pour réaliser cette analyse avec un isolat virulent (98AG31) et un isolat avirulent (93ID6) de *M. larici-populina* afin de pouvoir comparer les réponses transcriptionnelles du peuplier dans le cadre de ces deux interactions. Différentes approches tels que le séquençage d'ESTs, les banques soustractives, les puces à ADN et la RT-qPCR ont été utilisées afin d'identifier des

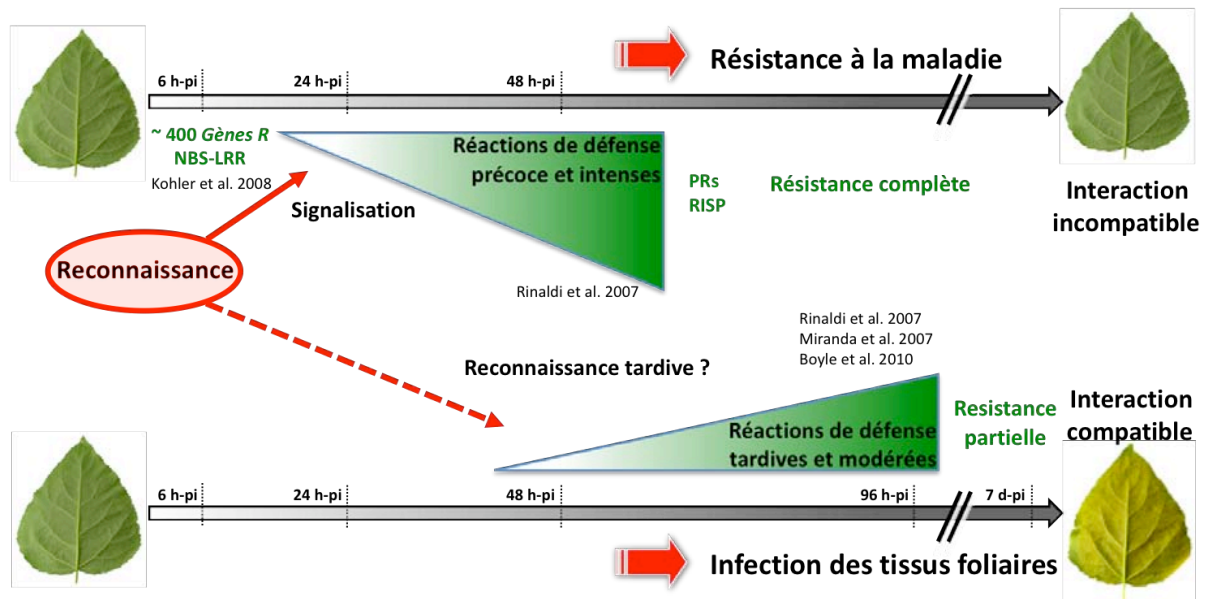


Figure 20: Réponses de défense du peuplier durant le processus infectieux de *Melampsora* Spp. (D'après Duplessis et al. 2009). Représentation schématique des interactions compatible et incompatible entre le peuplier et *Melampsora* spp.. Durant l'interaction incompatible, une reconnaissance précoce du pathogène est observée lorsque celui-ci colonise les cellules du parenchyme et tente de former les premiers haustoria. Cette reconnaissance spécifique du champignon (gènes *Avr*) fait intervenir des gènes de résistance (*R*) de la plante et conduit à de fortes réactions de défense (48 hpi) et une réaction d'hypersensibilité localisée au site d'infection. Les réactions de défense en découlant sont notamment marquées par l'accumulation des transcrits codant des protéines PR (pathogenesis-related) ainsi qu'une petite protéine sécrétée spécifique du peuplier (RISP). Durant l'interaction compatible, aucune réponse de défense n'est observée de manière précoce indiquant une absence de reconnaissance du pathogène lors des 1^{ères} étapes de l'infection. Néanmoins, un set similaire de gènes de défense est induit de manière tardive dans le cadre de l'interaction compatible (96 hpi et au delà).

transcrits induits par le peuplier en réponse au champignon. Alors que très peu de transcrits impliqués dans les réponses de défense de la plante ont été détectés dans le cadre de l'interaction compatible (48 hpi), nombreux ont été identifiés dans le cadre de l'interaction incompatible. Parmi ces transcrits les plus fortement induits au cours de l'interaction incompatible, plusieurs codent des protéines PR (Pathogenesis-Related) impliquées dans les réactions de défense ainsi qu'une petite protéine sécrétée induite par la rouille (RISP, Rust Induced Secreted Protein) (Rinaldi et al. 2007). L'un des résultats de cette analyse comparative est qu'il y aurait un délai dans la mise en place des réactions de défense de la plante au cours de l'interaction compatible résultant probablement de l'absence de reconnaissance de l'agent pathogène par la plante (Rinaldi et al. 2007; Duplessis et al. 2009). Une autre étude du transcriptome du peuplier a été menée dans le cadre de l'interaction entre *P. trichocarpa* x *P. deltoides* et *M. medusae* (interaction compatible) à des temps plus tardifs d'infection. Cette analyse a également révélé l'induction de transcrits codant des enzymes impliquées dans la voie des flavonoïdes ainsi que des transcrits codant des protéines de défense de type protéines PR suggérant une mise en place tardive des mécanismes de défense dans le cadre de l'interaction compatible (Miranda et al. 2007; Duplessis et al. 2009). Plus récemment, l'analyse comparative des transcrits de peuplier en réponse à *M. larici-populina* (interaction compatible), *M. medusae* (résistance partielle) ainsi qu'aux deux isolats inoculés de manière simultanée a confirmé l'induction tardive de nombreux gènes de défense lors de l'interaction Peuplier/*Melampsora* spp. avec des niveaux et des panels de gènes plus importants dans le cadre d'une résistance partielle (Azaiez et al. 2009; Boyle et al. 2010; Duplessis et al. 2009).

Ces réactions de défense, mises en place par la plante, sont gouvernées par la reconnaissance du pathogène par son hôte (§ I-1). Grâce au séquençage du génome du peuplier (Tuskan et al. 2006), de nombreuses familles de gènes impliquées dans la résistance et les réactions de défense de la plante en réponse au pathogène ont ainsi pu être identifiées. Plus de 400 gènes de résistance de type NB-LRR (§ I-4.1), pouvant participer aux mécanismes de reconnaissance (R-Avr) lors de l'interaction incompatible avec le Peuplier et *M. larici-populina*, ont ainsi pu être catalogués au sein de la séquence génomique de *P. trichocarpa* (Tuskan et al. 2006; Kohler et al. 2008; Yang et al. 2008). L'analyse fine de cette famille chez *P. trichocarpa* a notamment révélé l'expansion de certaines sous-classes possédant un domaine en doigt de zinc (BED finger) ou un signal d'adressage nucléaire (Kohler et al. 2008).

Alors que l'analyse fine des mécanismes moléculaires contrôlant la résistance à la rouille chez le peuplier a été initiée, les mécanismes moléculaires mis en jeu par le pathogène pour réaliser

l'infection des tissus foliaires restent à élucider. Le séquençage du génome de *M. larici-populina* (souche 98AG31), réalisé en 2007 par le JGI aux Etats-Unis en collaboration avec l'UMR 1136 du Centre INRA de Nancy et un consortium international de chercheurs, est un atout majeur pour mieux comprendre la biologie de l'interaction Peuplier/*Melampsora* et pour identifier précisément les mécanismes moléculaires qui contrôlent l'issue de cette interaction.

IV - La génomique

La génomique est une discipline qui a connu un essor considérable au cours de ces dernières décennies et qui vise à étudier le contenu des génomes des organismes vivants. Cette discipline est vaste et s'intéresse aux caractéristiques structurales des génomes (organisation des séquences codantes et non codantes) mais également à leurs caractéristiques fonctionnelles par l'identification de l'ensemble des gènes présents au sein de ces séquences. Ce dernier aspect vise plus particulièrement l'identification des fonctions qui sont codées par ces gènes et comprend l'étude de l'expression de l'ensemble des transcrits (transcriptomique), des protéines (protéomique) et des métabolites (métabolomique). La comparaison des fonctions contenues dans le génome de différents organismes (génomique comparative) permet de faire émerger les expansions et contractions de familles de gènes ainsi que la perte ou l'acquisition de nouvelles familles.

1- Etat des lieux sur le séquençage des génomes de champignons et oomycètes phytopathogènes

L'afflux continu de données génomiques issu du séquençage de plusieurs champignons et oomycètes pathogènes des plantes va permettre l'analyse comparative des séquences génomiques de ces différents organismes. L'identification de fonctions conservées par certains ainsi que les caractéristiques propres à chacun vont permettre une meilleure compréhension des différents modes de développement de ces différents organismes (Spanu et Kämper 2010).

Le séquençage des génomes des champignons phytopathogènes est récent puisqu'il a débuté en 2005 avec la publication du génome de *Magnaporthe oryzae* (*M. oryzae*), le champignon responsable de la pyriculariose du riz (Dean et al. 2005). Depuis, environ 40 génomes de champignons phytopathogènes ont été séquencés ou sont en cours de décryptage (**Tableau 4**). Ces génomes correspondent à des espèces représentatives des Basidiomycètes et des Ascomycètes. Parmi les Basidiomycètes, on retrouve notamment *U. maydis*, champignon responsable de la maladie du charbon du maïs (Kämper et al. 2006) ainsi que *P. graminis* f. sp. *tritici*, l'agent responsable de la rouille foliaire du blé. Chez les Ascomycètes, de nombreux génomes de champignons pathogènes des plantes ont également été séquencés tels que *M. oryzae* (Dean et al. 2005), différentes souches de *Fusarium* telles que *F. verticillioides*, *F. oxysporum* f. sp. *lycopersici* et *F. graminearum* (Cuomo et al. 2007, Ma et al. 2010), *B. graminis* (P. Spanu, Imperial College London), *Botrytis cinerea*, l'agent responsable de la pourriture grise

Tableau 4: Champignons phytopathogènes dont le génome a été séquencé ou est en cours de séquençage au 1^{er} juillet 2010. Les centres qui ont réalisé le séquençage de chaque organisme sont également présentés. Joint Genome Institute (JGI, USA), Broad Institute (USA), Génoscope (France), Centre National de la Recherche Scientifique (CNRS France), Université de Campinas (Unicamp, Brésil), Université de Western Cape (Afrique du Sud), Université de Washington à St Louis (WUSTL USA). Les champignons pathogènes des arbres sont présentés en gras.

Phylum/ <u>classe</u>	Espèce	Instituts de séquençage
Basidiomycota		
<u>Agaricostilbomycetes</u>	<i>Moniliophthora perniciosa</i>	Unicamp
	<i>Schizophyllum commune</i>	JGI
	<i>Heterobasidion annosum</i>	JGI
<u>Ustilaginomycetes</u>	<i>Ustilago maydis</i>	Broad
<u>Pucciniomycetes</u>	<i>Melampsora larici-populina</i>	JGI
	<i>Sporobolomyces roseus strain</i>	JGI
	<i>Puccinia graminis f. sp. tritici</i>	Broad
	<i>Phakopsora pachyrhizi</i>	JGI
	<i>Phakopsora meibomia</i>	JGI
	<i>Cronartium quercuum</i>	JGI
	<i>Microbotryum</i>	Genoscope et Broad
Archiascomycota	<i>Taphrina deformans</i>	JGI
Euascomycota		
<u>Sordariomycetes</u>	<i>Fusarium graminearum</i> (teleomorph <i>Gibberella zeae</i>)	Broad
	<i>Fusarium verticillioides</i> (teleomorph <i>Gibberella moniliformis</i>)	Broad
	<i>Fusarium oxysporum</i>	Broad
	<i>Podospora anserina</i>	CNRS et Genoscope
	<i>Cryphonectria parasitica</i>	JGI
	<i>Glomerella graminicola</i> (anamorph <i>Colletotrichum graminicola</i>)	Broad
	<i>Magnaporthe grisea</i>	Broad
	<i>Magnaporthe poae</i>	Broad
	<i>Gaeumannomyces graminis var. tritici</i>	Broad
	<i>Verticillium dahliae</i>	Broad
	<i>Verticillium albo-atrum</i>	Broad
	<i>Verticillium longisporum</i>	Broad
	<i>Ophiostoma clavigerum</i> (anamorph <i>Grosmannia clavigera</i>)	BCGSC
<u>Leotiomycetes</u>	<i>Sclerotinia sclerotiorum</i>	Broad
	<i>Botrytis cinerea</i> (teleomorph <i>Botryotinia fuckeliana</i>)	Genoscope et Broad
<u>Dothideomycetes</u>	<i>Alternaria brassicicola</i>	WUSTL
	<i>Stagonospora nodorum</i> (teleomorph <i>Phaeosphaeria nodorum</i>)	Broad
	<i>Pyrenophora tritici-repentis</i>	Broad
	<i>Mycosphaerella fijiensis</i> (anamorph <i>Pseudocercospora fijiensis</i>)	JGI
	<i>Cladosporium herbarum</i>	JGI
	<i>Mycosphaerella graminicola</i> (anamorph <i>Septoria tritici</i>)	JGI
	<i>Mycosphaerella populorum</i> (<i>Septoria musiva</i>)	JGI
	<i>Leptosphaeria maculans</i>	Genoscope
	<i>Cercospora zeae-maydis</i>	JGI
	<i>Dothistroma septosporum</i>	JGI
	<i>Setosphaeria turcica</i>	JGI
<i>Venturia inaequalis</i>	Univ. Western Cape	

(M-H. Lebrun, INRA Versailles), *L. maculans* (T. Rouxel, INRA Versailles) ou encore *Mycosphaerella graminicola*, l'agent responsable de la septoriose du blé (S. Goodwin, Purdue University). La plupart des pathogènes dont le génome est séquencé s'attaquent à des espèces annuelles et notamment des plantes de grande culture. Néanmoins, les génomes d'organismes phytopathogènes s'attaquant à des plantes pérennes sont également visés. C'est notamment le cas d'*Heterobasidion annosum*, champignon parasite des racines et responsable de la « maladie du rond » ayant des effets dévastateurs sur les résineux (J. Stenlid, Swedish university of agriculture sciences) ou encore de *Grosmannia clavigera*, champignon pathogène du pin responsable du bleuissement et associé au dendroctone pondero (Diguistini et al. 2009). Alors que chez les agents de rouilles, le séquençage du génome de *M. larici-populina* est achevé (S. Duplessis et F. Martin, INRA Nancy), celui des agents des rouilles du chêne *C. quercuum* f. sp. *fusiforme* (T. Kubisiak, US Forest service) et du pin *C. ribicola* (R. Hamelin, communication personnelle) est actuellement en cours de séquençage. C'est également le cas de *Mycosphaerella populorum*, un autre pathogène du peuplier responsable du chancre septorien (S Goodwin, Purdue University).

Le séquençage de plusieurs génomes d'oomycètes a également ouvert de nouvelles perspectives dans la compréhension de l'origine évolutive de la pathogénie ainsi que dans l'identification des mécanismes moléculaires impliqués dans la virulence du pathogène. Plusieurs espèces d'oomycètes pathogènes des plantes de grande culture ont par exemple été séquencées tels que *Phytophthora capsici*, *P. infestans* et *P. sojae* infectant respectivement le poivron, la pomme de terre et le soja (Lamour et al. 2007). Les séquences génomiques de *Phytophthora ramorum*, pathogène affectant le chêne ainsi que certaines plantes ornementales comme le rhododendron, et *H. parasitica*, pathogène des crucifères sont également disponibles (Tyler et al. 2006; McDowell, non publié). L'analyse des génomes de *P. sojae* et *P. ramorum*, premiers oomycètes à avoir été séquencés, a été publiée en 2006 (Tyler et al. 2006) suivie en 2009 par la publication de l'analyse du génome de *P. infestans* (Haas et al. 2009). Très récemment, le décryptage de la séquence génomique de *Pythium ultimum*, un oomycète nécrotrophe s'attaquant à un large spectre d'hôtes a permis de révéler des signatures caractéristiques de l'espèce *Pythium* ainsi qu'un large répertoire d'effecteurs potentiels (Lévesque et al. 2010).

Le séquençage de ces génomes nécessite des moyens scientifiques, techniques et financiers conséquents. Celui-ci s'effectue le plus souvent dans des grands centres de séquençage tels que le Broad Institute (USA), le JGI (USA) ou le Génoscope (France), capables de séquencer à haut débit de nombreux génomes. Néanmoins, les techniques de séquençage nouvelle génération sont en train de changer la donne puisque de nombreuses start-up sont maintenant équipées de

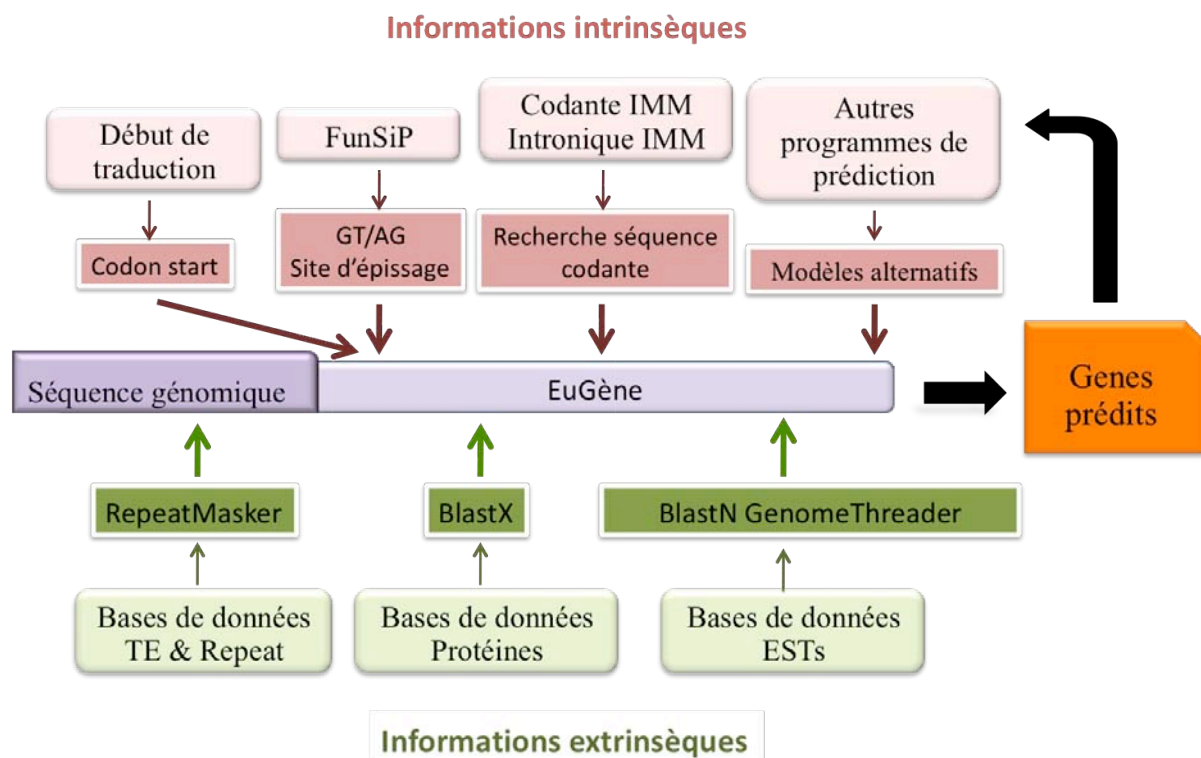


Figure 21: Exemple de plateforme bioinformatique utilisée pour la prédiction automatique de gène (EuGene). La prédiction de gène est basée sur des informations intrinsèques et extrinsèques de la séquence génomiques. Les informations intrinsèques correspondent aux propriétés locales de la séquence: la présence des codons start et stop, de sites d'épissage (début et fin d'intron) ou encore la signature exanucléotidique de la séquence (%GC) vont permettre de rechercher la séquence codante du gène. Les informations extrinsèques vont permettre la comparaison de la séquence avec des bases de données de séquences (protéines ou ESTs ...) afin d'affiner la prédiction. Les modèles de gène ainsi générés sont ensuite utilisés pour optimiser en retour les méthodes et algorithmes utilisés pour la prédiction automatique

nouveaux automates (§ IV-3), permettant un séquençage plus rapide des fragments d'ADN et à moindre coût. Les données génomiques générées à la suite du séquençage d'un organisme nécessitent un traitement bioinformatique indispensable pour leur interprétation. Enfin l'annotation des éléments codants identifiés dans ces séquences est essentielle pour son analyse fonctionnelle. Cet effort d'annotation est généralement réalisé par des biologistes experts au sein de consortiums internationaux.

2 - L'annotation des Génomes

La prédiction bioinformatique des gènes à partir d'une séquence génomique vise à repérer au sein de celle-ci les régions ayant une forte probabilité de coder un gène ainsi que l'assignation d'une fonction à ce gène, c'est l'annotation automatique des génomes (Nicolas et Chiapello 2004; Gibrat et Marin 2004). Cette annotation automatique nécessite de puissants algorithmes (Eugene, Fgenesh, GAZE ...) capables de traiter la séquence génomique pour identifier les caractéristiques typiques des gènes et de les comparer à de larges bases de données de séquences codantes. Cette annotation automatique est composée de plusieurs approches (**Figure 21**). Une première approche dite intrinsèque consiste en l'identification de propriétés locales de la séquence telles que la présence de nucléotides spécifiques définissant le début (GT) et la fin (AG) d'un intron, la présence des codons start (ATG) et stop (TAA, TGA, TAG) ou encore la signature exanucléotidique de la séquence (%GC). Une seconde, dite extrinsèque repose sur la comparaison avec des bases de données de séquences issues des tissus de l'organisme séquencé ou des bases de données internationales (transcrits, ESTs, ADNc pleine longueur, tags peptidiques) afin de révéler des homologies. Cette annotation automatique des génomes est essentielle pour générer rapidement une vue générale de l'espace potentiellement codant au sein d'un génome. Néanmoins, les modèles de gènes générés par cette approche automatique restent des prédictions qui doivent être considérées avec précaution et qui nécessitent une validation ou une correction par des biologistes experts. Ce second niveau d'annotation dite expertisée est une étape clé dans le processus d'annotation des génomes puisqu'il permet d'avoir une vision plus juste de l'espace codant de l'organisme considéré mais aussi pour optimiser en retour les méthodes et algorithmes utilisés pour la prédiction automatique. L'annotation manuelle repose sur la recherche d'homologues dans les bases de données, l'alignement du modèle de gène prédits avec ces homologues, la vérification des codons start et stop ainsi que les pieds d'introns. L'alignement du modèle de gène avec les séquences de transcrits (ESTs) de l'organisme séquencé, lorsque ceux-ci sont disponibles, est le plus souvent essentiel pour valider la structure du gène (**Figure 22**). Enfin, dans le cas de gènes appartenant à des familles multigéniques, la comparaison des différents membres entre eux et avec leurs

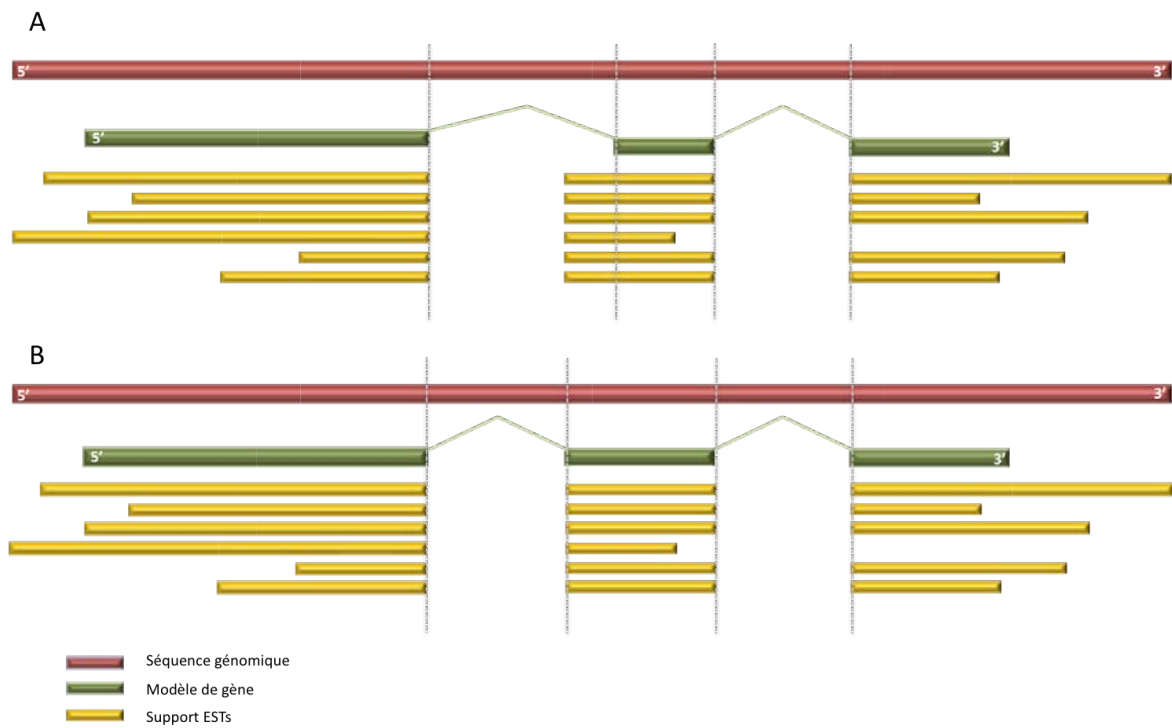


Figure 22: Visualisation d'un modèle de gène **A**, avant et **B**, après annotation expertisée. **A**, Le modèle de gène (en vert) prédit de manière automatique n'a pas été prédit correctement puisque l'exon 2 est tronqué. Cette mauvaise prédiction est ici facilement identifiable du fait de la présence d'ESTs (en jaune) s'alignant avec la séquence génomique (violet).

séquences homologues chez des organismes phylogénétiquement proches facilite les éventuelles corrections à apporter.

3 - La transcriptomique

La transcriptomique est une discipline qui a pour objet l'étude de l'ensemble des transcrits d'un organisme (organe tissu, cellule) à un temps précis et dans une condition donnée. Les techniques les plus communément utilisées pour réaliser de telles analyses sont les puces à oligonucléotides de type Affymetrix et NimbleGen ou les méthodes de séquençage de nouvelle génération telles que les technologies 454 de Roche ou CRT d'Illumina (Cyclic Reversible Termination). Dans cette partie nous nous focaliserons sur les puces à oligonucléotides NimbleGen et le pyroséquençage 454 qui ont été les techniques utilisées au cours de ce travail de thèse.

Les puces à oligonucléotides sont des supports solides sur lesquels sont fixés par synthèse *in situ* des oligonucléotides qui sont spécifiques de chaque gène prédit au sein du génome de l'organisme considéré. Deux types de puces, développés par les sociétés Affymetrix et NimbleGen, sont actuellement largement utilisés par la communauté scientifique. Les puces ADN développées par NimbleGen comportent l'ensemble des gènes prédits pour un organisme et chacun de ces gènes est représenté par 1 à 5 oligonucléotides différents et indépendants dessinés à partir de la séquence codante de chacun des gènes présents chez l'organisme. La synthèse de ces oligonucléotides (60 paires de bases), s'effectue grâce à un système de miroirs microscopiques capables de focaliser la lumière sur les sites de synthèse des oligonucléotides. Afin d'étudier le niveau d'expression de l'ensemble des transcrits d'un organisme dans une condition donnée par rapport à une condition de référence, les ARN totaux sont tout d'abord extraits pour chacune de ces conditions. Les ADNc sont ensuite synthétisés à partir de ces ARN totaux, marqués avec un fluorochrome puis hybridés avec la puce. Plus un ADNc sera en proportion importante au sein de la population, plus il va s'hybrider aux oligonucléotides qui lui sont spécifiques et plus l'intensité du signal détecté sera importante. Les intensités mesurées pour chaque gène et dans chaque condition sont ensuite normalisées afin de révéler des différences d'intensités qui reflètent les variations du niveau d'expression des gènes entre la condition testée et la condition de référence. L'inconvénient de cette technique repose sur le fait que ces puces à oligonucléotides ne comportent que les gènes prédits automatiquement et/ou annotés de l'organisme. Les gènes non détectés lors des étapes d'annotation ne seront donc pas représentés sur la puce.

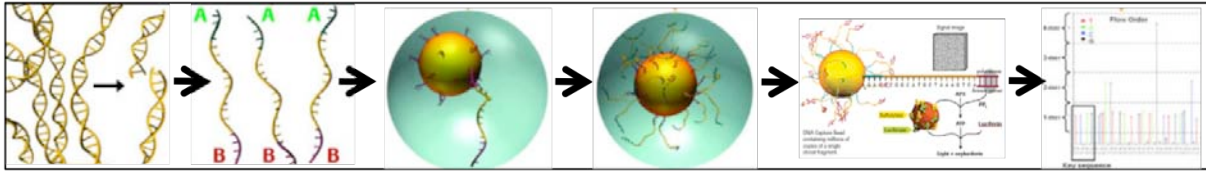


Figure 23: Principe du pyroséquençage 454 (Roche). Des adaptateurs sont liés aux extrémités 5' et 3' des molécules d'ADN. L'adaptateur A va permettre l'accrochage aux billes qui sont en large excès et qui ne capteront ainsi qu'un seul fragment d'ADN simple brin. Les billes sont ensuite placées dans une émulsion contenant le mélange réactionnel nécessaire à l'amplification des fragments. A la fin de l'amplification, chaque bille contient des milliers de séquences d'ADN simple brin identiques amplifiées à partir d'une seule séquence. Les billes sont ensuite placées dans des microplaques contenant des micropuits ne pouvant accueillir qu'une seule bille. Les fragments amplifiés sur chaque bille sont ainsi séquencés séparément des autres fragments par fixation d'une amorce complémentaire de l'adaptateur B puis par ajout successif d'une base dans un ordre précis (détail Figure 24). Pour chaque base complétant l'ADN cible, un signal luminescent est émis et enregistré par une caméra CCD.

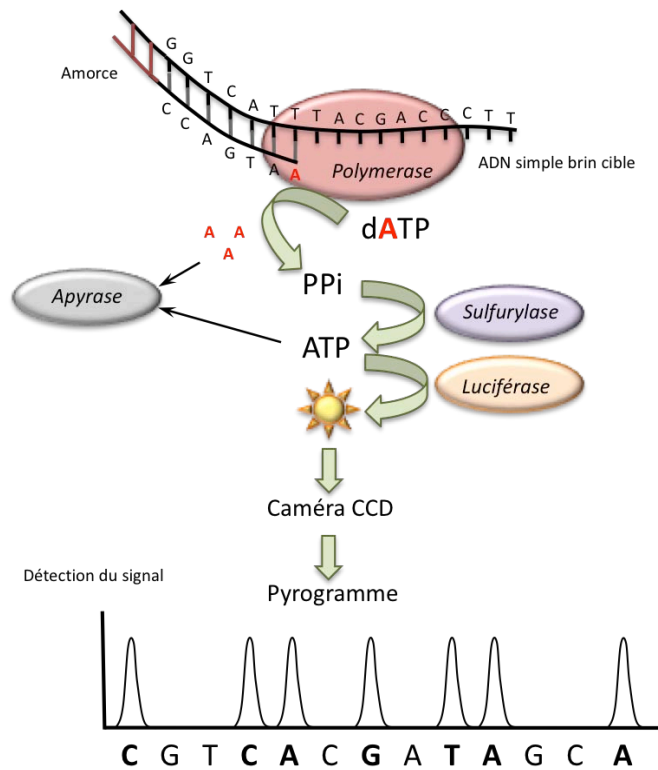


Figure 24: Réaction de pyroséquençage (Adapté de Lamoril et al. 2008; Armougom et Raoult 2009). Le principe de base de cette réaction de séquençage consiste à hybrider une amorce (en rouge) complémentaire de l'adaptateur de l'ADN cible puis à ajouter le mélange réactionnel contenant les enzymes et les substrats nécessaires à la réaction. Les nucléotides sont ensuite ajoutés l'un après l'autre (bases A puis bases C puis bases G puis bases T ...) dans le milieu réactionnel et chaque base complétant la séquence cible aboutit à la libération d'un pyrophosphate (PPi). L'ATP sulfurylase transforme ensuite ce PPi en ATP qui est directement utilisé par une luciférase afin de transformer la luciférase en oxyluciférase. Cette réaction libère de la lumière qui est captée par une caméra CDD (Charge-Coupled Device) reproduisant l'intensité émise sous forme d'un pic sur le pictogramme. L'apyrase dégrade ensuite les nucléotides non incorporés permettant ainsi le séquençage de nouvelles bases (Lamoril et al 2008).

Contrairement aux puces à ADN, le séquençage à très haut débit permet d'avoir accès à l'ensemble des gènes exprimés chez un organisme dans une condition donnée. Seules la profondeur de séquençage et la qualité de l'assemblage peuvent ici limiter l'identification de transcrits d'intérêt. Ce type de séquençage présente également un intérêt budgétaire considérable par rapport au séquençage Sanger de première génération. Par exemple, pour une même couverture de génome humain, le coût d'un séquençage Sanger est estimé à plus de 70 fois le coût d'un séquençage nouvelle génération réalisé avec les automates 454 ou Illumina (Metzer 2010). Pour ces raisons, les nouvelles techniques de séquençage sont en plein essor et sont actuellement en train de supplanter l'utilisation des puces à ADN et le séquençage Sanger. Parmi ces techniques de séquençage nouvelle génération, le principe du pyroséquençage a été décrit pour la première fois en 1985 (Ahmadian et al. 2006). Il permet l'analyse d'une large variété d'échantillons tels que l'ADN génomique, les produits PCR, ou encore les ADN complémentaires. Il s'agit d'une méthode permettant d'analyser la synthèse d'ADN cible en temps réel; on parle ainsi de séquençage par synthèse d'ADN (Ronaghi et al. 1998). Certaines sociétés commercialisent des automates de pyroséquençage permettant le séquençage de fragments d'ADN de plus en plus longs. C'est notamment le cas de l'automate 454 développé par la société Roche qui permettait au départ le séquençage de courts fragments d'ADN d'environ 100 puis 250 paires de bases (technologie GS-FLX) et qui aujourd'hui permet le séquençage de fragments d'ADN de 350-400 paires de bases (Technologie GS-FLX-Titanium). Le pyroséquençage réalisé avec l'automate 454 est illustré dans la **figure 23** et la réaction de séquençage est présentée dans la **figure 24**.

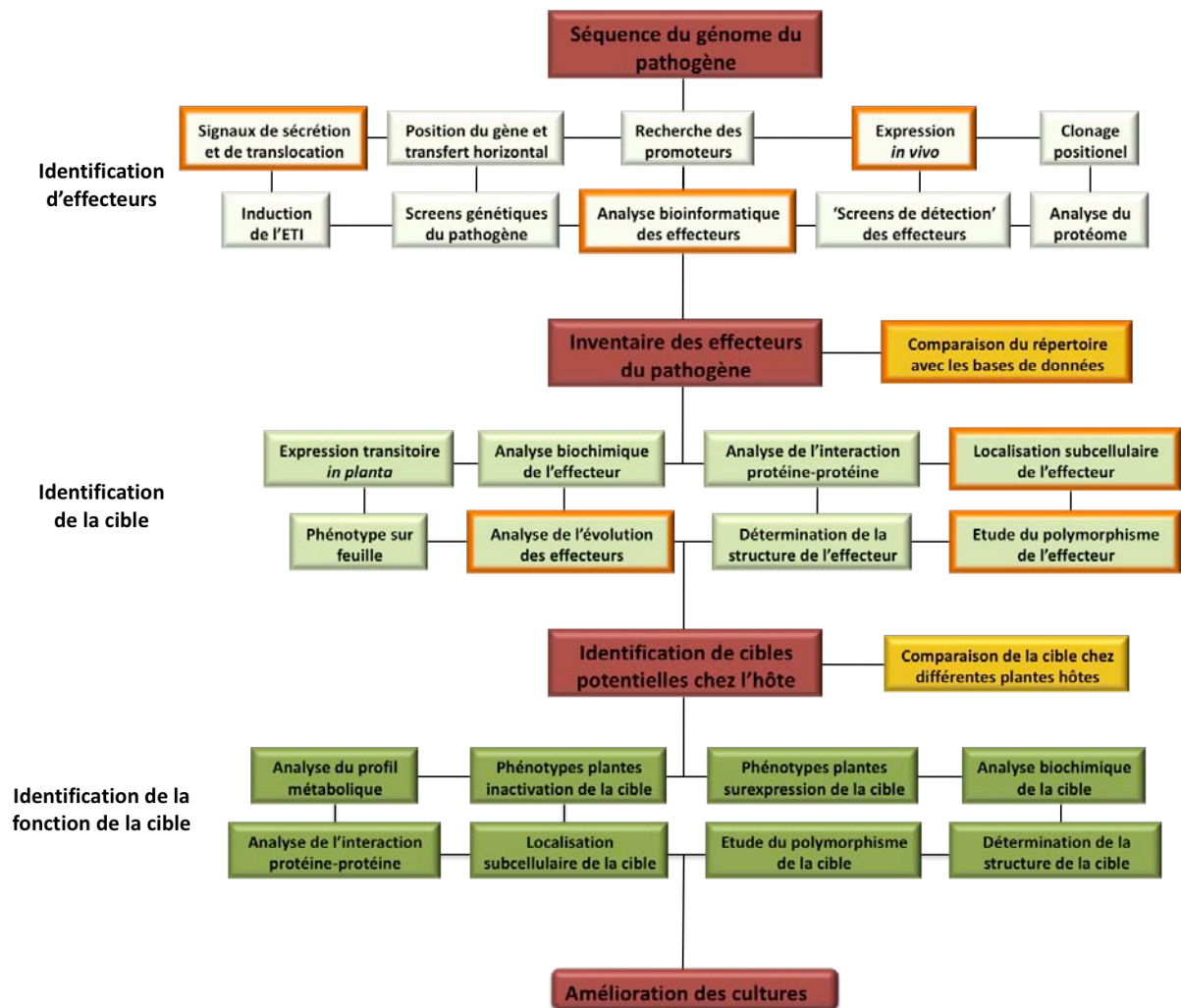


Figure 25: « Feuille de route » pour l'identification d'effecteurs et de leurs cibles (Adapté de Alfano 2009). L'organigramme montre les différentes approches utilisables pour identifier des effecteurs (blanc), leurs cibles (vert clair) et la fonction de leurs cibles (vert foncé). Les étapes (initiale, intermédiaire et finale) sont indiquées en rouge sur l'organigramme. Les rectangles jaunes représentent les comparaisons qui peuvent être réalisées après avoir dressé le catalogue des effecteurs potentiels ou après avoir identifié la cible de l'effecteur. Les lignes reliant les différents rectangles indiquent qu'une combinaison de plusieurs approches peut être utilisée afin d'identifier les effecteurs et leurs cibles. Le rectangle 'Screen de détection des effecteurs' correspond à la fusion d'un effecteur candidat avec le signal de sécrétion de type III d'un effecteur bactérien permettant ainsi l'injection de l'effecteur dans la cellule hôte via la bactérie et son système de sécrétion de type III (Sohn et al. 2007). **Les approches utilisées au cours de cette thèse sont repérées par un cadre orange.**

V - Objectifs de la thèse

1 - Contexte

Au cours de ces dernières années, de nombreuses connaissances ont été acquises sur le pathosystème peuplier-*M. larici-populina*, notamment sur l'épidémiologie de la maladie. Le développement de marqueurs moléculaires (microsatellites) de *M. larici-populina* a permis de répondre à des questions de génétique des populations telles que la quantification des flux de gènes entre les compartiments peupleraies sauvages et peupleraies cultivées ou encore la structure génétique des populations du champignon à l'échelle continentale et intercontinentale (Barrès et al. 2006; Barrès et al. 2008; Xhaard et al. 2009). De plus, le suivi temporel de l'évolution des pathotypes de *M. larici-populina* a également permis de révéler une évolution rapide des populations ainsi qu'une augmentation de la fréquence des virulences nécessaires à l'infection de certains cultivars (Pinon et Frey 2004). Par ailleurs, certains des déterminants génétiques impliqués dans les réactions de défense de l'hôte au cours de l'interaction Peuplier/*M. larici-populina* ont pu être identifiés (§ III-3) (Rouhier et al. 2004; Rinaldi et al. 2007; Miranda et al. 2007; Kohler et al. 2008; Duplessis et al. 2009; Asaiez et al. 2009; Boyle et al. 2010). La disponibilité du génome du peuplier a notamment permis de dresser le catalogue des gènes de résistance de l'hôte (Kohler et al. 2008) ainsi que d'identifier des gènes clés impliqués dans la résistance du peuplier à *M. larici-populina* (Rinaldi et al. 2007). Après le séquençage du génome du peuplier, le projet de séquençage de la souche 98AG31 de *M. larici-populina*, proposé au JGI et accepté en mai 2005, a ouvert de nouvelles perspectives pour la compréhension des mécanismes moléculaires qui régissent l'interaction avec le peuplier.

2 - Projet de thèse

Mon projet de thèse, réalisé au sein de l'équipe Ecogénomique des Interactions de l'UMR 1136 Interactions Arbres/Microorganismes, a consisté en la caractérisation des petites protéines sécrétées (SSPs) de *M. larici-populina*. Les fonctions de ces SSPs pourraient être variées et impliquées dans différents processus de développement du champignon. Certaines pourraient avoir un rôle structural, d'autres (protéines effectrices) serviraient de leurres afin de contourner les mécanismes de défense de la plante et quelques unes seulement pourraient être reconnues par le système immunitaire de l'hôte et déclencher des réactions de défense (facteurs d'avirulence). En plus de rechercher des SSPs potentiellement impliquées dans les différentes étapes du processus infectieux du champignon, nous avons porté une attention particulière à l'identification de ces protéines effectrices au sein du sécrétome de *M. larici-populina* (**Figure 25**).

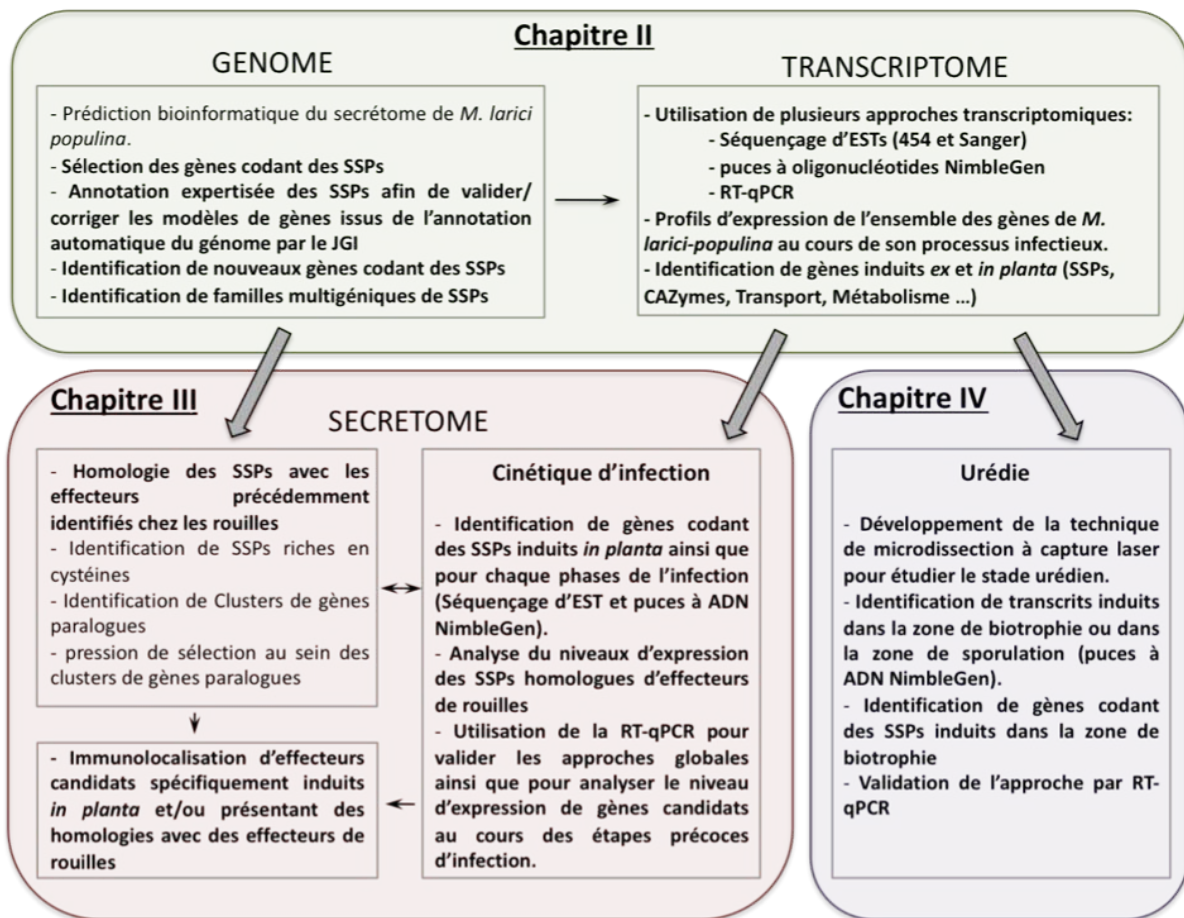


Figure 26: Représentation schématique de mon projet de thèse. Les tâches qui m'ont été confiées sont indiquées en gras dans les encadrés. Les trois sous-parties qui apparaissent en couleur correspondent aux trois chapitres de résultats présentés dans cette thèse. SSP = petites protéines sécrétées.

Les objectifs principaux de mon projet de thèse ont été (**Figure 26**):

(1) Identifier, annoter et caractériser le catalogue de gènes codant des protéines sécrétées de *M. larici-populina*.

Ce premier volet de ma thèse a consisté en l'analyse systématique des petites protéines sécrétées (SSPs, Small Secreted Proteins) de *M. larici-populina*. Avec la séquence du génome de *M. larici-populina* à disposition, l'identification de l'ensemble des gènes codant des SSPs a été réalisée par l'intermédiaire d'outils informatiques repérant des structures spécifiques de ces protéines (présence d'un signal peptide, absence de domaines transmembranaires, taille de la protéine). J'ai participé à l'annotation expertisée de l'ensemble de ces SSPs afin de valider ou corriger les modèles de gènes issus de l'annotation automatique du génome. A l'aide de séquences de transcrits du champignon (ESTs) exprimés dans les spores ou lors de l'infection des tissus de l'hôte, des structures géniques ont pu être confirmées, de nouveaux gènes candidats ont été identifiés et une analyse fine de l'organisation de ces gènes au sein du génome a été établie (familles multigéniques, clusters de gènes paralogues, pression de sélection au sein des clusters de gènes paralogues). J'ai porté une attention particulière aux gènes codant des fonctions potentiellement impliquées dans le processus infectieux à partir des séquences déjà décrites chez d'autres agents de rouilles comme *M. lini* ou *U. fabae*. J'ai notamment ciblé les gènes et protéines exprimés au niveau de l'haustorium, la structure d'infection mise en place par ces pathogènes biotrophes obligatoires et qui joue vraisemblablement un rôle clé pour la sécrétion de protéines effectrices.

(2) Analyser le niveau d'expression des gènes codant des SSPs de *M. larici-populina*

Dans un second volet, j'ai étudié le niveau d'expression de l'ensemble des gènes codant ces SSPs par différentes approches transcriptomiques afin d'identifier des gènes marqueurs de chaque stade de développement du champignon ainsi que pour mettre à jour des transcrits spécifiquement et fortement accumulés durant le processus infectieux pouvant représenter des effecteurs potentiels.

- Une première approche a consisté en l'analyse d'ESTs produites à partir d'une banque d'ADNc issue d'un mélange d'urédospores dormantes et germées séquencées par le JGI et l'analyse de séquences exprimées dans des tissus foliaires à 96 et 168 hpi obtenues par la technique de pyroséquençage 454 (§ IV-3) à partir d'ADNc de feuilles de peuplier infectées. Cette approche, pour laquelle j'ai préparé le matériel biologique, m'as permis suite à l'analyse

bioinformatique réalisée au sein de l'équipe (A. Kohler, E. tisserant, S. Duplessis), d'établir les listes de transcrits codant des SSPs spécifiquement accumulés *in planta* ou exprimés dans les urédospores dormantes ou en germination du champignon.

- Afin d'analyser plus finement les modifications de l'expression génique des SSPs fongiques, j'ai réalisé une cinétique d'infection du peuplier Beaupré par la souche de *M. larici-populina* en collaboration avec Pascal Frey de l'équipe Ecologie des Champignons Pathogènes Forestiers de l'UMR1136 IAM (urédospores dormantes, urédospores germées, 2, 6, 12, 24, 48, 96 et 168 heures post inoculation). J'ai extrait les transcrits de ces tissus et produit les ADNc correspondants. Les ADNc de ces différents stades (urédospores dormantes, urédospores germées, 24 hpi, 48 hpi, 96 hpi et 168 hpi) ont été hybridés avec des puces à oligonucléotides (NimbleGen) comportant la quasi totalité des gènes de *M. larici-populina*.

- Afin de confirmer les profils d'expression obtenus par ces approches globales, j'ai utilisé la RT-qPCR pour mesurer les différences d'expression de gènes candidats d'intérêt lors de la cinétique d'infection. Cette approche a nécessité une étape préalable de validation de gènes de référence dont l'expression est constante aux différentes étapes d'infection et qui permettent de rapporter le niveau d'expression des gènes candidats à un standard pour établir des profils d'expression comparables au cours de la cinétique d'infection.

Au delà de l'analyse des transcrits codant des SSPs, ces données d'expression (ESTs et puces à oligonucléotides NimbleGen) ont permis, au sein du *Melampsora* Genome Consortium, de réaliser le décryptage d'autres aspects du développement et de la biologie du champignon comme le transport ou le métabolisme

(3) Localiser les effecteurs candidats *in planta*

Afin de me familiariser avec les structures d'infection formées *in planta* par le champignon, j'ai dans un premier temps entrepris par microscopie confocale, l'étude des structures du pathogène dans les tissus de la plante. J'ai ensuite sélectionné, à partir des approches (1) et (2), des gènes codant des effecteurs candidats au sein du secrétome de *M. larici-populina*. Dans un second temps, j'ai utilisé la technique d'immunolocalisation afin de visualiser l'accumulation de ces SSPs au niveau des structures fongiques formées *in planta* au cours de l'interaction avec le peuplier.

(4) Développer la technique de microdissection à capture laser dans le cadre de la formation de l'urédie

Le stade urédien (168 hpi) correspond à un mélange de structures fongiques distinctes formées *in planta* (hyphes d'infection, haustoria, tissu sporifère, spores nouvellement formées). Il apparaissait donc essentiel de pouvoir discriminer les gènes du champignon (notamment ceux codant des SSPs) accumulés dans la zone de biotrophie (hyphes d'infection et haustoria) de ceux accumulés dans la zone de sporulation (tissu sporifère et spores) à ce stade tardif d'infection. Après avoir optimisé les protocoles de fixation du matériel biologique et d'extraction des acides nucléiques à partir de matériel fixé, j'ai utilisé la technique de microdissection à capture laser, couplée avec l'utilisation des puces à oligonucléotides NimbleGen afin d'établir le profil d'expression des transcrits de *M. larici-populina* au sein de tissus microdisséqués des zones de biotrophie et de sporulation du champignon.

Chapitre II : Analyse globale du génome et du transcriptome de *Melampsora larici-populina*

Chapitre II : Analyse globale du génome et du transcriptome de *Melampsora larici-populina*

M. larici-populina est un champignon basidiomycète de la classe des *Pucciniales* responsable la maladie de la rouille foliaire du peuplier. Ce pathogène entraîne une réduction de croissance et un affaiblissement général des plants provoquant des pertes économiques pouvant aller jusqu'à 60% (Miot et al. 1999). Les programmes de sélections variétales conduits depuis les années 70 ont permis la sélection de peupliers hybrides issus du croisement entre *P. trichocarpa* et *P. deltoides* (hybrides interaméricains) et entre *P. deltoides* et *P. nigra* (hybrides euraméricains). Ces cultivars, en plus de présenter une résistance complète à *M. larici-populina*, ont également été sélectionnés pour leur croissance rapide et leur vigueur. Néanmoins, cette résistance complète au pathogène présente en général un déterminisme monogénétique et dominant (relation gène-pour-gène) qui le rend facilement contournable par l'agent pathogène (Frey et Pinon 2004). Certaines résistances complètes se sont en effet totalement effondrées après seulement quelques années de culture. Le contournement de la résistance R7 du peuplier par *M. larici-populina*, entraînant des épidémies de rouille sans précédent dans les années 90, a par exemple eu de lourdes conséquences sur la populiculture française. Devant un tel constat, la stratégie d'amélioration variétale du peuplier s'oriente désormais vers la sélection de clones présentant des résistances quantitatives à la maladie dont le contrôle est supposé polygénique et donc plus difficilement contournable par le champignon. A ce jour, la totalité des cultivars de peuplier disponibles sur le marché restent néanmoins sensibles à *M. larici-populina* et il apparaît essentiel de comprendre les mécanismes d'interaction entre le champignon et sa plante-hôte afin de développer des moyens de lutte efficaces contre la maladie. Le séquençage du génome du peuplier par le Joint Genome Institute (JGI, Department Of Energy, USA) a permis de dresser le catalogue des gènes de résistance de la plante ainsi que d'identifier les gènes clés impliqués dans la défense de l'arbre en réponse au champignon (Tuskan et al. 2006; Rinaldi et al. 2007; Kohler et al. 2008). Bien que l'identification des mécanismes moléculaires contrôlant la résistance à la rouille foliaire chez le peuplier ait permis une avancée considérable dans la compréhension de la maladie, il apparaît néanmoins essentiel d'identifier les mécanismes moléculaires mis en jeu par le champignon afin d'infecter les tissus foliaires. Le projet de séquençage du génome de *M. larici-populina* par le JGI est un atout majeur pour identifier les signatures caractéristiques du champignon pouvant refléter son statut de biotrophe obligatoire ainsi que pour mettre à jour les fonctions impliquées dans le processus infectieux de l'agent pathogène. L'analyse de la séquence génomique de *M. larici-populina*, coordonnée par l'UMR1136 Interactions Arbres-Microorganismes de l'INRA Nancy (F. Martin et S. Duplessis) est réalisée au sein d'un consortium

international rassemblant une trentaine de membres (<http://genome.jgi-psf.org/Mellp1/Mellp1.home.html>). La disponibilité de ce génome a permis de dresser le catalogue de l'ensemble des gènes présents au sein de la séquence génomique du champignon, d'identifier les familles multigéniques ainsi que les gènes codant des petites protéines sécrétées (SSPs), des enzymes de dégradation ou encore des transporteurs de sucres ou d'acides aminés dont le rôle pourrait être essentiel durant le processus infectieux. L'annotation experte de ces familles a été réalisée au sein du consortium afin de confirmer/d'infirmier l'annotation automatique. Deux banques d'ESTs issus de spores/spores germés ainsi que de feuilles infectées (96 et 168 hpi) ont permis d'améliorer considérablement ce processus d'annotation et d'identifier des transcrits spécifiques de la phase de biotrophie. Afin d'identifier plus finement les gènes clés du processus infectieux, le niveau d'expression de l'ensemble des gènes du champignon a été suivi *ex planta* et *in planta* au cours de la cinétique d'infection de *M. larici-populina* par l'utilisation de puces à ADN NimbleGen.

Le manuscrit intitulé : « **Obligate biotrophy features unraveled by the genomic analysis of the rust fungi** », actuellement en cours de préparation, est inclus dans ce chapitre de thèse et présente les résultats de cette analyse du génome. Au cours de cette analyse, j'ai réalisé l'annotation des SSPs ainsi que la préparation des échantillons nécessaires à l'analyse transcriptomique (puces NimbleGen).

Le manuscrit intitulé : « **Transcript profiling of *Melampsora larici-populina* during timecourse infection of poplar leaves reveals major insights into rust biotrophy** », actuellement en cours de préparation, est également présenté dans ce chapitre de thèse et décrit l'analyse détaillée du transcriptome de *M. larici-populina* au cours de la cinétique d'infection sur feuilles de peuplier. Dans le cadre de cette étude, j'ai réalisé la préparation des échantillons nécessaires à l'analyse transcriptomique (puces NimbleGen et RT-qPCR) et j'ai contribué à l'analyse des données ainsi qu'à la rédaction du manuscrit.

L'analyse des profils d'expression par RT-qPCR de gènes candidats (codant des SSPs ou d'autres fonctions) a nécessité au préalable la validation de gènes de référence exprimés de manière stable durant le processus infectieux. Le manuscrit intitulé : « **Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection of poplar leaves** », publié dans la revue **Physiological and Molecular Plant Pathology** est aussi présenté dans ce chapitre.

**Obligate Biotrophy Features Unraveled by the Genomic Analysis
of the Rust Fungi**

Duplessis S, Cuomo CA, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, Chiu R, Couthinho P, Feau N, Field M, Frey P, Gelhaye E, Goldberg J, Grabherr M, Kodira CD, Kohler A, Kues U, Lindquist EA, Lucas S, Mago R, Mauceli E, Morin E, Murat C, Pangilinan JL, Park R, Pearson M, Quesneville H, Rouhier N, Sakthikumar S, Salamov AA, Schmutz J, Selles B, Shapiro H, Tangay P, Tuskan GA, Henrissat B, Van de Peer Y, Rouzé P, Ellis JG, Dodds PN, Schein J, Zhong S, Hamelin RC, Grigoriev IV, Szabo LJ, Martin F.

Publication soumise à la revue Proceedings of the National Academy of Sciences, U.S.A.

Les données supplémentaires de cette publication sont disponibles sur le site web:
http://mycor.nancy.inra.fr/pub/SOM_these_Hacquard

Obligate Biotrophy Features Unraveled by the Genomic Analysis of Rust Fungi

Sébastien Duplessis^{a,1,*}, Christina A. Cuomo^{b,1,*}, Yao-Cheng Lin^c, Andrea Aerts^d, Emilie Tisserant^a, Claire Veneault-Fourrey^a, David L. Joly^e, Stéphane Hacquard^a, Joelle Amselem^f, Brandi L. Cantarel^g, Readman Chiu^h, Pedro Couthinho^g, Nicolas Feu^{e2}, Matthew Field^h, Pascal Frey^a, Eric Gelhaye^a, Jonathan Goldberg^b, Manfred Grabherr^b, Chinnappa D. Kodira^{b,3}, Annegret Kohler^a, Ursula Kües^j, Erika A. Lindquist^d, Susan Lucas^d, Rohit Magoi, Evan Mauceli^b, Emmanuelle Morina^a, Claude Murat^a, Jasmyn L. Pangilinan^d, Robert Park^k, Matthew Pearson^b, Hadi Quesneville^f, Nicolas Rouhier^a, Sharadha Sakthikumar^b, Asaf A. Salamov^d, Jeremy Schmutz^d, Benjamin Selles^a, Harris Shapiro^d, Philippe Tangay^e, Gerald A. Tuskan^l, Bernard Henrissat^g, Yves Van de Peer^c, Pierre Rouzé^c, Jeffrey G. Ellis^j, Peter N. Dodds^j, Jacqueline E. Schein^h, Shaobin Zhong^{m,4}, Richard C. Hamelin^e, Igor V. Grigoriev^d, Les J. Szabo^{n,m,1}, Francis Martin^{a,1}

^aINRA, UMR 1136 INRA/Nancy Université, Interactions Arbres/Micro-organismes, Centre INRA de Nancy, 54280 Champenoux, France. ^bThe Broad Institute of MIT and Harvard, Cambridge, MA 02142. ^cDepartment of Plant Systems Biology, VIB, Ghent, Belgium. ^dUS Department of Energy Joint Genome Institute, Walnut Creek, CA 94598. ^eNatural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du PEPS, PO Box 10380, Stn Sainte-Foy, Québec, QC, G1V 4C7 Canada. ^fINRA URGI, Route de St Cyr 78026 Versailles Cedex, France. ^gUMR 6098 CNRS-Universités Aix-Marseille I & II, Marseille, France. ^hGenome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada. ⁱUniversity of Goettingen, Buisgen-Institute, Division of Molecular Wood Biotechnology and Technical Mycology, Buisgenweg, 37077 Göttingen, Germany. ^jDivision of Plant Industry, Commonwealth Scientific and Industrial Research Organization, Canberra, ACT 2601, Australia. ^kUniversity of Sydney, Plant Breeding Institute Cobbitty, PMB 11, Camden, NSW 2570, Australia. ^lEnvironmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6422. ^mDepartment of Plant Pathology, University of Minnesota, St. Paul, MN 55108. ⁿCereal Disease Laboratory, USDA-Agricultural Research Service, St. Paul, MN 55108.

Authors contribution: S.D., C.A.C., L.J.S., I.V.G., G.A.T. and F.M. designed research; S.D., C.A.C., Y.C.L., A.E., E.T., C.V.F., D.L.J., S.H., J.A., B.L.C., R.C., P.C., N.F., M.F., E.G., J.G., M.G., C.D.K., A.K., U.K., E.A.L., S.L., R.M., E.Ma., E.Mo., C.M., J.L.P., M.P., H.Q., N.R., S.S., A.A.S., J.S., B.S., H.S., P.T., B.H., J.E.S., J.G.E., P.N.D., R.C.H., L.J.S., I.V.G., and F.M. performed research; R.P., P.R., Y.V.D.P., and S.Z. contributed new reagents/analytic tools; P.F. contributed the *Melampsora larici-populina* strain that was sequenced; L.J.S. contributed the *Puccinia graminis* f.sp. *tritici* strain that was

sequenced; A.A., I.V.G., S.D., and F.M. supervised the *Melampsora* genome project; C.A.C. and L.J.S. supervised the *Puccinia* genome project; S.D., F.M., C.A.C., and C.D.K. supervised the genome annotation teams; S.D., C.V.F., C.A.C., L.J.S., I.V.G., F.M., Y.C.L., D.L.J., S.H., N.F., P.F., A.K., U.K., C.M., N.R., P.T., B.H., P.N.D. and R.C.H. wrote the paper.

The authors declare no conflict of interest.

This article is a direct PNAS submission.

Data deposition: Assembly accessions: AECX00000000 (for *M. larici-populina* 98AG31) and AAWC01000000 (for *Puccinia graminis* f. sp. *tritici*); Expression data in GEO: GSE23097 (for *M. larici-populina* 98AG31) and GSE25020 (for *Puccinia graminis* f. sp. *tritici*).

¹To whom correspondence should be addressed. E-mail: duplessi@nancy.inra.fr (S.D.), cuomo@broadinstitute.org (C.A.C.), les.szabo@ars.usda.gov (L.J.S.) & fmartin@nancy.inra.fr (F.M.)

²Present address: UMR 1202 BIOGECO, INRA 69 Route d'Arcachon, 33612 Cestas Cedex, France

³Present address: 454 Life Sciences, a member of the Roche Group, Branford, CT 06405

⁴Present address: Department of Plant Pathology, North Dakota State University, Fargo, ND 58102.

*These authors contributed equally to this work.

Supporting Information

SI Text

Fig. S1 to S8

Dataset S1 (XLS)

Abstract. Rust fungi are some of the most devastating pathogens of crop plants. They are obligate biotrophs, which extract nutrients only from living plant tissues and cannot grow apart from their hosts. Their lifestyle has slowed the dissection of molecular mechanisms underlying host invasion and avoidance or suppression of plant innate immunity. We sequenced the 101-mega-base pair genome of *Melampsora larici-populina*, the causal agent of poplar leaf rust, and the 89-mega-base pair genome of *Puccinia graminis* f. sp. *tritici*, the causal agent of wheat and barley stem rust. We then compared the 16,841 predicted proteins of *M. larici-populina* to the 18,241 predicted proteins of *P. graminis* f. sp. *tritici*. Genomic features related to their obligate biotrophic life-style include expanded lineage-specific gene families, a large repertoire of effector-like small secreted proteins (SSPs), impaired nitrogen and sulfur assimilation pathways, and expanded families of amino-acid, oligopeptide and hexose membrane transporters. The dramatic upregulation of transcripts coding for SSPs, secreted hydrolytic enzymes, and transporters *in planta* suggests that they play a role in host infection and nutrient acquisition. Some of these genomic hallmarks are mirrored in the genomes of other microbial eukaryotes that have independently evolved to infect plants, indicating convergent adaptation to a biotrophic existence inside plant cells.

Rust fungi (Pucciniales, Basidiomycota), is a diverse group of plant pathogens composed of more than 120 genera and 6,000 species and are one of the most economically important groups of pathogens of native and cultivated plants (1, 2). *Puccinia graminis*, the causal agent of stem rust, has caused devastating epidemics wherever wheat is grown (3) and a new highly virulent strain (Ug99) threatens wheat production worldwide (4). Similarly, epidemics of poplar leaf rust, caused by *Melampsora* spp., is a major constraint on the development of bioenergy programs based on domesticated poplars (5) due to the lack of durable host resistance (6,7). Rust fungi are obligate biotrophic parasites with a complex life cycle that often includes two phylogenetically unrelated hosts (2). They have evolved specialized structures, haustoria, formed within host tissue to efficiently acquire nutrients and suppress host defense responses (8). Molecular features driving adaptations to an obligate biotrophic association with plant hosts are unknown. Whether the convergent biotrophic adaptation observed in bacterial parasites (9) and other lineages of microbial eukaryotes (e.g. microsporidia) (10) has lead to functional specializations at the genome level (i.e. gene gain or loss, regulation of gene expression) remains to be determined. The recent report of the genome sequence of *Blumeria graminis*, an ascomycete biotroph pathogen responsible for Barley powdery mildew revealed a genome size expansion due to transposons proliferation concomitant with dramatic reduction in gene content, i.e. genes encoding sugar-cleaving enzymes, transporters and assimilatory enzymes for inorganic nitrate and sulfur (11). Similar gene losses were observed in the genome of the oomycete

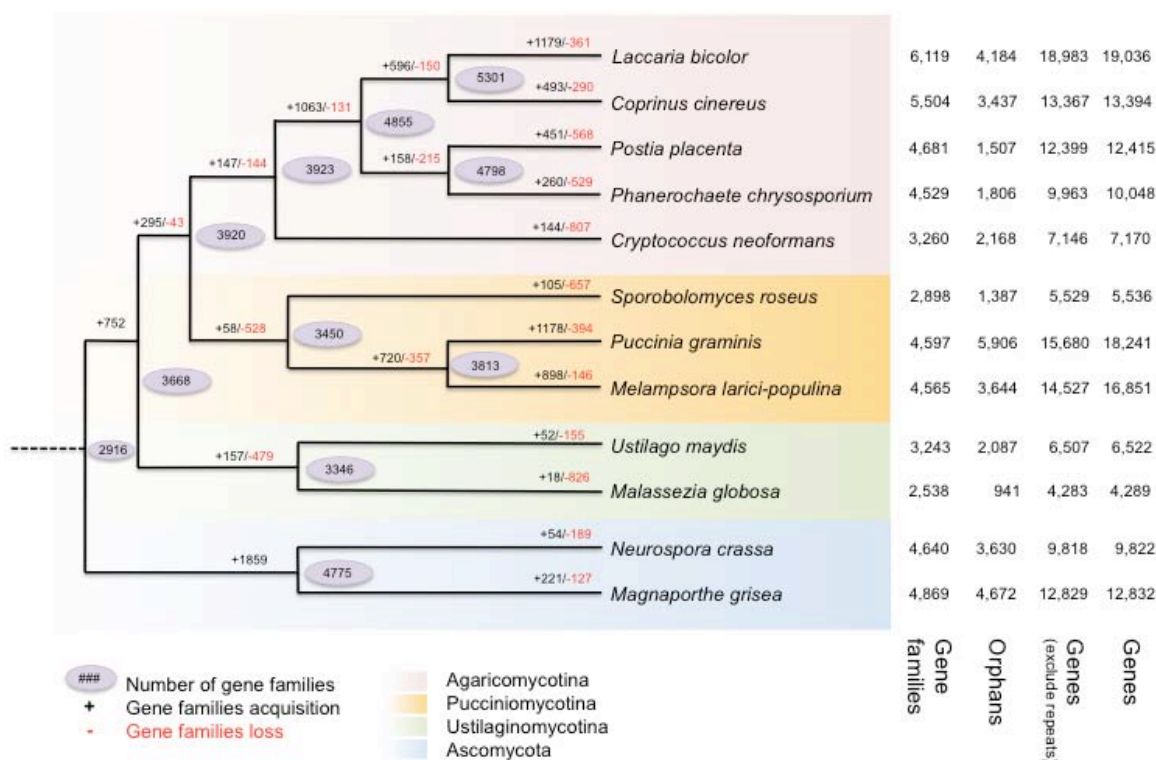


Fig. 1. Predicted pattern of gene families gain and loss in representative fungal genomes. The figure represents the total number of protein families in each species or node estimated by Dollo parsimony principle. The numerals on the phylogenetic tree branches show numbers of expanded (left, black), contracted (right, red) or inferred ancestral (oval) protein families along lineages by comparison to the putative pan-proteome. Genes excluding repeats were considered for gene family analysis in order to avoid creation of artifactual families due to the overlap between genes and repetitive elements that occurs in large numbers in rust fungi. For each species, the number of gene families, orphan genes, gene number (exclude repeats), and the total gene number are indicated on the right.

Hyaloperonospora arabidopsidis, a biotroph parasite infecting *Arabidopsis thaliana*, and the diversification of genes encoding RXLR-effector-like secreted proteins (12). Despite their phylogenetic distance, these two pathogens forming haustoria seems to share striking adaptation convergences to biotrophy. To determine the genetic features underlying pathogenesis and biotrophic ability of rust pathogens, we report here the genome sequences of the rust fungi *M. larici-populina* and *P. graminis* f.sp. *tritici*.

Results and Discussion

Genome sequencing, gene family annotation and expression analysis. We have sequenced the diploid genomes of the poplar leaf rust fungus, *Melampsora larici-populina* and of the wheat stem rust fungus, *Puccinia graminis* f. sp. *tritici*, by Sanger whole-genome shotgun strategy (*SI Text*). The overall assembly sizes of the haploid genomes of *M. larici-populina* and *P. graminis* f. sp. *tritici* are 101.1 Mb and 88.6 Mb, respectively (Dataset S1, Table S1). These genomes are much larger than the other sequenced basidiomycete genomes (13,14), but no evidence for whole-genome duplication or large scale dispersed segmental duplications was observed. The expanded size results from a massive proliferation of transposable elements (TEs), which account for nearly 45% in both assembled genomes (Dataset S1, Tables S4 and S5). Class I long-terminal-repeat (LTR) retroelements are more abundant in *M. larici-populina*, whereas class II TIR (Terminal Inverted Repeat) DNA transposons are prominent in *P. graminis* f. sp. *tritici*. Timing of TE activity using sequence divergence of extant copies suggests that a major wave of retrotransposition in the *M. larici-populina* and *P. graminis* f. sp. *tritici* lineages occurred <1 million years ago (*SI Text*).

We predicted 16,841 and 18,241 protein-coding genes in *M. larici-populina* and *P. graminis* f. sp. *tritici*, respectively. The size of these proteomes is similar to the symbiotic basidiomycete *Laccaria bicolor* (14), but strikingly larger than the corn smut fungus, *Ustilago maydis*, a pathogenic biotroph that only possesses ~6,500 proteins (15). Among the predicted proteins, only 41 and 44% in *M. larici-populina* and *P. graminis* f. sp. *tritici*, respectively, showed significant sequence similarity to documented proteins (BLASTP \leq e-value 10^{-5}) (*SI Text*). *M. larici-populina* and *P. graminis* f. sp. *tritici* possess a large set of lineage-specific gene pairs showing high similarity levels (80-100%) (*SI Text*). To investigate protein evolution in *M. larici-populina* and *P. graminis* f. sp. *tritici*, we constructed families containing both orthologs and paralogs from a diverse set of ascomycetous and basidiomycetous fungi (*SI Text*). Large numbers of rust genes presented overlaps with repetitive elements and were excluded to perform the gene family analysis (*SI Text*). The two genomes shared 3,273 orthologous Tribe-MCL families

Table 1. Selection of *Melampsora larici-populina* genes strongly upregulated during poplar leaf infection (see Table S13 for complete list)

Mlp ID	Function	Best blast hit		Expression levels		96 hpi / USp	
		Pgt ID	GenBank Accession #	96 hpi	USp	FC	P-value
89465	Aspartic peptidase A1, secreted	PGTG_10570	XP_001881739	44063	38	1159.6	3.42 E-05
94889	Lipase, secreted	PGTG_15782	XP_749106	27318	36	758.9	1.72 E-04
123524	Small secreted protein, RTP homolog	PGTG_18022	ABS86408	49354	68	725.8	8.53 E-04
106755	Small secreted protein, glycosyl hydrolase 16	no hit	no hit	25530	57	447.9	7.42 E-05
88574	Oligopeptide transporter, OPT	PGTG_17016	XP_001394363	38726	88	440.1	1.40 E-04
86448	Transporter, AEC (Auxin Efflux Carrier) family	PGTG_06747	XP_759229	17984	42	428.2	1.33 E-04
112330	a-glycosidase related to a-mannosidases, secreted, glycosyl hydrolase 47	PGTG_09507	XP_001881296	14561	41	355.2	3.92 E-05
36184	Amino acid permease, PIG2 homolog	PGTG_15547	XP_001873273	10319	34	303.5	2.10 E-04
95696	Alanine amino-transferase	PGTG_07510	XP_001837651	11018	37	297.8	3.84 E-04
53832	Thiazole biosynthetic enzyme, THI4 homolog	PGTG_01304	Q9UVF8	52910	194	272.8	1.14 E-04
39287	Small secreted protein, Cro r I homolog	no hit	AAF87492	7916	30	263.9	0.026
64764	Small secreted protein, hesp-376 homolog	no hit	no hit	7596	35	217.1	1.26 E-03
89463	Subtilisin protease, secreted, peptidase subfamily S8A	PGTG_18581	XP_001877576	18072	87	207.8	1.15 E-04
40379	Sugar transporter HXT1, MFS	PGTG_15147	XP_001874568	12387	61	203.1	2.64 E-04
91040	b-glycosidase, endoglucanase, glycosyl hydrolase 5	PGTG_17056	XP_001875020	7212	36	200.4	5.13 E-04
124202	Secreted protein, AvrM-B homolog	no hit	ABB96259	3764	27	139.5	4.12 E-04
67013	Thiamin biosynthesis enzyme Thi1p homolog	PGTG_10151	ABK96768	35825	274	130.8	1.51 E-04
48366	Carotenoid ester lipase, secreted	PGTG_13346	XP_001875752	14890	121	123.1	1.26 E-03
40488	Chitin deacetylase, carbohydrate esterase 4	PGTG_09635	XP_774611	3704	39	95	1.10 E-03
109896	Secreted protein related to plant expansins	PGTG_19856	XP_771894	4998	52	96.2	4.34 E-03
60884	Glycosyltransferase, glycosyltransferase 18	PGTG_01151	XP_001884748	3889	41	94.9	8.36 E-04
87910	Oligopeptide Transporter, OPT	PGTG_15138	XP_001834544	12366	160	77.3	6.03 E-05
39227	Zinc transporter, CDF (Cation Diffusion Facilitator) family	PGTG_14264	CAE00445	3210	43	74.7	5.67 E-03
25498	Chitin deacetylase, carbohydrate esterase 4	PGTG_09635	XP_774611	4541	61	74.5	3.41 E-03
55212	Small secreted protein, hesp-735 homolog	no hit	ABB96276	2221	33	67.4	5.20 E-04

Up-regulation in poplar infected leaves is assessed by comparing transcripts profiles to those from resting urediniospores (USp). Poplar leaves were infected by *M. larici-populina* urediniospores and left for 96 hpi under controlled conditions. At this stage, poplar rust pathogen has formed many haustoria in planta and sporulation has not yet occurred. Expression values are the means of three biological replicates for 96 hpi and USp. Based on statistical analysis of normalized fluorescence levels, a gene was considered significantly regulated if it met two criteria: (1) t-test Pvalue, 0.05 (ArrayStar, DNASTar); infected poplar leaves at 96 hpi versus urediniospores fold-change > 10. Genes were selected on the basis of homology to a function, and hypothetical proteins or genes without homology of unknown function (exception of small secreted proteins representing candidate rust pathogen effectors) were discarded. The complete list of significantly regulated genes is detailed in Table S13. Genes presented here are a selection of in planta highly up-regulated rust genes falling in the small secreted proteins, lipases, proteases, transporters, carbohydrate active enzymes, metabolism-related genes categories. A grey cell indicates an expression value below background expression level.

which comprised 5,709 *P. graminis* f. sp. *tritici* genes and 6,117 *M. larici-populina* genes; ~30% of the predicted proteins are lineage-specific, whereas 854 gene families were unique to these two rust fungi. Expansion of protein family sizes was prominent in both *M. larici-populina* and *P. graminis* f. sp. *tritici* (Fig. 1; Fig. S4; Dataset S1, Tables S7-9); several expanded gene families are lineage-specific, suggesting that important protein-coding innovation occurred in these lineages. Of the 4,858 *M. larici-populina* genes that have an orthologue in *P. graminis* f. sp. *tritici* (Best Reciprocal Hit, e-value $\leq 10^{-3}$), very few show conservation of neighbouring orthologues (synteny) (*SI Text*; Fig. S5). This is likely due to the expansion of the TE and massive reshuffling of the genome as a result. In addition, within the rust fungi, *M. larici-populina* and *P. graminis* f. sp. *tritici* represent very divergent phylogenetic lineages (1). Marked gene family expansions also occurred in those genes coding for a-kinases (Fig. S7), oligopeptide membrane transporters (OPT) (Dataset S1, Table S19), copper/zinc superoxide dismutase (Dataset S1, Table S24), and several groups of predicted transcription factors (Fig. S4).

Seventy and 54% of the predicted genes of *M. larici-populina* and *P. graminis* f. sp. *tritici*, respectively, were detected by custom microarray transcript profiling of resting and germinating urediniospores, as well as infected leaves (*SI Text*). A significant proportion of the detected transcripts (18%) is differentially expressed (fold-ratio ≥ 10.0 , $P < 0.05$) in infected leaves, whereas only ~8.0% are specifically expressed *in planta* (*SI Text*). Transcripts coding for secreted peptidases and lipases, transporters of hexoses, amino-acids and oligopeptides, and carbohydrate-cleaving enzymes, such as chitin deacetylase and cutinase (Tables 1 and 2, Dataset S1, Tables S13 and S17), are strikingly enriched (≥ 10 -fold) *in planta*. However, the most highly upregulated transcripts *in planta* (≥ 100 -fold) are mainly comprised of species-specific transcripts, including those coding for small secreted proteins (SSPs) (Dataset S1, Tables S11 and S15). These *in planta*-induced, lineage-specific genes are likely involved in the specific relationship established between these rusts and their respective hosts.

Rust fungi secretomes contain candidate novel rust effectors. Microbial pathogens have evolved highly advanced mechanisms to engage their hosts in intimate contact and sabotage host immune responses by secreting effector proteins into host cells to target regulators of defense (16,17,18). Most SSPs that are specifically produced during plant infection are likely to be effectors that manipulate host cells to facilitate parasitic colonization, such as by suppressing plant innate immunity or enhancing nutrient availability (17). *In silico* gene prediction and manual annotation of SSP genes in *M. larici-populina* genome identified a set of 1,184 SSPs (Dataset S1, Table S18), of which 74% are species-specific. Homologs of known effectors from *M. lini*, such as haustorially expressed secreted proteins (HESPs) and the avirulence factors

Table 2. Selection of *Puccinia graminis* f. sp. *tritici* genes strongly upregulated during wheat infection (see Table S17 for complete list)

Pgt ID	Function	Best blast hit		Expression levels		Wheat/USp	
		Mlp ID	GenBank Accession #	Wheat	USp	FC	P-value
PGTG_12502	Amino acid permease	113062	no hit	31670	68	467.2	0.004
PGTG_15174	Differentiation-related protein Infp	no hit	AAD38996	23002	50	466.3	0.002
PGTG_07532	Amino acid permease	113062	no hit	13666	47	293.8	0.005
PGTG_07938	Invertase 1 precursor	44167	CAG26671	18901	70	271	3.63 E-04
PGTG_17720	Zinc finger, C2H2 type	no hit	no hit	31604	175	180.9	0.004
PGTG_16569	Multicopper oxidase, secreted	112024	BAG50320	18825	114	166.6	0.012
PGTG_15026	Lipase, putative	96073	XP_001273241	21088	229	92.4	1.22 E-06
PGTG_10570	Aspartyl protease, secreted	89871	no hit	3493	46	76.1	0.04
PGTG_05667	Cu/Zn superoxide dismutase, secreted	73483	XP_002418001	10257	138	74.7	0.004
PGTG_11683	Major intrinsic protein	106246	no hit	8738	118	74.6	4.73 E-04
PGTG_19191	Serine carboxypeptidase, secreted	49959	EEY14780	6156	86	71.8	0.017
PGTG_11725	Endo-1,4-b-glucanase, secreted	47207	AAR29981	6503	100	65.3	0.038
PGTG_08842	Thiamine monophosphate synthase/TENI	63716	no hit	7343	117	63.1	7.47 E-04
PGTG_10915	Major intrinsic protein	89561	no hit	41747	686	61	0.006
PGTG_05491	MFS sugar transporter, putative	86594	XP_002480590	28494	478	59.8	0.006
PGTG_15162	Endo-b-mannanase	86044	ABR27262	6992	123	57.3	0.009
PGTG_02527	Chitin synthase N-terminal	73345	ABB70409	33954	766	44.4	8.40 E-04
PGTG_06309	Plasma membrane (H+) ATPase	44104	CAA05841	10443	272	38.5	0.003
PGTG_01889	Lipase, secreted	91294	no hit	13249	348	38.2	9.55 E-04
PGTG_15889	Aspartic peptidase A1	34644	XP_001880663	4880	128	38.2	0.019
PGTG_15122	Chitinase	75188	CAQ51152	15175	415	36.6	2.63 E-04
PGTG_12200	MFS monocarboxylate transporter	86626	XP_001267950	1636	49	34	0.012
PGTG_15888	Aspartic peptidase A1	34644	XP_001880663	2159	77	28.2	0.021
PGTG_18584	Hexose transporter HXT1p	38418	CAC41332	8629	378	22.9	0.006

Up-regulation in infected wheat is assessed by comparing transcripts profiles to those from resting urediniospores (USp). Wheat leaves were infected by *P. graminis* f.sp. *tritici* urediniospores and left for 8 days post-inoculation under controlled conditions. At this stage, wheat rust pathogen has started to sporulate and macroscopic flecking are visible. Expression values are the means of three biological replicates for 8 dpi and USp. Based on statistical analysis of normalized fluorescence levels, a gene was considered significantly regulated if it met two criteria: (1) t-test *P*-value, 0.05 (using *matlab* in *Matlab*); infected wheat at 8 dpi versus USp fold-change > 10. Genes were selected on the basis of homology to a function, and hypothetical proteins or genes without homology of unknown function (exception of small secreted proteins representing candidate rust pathogen effectors) were discarded. The complete list of significantly regulated genes is detailed in Table S17. Genes presented here are a selection of in planta highly up-regulated rust genes falling in the small secreted proteins, lipases, proteases, transporters, carbohydrate active enzymes, metabolism-related genes categories. A grey cell indicates an expression value below background expression level.

AvrM, *AvrL567*, *AvrP123*, *AvrP4* from the flax rust fungus *M. lini* (8,17), and the rust-transferred protein RTP1 from the bean rust pathogen (18), are present among highly upregulated *M. larici-populina* transcripts (Table 1, Dataset S1, Tables S11-S13). At least 43% of *M. larici-populina* SSPs are expressed in infected leaves at 96 hours post infection. *P. graminis* f. sp. *tritici* contains a similar number of 1,103 SSP genes, of which 85% are species-specific. In *P. graminis* f. sp. *tritici*, PGTG_17547 matches the highest number of haustorial ESTs, and is similar in sequence to a predicted secreted protein (ADA54575) from the wheat stripe rust fungus, *P. striiformis* (19). In both rust species, one protein in this group (PGTG_13212, JGI ID# 85525), is similar in sequence to a haustorially expressed protein from the flax rust pathogen, HESP-735 (17) (Dataset S1, Tables S11 and S15). Fifty and 29 SSPs belong to the top 100 most highly transcriptionally up-regulated in infected poplar and wheat leaves compared to *M. larici-populina* and *P. graminis* f. sp. *tritici* urediniospores, respectively (Dataset S1, Tables S11 and S15). Most upregulated SSP transcripts *in planta* were species-specific, as only 16% have an ortholog in both rust species, suggesting that these sequences are evolving at a very high rate. It remains to be determined whether upregulated SSPs are expressed in infection hyphae and/or haustoria, and whether they remain in the cell-wall, the extra-haustorial matrix, or are addressed to specific compartments of the host cell where they interact with their target proteins as shown for avirulence proteins in *M. lini* (8,17). In *M. larici-populina*, a total of 812 SSPs are organized in 169 families of 2 to 111 members (Dataset S1, Table S18); the largest family contains a highly conserved ten-cysteine pattern (Fig. S6). In *P. graminis* f. sp. *tritici*, a total of 1,105 SSPs are organized in 164 families of 2 to 38 members; the largest family contains a highly conserved eight-cysteine pattern. Four of these proteins show evidence of haustorial expression, suggesting they could be potential effectors.

Rust fungi Carbohydrate-Active Enzymes set. Gene families encoding host-targeted, hydrolytic enzymes acting on plant biopolymers, such as proteinases, lipases, and several sugar-cleaving enzymes (CAZymes)(20), are highly upregulated in both rust pathogen transcriptomes *in planta* (Tables 1 and 2, Dataset S1, Tables S13 and S17), suggesting that the invading hyphae is penetrating the host cells by using these degrading enzymes. The comparison of the glycoside hydrolase (GH), glycosyltransferases, polysaccharide lyase (PL) and carbohydrate esterase (CE) of 21 sequenced fungi (Fig. 2) however revealed that *M. larici-populina* and *P. graminis* f. sp. *tritici* have a relatively smaller set of GH-encoding genes (173 and 158 members, respectively) (SI Text, Dataset S1, Table S20); similar to the basidiomycete symbiont *L. bicolor* (14), but much fewer than hemibiotrophic or necrotrophic phytopathogens (e.g., *Magnaporthe grisea*) and saprotrophs (e.g., *Neurospora crassa*; *Coprinopsis cinerea*; *Schizophyllum commune*)(21). This set of CAZymes is strikingly larger than the repertoire of the biotroph *Ustilago maydis* (100

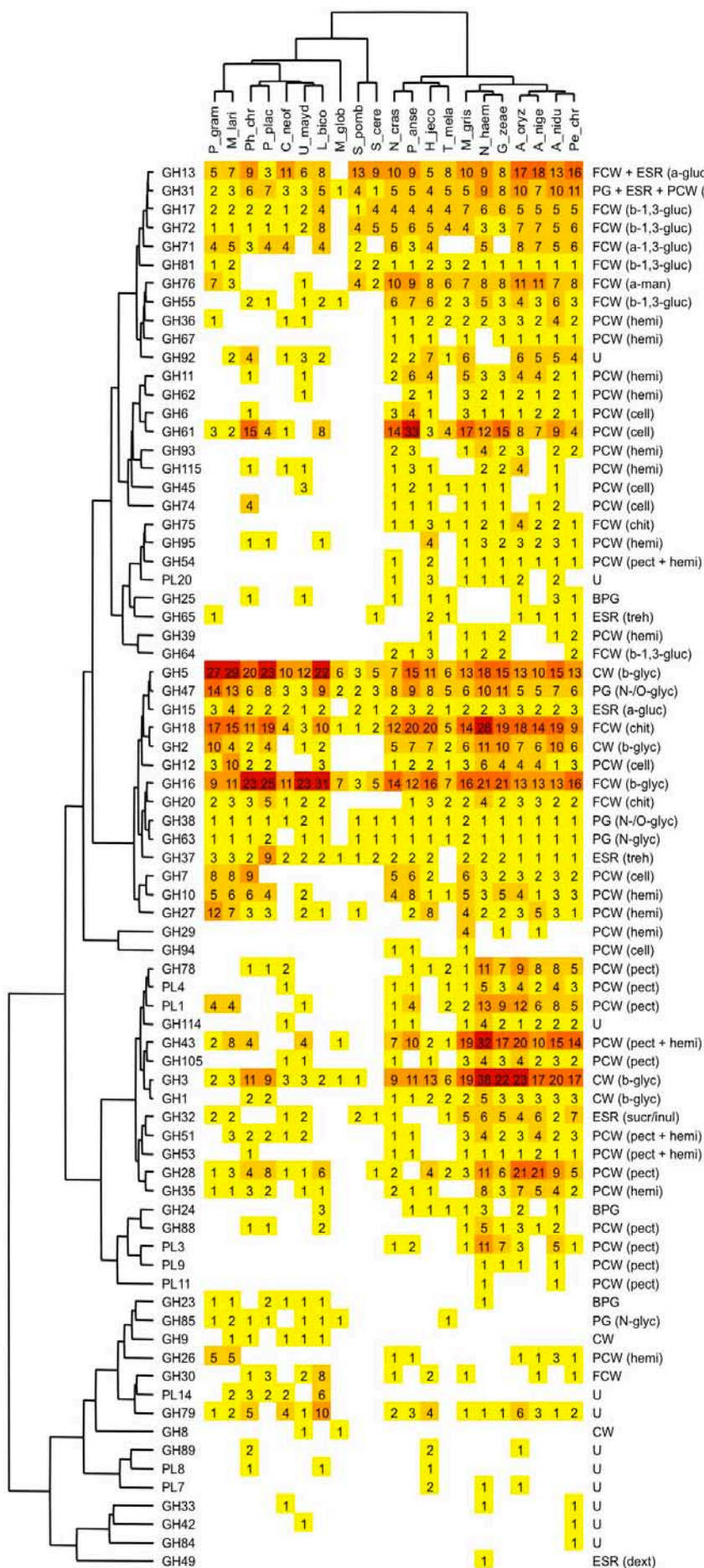


Fig. 2. Double clustering of the carbohydrate-cleaving families (19) from representative fungal genomes. Top tree: the fungi named are *Aspergillus nidulans* (A_nidu), *Aspergillus niger* (A_nige), *Aspergillus oryzae* (A_oryz), *Cryptococcus neoformans* (C_neof), *Gibberella zeae* (G_zeae), *Hypocrea jecorina* (H_jeco), *Laccaria bicolor* (L_bico), *Magnaporthe grisea* (M_gris), *Malassezia globosa* (M_glob), *Melampsora larici-populina* (M_lari), *Nectria haematococca* (N_haem), *Neurospora crassa* (N_cras), *Penicillium chrysogenum* (Pe_chr), *Phanerochaete chrysosporium* (Ph_chr), *Podospira anserina* (P_anse), *Postia placenta* (P_plac), *Puccinia graminis* f. sp. *tritici* (P_gram), *Saccharomyces cerevisiae* (S_cere), *Schizosaccharomyces pombe* (S_pomb), *Tuber melanosporum* (T_mela) and *Ustilago maydis* (U_mayd). Left tree: the enzyme families are represented by their class (GH, glycoside hydrolase; PL, polysaccharide lyase) and family number according to the carbohydrate-active enzyme database (27). Right side: known substrate of CAzy families (most common forms in brackets): BPG, bacterial peptidoglycan; CW, cell wall; ESR, energy storage and recovery; FCW, fungal cell wall; PCW, plant cell wall; PG, protein glycosylation; U, undetermined; a-gluc, a-glucans (including starch/glycogen); a-man, a-mannan, b-glyc, b-glycans; b-1,3-gluc, b-1,3-glucan; cell, cellulose; chit, chitin/chitosan; dext, dextran; hemi, hemicelluloses; inul, inulin; N-glyc, N-glycans; N-/O-glyc, N- / O-glycans; pect, pectin; suc, sucrose; and treh, trehalose. Abundance of the different enzymes within a family is represented by a colour scale from 0 (white) to 33 occurrences (red) per species.

members) (15). In evolving a biotrophic lifestyle, the rust fungi have lost several secreted hydrolytic GH and PL enzymes acting on plant cell wall (PCW) polysaccharides (Fig. 2) and they are lacking the cellulose-binding CBM1 module. However, they show a moderate expansion of a few GHs cleaving plant celluloses and hemi-celluloses (e.g., GH7, GH10, GH12, GH26 and GH27) compared to the biotroph *U. maydis* or the hemibiotroph *M. grisea*. These enzymes, together with *in planta* upregulated and expanded α -mannosidase (GH47) and β -1,3-glucanase (GH5) transcripts (Dataset S1, Tables S13 and S17), may play a key role in the initial stages of host colonization, i.e. penetration of the parenchyma cells. On the other hand, induced chitin deacetylases (CE4) present in *P. graminis* f. sp. *tritici*, *M. larici-populina* and the symbiont *L. bicolor* (14) are likely involved in fungal cell wall remodelling and may play a role in the alteration of the fungal cell wall surface during infection to conceal the hyphae from the host (22).

Expanded rust transporters gene families are expressed during host infection. A process that is crucial to the success of rust pathogen biotrophic interactions is the acquisition of nutrients (carbohydrates and amino acids) by invading hyphae from its host plant through the haustoria (17,23,24). The repertoire of membrane transporters (*SI Text*, Dataset S1, Table S19) in *M. larici-populina* and *P. graminis* f. sp. *tritici* contains homologs of the hexose transporter HXT1, amino-acid transporters AAT1, AAT2 and AAT3 and H⁺-ATPases from the bean rust pathogen (*Uromyces fabae*), known to be highly upregulated during the interaction with its host plant. In addition, *M. larici-populina* and *P. graminis* f. sp. *tritici* genomes display an increased genetic potential for peptide uptake with 22 and 21 oligopeptide transporter (OPT) genes, respectively, whereas other basidiomycete genomes only contain five to 16 OPT genes (Dataset S1, Table S19). OPT genes that are transcriptionally upregulated *in planta* (Dataset S1, Tables S13 and S17), are likely involved in the transport of peptides released by the action of the induced proteinases (aspartic peptidase, subtilisin) expressed in infected leaf tissues. The Major Facilitator Superfamily (MFS) gene family is reduced in the *M. larici-populina* and *P. graminis* f. sp. *tritici* genomes compared to other basidiomycetes (Dataset S1, Table S19), but many MFS transcripts are however highly expressed *in planta* including two HXT1 homologs. Consistent with *in planta* expression of *M. larici-populina* and *P. graminis* f. sp. *tritici* invertase genes (Dataset S1, Tables S13 and S17), no homolog of the sucrose transporter Srt1 recently described in *U. maydis* (23) was identified, supporting the preferential uptake of host hexoses by invading rust pathogen hyphae (24). The increased activity of membrane transporters provides the needed fuel for the high primary metabolism activity observed in the invading rust fungi (Dataset S1, Tables S13 and S17).

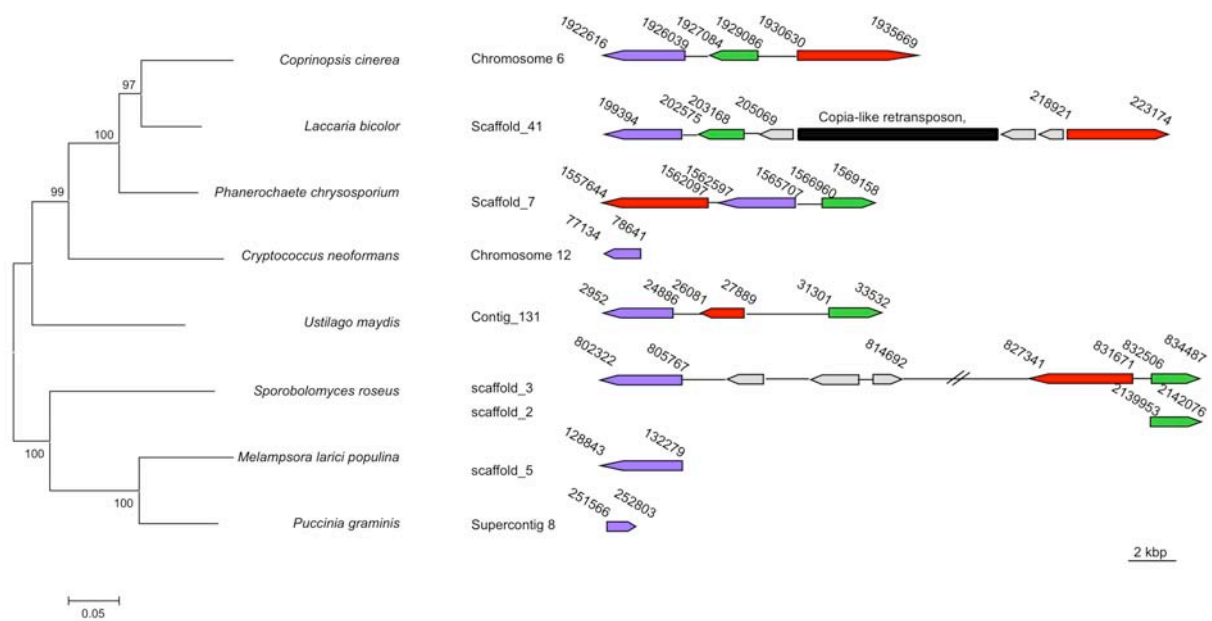


Fig. 3. Structure of the nitrate assimilation cluster among Basidiomycetes. Phylogram based on the MS277 and MS456 genes (26) from eight Basidiomycete fungi was obtained using the minimum evolution method implemented in MEGA4 (27), with the complete deletion option for handling alignment gaps, and with the Poisson correction model for distance computation. Bootstrap tests were conducted using 1000 replicates. Branch lengths (drawn in the horizontal dimension only) are proportional to phylogenetic distances. Description of open reading frames coding nitrate/nitrite porter (green), nitrate reductase (purple) and nitrite reductase (red) is given. Numbers indicate the start and stop codons for each ORF. Grey rectangles indicate proteins that are not functionally related to nitrate assimilation.

Nitrate and sulfate assimilation pathways deficiencies in rust fungi. Based on the inability of rust fungi to grow *in vitro* we hypothesized that the *M. larici-populina* and *P. graminis* f. sp. *tritici* genomes may lack genes typically present in saprotrophic basidiomycetes. Major anabolic pathways of primary metabolism were manually inspected for potential deficiencies. Although the enzymes of the NH_4^+ assimilation pathway were identified, several genes involved in nitrate assimilation were lacking in both rust pathogen gene repertoires. The nitrate/nitrite porter and the nitrite reductase (NiR) are missing from the nitrate assimilation gene cluster found in other fungi (Fig. 3) (25). Genes required to perform the primary sulfate assimilation were identified in *M. larici-populina* whereas they were lacking in *P. graminis* f. sp. *tritici*. The latter fungus lacks both a- and b-subunits of sulfite reductase (SiR), whereas the *M. larici-populina* b-subunit of SiR is missing the transketolase domain present in other fungal SiRs. The apparent lack of nitrate and sulfate assimilation enzymes in both rust fungi is consistent with their obligate biotrophic life style, as they depend on reduced nitrogen (either NH_4^+ or amino acids) and sulfur from plant cells. These metabolic deficiencies have also been found in plant pathogens that represent two independent evolutionary lineages of obligate biotrophy in the oomycete (*H. arabidopsidis*) and ascomycete (*Blumeria graminis*) lineages (11,12).

Conclusions

The obligate biotroph status of rust fungi has limited studies to understand how they invade their hosts and avoid or suppress defense responses. The genome sequences of the poplar leaf and wheat stem rust fungi are an unparalleled opportunity to address questions related to the obligate biotrophy lifestyle. The genetic changes that brought about the evolution of obligate biotrophy from biotrophic progenitors remain obscure. Our comparisons of *M. larici-populina* and *P. graminis* f. sp. *tritici* to other saprotrophic, pathogenic and symbiotic basidiomycetes indicate that the developmental innovations in the lineages of rust fungi did not involve major changes in the ancestral repertoire of proteins with known function. On the other hand, the large set of lineage-specific, expanding gene families may provide a key source of developmental innovation and adaptation. Our analysis shows that the colonization of the host leaf, differentiation of pathogenic structures and control of the plant immune system can be associated with a large-scale invention of lineage-specific proteins. For example, the rich repertoire of candidate effector-like SSPs could underline the co-evolution and adaptation of these obligate pathogens to the plant immune system. Contrary to obligate bacterial biotrophs and microsporidian fungal parasites which oftenly undergo gene loss and genome compaction (9,10), the rust pathogen genomes are amongst the largest fungal genomes sequenced so far showing expanded gene families and massive proliferation of TEs. No massive gene loss was

observed in *M. larici-populina* and *P. graminis* f. sp. *tritici*, but irreversible deletion of genes not essential for the obligate biotrophic life-style (e.g., N and S assimilation), together with a lower set of plant cell wall polysaccharide degrading enzymes are genomic hallmarks of rust fungi and other biotrophic pathogens (11,12). A deeper understanding of the complex array of factors, such as effector-like SSPs, affecting host–pathogen interactions and co-evolution could ensure efficient targeting of parasite-control methods in agricultural and forest ecosystems.

Material and Methods

Genome sequencing, assembly and annotation. The dikaryotic *M. larici-populina* 98AG31 and *P. graminis* f.sp. *tritici* CDL 75-36-700-3 (race SCCL) strains were sequenced by whole-genome sequencing (WGS) and were assembled into predicted 101.1Mb and 88.6 Mb genomes respectively (*SI Text*). The protein coding-genes (16,841 for *M. larici-populina* and 18,241 for *P. graminis* f.sp. *tritici*) were predicted with a combination of gene callers using ESTs produced from each rust fungus (*SI Text*). Reduced gene sets were considered to perform multigene families analyses (14,527 and 15,680 predicted genes for *M. larici-populina* and *P. graminis* f.sp. *tritici* respectively) by removing gene models presenting overlaps with specific repeats/TE fragments to avoid creation of biased gene families (*SI Text*). The *M. larici-populina* genome sequence can be accessed at <http://genome.jgi-psf.org/Mellp1/Mellp1.home.html> and the *P. graminis* f.sp. *tritici* genome sequence can be accessed at http://www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.html.

Microarray analysis of gene expression in urediniopores and rust-infected plants. For both *M. larici-populina* and *P. graminis* f.sp. *tritici*, gene expression was assessed in resting and *in vitro* germinating urediniospores of the sequenced rust strains as well as in respective host plant tissues at late stages of infection using specific custom 70-mer oligoarrays (*SI Text*). Methods for RNA isolation, probes synthesis and hybridization, data capture and analysis are described in *SI Text* and data have been deposited in GEO (GSE23097 for *M. larici-populina* and GSE25020 for *P. graminis* f.sp. *tritici*).

ACKNOWLEDGEMENTS. The work conducted on *M. larici-populina* by the Joint Genome Institute of the U.S. Department of Energy is supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC02-05CH11231. This project was also funded by grants from the INRA and the Région Lorraine Council to F.M. and S.D.; and a grant from Natural Resources Canada to R.C.H. The sequence of *P. graminis* f. sp. *tritici* was funded by the US National Science Foundation and conducted by the Broad Institute Sequencing Platform. The

work of Y-CL, P.R. and Y.VdP was supported by IUAP P6/25 (BioMaGNet). We thank Marie-Pierre Oudot-LeSeq for the initial *M. larici-populina* TE annotation, Benoît Hilselberger for database construction, Carine Commun and H  l  ne Niculita-Hirzel for the annotation of the *M. larici-populina* secretome and mating-type genes, respectively, and Jerry Johnson for technical assistance.

References

1. Aime MC, *et al.* (2006) An overview of the higher level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. *Mycologia* 98:896-905.
2. Cummins GB, Hiratsuka Y (2004) in *Illustrated genera of rust fungi, 3rd edition* (APS Press, St. Paul), pp1-240.
3. Leonard KJ, Szabo LJ (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol Plant Pathol* 6:99-111.
4. Stokstad E (2007) Deadly wheat fungus threatens world's breadbaskets. *Science* 315:1786-1787.
5. Rubin EM (2008) Genomics of cellulosic biofuels. *Nature* 454:841-845.
6. Duplessis S, Major I, Martin F, S  guin A (2009) Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit Rev Plant Sci* 28:309-334.
7. G  rard PR, Husson C, Pinon J, Frey P (2006) Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathol* 96:1027-1036.
8. Dodds *et al.* (2009) Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytol* 183:993-1000.
9. Ochman H, Moran NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science*, 292:1096-1099.
10. Corradi N, Pombert JF, Farinelli L, Didier E, Keeling PJ (2010) The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nature Com* 1:77.
11. Spanu PD *et al.* (2010) Genome expansion and gene loss in Powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* 330:1543-1546.
12. Baxter L *et al.* (2010) Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330:1549-1551.

13. Cuomo CA, Birren B (2010) The fungal genome initiative and lessons learned from genome sequencing. *Meth Enzymol* 470:833-855.
14. Martin F *et al.* (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88-92.
15. Kämper J *et al.* (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 444:97-101.
16. Panstruga R, Dodds PN (2009) Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens. *Science* 324:748-750.
17. Ellis JG, Rafiqi M, Gan P, Chakrabarti A, Dodds PN (2009) Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. *Curr Opin Plant Biol* 12:399-405.
18. Voegelé RT, Hahn M, Mendgen K (2009) in *The Mycota V: Plant relationships*, ed Deising HB (Springer, Berlin), pp69-98.
19. Godfrey D *et al.* (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics* 11:317.
20. Cantarel BL *et al.* (2009) The Carbohydrate-Active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* 37:D233-D238.
21. Ohm RA *et al.* (2010) Genome sequence of the model mushroom *Schizophyllum commune*. *Nat Biotechnol* 28:957-963.
22. El Gueddari NE, Rauchhaus U, Moerschbacher BM, Deising HB (2002) Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytol* 156:103-112.
23. Wahl R, Wippel K, Goos S, Kämper J, Sauer N (2010) A novel high-affinity sucrose transporter is required for virulence of the plant pathogen *Ustilago maydis*. *PLoS Biology* 8:e1000303.
24. Voegelé RT, Struck C, Hahn M, Mendgen K (2001) The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proc Natl Acad Sci U S A* 98:8133-8138.
25. Slot JC, Hibbett DS (2007) Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. *PLoS One* 10:e1097.
26. Aguilera G *et al.* (2008) Assessing the performance of single-copy genes for recovering robust phylogenies. *Syst Biol* 57:613-627.
27. Tamura K *et al.* (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-1599.

**Transcript profiling of *Melampsora larici-populina* during
time course infection of poplar leaves reveals major insights
into rust biotrophy**

Duplessis S, Hacquard S, Delaruelle C, Tisserant E, Fourrey
C, Frey P., Martin F, Kohler A.

Publication soumise à la revue Molecular Plant-Microbe Interactions

Les données supplémentaires de cette publication sont disponibles sur le site web:
http://mycor.nancy.inra.fr/pub/SOM_these_Hacquard

Transcript profiling of *Melampsora larici-populina* genes during timecourse infection of poplar leaves reveals major insights into rust biotrophy

Sébastien Duplessis^{*†}, Stéphane Hacquard[†], Christine Delaruelle, Emilie Tisserant, Claire Veneault-Fourrey, Pascal Frey, Francis Martin and Annegret Kohler

INRA, Institut National de la Recherche Agronomique, UMR 1136 INRA/Université Nancy, Interactions Arbres/Micro-organismes, Centre INRA de Nancy, 54280 Champenoux, France

[†] These authors contributed equally to this work

Author for correspondence:

Sébastien Duplessis

duplessi@nancy.inra.fr

Phone: +33 383 39 40 13

Fax: +33 383 39 40 69

Abstract

Melampsora larici-populina is the basidiomycete responsible for the poplar leaf rust disease that causes severe epidemics in poplar plantations in Europe. The poplar rust genome has been recently sequenced and in order to find the genetic determinants associated with its biotrophic lifestyle within the 14,527 rust genes, we generated whole-genome custom oligoarrays and analyzed transcripts profiles of *M. larici-populina* during time-course infection of poplar leaves. Nine different stages were investigated during the asexual development of rust in its uredinial host including 7 stages *in planta* and resting and germinating urediniospores. In total, 76% of the genes were detected as expressed during leaf infection between 24 and 168 hours post-inoculation as well as in urediniospores, whereas 20% were only expressed *in planta* including several transporters and many small secreted proteins (SSPs). We focused our analysis on gene categories commonly related to plant colonization and biotrophic growth in fungal pathogens such as SSPs, carbohydrate active enzymes (CAZymes), transporters, lipases and proteases. Distinct sets of genes encoding SSPs were expressed all along the infection process of the poplar leaves as well as during germination suggesting that expression of probable candidate rust effectors is highly dynamic throughout the whole biotrophic growth. Contrary, transporters were mostly expressed after 48 hpi when numerous haustoria are already formed in the leaf mesophyll supporting a central role in nutrients acquisition during the biotrophic growth. Finally, CAZymes, lipases and proteases were predominantly expressed at later stages of infection during uredinia formation and sporulation.

Key-words: virulence, obligate biotroph, transcriptomics, small-secreted proteins

Introduction

Melampsora larici-populina (Basidiomycota, Pucciniales) is an obligate biotrophic fungus responsible for foliar rust disease in poplar. This is the most devastating and widespread pathogen for this tree and it has limited the use of poplars for environmental and wood production goals in many part of the world (Steenackers et al. 1996). In Europe, hybrids between *Populus x euramericana* and *Populus x interamericana* were selected in the mid-20th century for their complete resistance to the rust fungus (Steenacker 1988). Cultivation of these hybrids was successful for many years but led to the emergence of new fungal virulence and the breakdown of several major resistance genes (Frey et al. 2005; Pinon and Frey 2005). To date, most poplar cultivars are susceptible to *M. larici-populina*, and new virulent strains are continuously developing. Like other rust fungi, *M. larici-populina* has a complex life cycle. Five

different spore types are successively produced and two distinct host plants (larch and poplar) are required to complete its cycle. In contrast with the limited symptoms developing on larch, *M. larici-populina* causes severe epidemics during its asexual phase on poplar. The infection process, previously described during poplar leaves colonization (Laurans and Pilate 1999; Rinaldi et al. 2007; Hacquard et al. 2011), starts when urediniospores land on the abaxial surface of poplar leaves. A primary germ tube is formed within 2 hours and extends to allow fungal penetration through stomata. Substomatal vesicles, from which first infection hyphae are formed, are then produced *in planta* (Rinaldi et al. 2007) and allow rapid differentiation of the first haustoria at 17 hours post inoculation (hpi) (Laurans and Pilate 1999). This structure, essential for the uptake of host metabolism and the suppression of host defences (Voegelé and Mengesha 2003; Catanzariti et al. 2007) enables rapid proliferation of the rust fungus i.e. between 24 and 96 hpi (Rinaldi et al. 2007; Hacquard et al. 2011). By 96 hpi, infection structures have colonized the whole plant mesophyll and differentiation of the first sporogenous hyphae are observed (Hacquard et al. 2010). Only one week after spores landing, uredinia are observed and newly formed dikaryotic spores are released to achieve further infection of poplar leaves (Hacquard et al. 2010).

While molecular mechanisms underlying poplar defense reactions in response to *Melampsora* spp. have been addressed experimentally using transcriptomic approaches (Rinaldi et al. 2007; Miranda et al. 2007; Azaiez et al. 2009; Boyle et al. 2010), still little is known about the molecular mechanisms allowing successful invasion of rust fungi in the host (Duplessis et al. 2009). The recent sequencing of the *M. larici-populina* genome (Duplessis et al. submitted) now opens the possibility to dissect the infection process in order to identify key genes underlying fungal virulence. Recently, laser capture microdissection was used to collect poplar infected tissues associated with *M. larici-populina* at a late stage of rust infection (i.e. 168 hpi; uredinal stage), and microarray analysis led to the identification of rust transcripts associated with biotrophy and sporulation of the rust fungus (Hacquard et al. 2010). Nevertheless, genes expression profiling at earliest stages of rust infection during timecourse infection of poplar leaves remains to be determined.

Obligate biotrophic fungi such as *M. larici-populina* establish an intimate relationship within host cell, which is essential for pathogen growth and reproduction. This particular status precludes fungal culture *ex planta* and limits some routine manipulation. Previous attempts to understand the biological processes that control biotrophic rust development mostly focused on Expressed Sequence Tags from isolated rust haustoria or infected leaves (Hahn and Mendgen 1997; Thara et al. 2003; Broeker et al. 2006; Catanzariti et al. 2006; Feau et al. 2007; Puthoff et al. 2008; Yin et al. 2009; Joly et al. 2010). However, the development of DNA arrays offers the possibility to

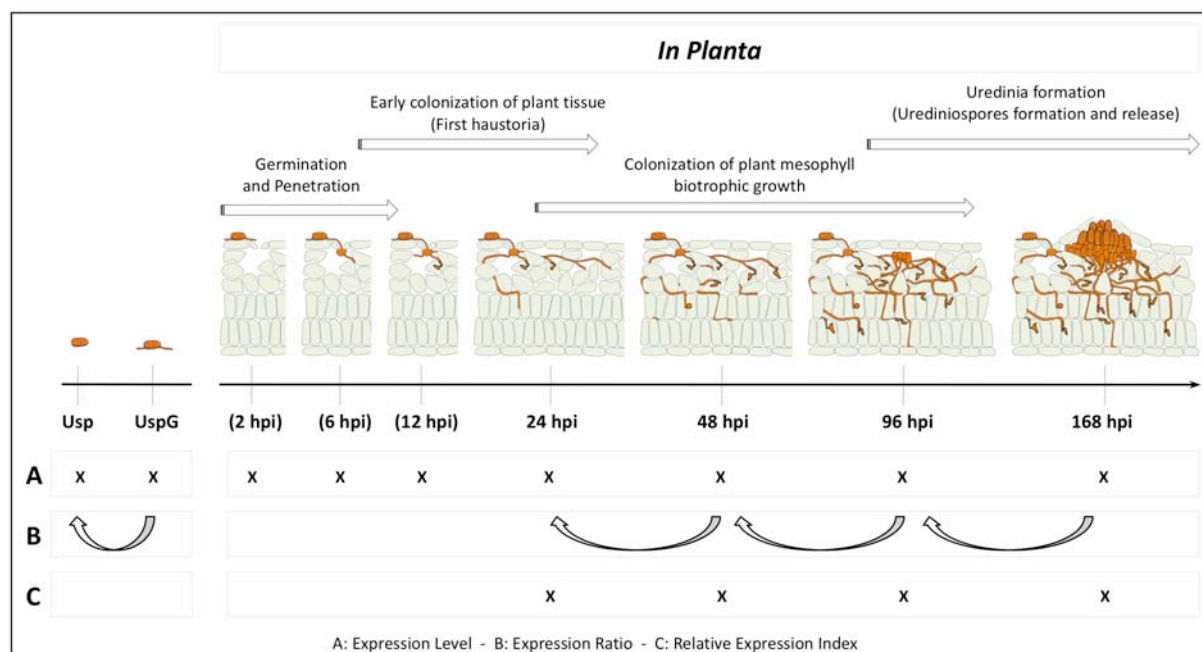


Figure 1. Schematic drawing presenting the major phases of poplar leaf infection by *Melampsora larici-populina* and the transcript expression analyses performed in the study. RNA used for transcript profiling were isolated from three independent replicate experiments corresponding to *M. larici-populina* resting (Usp) and germinating (UspG) urediniospores and *Populus trichocarpa* x *Populus deltoides* infected leaves at 2, 6, 12, 24, 48, 96 and 168 hpi (*in planta*). Gene expression levels were analysed in Usp, UspG and at 2, 6, 12, 24, 48, 96, 168 hpi to identify genes specifically or highly expressed in spores or *in planta*. Low level of expression measured at 2, 6 and 12 hpi did not allow comparison with later time-points and were not further considered in the analysis. In order to identify groups of genes significantly regulated during spore germination and during the infection process, expression ratios were calculated between UspG/Usp, 48 hpi/24 hpi, 96 hpi/48 hpi and 168 hpi/96 hpi. A relative expression index (expression level at each time-point relative to mean expression level across time-course) was calculated to derive coordinated patterns of gene expression during time-course infection.

analyse transcriptional changes of very large numbers of genes and this is an approach of choice to decipher plant-pathogen interactions. Transcript profiling has been successfully used to identify candidate genes involved in fungal virulence as well as to monitor dynamic changes during *in planta* development of *Blumeria graminis*, the ascomycete biotroph responsible for powdery mildew of cereals (Both et al. 2005a; Both et al. 2005b). Similar studies were also conducted in *Magnaporthe oryzae* the hemi-biotroph pathogen responsible for rice blast (Dean et al. 2005; Oh et al. 2008) and the biotroph corn smut fungus *Ustilago maydis* during colonization of plant tissues (Kämper et al. 2006) and also in the basidiomycetes *Pisolithus microcarpus* and *Paxillus involutus* during the development of the ectomycorrhizal symbiosis in their respective hosts (Duplessis et al. 2005; Le Quéré et al. 2005).

In this paper, we report the transcriptome analysis of the rust fungus *M. larici-populina* during poplar leaf infection using whole-genome custom oligoarrays. Stages related to the asexual development in the uredinial host were analysed including dormant and germinating urediniospores outside of the host and *in planta* infection fungal structures and cell types from early colonisation of poplar leaf tissues to the formation of mature uredinia. Transcript profiles were validated using quantitative real-time PCR (RT-qPCR). Our results indicate that plant infection is marked by dynamic and coordinated temporal patterns of rust gene expression. We identified gene families potentially associated with virulence, biotrophic growth and urediniospore formation. This first genome-scale analysis of rust gene expression during host infection brings new highlights into rust biology.

Results

Experimental design of time course infection and data normalisation

RNA were isolated from 9 different timepoints in a time-course infection of leaves from 'Beaupré' poplar cultivar by *M. larici-populina* strain 98AG31 (*in planta* stages) from 2 hours post-inoculation (hpi) when urediniospores germinated on poplar leaf abaxial epidermis to 168 hpi when new urediniospores were produced and released from uredinia, as well as from resting and germinating urediniospores (stages outside plant tissues), covering the main developmental transitions of the rust pathogen (Figure 1). Sufficient amounts of cDNA were recovered from all biological stages to perform hybridization onto oligoarrays. However, the overall expression levels (i.e. fluorescence levels) measured at early stages of the time-course infection were too low for comparison with later stages. Less than 500 transcripts were detected above background fluorescence level at 2, 6 and 12 hpi and most transcripts corresponded to

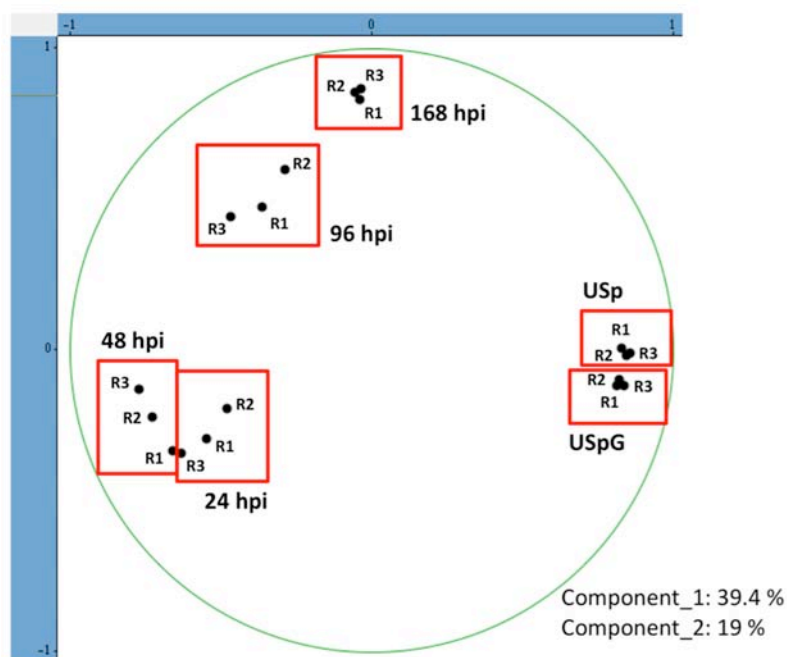


Figure 2. Principal Component Analysis (PCA) of *Melampsora larici-populina* transcript levels measured in infected poplar leaves at 24, 48, 96 and 168 hours post-inoculation (hpi), resting urediniospores (USp) and germinating urediniospores (USpG) with custom oligoarrays. Three biological replicates (R1, R2 and R3) per stage were used for the PCA and the correlation circle is presented. Expression level of each gene assessed in a given biological situation and biological replicate was reported to the mean expression level calculated for the 18 hybridizations on oligoarrays (6 conditions x 3 replicates) and was log10 normalized prior to proceed with PCA. The PCA plot places biological conditions along the two axes (components 1 and 2) explaining 39.9 % and 19.02 % of the variance observed within samples.

Table 1: Number of *Melampsora larici-populina* genes expressed during timecourse infection of poplar leaves and in resting or germinating urediniospores. Genes were classified according to their normalized expression levels measured with custom oligoarrays.

Expression level ^a	24 hpi ^b	48 hpi ^b	96 hpi ^b	168 hpi ^b	Usp ^b	USpG ^b
background to 1,000	2,587	3,798	4,113	4,266	1,726	1,845
1,001 to 5,000	1,366	1,921	2,832	3,040	2,014	2,104
5,001 to 10,000	210	314	590	693	1,322	1,346
10,001 to 20,000	101	144	243	257	1,405	1,345
> 20,001	15	39	78	70	1,268	1,232
Total number of genes	13,093	13,093	13,093	13,093	13,093	13,093
not detected (<background)	8,814	6,877	5,237	4,767	5,358	5,221
Expressed (> Background)	4,279	6,216	7,856	8,326	7,735	7,872

^a A transcript was deemed expressed when its fluorescence signal intensity was three-fold higher than the mean signal-to-noise threshold (cut-off background value) of 1,063 random oligonucleotide probes.

^b 24, 48, 96, 168 hours post inoculation (hpi), resting urediniospores (Usp) and germinating urediniospores (UspG).

ribosomal protein encoding genes or hypothetical/unknown genes (data not shown). Thus, these timepoints were not further considered in the analysis. Component analysis of expression levels measured at the different stages (Figure 2) showed good consistency between biological replicates and distinct profiles could be distinguished between *in planta* stages and urediniospores, as well as between stages of timecourse infection. Biological replicates of resting and germinating urediniospores presented a low variability. Similarly, replicates at 168 hpi which corresponds to the highest proportion of fungal tissues *in planta* (Hacquard et al. 2011) also showed a greater reproducibility than earlier timepoints, suggesting a correlation between fungal transcripts content and fluorescent levels measured on oligoarrays. In order to avoid technical biases due to dilution of labelled fungal transcripts in fluorescent probes, *in planta* stages and urediniospores were normalized separately. Different means were used to analyse the data (Figure 1). First, gene expression levels were analysed to identify most highly expressed genes for each time point and to highlight *ex-* or *in-* *planta* specific genes. Then, two types of comparison were performed between transcripts levels measured at the different stages of timecourse infection. First, successive timepoints were compared by calculating direct ratios between one stage to another (i.e. UspG/Usp; 48 hpi/24 hpi; 96 hpi/48 hpi and 168 hpi/96 hpi). Then, calculated REI (Relative Expression Index, see methods) were used to derive expression patterns of genes involved in plant colonization and the fungal biotrophic growth. This does not allow direct comparison between fungal stages in and outside the host as previously done (Duplessis et al., submitted), but it eliminates possible over- or under-estimation of transcript concentration in extreme expression situations (i.e. 24 hpi versus urediniospores).

Genes specifically expressed in planta and in urediniospores

Table 1 presents proportions of genes expressed *in planta* and in urediniospores and highlights the contrasted expression profiles between stages surveyed. Indeed, 51 and 50% of the transcripts expressed were detected at a normalized level above 5,000 in resting and germinating urediniospores respectively, whereas 7 to 11% of the expressed transcripts were detected at similar levels between 24 and 168 hpi. Interestingly, a peak of expression is noticed during timecourse infection at 168 hpi with more than 8,326 transcripts detected and a high number of genes are already expressed at 96 hpi compared to 24 or 48 hpi which suggests that expression measured throughout the timecourse correlates with fungal growth in the plant (Rinaldi et al. 2007; Hacquard et al. 2011). In total, 76% of the genes expressed at least once *in planta* are also expressed in urediniospores (Figure 3) whereas only 4% (Supplementary table 1) and 20% (Supplementary table 2) are only expressed in urediniospores or *in planta*, respectively. Interestingly, a coreset of 32% of the genes is always expressed *in planta*. Beside,

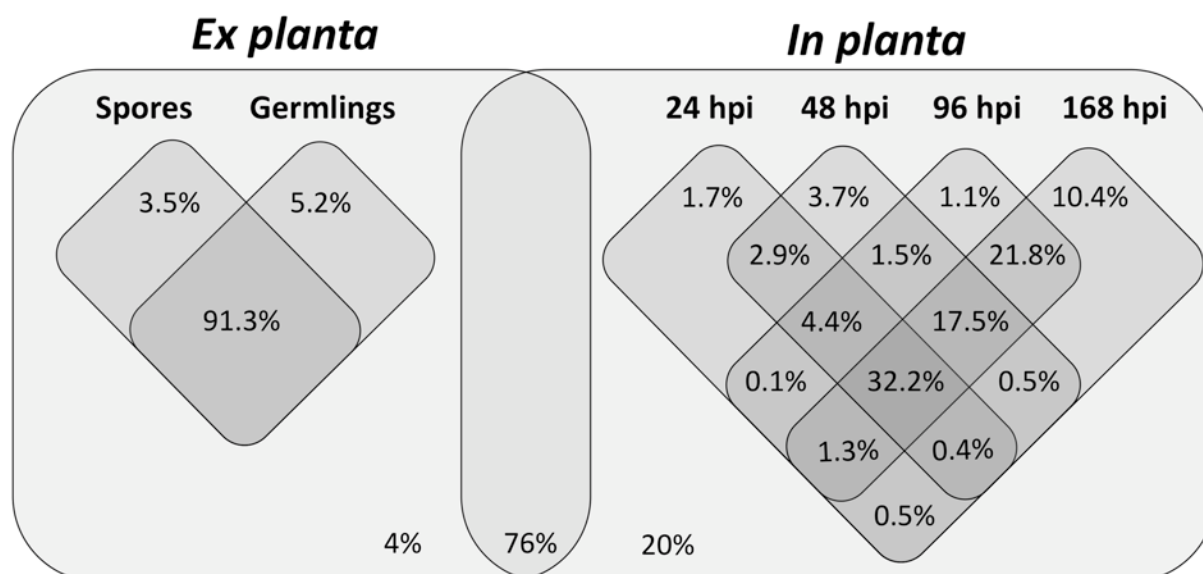


Figure 3. Venn diagram of *M. larici-populina* expressed genes during poplar leaf infection (*in planta*) or germination (spore and germlings). Expression was assessed with custom oligoarrays at 24, 48, 96 and 168 hours post-inoculation in infected poplar leaves and in resting or germinating urediniospores. In the main Venn diagram, 4% of the genes were specifically detected *ex planta* whereas 20% were specifically detected *in planta* at least at one of the timepoints and 76% were detected in both *ex planta* and *in planta* stages. Smaller Venn diagrams for *ex planta* and *in planta* details genes expressed at any particular stage.

Table 2: Number of *Melampsora larici-populina* genes significantly regulated during time course infection of poplar leaves and during germination using custom oligoarrays. Down- and up-regulated genes were classified according to their fold-change ratio.

Regulated genes ^a	48/24 ^b	96/48 ^b	168/96 ^b	UspG/Usp ^b
significantly regulated genes	7	1848	3193	3347
up-regulated	0	855	1649	1737
> 3-fold	0	783	1100	641
> 10-fold	0	330	499	177
> 100-fold	0	3	65	19
down-regulated	7	993	1544	1610
< -3-fold	5	806	638	545
< -10-fold	4	76	59	98
< -100-fold	0	0	0	1

^a Transcripts with a significant *p*-value (<0.05) and more than a 3-fold change in transcript level were considered as significantly differentially expressed.

^b Ratio 48 vs. 24 hours post-inoculation (hpi), 96 vs. 48 hpi; 168 vs. 96 hpi and urediniospores germinating on agar (UspG) vs. resting urediniospores (UspG).

smaller sets of genes are also expressed specifically at any of the timepoints including the earliest ones (Figure 3). In detail, genes highly expressed *in planta* at early timepoints mostly corresponded to ribosomal proteins and hypothetical proteins of unknown function (supplementary table 3). Interestingly, many of the hypothetical proteins encode small secreted proteins (SSPs) of unknown function that could represent putative rust effectors. Some SSP transcripts were highly expressed since 24 hpi and all along the plant colonization process (JGI protein IDs no. 74599 and 124067). Contrary, less ribosomal proteins encoding transcripts were detected at 168 hpi.

Among the 13,093 genes analysed on the oligoarray, 2,078 (20%) were specifically detected *in planta* (Figure 3; Supplementary table 2) at least at one of the timepoint (24, 48, 96, 168 hpi) and showed expression levels below background in Urediniospores. Strikingly, almost 80% (in detail 1,617) of the genes belonging to this set did not show homologies with the non-redundant database at NCBI and only 481 of these unknown proteins displayed similarities with genes predicted in *Puccinia graminis* f. sp. *Tritici* genome sequence. Genes encoding SSPs, transporters, proteases, Cazymes and lipases were specifically detected *in planta* during rust infection, highlighting their important role during the biotrophic growth (Supplementary table 2). Expression values of genes specifically expressed *in planta* were converted to REI in order to monitor expression profiles during fungal growth in poplar leaves. K-means partitioning methods revealed different profiles during plant infection (Supplementary figure 1). Almost 70% (1,400/2,078) of the genes specifically expressed *in planta* are preferentially expressed at early stages of rust infection (i.e. 24 hpi, 24 and 48 hpi or 48 hpi) whereas only 24% (502/2,078) were preferentially expressed when the biotrophy is well established (i.e 48 and 96 hpi, 96 hpi or 96 and 168 hpi). A large proportion of SSPs was detected for those timepoints (33%, in detail 141/502) whereas only 8% were expressed at others timepoints (in detail 124/1576). Last, 8% (176) of the genes specifically detected *in planta*, including several cytochromes P450 and carbohydrate active enzymes, showed preferential transcript accumulation when sporulation occurs (168 hpi) (Supplementary table 2). In contrast with the strong proportion of genes specifically detected *in planta*, only 401 genes (4%) of the 13,093 genes analysed on the oligoarray were specifically detected in urediniospores (Figure 3; Supplementary table 1). Among these, 101 were only detected during germination including several carbohydrate active enzymes and detoxification enzymes (Supplementary Table 1).

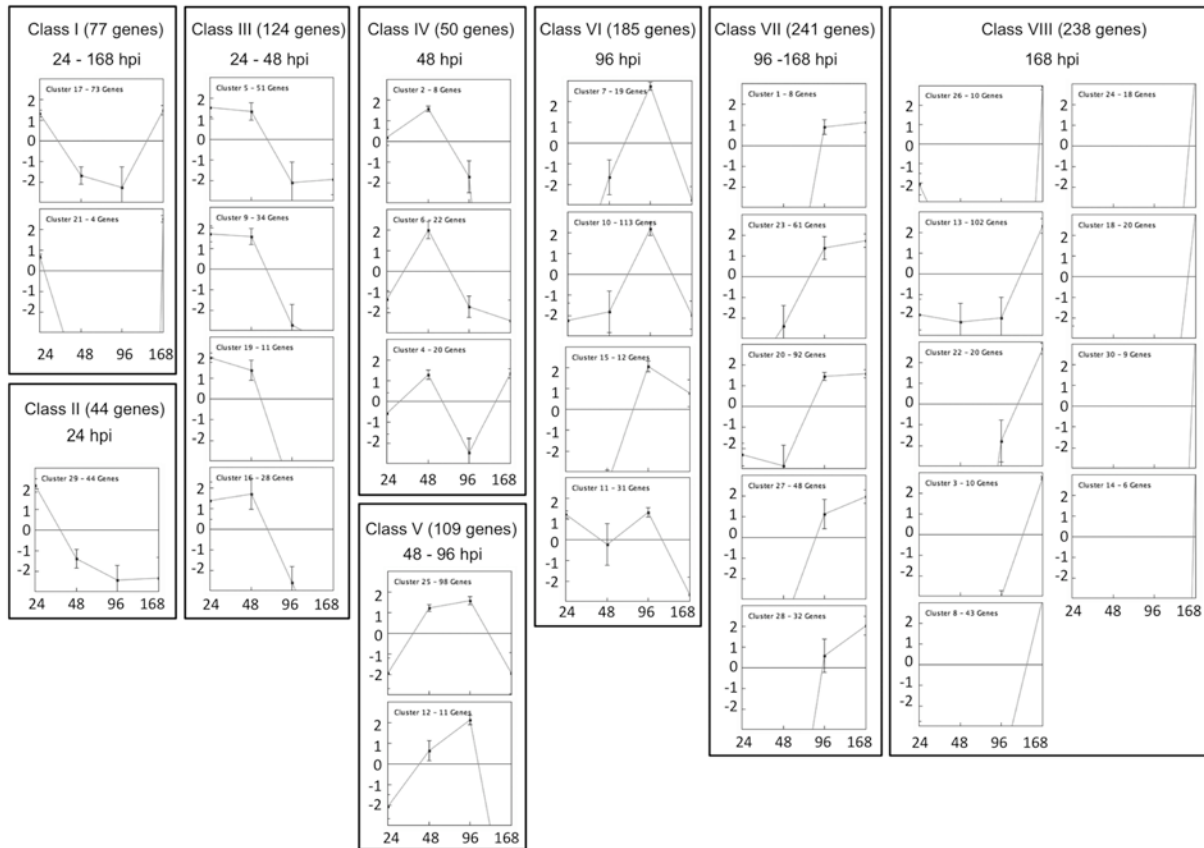


Figure 4. Expression patterns of 1,058 *M. larici-populina* genes significantly regulated during poplar leaf infection encoding small secreted proteins (SSPs), transporters, proteases, carbohydrate active enzymes and lipases. Clusters of genes were derived by partitioning log₂ REI values calculated at 24, 48, 96, 168 hours post-inoculation (hpi) in 30 k-means profiles. Relative Expression Index (REI) corresponds to the mean expression level measured at any timepoint divided by the mean expression level overall the timecourse serie. The 30 k-means profiles were grouped in eight classes according to their peak of expression (Class I to VIII). Class I: Genes preferentially expressed at 24 and 168 hpi, Class II: 24 hpi, Class III: 24 and 48 hpi, Class IV: 48 hpi, Class V: 48 and 96 hpi, Class VI 96 hpi, Class VII: 96 and 168 hpi and Class VIII: 168 hpi.

Rust transcript induced during poplar leaf infection

The Supplementary Table 4 details the numbers of rust genes significantly regulated between distinct stages of plant infection. Interestingly almost no genes were deemed as significantly regulated between 24 and 48 hpi. These timepoints correspond to rather low fungal amounts, however such an absence of significantly regulated genes might only be related to the variability between biological replicates and similarity between some replicates at 24 and 48 hpi (Figure 2). Interestingly, although a higher number of regulated genes was observed during urediniospore germination than during leaf infection, a larger proportion of highly regulated genes (> 3-fold change and > 10-fold change) were observed at 168 versus 96 hpi or at 96 hpi versus 48 hpi (Table 2). This highlights that although a few genes are highly expressed *in planta* (Table 1), expression is highly dynamic during plant colonization. Indeed, details of the genes that showed the highest significant fold-change expression between timepoints (up-regulated and down-regulated) reflect that distinct categories of genes are related to plant colonization (Supplementary table 4). Among the 855 genes significantly induced at 96 hpi compared to 48 hpi, the most highly up-regulated transcript (191-fold induction) encodes a CFEM domain SSPs (JGI protein ID no. 114961). Strikingly, two genes related to mating and encoding a mating pheromone A receptor (JGI protein ID no. 41282) and a mating factor A secretion protein (ABC transporter, JGI protein ID no. 493844) were also identified among the most highly up regulated transcripts at 96 hpi (85-fold and 55-fold, respectively). Among the 1,649 genes significantly up-regulated at the sporulation stage (168 hpi) compared to 96 hpi, several of the most highly induced genes (Prot IDs no. 72264, 72999, 124001) displayed homologies to genes previously described in *Lentinula edodes* (shiitake mushroom) and *Agaricus bisporus* as involved in fruiting-body formation such as the developmentally regulated gene *priA* (Kajiwara et al. 1992), the riboflavin-aldehyde-forming (RAF) enzyme (Sreenivasaprasad et al. 2006), or the *Drmip_hesp* family proteins (Szeto et al. 2007). Several transcripts encoding a beta mannanase (JGI protein ID no. 60617), a lysophospholipase (JGI protein ID no. 94935) and tripeptidyl peptidases (JGI protein IDs no 86030 and 45226) also showed drastic inductions ranking between 225- and 730-fold compared to 96 hpi suggesting their involvement in poplar rust sporulation (Supplementary table 4).

Dynamic expression of rust genes during plant colonization and biotrophic growth

We focused our analysis on gene categories commonly associated to plant colonization and biotrophic growth in rust fungi (Dodds et al. 2009; Voegelé et al. 2009) encoding small secreted proteins that likely contain candidate rust effectors necessary to set a compatible interaction

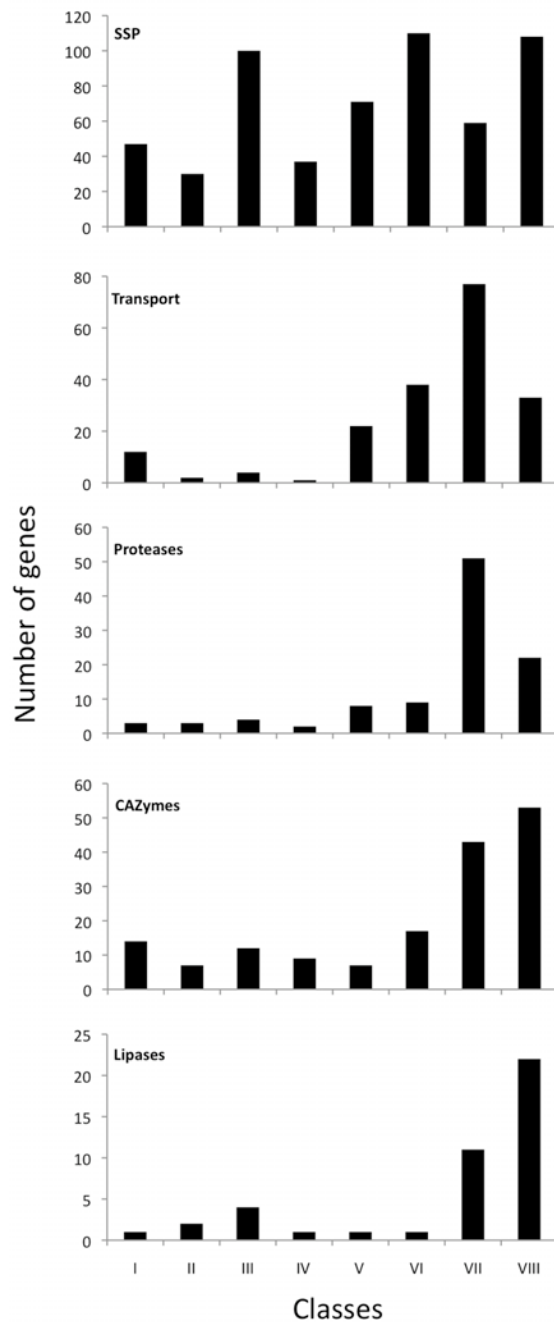


Figure 5. Number of *Melampsora larici-populina* genes identified in the eight cluster classes from Fig. 3 for the following selected categories: small secreted proteins (SSP), transport, proteases, carbohydrate active enzymes (CAZymes) and lipases. Cluster classes indicate rust genes preferentially expressed at distinct stages of poplar leaf infection: Class I: 24 and 168 hpi, Class II: 24 hpi, Class III: 24 and 48 hpi, Class IV: 48 hpi, Class V: 48 and 96 hpi, Class VI 96 hpi, Class VII: 96 and 168 hpi and Class VIII: 168 hpi.

with poplar, transporters, carbohydrate active enzymes, proteases and lipases. In these gene categories, only a total of 1,058 genes were significantly regulated at least once during leaf infection and were considered. *In planta* gene expression profiles of selected genes were derived using the k-means partitioning method implemented in the Genesis expression analysis package (Sturn et al. 2002) based on REI values (Figure 4). A total of 30 k-means were necessary to extract clusters of rust genes with distinct profiles during leaf infection, and those clusters were grouped in eight classes of patterns (class I to VIII, Figure 4) by considering the timepoint(s) where the REI peaked. Interestingly, although surveyed genes clearly showed a peaked of expression at 168 hpi, when the fungus is highly abundant in the leaf tissue, important proportions of genes were also detected as early as 24 hpi (classes I, II and III) or 48 hpi (classes III, IV and V). Strikingly, among the categories surveyed (Figure 5), distinct and important groups of genes encoding SSPs were clearly expressed in all classes, i.e. all along the infection process, which indicates that expression of putative rust effectors was not restricted to the earliest stages of colonization but was highly dynamic and maintained all along the interaction with the plant. Interestingly, several homologs of genes encoding haustorially expressed secreted proteins (HESPs) and avirulence factors identified in the flax rust *M. lini* (JGI prot IDs no. 33341, 36204, 71404, 93569, 74496, 39396, 37347, 58916) were preferentially expressed between 48 and 96 hpi (Class III, IV and IV). In contrast, transporters encoding genes were mostly expressed after 48 hpi when numerous haustoria are detected in the leaf mesophyll supporting their role in the biotrophic growth. Many of them were preferentially expressed between 96 and 168 hpi (Class V and VI) including several ammonium (Prot IDs no. 48332, 24225, 42854), amino acid (JGI protein IDs no. 49500, 90273, 72195), oligopeptide (JGI protein IDs no. 93456, 88746, 77624, 118568, 34758, 85440), and sugar (JGI protein IDs no. 38418, 40410, 40379) transporters. Contrary, CAZymes and lipases showed a higher expression at later stages of rust infection and particularly during sporulation (168 hpi, Class VII), although distinct genes were also highly expressed as early as 24 hpi. Interestingly, several carbohydrate active enzymes, previously identified as strongly induced in the sporulation area compared to the biotrophic area of the poplar rust uredinial stage using laser capture microdissection (Hacquard et al. 2010), were also strongly accumulated at 168 hpi such as an alpha-glycosidase (GH71, JGI prot ID no. 92210), a beta-glycosidase (GH16, JGI prot ID no. 55983), an alpha-L-arabinofuranosidase (GH51, JGI prot ID no. 107307) and a chitin deacetylase (JGI prot ID no. 87698). This study revealed also the strong induction of other genes encoding alpha-glycosidases (JGI prot IDs no. 73317 (GH47), 104873 (GH71)), beta glycosidases (JGI prot IDs no. 51995 (GH26), 86044 (GH5), 90252 (GH5)) and chitinases (JGI prot IDs no. 46588, 90663) during sporulation. Lipid metabolism also showed drastic changes during fungal sporulation. For exemple, several genes encoding an epoxide hydrolase (JGI Protein ID no. 115158),

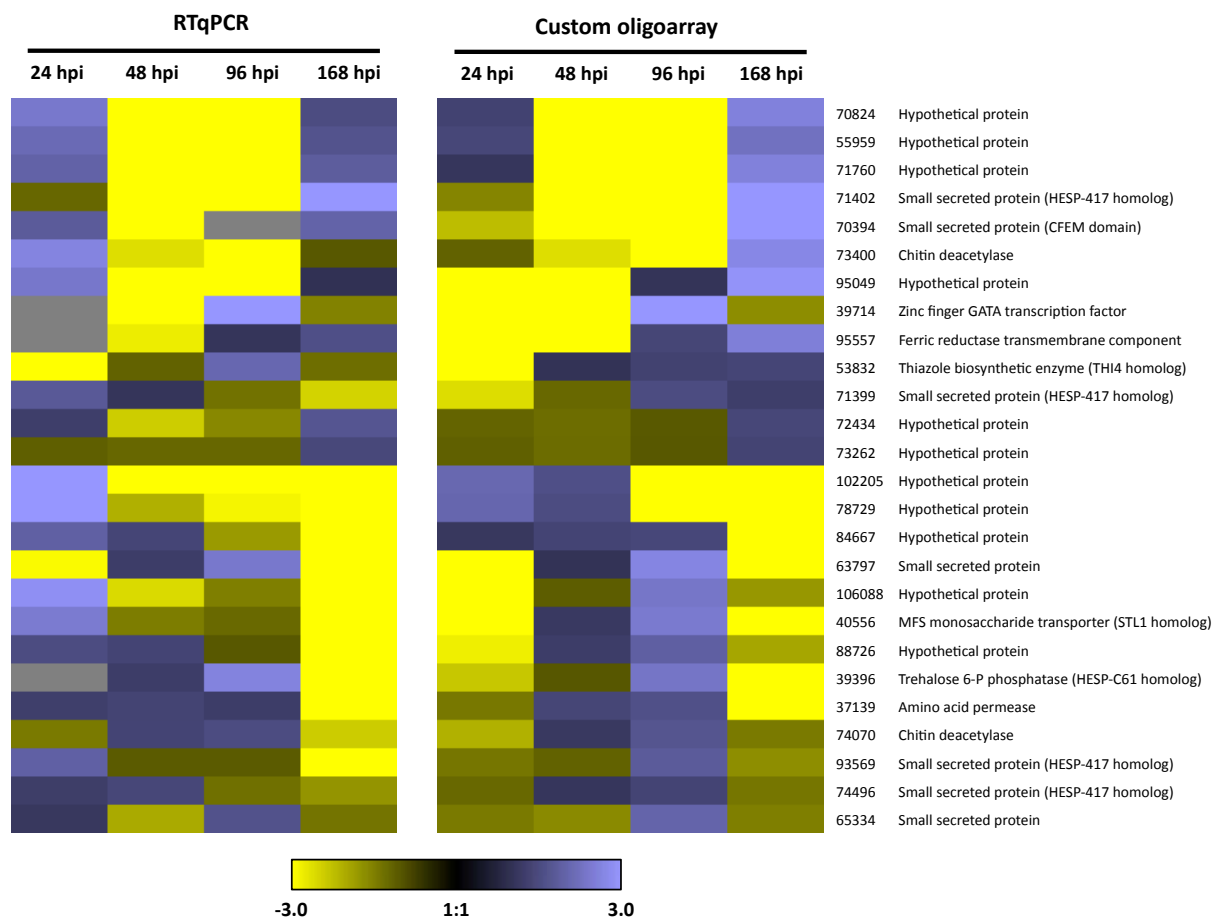


Figure 6. Expression profiles of 26 selected genes obtained by RTqPCR and custom oligoarrays. Expression levels measured by RT-qPCR or with oligoarrays were converted to Relative Expression Index (REI; expression level at any timepoint divided by the mean level calculated overall the timecourse serie) to help direct comparison of expression profiles of selected genes. REI were converted to log₂ to allow clustering. REI levels range from pale to saturated colors (yellow for negative values and blue for positive value).

cutinases (JGI Protein ID no. 86941, 89616) and lipases (JGI Protein IDs no. 117051, 78777) were preferentially expressed at this late stage of rust infection (168 hpi, Class VII) suggesting a metabolic switch between biotrophy and sporulation phases.

Validation of temporal expression patterns of rust genes by RTqPCR

A total of 26 genes with distinct patterns of expression *in planta* were randomly selected in the expression dataset to assess their transcript profiles by RTqPCR. Gene expression measured by RT-qPCR was normalized to the expression of reference rust genes alpha-tubulin and elongation factor 1-alpha (Hacquard et al. 2011), whereas expression was normalized through quantile normalization for oligoarrays. Expression values measured by RTqPCR and with oligoarrays were converted to REI to help direct comparison of expression profiles for the selected genes (Figure 6). Although some differences could be observed at 24 hpi, most profiles were confirmed by the two techniques and assessed the quality of the transcriptome analysis conducted with *M. larici-populina* custom oligoarray.

Discussion

In this study, we analyzed the transcriptome of the rust fungus *M. larici-populina* during the asexual phase of its biological cycle, when the fungus colonize its main host poplar and produces in a matter of 7-10 days newly formed asexual urediniospores (Frey et al. 2005). We covered stages such as germination with resting and germinating urediniospores which are two *ex planta* situations; and leaf colonization, biotrophic growth and finally sporulation in a timecourse serie by analyzing distinct timepoints *in planta* all along the leaf infection process. As an obligate biotroph fungus, *M. larici-populina* must overcome plant surveillance systems (i.e., Pattern recognition receptor and Effector triggered immunity) likely by secreting a large array of effectors in order to manipulate plant defense system (Dodds and Rathjen, 2010). Then the fungus must set specialized structures to derive nutrients from the host, the so-called haustoria, before to achieve spore production and sporulation (Voegelé et al. 2009). Thus, studying the rust transcriptome during host infection could help to unravel the genetic programs required to properly achieve these crucial steps in the fungal development. Due to the difficulty to study expression from infected plant tissues, analysis of *in planta* expressed rust genes was so far mostly addressed by studying transcripts expressed in resting and germinating urediniopores out of the plant tissue (Zhang et al. 2008; Joly et al. 2010) and *in planta* in the haustorial structures (Hahn and Mendgen 1997; Jakupovic et al. 2006; Catanzariti et al. 2006; Voegelé et al, 2009). A major limitation to study *in planta* situations is the low amount of fungal material

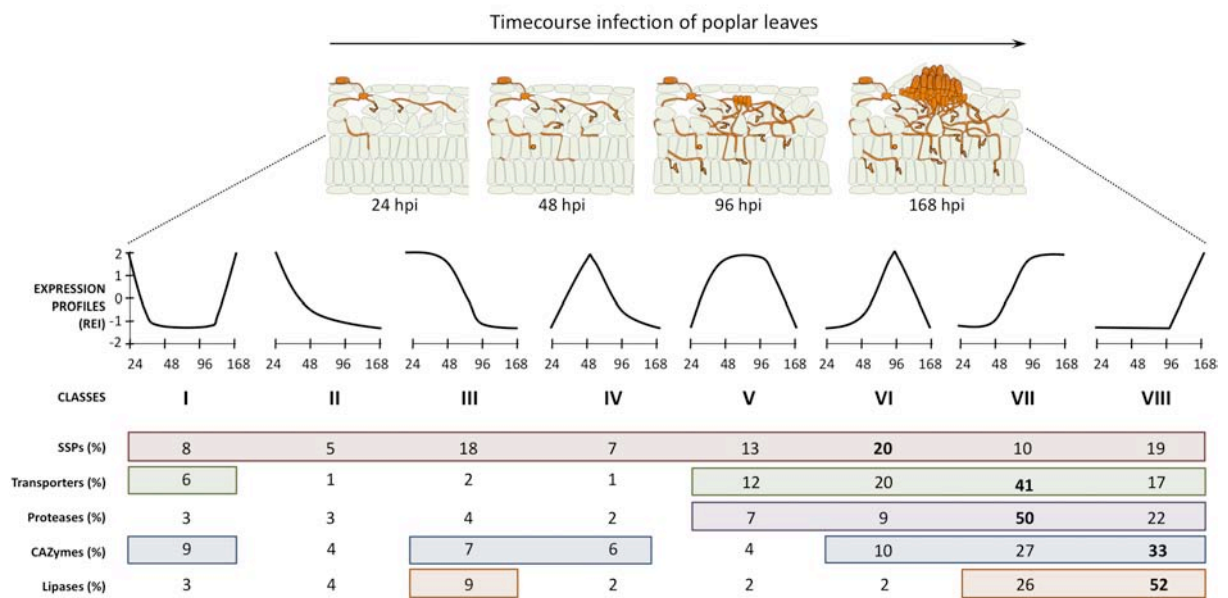


Figure 7. Major patterns of expression of *Melampsora larici-populina* genes encoding small secreted proteins (SSPs), transporters, proteases, carbohydrate active enzymes (CAZymes) and lipases during poplar leaf infection (24, 48, 96 and 168 hpi). Expression profiles correspond to the different classes identified in Figures 3 and 4 (Classes I to VIII) and for each gene category, the percentage of genes in each class is indicated. Boxes indicate a classe with more than 5% of the genes. A percentage in bold highlights the class with the largest number of genes in the category.

among RNA isolated from host tissues. Purification of haustoria from infected leaves of bean by Hahn and Mendgen (1992) by affinity chromatography on a ConA-Sepharose macrobead column was a revolution to perform study of rust biology since it allowed isolation of RNA for further expression analysis. Such an approach was used to produce cDNA libraries and ESTs of *U. fabae* (Hahn and Mendgen 1997), *U. appendiculatus* (Puthoff et al. 2008), *P. striiformis* (Yin et al. 2009), *M. lini* (Catanzariti et al. 2006) and *M. larici-populina* (Feau et al. 2007; Joly et al. 2010) or to perform microarray analysis (Jakupovic et al. 2006). These studies helped in the identification of genes expressed during germination or during the biotrophic growth which are two critical steps of plant infection. By selecting distinct timepoints during host leaf infection, our study opens possibilities to identify function related to very precise steps in host infection. To our knowledge, this is the first expression study that is conducted during timecourse infection of a host for a rust fungus and at the genome scale. Other pathogens were previously studied at the genome scale, such as the biotroph *U. maydis* (Kämper et al. 2006) or the hemibiotroph *Magnaporthe oryzae* (Oh et al. 2008). Gene expression profiling of the ascomycete *Blumeria graminis* f.sp. *hordei*, an obligate biotroph responsible for powdery mildew of cereals was achieved during plant colonization using microarrays and helped in identifying genes related to virulence (Both et al. 2005a) and specific expression of metabolic pathways before and after the penetration of Barley leaves (Both et al. 2005b).

As stressed above, *in planta* expression studies could be impaired by low amounts of fungal tissues developed in plant tissues at earlier stages of infection. Indeed, although high doses of inoculum could be used to infect plant tissues, fungal transcripts finally only correspond to a small fraction of the total RNA extracted. Highly sensitive transcript methods such as RTqPCR can help in detecting transcripts at very low concentration, however, large-scale expression studies using DNA-arrays are limited in that way. Indeed, since RNA are isolated from plant tissues containing only a few transcripts from the pathogen and very abundant plant transcripts, during synthesis of cDNA probes for hybridization to the array, plant cDNA will drain most of the fluorescence and will therefore conduct to measurement of very low fluorescent levels for fungal transcripts on the array. In the poplar-rust pathosystem, the rust fungus penetrates the plant mesophyll through stomata and forms its first haustoria in the spongy mesophyll between 17 and 24 hpi (Laurans and Pilate 1999). In a first attempt to study early stages of poplar leaf colonization, we included stages such as 2, 6 and 12 hpi that encompass germination on leaf epidermis, penetration through stomata eventually by forming appressoria and formation of substomatal vesicles (Rinaldi et al. 2007). Unfortunately, levels of fluorescence measured at those timepoints were so low that we were not able to compare them to other stages of plant infection. Fluorescence levels of early timepoints replicates were artificially increased whereas

levels at later timepoints were diminished. In order to attempt to detect functions expressed very early by the fungus, we normalized timepoints 12 and 24 hpi together. By doing so, less than 4% of the oligoprobes had a signal higher than the background level and only 14 genes presented a level sufficient to consider the gene as consistently expressed (data not shown). Among the 14 genes, two encode Rho GTPases and 2 encode small secreted proteins containing a CFEM domain or a SCP domain. We decided to discard earlier timepoints to focus our attention to 24, 48, 96 and 168 hpi. Although 24 hpi was presenting a lower amount of expressed genes compared to other timepoints (Table 1) and showed almost no difference of transcript accumulation compared to 48 hpi (Table 2), we were able to detect sets of genes that had a higher expression than at any other timepoints (Figure 4, clusters in classes I, II and III). This observation indicates that normalization was not strongly biased difference of genes expression between *in planta* samples. Indeed the proportion of low and highly expressed transcripts at the 4 *in planta* timepoints were rather comparable and were only distinct compared to the *ex planta* situations. Expression of selected genes that exhibited different expression patterns during time-course infection (i.e. belonging to the different classes revealed in Figure 4) assessed by RTqPCR showed consistency with expression measured on oligoarrays. RTqPCR gene expression levels were normalized to the expression of two reference genes encoding alpha-tubulin and elongation factor 1-alpha that have expression profiles fitting to the fungal ITS growth curve (Hacquard et al. 2011). On the other hand, quantile normalization was used for oligoarray expression values. Finally, REI of expression profiles using the two techniques that rely to distinct normalization process were almost comparable for genes expressed at 24, 48, 96 and 168 hpi, supporting the proper normalization of *in planta* situations. Beside, expression levels presented in Table 1 clearly show that resting and germinating urediniospores are too different to allow normalization of *in planta* and *ex planta* situations together.

We used different means to compare *in planta* timepoints. On one hand, identification of significantly regulated transcripts by direct comparison of two timepoints was highly informative in revealing functions timely highly regulated during the infection process. Among the genes significantly regulated at 96 and 168 hpi were many genes only expressed *in planta* and not detected in urediniospores or during germination. A strikingly important number of genes did not present any homology to other fungi, including the rust *P. graminis* f.sp. *tritici*. This result indicates that a specific genetic program that comprises several unknown genetic determinants is set by *M. larici-populina* to achieve its growth in poplar leaves. Many were specific to *M. larici-populina* and not to the wheat stem rust and might represent candidate for host specialization.

REI was previously successfully used to study *Blumeria graminis* gene expression patterns with cDNA microarray (Both et al. 2005a and 2005b). Although the hybridization design was different to the one reported here, calculation of REI greatly help to compare overall patterns of expression during timecourse studies. By allowing direct comparison of all timepoints of the infection process, REI clearly helped to extract informative patterns of gene expression from the dataset and allowed to derive 8 classes of genes associated with early, intermediate or late timepoints in the infection process. Our results showed that the rust transcriptome is marked by the dynamic expression of different sets of SSPs (Figures 5 and 7). SSPs were previously described in pathogenic fungi and oomycetes (Schornack et al. 2009; Stergiopoulos and de Wit 2010) as well as in rust fungi (RTP1, 220 amino acids (AA), Kemen et al. 2005; AvrL567, 150 AA, Dodds et al. 2004; AvrM, 314 AA, Catanzariti et al. 2006; Rafiqi et al. 2010) and are candidate effectors to help the rust fungus to establish a compatible interaction with its host. Those effectors were previously identified from rust EST libraries constructed from purified haustorial structures (Hahn and Mendgen 1997; Catanzariti et al. 2006). The dynamic expression of the poplar rust during leaf infection is surprising and suggests that during the 7 days from germination to sporulation the rust fungus express hundreds of distinct SSPs. Almost all these genes do not have a known function, and these proteins might not only correspond to unknown effectors but also to masking or protecting cell-wall proteins to prevent detection by the host defense system. A recent exemple of how fungal effector could suppress PAMP-triggered immunity has been reported during tomato/*Cladosporium fulvum* interaction (de Jonge et al. 2010). Authors showed that fungal Ecp6 effector sequesters chitin oligosaccharides that are released from the cell walls of invading hyphae to prevent elicitation of host immunity (de Jonge et al. 2010). Interestingly, many of these SSPs were still expressed at late stages of infection which could indicate that the poplar rust maintain a late active virulence in its host. In compatible *Populus-Melampsora* interactions, the plant expresses defense genes at late stages of infection (i.e. sporulation). The set of SSPs expressed at this stage might be involved in protecting the fungal structures during asexual spore formation. Among the secreted proteins transcripts showing homology to previously described HESP or avirulence factor identified in *M. lini*, several were specifically detected *in planta* (JGI protein IDs no. 37347, 124202, 124206, 124207, 124208, 124209, 37332, 53986, 55212, 110232, 101611; supplementary table 2) and some belonged to multigene families such as the HESP-417. Interestingly, distinct genes of this family were identified in distinct expression profiles confirmed by RTqPCR (Figure 6).

Among the *in planta* induced genes highly expressed in haustoria identified in *U. fabae*, several encodes amino acid (*AAT2p*, Hahn et al. 1997; *AAT1p*, Struck et al. 2002; *AAT3p*, Struck et al; 2004) and sugar transporters (*HXT1*, Voegelé et al. 2001). During poplar leaf infection, a

homolog of *Uf-AAT3* was specifically detected *in planta* (Prot ID no. 36162, Supplementary table 2) whereas the closest homologs of *Uf-AAT2* (JGI prot ID no. 37139) and *Uf-HXT1* (JGI prot ID no. 38418) were preferentially expressed after haustoria formation i.e. 48 and 96 hpi, respectively. As previously described in *U. fabae*, those transporters could be essential for the uptake of host metabolism (Voegelé et al. 2009). Genes associated with carbon and amino acid metabolism significantly expressed during plant infection were mostly showing higher levels of expression at 96 and 168 hpi supporting an intense metabolic activity after the rust fungus has formed many haustoria in its host (data not shown). Recently the sucrose transporter Srt1 was described in the biotroph *Ustilago maydis* for direct acquisition of sucrose from maize (Wahl et al. 2010). Contrary to the maize smut fungus, *M. larici-populina* does not possess an homolog of Srt1 and thus must acquire sugars in other forms. Voegelé et al (2001) showed that the bean rust *U. fabae* acquires hexoses through the monosaccharide transporter HXT1. Here several hexose transporter genes were induced *in planta* at 96 and 168 hpi (JGI protein IDs no. 38418, 40410, 40379) whereas transcripts coding for an invertase were preferentially accumulated after 48 hpi (GH32, JGI protein ID no. 44167). Concomitant expression of genes encoding GH32 and hexose transporters strongly support carbon acquisition from the host after the action of invertases.

In the bean rust *U. fabae*, many lytic enzymes were identified concomitantly with appressorium formation at early stages of rust infection such as chitin deacetylase, proteases and cellulases (Voegelé et al. 2009). Whereas few lytic enzymes were detected at early stages of infection i.e. 24 hpi (Figure 5 and 7), a large proportion of carbohydrate active enzymes, proteases and lipases, were preferentially expressed at later stages of rust infection and particularly when sporulation occurs (Figure 5 and 7). These enzymes could be involved in the degradation and/or softening of the plant cell wall at uredinial stage to allow formation of the uredinia (Hacquard et al. 2010), the degradation of plant defense proteins lately produced in poplar leaf (Duplessis et al. 2009) or could indicate a shift in lipid metabolism for storage in newly formed spores.

One week after inoculation, uredinia are formed and newly urediniospores are released to achieve further fungal infection of poplar. Gene expression profiles measured using NimbleGen whole genome exon oligoarrays during time-course infection of poplar leaves led to the identification of rust genes strongly induced at this particular stage of rust development (168 hpi) compared to the earliest timepoint (96 hpi) (Supplementary table 4). Nevertheless, the complexity of fungal structures at this late stage of infection (i.e. infection hyphae, haustoria and uredinia containing sporogenous hyphae and newly formed dicaryotic urediniospores) do not allow to show if those genes are related to late biotrophy or sporulation of the rust fungus. In

this purpose, expression levels of those genes were checked in different areas corresponding to uredinia (sporulation area) and subjascent zones (biotrophic areas) previously isolated using laser capture microdissection at uredinial stage (Hacquard et al. 2010). Interestingly, among the 20 most highly induced genes at 168 hpi compared to 96 hpi (Supplementary table 4), 18 genes were preferentially expressed in the sporulation microdissected area suggesting their implication in fungal sporogenesis (Supplementary table 5).

Material & Methods

Growth conditions of plants and fungi and inoculation procedures

Experiments were performed using strain 98AG31 (pathotype 3-4-7) of *M. larici-populina* on *Populus trichocarpa* × *Populus deltoides* ‘Beaupré’ leaves (compatible interaction). Urediniospores of *M. larici-populina* were propagated on detached leaves of susceptible *P. deltoides* × *Populus nigra* ‘Robusta’ as previously reported (Rinaldi et al. 2007). Germlings were obtained from 1 mg of urediniospores grown on water agar medium (2%) in Petri dishes for 3 hours at 19±1°C. Plant inoculation procedures were performed as previously described using the same inoculum dose of 100,000 urediniospores/ml and strictly identical culture conditions (Rinaldi et al. 2007). The samples harvested at different time points (2, 6, 12, 24, 48, 96, 168 hours post inoculation, hpi) in the different treatments were immediately snap-frozen in liquid nitrogen and kept at -80°C for further nucleic acid isolation.

RNA extraction and cDNA synthesis

Isolation of total RNA was performed with the RNeasy Plant Mini kit (Qiagen, Courtaboeuf, France) from 1 mg of ungerminated and germinated spores, and from 100 mg of infected leaf tissues during compatible interaction (2, 6, 12, 24, 48, 96, 168 hpi), including a DNase I (Qiagen) treatment according to the manufacturer’s instructions to eliminate traces of genomic DNA. Electrophoretic RNA profiles were assessed with an Experion analyzer using the Experion RNA Standard-sens analysis kit (Bio-Rad, Marnes la Coquette, France). For microarray experiments, total RNA from each sample (spores, germlings, 2, 6, 12, 24, 48, 96, 168 hpi) was subjected to a single round of amplification using the MessageAmp™ II aRNA amplification kit (Ambion, Austin, USA) following manufacturer’s instructions. RNA amplification generated more than 12 µg of amplified RNA (aRNA) for each sample and aRNA profiles were verified using the Experion analyzer and Experion RNA Standard-sens analysis kit (Bio-Rad). Double-stranded cDNA were synthesized from 2.5 µg of aRNA using the Superscript™ Double-Stranded cDNA Synthesis Kit

(Invitrogen, Cergy Pontoise, France) according to the NimbleGen user protocol. Single dye labeling of samples, hybridization procedures and data acquisition were performed at the NimbleGen facilities (NimbleGen Systems, Reykjavik, Iceland) following their standard protocol. For quantitative PCR, first-strand cDNA were synthesized by RT from 1 µg DNase-treated total RNA using the iScript cDNA synthesis kit (Bio-Rad) in a total volume of 20 µL according to the manufacturer's instructions.

Construction of M. larici-populina exon oligoarray

The *M.larici-populina* custom-exon expression oligoarray (4 x 72K) manufactured by Roche NimbleGen Systems Limited (Madison, WI) (<http://www.nimblegen.com/products/exp/index.html>) contained four independent, non-identical, 60-mer probes per gene model coding sequence. Included in the oligoarray were 17,556 coding sequences, 1063 random 60-mer control probes and labelling controls (Hacquard et al. 2010). For 440 gene models, technical duplicates were included on the array for internal quality control. The 17,556 coding sequences were selected based on an earlier version of the *M. larici-populina* gene catalog set at the Joint Genome Institute before manual curation of gene and transposable elements and also contained rust ESTs and transposable elements (Duplessis et al. submitted). The actual *M. larici-populina* gene complement is of 14,527 genes (01/09/2010), of which 13,093 were represented on the array used in the study. Indeed, 281 oligonucleotide probes that presented a risk of cross-hybridization with poplar transcripts (i.e. fluorescence signal over the background level when arrays were hybridized with non-inoculated poplar leaf cDNA) and 859 oligonucleotide probes that presented a risk of cross-hybridization between transcript species expressed by different genes from a same gene family (i.e. probes with more than 90% homology between two transcripts) were not considered in our analysis. Finally, an almost entire transcriptome was considered in this study with more than 90% of the *M. larici-populina* predicted genes.

Microarray data analysis

Microarray probe intensities were quantile normalized across chips. Average expression levels were calculated for each gene from the independent probes on the array and were used for further analysis. Raw array data were then renormalized using the ARRAYSTAR software (DNASTAR, Inc. Madison, WI, USA). A transcript was deemed expressed when its signal intensity was three-fold higher than the mean signal-to-noise threshold (cut-off value) of 1,063 random oligonucleotide probes present on the array. Gene models with an expression value higher than

three-fold the cut-off level were considered as transcribed. All expression assays were conducted on three independent biological replicates. A Student *t*-test with false discovery rate (FDR) (Benjamini-Hochberg) multiple testing correction was applied to the data using the ARRAYSTAR software (DNASTAR). Transcripts with a significant *p*-value (<0.05) and more than a 3-fold change in transcript level were considered as differentially expressed. The complete expression datasets are available at the Gene Expression Omnibus (NCBI) as serie # GSE23097. The following link has been created to allow review of record GSE23097:<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=brsldmoacugumtm&acc=GSE23097>. In order to derive rust gene expression profiles along the timecourse infection of poplar leaves that allow comparison of expression levels at the different stages (i.e., 24, 48, 96 and 168 hpi) we calculated Relative Expression Indexes (REI) similar to those described in Both et al (2005b). For each gene, a mean expression level was calculated along the timecourse serie and the expression level measured at each stage was compared to this reference. REI were then normalized by log2 transformation to allow the use of k-means partitioning methods implemented in the Genesis software (Sturn et al. 2002; http://genome.tugraz.at/genesisclient/genesisclient_description.shtml).

Reverse Transcription quantitative-PCR

Specific primers amplifying fragments ranging from 102 to 262 nucleotides were designed with the Primer 3 and Amplify 3X programs (Supplementary table 6). Primers efficiency was assessed for each target sequence and ranged between 76 and 118%. A BLASTN search was performed against the *P. trichocarpa* genome sequence (Tuskan et al. 2006) with each primer sequence to verify the absence of cross annealing in the poplar genome sequence. For amplification, 2 µL of one tenth diluted RT products were amplified in 1X iQ SYBR Green Supermix (Bio-Rad) with 1.6 µM of specific 5'- and 3'-primers using a MJ-opticon2 DNA engine (Bio-Rad). Normalisation of gene expression was assess using two reference genes, a-tubulin (*Mlp-aTUB*) and elongation factor-1-alpha (*Mlp-ELF1a*) previously described (Hacquard et al. 2011). Expression of the *M. larici-populina* candidate genes was determined using the 2^{-ΔCt} calculation (Livak and Schmittgen 2001). In order to allow comparison of expression levels measured with oligoarrays and by RTqPCR, expression of selected genes were compared to their REI calculated for each technique.

Acknowledgements

We greatly thank the Melampsora Genome Consortium and particularly David Joly, Nicolas Feau, Philippe Tanguay and Richard Hamelin for the quality of their expert annotation of *M. larici-*

populina predicted genes. We thank the Joint Genome Institute for allowing early access to the *M. larici-populina* genome sequence. JGI sequencing is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This work was supported by the Région Lorraine and the Institut National de la Recherche Agronomique (Doctoral Scholarship to Stéphane Hacquard and Junior Scientist Research Grants to Sébastien Duplessis).

References

- Azaiez, A., Boyle, B., Levée, V., and Séguin, A. 2009. Transcriptome profiling in hybrid poplar following interactions with *Melampsora* rust fungi. *Mol. Plant-Microbe Interact* 22:190-200.
- Both, M., Csukai, M., Stumpf, M. P., and Spanu, P. D. 2005a. Gene expression profiles of *Blumeria graminis* indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. *Plant Cell* 17:2107-2122.
- Both, M., Eckert, S. E., Csukai, M., Mueller, E., Dimopoulos, G., and Spanu, P. D. 2005b. Transcript profiles of *Blumeria graminis* development during infection reveal a cluster of genes that are potential virulence determinants. *Mol. Plant-Microbe Interact.* 18:125-133.
- Boyle, B., Levée, V., Hamel, L. P., Nicole, M. C., and Séguin, A. 2010. Molecular and histochemical characterisation of two distinct poplar *Melampsora* leaf rust pathosystems. *Plant Biol.* 12:364-376.
- Broeker, K., Bernard, F., and Moerschbacher, B. M. 2006. An EST library from *Puccinia graminis* f. sp. *tritici* reveals genes potentially involved in fungal differentiation. *FEMS Microbiol. Lett.* 256:273-281.
- Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A., and Ellis, J. G. 2006. Haustorially-expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18:243-256.
- Catanzariti, A. M., Dodds, P. N., and Ellis, J. G. 2007. Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiol. Lett.* 269:181-188.
- Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., Thon, M., Kulkarni, R., Xu, J. R., Pan, H., Read, N. D., Lee, Y. H., Carbone, I., Brown, D., Oh, Y. Y., Donofrio, N., Jeong, J. S., Soanes, D. M., Djonovic, S., Kolomiets, E., Rehmeier, C., Li, W., Harding, M., Kim, S., Lebrun, M. H., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L. J., Nicol, R., Purcell, S., Nusbaum, C., Galagan, J. E., and Birren, B. W. 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434:980-986.
- De Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H., and Thomma, B. P. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science* 329:953-955.

- Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Ayliffe, M. A., and Ellis, J. G. 2004. The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* 16:755-768.
- Dodds, P. N., Rafiqi, M., Gan, P. H. P., Hardham, A. R., Jones, D. A., and Ellis, J. G. 2009. Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytol.* 183:993-1000.
- Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat. Rev. Genet.* 11:539-548.
- Duplessis, S., Courty, P. E., Tagu, D., and Martin, F. 2005. Transcript patterns associated with ectomycorrhiza development in *Eucalyptus globulus* and *Pisolithus microcarpus*. *New Phytol.* 165:599-611.
- Duplessis, S., Major, I. T., Martin, F., and Séguin, A. 2009. Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit. Rev. Plant Sci.* 28:309-334.
- Feau, N., Bergeron, M. J., Joly, D., Roussel, F., and Hamelin, R. 2007. Detection and validation of EST-derived SNPs for poplar leaf rust *Melampsora medusae* f. sp. *deltoidae*. *Mol. Ecol. Notes* 7:1222-1228.
- Frey, P., Gérard, P., Feau, N., Husson, C., and Pinon, J. 2005. Variability and population of *Melampsora* rusts on poplars. In *Rust diseases on willow and poplar*. Pei, M. and McCracken, A. R., eds. Wallingford, UK, CAB International: p. 63-72.
- Hacquard, S., Delaruelle, C., Legué, V., Tisserant, E., Kohler, A., Frey, P., Martin, F., and Duplessis, S. 2010. Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Mol. Plant-Microbe Interact.* 23:1275-1286.
- Hacquard, S., Veneault-Fourrey, C., Delaruelle, C., Frey, P., Martin, F., and Duplessis, S. 2011. Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection of poplar leaves. *Physiol. Mol. Plant Pathol.* doi: 10.1016/j.pmpp.2010.10.003.
- Hahn, M., and Mendgen, K. 1992. Isolation by ConA binding of haustoria from different rust fungi and comparison of their surface qualities. *Protoplasma* 170:95-103
- Hahn, M., and Mendgen, K. 1997. Characterization of *in planta*-induced rust genes isolated from a haustorium-specific cDNA library. *Mol. Plant-Microbe Interact.* 10:427-437.
- Hahn, M., Neef, U., Struck, C., Göttfert, M., and Mendgen, K. 1997. A putative amino acid transporter is specifically expressed in haustoria of the rust fungus *Uromyces fabae*. *Mol. Plant-Microbe Interact.* 10:438-445.

- Jakupović, M., Heintz, M., Reichmann, P., Mendgen, K., and Hahn, M. 2006. Microarray analysis of expressed sequence tags from haustoria of the rust fungus *Uromyces fabae*. *Fungal Genet. Biol* 43:8-19.
- Joly, D. L., Feau, N., Tanguay, P., and Hamelin, R. C. 2010. Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (*Melampsora* spp.). *BMC Genomics* 11:422.
- Kajiwara, S., Yamaoka, K., Hori, K., Miyazawa, H., Saito, T., Kanno, T. and Shishido K. 1992. Isolation and sequence of a developmentally regulated putative novel gene, *priA*, from the basidiomycete *Lentinus edodes*. *Gene* 114:173-178.
- Kämper, J., Kahmann, R., Bölker, M., Ma, L. J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J; W., Gold, S. E., Müller, O., Perlin, M. H., Wösten, H. A. B., de Vries, R., Ruiz-Herrera, J., Reynaga-Peña, C. G., Snetselaar, K., McCann, M., Pérez-Martín, J., Feldbrügge, M., Basse, C. W., Steinberg, G., Ibeas, J. I., Holloman, W., Guzman, P., Farman, M., Stajich, J. E., Sentandreu, R., González-Prieto, J. M., Kennel, J. C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Münch, K., Rössel, N., Scherer, M., Vraneš, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E. C. H., Cahill, M. J., Boyce, K. J., Weixi Li, J., Sanchez-Alonso, P., Schreier, P. H., Häuser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schlüter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Güldener, U., Münsterkötter, M., Haase, D., Oesterheld, M., Mewes, H. W., Mauceli, E. W., DeCaprio, D., Wade, C. M., Butler, J., Young, S., Jaffe, D. B., Calvo, S., Nusbaum, C., Galagan, J. E. and Birren, B. W. (2006) Insights from the genome of the biotrophical fungal plant pathogen *Ustilago maydis*. *Nature* 444:97-101.
- Kemen, E., Kemen, A. C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., Voegelé, R. T. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol. Plant-Microbe Interact.* 18:1130-1139.
- Laurans, F., and Pilate, G. 1999. Histological aspects of a hypersensitive response in poplar to *Melampsora larici-populina*. *Phytopathology* 89:233-238.
- Le Quéré, A., Wright, D. P., Söderström, B., Tunlid, A., and Johansson, T. 2005. Global patterns of gene regulation associated with the development of ectomycorrhiza between birch (*Betula pendula* Roth.) and *Paxillus involutus* (Batsch) Fr. *Mol. Plant-Microbe Interact.* 18:659-673.
- Livak, K. J., and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC(T)} method. *Methods* 25:402-408.
- Miranda, M., Ralph, S. G., Mellway, R., White, R., Heath, M.C., Bohlmann, J., and Constabel, C. P. 2007. The transcriptional response of hybrid poplar (*Populus trichocarpa* x *P. deltoides*) to infection by *Melampsora medusae* leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. *Mol. Plant Microbe Interact.* 20:816-831.
- Oh, Y., Donofrio, N., Pan, H., Coughlan, S., Brown, D. E., Meng, S., Mitchell, T., and Dean, R. A. 2008.

Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus *Magnaporthe oryzae*. *Genome Biology* 9:R85.

Pinon, J., and Frey, P. 2005. Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In Pei M. H., McCracken A. R., eds. *Rust diseases of Willow and Poplar*. CAB International, Wallingford, UK, 139-154.

Puthoff, D. P., Neelam, A., Ehrenfried, M. L., Scheffler, B. E., Ballard, L., Song, Q., Campbell, K. B., Cooper, B., and Tucker, M. L. 2008. Analysis of expressed sequence tags from *Uromyces appendiculatus* hyphae and haustoria and their comparison to sequences from other rust fungi. *Phytopathology* 98:1126-1135.

Rafiqi, M., Gan, P. H., Ravensdale, M., Lawrence, G. J., Ellis, J. G., Jones, D. A., Hardham, A. R., and Dodds, P. N. 2010. Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell* 22:2017-2032.

Rinaldi, C., Kohler, A., Frey, P., Duchaussoy, F., Ningre, N., Couloux, A., Wincker, P., Le Thiec, D., Fluch, S., Martin, F., and Duplessis, S. 2007. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol.* 144:347-366.

Schornack, S., Huitema, E., Cano, L. M., Bozkurt, T. O., Oliva, R., Van Damme, M., Schwizer, S., Raffaele, S., Chaparro-Garcia, A., Farrer, R., Segretin, M. E., Bos, J., Haas, B; J., Zody, M. C., Nusbaum, C., Win, J., Thines, M., and Kamoun, S. 2009. Ten things to know about oomycete effectors. *Mol. Plant Pathol.* 10:795-803.

Sreenivasaprasad, S., Eastwood, D. C., Browning, N., Lewis, S. M., and Burton, K. S. 2006. Differential expression of a putative riboflavin-aldehyde-forming enzyme (raf) gene during development and post-harvest storage and in different tissue of the sporophore in *Agaricus bisporus*. *Appl. Microbiol. Biotechnol.* 70:470-476.

Steenackers, M. 1988. Breeding poplars for rust resistance – recent advances. *Communications of the Faculty of Agricultural Sciences of the State University of Ghent, Belgium.* 53:417-422.

Steenackers, J., Steenackers, M., Steenackers, V., and Stevens, M. 1996. Poplar diseases, consequences on growth and wood quality. *Biomass Bioenergy* 10:267-274.

Struck, C., Ernst, M., and Hahn, M. 2002. Characterization of a developmentally regulated amino acid transporter (AAT1p) of the rust fungus *Uromyces fabae*. *Mol. Plant Pathol.* 3:23-30.

Struck, C., Mueller, E., Martin, H., and Lohaus, G. 2004. The *Uromyces fabae* UfAAT3 gene encodes a general amino acid permease that prefers uptake of in planta scarce amino acids. *Mol. Plant Pathol.* 5:183-189.

Sturn, A., Quackenbush, J., and Trajanoski, Z. 2002. Genesis: cluster analysis of microarray data. *Bioinformatics* 18:207-208.

- Szeto, C. Y., Leung, G. S., and Kwan, H. S. 2007. Le.MAPK and its interacting partner, Le.DRMIP, in fruiting body development in *Lentinula edodes*. *Gene* 393:87-93.
- Thara, V. K., Fellers, J. P., and Zhou, J. M. 2003. *In planta* induced genes of *Puccinia triticina*. *Mol. Plant Pathol.* 4:51-56.
- Voegelé, R. T., Struck, C., Hahn, M., and Mendgen, K. 2001. The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proc. Natl. Acad. Sci. U.S.A.* 98:8133-8138.
- Voegelé, R. T., and Mendgen, K. 2003. Rust haustoria: nutrient uptake and beyond. *New Phytol.* 159:93-100.
- Voegelé, R. T., Hahn, M., and Mendgen, K. 2009. The Uredinales: Cytology, Biochemistry, and Molecular Biology. First publ. in: *The Mycota, 5. Plant relationships / Vol. ed.*: H. B. Deising. Berlin: Springer, 2. ed., pp. 69-98.
- Wahl, R., Wippel, K., Goos, S., Kämper, J., and Sauer, N. 2010. A novel high-affinity sucrose transporter is required for virulence of the plant pathogen *Ustilago maydis*. *PLoS Biol.* 8:e1000303.
- Yin, C., Chen, X., Wang, X., Han, Q., Kang, Z., and Hulbert, S. H. 2009. Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus *Puccinia striiformis* f. sp. *Tritici*. *BMC Genomics* 10:626.
- Zhang, Y., Qu, Z., Zheng, W., Liu, B., Wang, X., Xue, X., Xu, L., Huang, L., Han, Q., Zhao, J., and Kang, Z. 2008. Stage-specific gene expression during urediniospore germination in *Puccinia striiformis* f. sp. *tritici*. *BMC Genomics* 9:203.

**Validation of *Melampsora larici-populina* reference genes for
in planta RT-quantitative PCR expression profiling during
time-course infection of poplar leaves**

Hacquard S, Veneault-Fourrey C, Delaruelle C,
Frey P, Martin F, Duplessis S

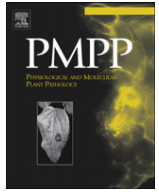
Article publié dans la revue Physiological and Molecular Plant Pathology (January 2011)

Les données supplémentaires de cette publication sont disponibles sur le site web:
http://www.elsevier.com/wps/find/journaldescription.cws_home/622932/description#description



Contents lists available at ScienceDirect

Physiological and Molecular Plant Pathology

journal homepage: www.elsevier.com/locate/pmpp

Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection of poplar leaves

Stéphane Hacquard, Claire Veneault-Fourrey, Christine Delaruelle, Pascal Frey, Francis Martin, Sébastien Duplessis*

INRA (Institut National de la Recherche Agronomique), Unité Mixte de Recherche 1136 INRA/Nancy Université 'Interactions Arbres/Micro-organismes', Centre INRA de Nancy, 54280 Champenoux, France

ARTICLE INFO

Article history:

Accepted 29 October 2010

Keywords:

Populus
Poplar rust
Biotrophic fungus
Gene expression
Housekeeping genes
Thiamine

ABSTRACT

The foliar rust caused by *Melampsora larici-populina* (*Mlp*) is the main disease affecting poplar plantations in Europe. The biotrophic status of this fungus is a major limitation to address *in planta* transcripts profiling. Thus, identification of reference rust genes steadily expressed during plant tissue colonization is a crucial point. A quantitative PCR approach to assess fungal ITS amplification profile and Reverse Transcription quantitative-PCR was set to compare candidate reference genes amplification profiles in poplar infected tissues. We selected two *M. larici-populina* genes encoding an alpha-tubulin and the elongation factor-1-alpha that showed the highest expression stability across biological samples and for which transcript levels were correlated with fungal ITS amplification during time-course infection of poplar leaves. We report the use of these reference genes to assess *in planta* expression profiles of two genes involved in thiamine biosynthesis (*THI1* and *THI2*) for which specific haustorium expression was previously described in the bean rust fungus *Uromyces fabae*. *Mlp-THI1* and *Mlp-THI2* showed similar expression profiles. Transcripts were barely detectable in urediniospores as well as during the early stages of infection compared to those reported in the bean rust, whereas a strong induction was observed after haustorial formation after 24 hpi. These data are in frame with the results obtained in *U. fabae* and consistent with a metabolic reorientation that likely occurs after the fungus derived nutrients from its host in the haustorial structure essential for fungal biotrophy.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Rust caused by *Melampsora larici-populina* is the main disease affecting poplar plantations in Europe with severe economic losses [1]. Epidemiology of the disease and poplar defence responses have been previously addressed [2,3] but still very little is known about fungal molecular mechanisms during the infection process. It is crucial to increase our understanding of the biology of the pathogen in order to develop strategies to control the disease. In the wake of the *Populus* genome sequencing [4], the ~100 Mb genome of *M. larici-populina* has been recently sequenced by the Joint

Genome Institute (JGI; <http://genome.jgi-psf.org/Mellp1/Mellp1.home.html>) and the analysis of this genome is a great opportunity to identify effector genes expressed by the rust fungus to achieve plant infection [3,5]. Availability of the genome sequences of both the plant host and the rust fungus and recent updates on *Populus* defence reactions upon rust infection [6,7,8], make the *Populus/Melampsora* pathosystem an emerging model in forest pathology to study the molecular mechanisms developed by rust to infect plant tissues and to feed on the host [3].

The biotrophic status of rust fungi precludes some routine manipulations in the laboratory. Up to now, most of the knowledge gained on rust gene expression was obtained after sequencing of Expressed Sequence Tags (EST) from urediniospores, haustorial tissues or infected plant tissues [5,9,10,11]. A synthetic system has been established to study the development of *Uromyces fabae* infection structure after the growth of urediniospores on polyethylene membranes [12]. By this means, it is possible to follow transcripts accumulation in most fungal structures usually formed

Abbreviations: Cq, quantification cycle; EST, expressed sequence tag; ITS, internal transcribed spacer; LPI, leaf plastochrony index; *Mlp*, *Melampsora larici-populina*; qPCR, quantitative polymerase chain reaction; rDNA, ribosomal DNA; RT, Reverse Transcription.

* Corresponding author. Tel.: +33 383 39 40 13; fax: +33 383 39 40 69.
E-mail address: duplessi@nancy.inra.fr (S. Duplessis).

0885-5765/\$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pmpp.2010.10.003

Please cite this article in press as: Hacquard S, et al., Validation of *Melampsora larici-populina* reference genes for ..., Physiological and Molecular Plant Pathology (2010), doi:10.1016/j.pmpp.2010.10.003

in planta after penetration in the leaf tissue, i.e. germ tubes, appressoria, primary infection hyphae and haustorial mother cells. Beside, semi-quantitative PCR was used to monitor *in planta* gene expression of *U. fabae* during bean infection [13,14]. Another method developed to specifically isolate haustoria from infected bean leaves also allowed measuring transcript levels in this structure [9,11,15]. Such approaches led to the description of specific haustorial expression profiles for several rust genes such as the thiamine biosynthesis genes *THI1*; *THI2*, the hexose transporter *HXT1* and the Rust Transferred Protein *RTP1* [15,16,17].

Great progresses have been made in epidemiology after the development of the quantitative-PCR (qPCR) technique that allows a better detection of fungal pathogens *in planta* [18,19,20,21]. Such technique was successfully applied to the poplar leaf rust disease and has allowed to assess *M. larici-populina* progression in the plant tissue through specific amplification of fungal internal transcribed spacer (ITS) by qPCR [7,19]. Due to its sensitivity and ease of use in routine in laboratory, qPCR is now a technique of choice to study gene expression, although it requires rigorous standardization to accurately interpret the data and generate reliable results [22]. Based on the specificity of primers to target fungal sequence and to its great sensitivity, qPCR is a perfect tool to follow-up accumulation of specific fungal transcripts *in planta*. However, in order to allow comparison of expression levels in time-course experiments to draw-up expression profiles, reference genes are required to take into account differences in fungal content between different time-points [23,24].

Here we present an expression assay of five reference fungal genes commonly used in molecular studies, by Reverse Transcription (RT) semi-quantitative-PCR (semi-qPCR) and qPCR during time-course infection of leaves from resistant and susceptible poplar cultivars challenged by *M. larici-populina* strain 98AG31. Based on their expression levels and stability and comparison to fungal ITS amplification profiles *in planta*, we selected alpha-tubulin (*Mlp-αTUB*) and elongation factor-1-alpha (*Mlp-ELF1a*) as candidate reference genes for *M. larici-populina* gene expression *in planta*. Finally, we report the use of these reference genes to derive expression profiles of genes for which specific haustorium expression was previously described in the bean rust fungus *U. fabae* during a compatible poplar-rust interaction.

2. Material and methods

2.1. Plant material, growth conditions and inoculation procedures

All experiments were performed on rooted cuttings of *Populus trichocarpa* × *Populus deltoides* 'Beaupré' (susceptible cultivar) and 'Hoogvorst' (resistant cultivar). For the analysis of incompatible and compatible poplar-rust interactions, plants were grown in greenhouse from dormant cuttings as previously described [7]. Isolate 98AG31 of *M. larici-populina* (pathotype 3-4-7, virulent on 'Beaupré', avirulent on 'Hoogvorst'), was used in this study and was grown on detached leaves of susceptible *P. deltoides* × *Populus nigra* 'Robusta' as previously reported [7]. Expanded leaves from leaf plastochrony index (LPI) 5 to 9 were detached from several 'Beaupré' and 'Hoogvorst' plants and were spray-inoculated on their abaxial surface with an urediniospore suspension (100,000 urediniospores/ml) or with water-agar as a control (mock-inoculated leaves) as described in Rinaldi et al. [7]. Inoculations were done by pooling leaves of different LPI from different plants for each treatment (i.e. compatible, incompatible and mock-inoculation) and each time point. The inoculated leaves were incubated with the abaxial surface uppermost, floating on deionized water in Petri dishes, at 19 ± 1 °C under continuous artificial illumination, for various durations. The material harvested at different time points in the different

treatments consisted of 30 mm diameter leaf discs randomly sampled on the overall leaf surface. The leaf discs were immediately snap-frozen in liquid nitrogen and transferred to -80 °C.

2.2. DNA and RNA extraction

Total DNA was extracted from leaf tissues with the DNeasy Plant Mini kit (Qiagen, Courtaboeuf, France) from 100 mg of frozen (-80 °C) material. RNA was removed by the addition of ribonuclease A during extraction. DNA quality was verified by electrophoresis on agarose gel, and DNA quantity was measured by spectrophotometry [25].

Isolation of total RNA was performed with the RNeasy Plant Mini kit (Qiagen) from 100 mg of pooled (-80 °C) foliar disks harvested from leaves of various LPI and various individual poplar plants for each treatment considered. Pooling of samples from different trees and LPI helped in minimizing the variations between individual RNA samples. A DNase I (Qiagen) treatment was included in the RNA isolation procedure according to the manufacturer's instructions to eliminate traces of genomic DNA. Quality and quantity of RNA samples were checked with the RNA analyzer Experion (Bio-Rad, Marnes la Coquette, France) following the manufacturer's recommendation.

2.3. ITS amplification by semi-quantitative and quantitative-PCR

In planta development of the rust fungus was followed in 'Beaupré' and 'Hoogvorst' by specific amplification of the nuclear ribosomal DNA (rDNA) ITS on total DNA isolated from inoculated leaf tissues [7,19]. Amounts of 100 ng and 10 ng DNA were used respectively for semi-qPCR and qPCR amplifications with ITS primers specific for poplar and for *M. larici-populina* (Table 1). Specificity of ITS primers was validated *in silico* with *M. larici-populina* ITS sequences available at the JGI *M. larici-populina* genome portal (<http://genome.jgi-psf.org/Mellp1/Mellp1.home.html>) and in international databases. In order to improve qPCR amplification efficiency, we designed a new set of *M. larici-populina* ITS primers amplifying a 169 bp fragment shorter than those previously described in [7,19]. These primers also matched other *Melampsora* spp. in the non-redundant nucleotide database at the National Center for Biotechnology Information (BLASTN), thus, although perfectly suited for molecular studies with controlled *M. larici-populina* inoculation, these primers should not be used with infected plants collected in field experiments due to the risk of cross-hybridization with other rust species. Semi-qPCR amplification was performed in 25 µL reaction volumes using 10 µM of primers with following amplification parameters: 95 °C for 5 min, then 95 °C for 1 min, 60 °C for 30 s, 72 °C for 1 min for a maximum of 35 cycles. For qPCR, amplifications were performed in 1X iQ SYBR Green Supermix (Bio-Rad) with 1.6 µM of specific 5' and 3' ITS primers with an MJ-opticon2 DNA engine (Bio-Rad). Assuming a signal intensity proportional to amplified ITS sequences, we considered the pathogen growth as the relative difference between cycle threshold or quantification cycle (Cq) of fungal ITS amplicons quantified in the compatible interaction compared to Cq of poplar ITS amplicons at 2, 6, 12, 24, 48, 96 and 168 h post-inoculation (hpi) ($2^{-\Delta Cq}$ calculation, [26]). PCR amplifications were carried out on two biological replicates and included two distinct technical replicates.

2.4. Transcript profiling through semi-quantitative and quantitative RT-PCR

To allow the amplification of specific transcripts by RT followed by either semi-qPCR and qPCR, we designed primers for the following *M. larici-populina* genes retrieved from the *M. larici-*

Table 1

Specific 5' and 3' primers designed for qPCR and semi-qPCR amplification of *Melampsora larici-populina* and *Populus* ITS and amplification of *Melampsora larici-populina* transcripts during controlled poplar-rust time-course infection. Primers of *Laccaria bicolor* MISSP8 gene used to test amplification inhibition are also detailed.

Accession no.	Target sequence	5' primer	3' primer	Amplicon length	Primer Efficiency (%)
NCBI ^a					
AJ006438	<i>Populus</i> ITS	GAAGCGCCAAGAAATTGA	GCGTTCAAAGACTCGATGGT	192	84
AY375268	<i>Melampsora</i> ITS	TCGGCCCTCATTCTTCTAA	GATAGTCGGAGCGACCAAAG	169	107
JGI proteinID ^b					
45600	Actin (ACT)	GGCACCTCTTAATCCAAAAGTC	GTGAGTAACACCGTCACCAGAA	163	91
73137	Alpha-tubulin (aTUB)	CAAGCGTACCATTACAGTTCGT	CACCTCTCCATACCTTCTC	239	94
73378	Elongation factor-1-alpha (ELF1a)	CGAGACTCCCAAACTCTCGTT	GTTCCAGAGTTTGACCATCTT	167	102
46086	Glyceraldehyde-3-phosphate-dehydrogenase (GAPDH)	CCATCAACGACCCCTTTATC	ATCGGCACCAGATTACCC	185	92
40152	Histone H4 (HISTH4)	GAGGAAAATCCGGTGGAAAG	TGGCATGTTCCGGTGTAGGTA	225	99
67013	Thiamine biosynthesis gene 1 (THI1)	GCTCTGCTCCATCCTTACAT	GATTTGGTTAAGTGGGCTGTC	179	97
53832	Thiamine biosynthesis gene 2 (THI2)	CTCTCGTATCGAAGGCAACTCT	CTCTCCATGTCGTTGAAGTATC	161	95
388224 ^c	Mycorrhiza-induced small secreted protein MISSP8	AGCTGATCTCAACGAGAGCGTTGAGGCC	TCAATCACTATCGCGCTCCAGTCAC	118	121

^a <http://www.ncbi.nlm.nih.gov/>

^b <http://genome.jgi-psf.org/Mellp1/Mellp1.home.html>

^c <http://genome.jgi-psf.org/Lacbi1/Lacbi1.home.html>

populina JGI genome portal (Table 1): actin, alpha-tubulin, elongation factor-1-alpha, glyceraldehyde-3-phosphate-dehydrogenase, histone H4, thiamine 1 *THI1* (*U. fabae* *THI1* homologue previously described as *PIG1*; [9]) and thiamine 2 *THI2* (*U. fabae* *THI2* homologue previously described as *PIG4* [9]). The primers were designed in the coding sequence, and amplified fragments showed expected lengths ranging between 161 and 239 nucleotides (Table 1). Primers efficiency was assessed for each target sequence and ranged between 91 and 102%. A BLASTN search was performed against the *P. trichocarpa* genome sequence [4] with each primer sequence to verify the absence of cross annealing in the poplar genome sequence. First-strand cDNAs were synthesized by RT from 1 µg DNase-treated total RNA using the iScript cDNA synthesis kit (Bio-Rad) in a total volume of 20 µL according to the manufacturer's instructions. For the semi-qPCR amplification, 1 µL of cDNA was amplified by PCR (95 °C for 5 min, then 95 °C for 1 min, 60 °C for 30 s, 72 °C for 1 min for a maximum of 35 cycles) in 25 µL reaction volume using 10 µM of primers. For the qPCR amplification, 2 µL of one-tenth diluted RT products were amplified in 1X iQ SYBR Green Supermix (Bio-Rad) with 1.6 µM of specific 5'- and 3'-primers with an MJ-opticon2 DNA engine (Bio-Rad). *M. larici-populina* gene expression was normalized to reference genes, a-tubulin (*Mlp-aTUB*) and elongation factor-1-alpha (*Mlp-ELF1a*) using the $2^{-\Delta Cq}$ calculation [26], and expressed relative to the highest expression level, set at 100%. The GeNorm VBA Applet for Microsoft Excel (<http://medgen.ugent.be/~jvdesomp/genorm/>) based on the principles and calculations described in Vandesompele et al. [27] helps to identify most stably expressed reference genes in a given panel of samples for qPCR normalization. This approach was used to compare the selected *M. larici-populina* candidate reference genes to select stably expressed housekeeping genes. In order to validate the absence of amplification inhibition in any of the rust infected plant samples tested in this study, we used the *Laccaria bicolor* gene encoding MISSP8, a protein specific of *L. bicolor* (<http://genome.jgi-psf.org/Lacbi1/Lacbi1.home.html>, protein ID 388224), as a spike control [28]. This gene was cloned in the plasmid TOPO-PCR2 (Invitrogen, Cergy-Pontoise, France) and primers were designed to specifically amplify a 118 bp region in the target gene (Table 1). Primers did not match any loci of the *P. trichocarpa* or the *M. larici-populina* genomes. An aliquot of 0.1 ng of the *pLbiMISSP8* plasmid was added to each biological sample of the time-course serie (three biological replicates) and amplification of *LbiMISSP8* was performed by qPCR as described above. *LbiMISSP8* Cq in poplar-

rust samples were then compared to Cq obtained from a 0.1 ng control plasmid sample.

3. Results

3.1. *M. larici-populina* growth in poplar leaves during compatible and incompatible interactions

In planta growth of *M. larici-populina* strain 98AG31 was limited during the early stages of 'Beaupré' leaves infection (2, 6 and 12 hpi; Fig. 1). A slow increase of fungal ITS amount was observed between 12 and 48 hpi (6-fold) while a drastic increase of about 70-fold was measured between 48 and 96 hpi, indicating a fast colonization of the plant mesophyll by the rust fungus during the compatible interaction. A fungal growth slackening was observed between 96 and 168 hpi, the later time-point corresponding to urediniospores formation. Interestingly, at the end of the time-course infection, fungal ITS amount in 'Beaupré' leaves was about 7000-fold higher than the amount measured at the earliest time-point (2 hpi). At the contrary, *M. larici-populina* strain 98AG31 growth was limited during the whole time-course infection of 'Hoogvorst' leaves, corresponding to an incompatible poplar-rust interaction (Fig. 1). These fungal growth profiles confirm and extend the results

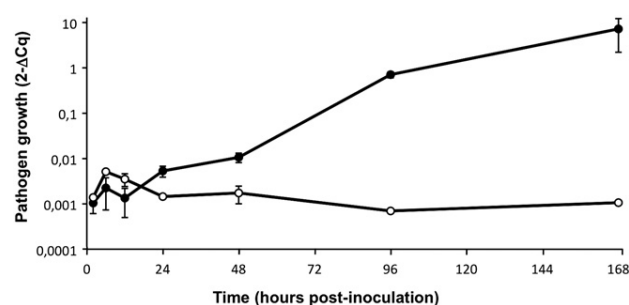


Fig. 1. Time-course infection of poplar 'Beaupré' (susceptible cultivar, black dots) and 'Hoogvorst' (resistant cultivar, white dots) leaves by *M. larici-populina* strain 98AG31. Development of the rust strain was monitored by specific quantitative PCR amplification of the rDNA intergenic ITS region from total DNA extracted from inoculated leaf tissues at 2, 6, 12, 24, 48, 96, 168 h post-inoculation. Pathogen growth curves correspond to fungal ITS amplicons compared to poplar ITS amplicons ($2^{-\Delta Cq}$). Note the log-scale.

described by Rinaldi et al. [7] during the poplar-rust interaction. Whereas Rinaldi et al. [7] compared compatible and incompatible poplar-rust interactions based on the growth curves of two distinct *M. larici-populina* strains (i.e. virulent and avirulent) on the same poplar cultivar ('Beaupré'), the results presented here corresponded to the same *M. larici-populina* strain infecting two poplar cultivars with distinct genetic resistance background. Interestingly, similar fungal growth profiles were observed independently of the combination of cultivars/strains between 2 and 96 hpi. In the present study, the time-course was extended to the uredinia formation and urediniospores production, indicating that the uredinial stage is marked by a drastic increase in fungal mass *in planta*.

3.2. RT-semi-qPCR expression profiles of *M. larici-populina* reference genes

In order to accurately measure the relative abundance of a target transcript, normalization to stable reference transcripts is required [26,27,28,29]. Expression level of the typical fungal reference genes actin (*Mlp-ACT*), glyceraldehyde-3-phosphate-dehydrogenase (*Mlp-GAPDH*), alpha-tubulin (*Mlp-aTUB*), elongation factor-1-alpha (*Mlp-ELF1a*) and histone H4 (*Mlp-HISTH4*) in *M. larici-populina* was assessed by semi-qPCR during incompatible and compatible

interactions (Fig. 2A) and compared to fungal ITS semi-qPCR amplification profiles (Fig. 2B). Transcripts of these genes were not detected or detected at lower levels during the incompatible interaction compared to the compatible one, consistent with the lower growth noticed by fungal ITS amplification (Figs. 1 and 2). During the compatible interaction, *Mlp-ACT* and *Mlp-GAPDH* were not or barely detectable at early stages of infection and thus might not be proper reference genes in the corresponding biological situations. In contrast, *Mlp-aTUB*, *Mlp-ELF1a* and *Mlp-HISTH4* were detected at all time-points of the infection process in the compatible interaction. Semi-qPCR profiles of these candidate genes were comparable to *M. larici-populina* ITS profile (Fig. 2) and might represent candidate reference genes. Accuracy of these *M. larici-populina* reference genes was further tested by RT-qPCR.

3.3. qPCR amplification inhibition assay

The absence of amplification inhibition was verified in our experimental system using the SPUD assay described by Nolan et al. [28] with *MISSP8* a specific gene of the symbiotic basidiomycete *L. bicolor* as a spike control gene. For this, 0.1 ng of plasmid *pLbiMISSP8* were added to cDNA and genomic DNA samples corresponding to the time-course infection series used in the study for semi-qPCR and qPCR, then qPCR amplification cycle Cq of the 118 bp *MISSP8* cDNA were compared with the sole plasmid. Identical Cq values were recovered for the different biological situations and the control plasmid validating the absence of amplification inhibition in the poplar-rust interactions samples (Supplemental Fig. 1).

3.4. Validation of *M. larici-populina* reference genes expression profiles by RT-qPCR

To validate the use of *M. larici-populina* candidate reference genes as relative controls in expression studies of compatible poplar-rust interaction, we tested their amplification by RT-qPCR and compared the *M. larici-populina* ITS Cq to reference genes Cq during time-course infection of 'Beaupré' leaves. The low expression of *Mlp-GAPDH* observed by semi-qPCR was confirmed, indicating that this gene is not suitable for normalization during time-course infection. Remaining candidate reference genes were all detected and Cq could be determined and compared to fungal ITS Cq (Fig. 3A). Although all tested genes presented similar profiles at late stages of infection, differences were observed between 2 and 24 hpi. Indeed, *Mlp-HISTH4* and *Mlp-ACT* both showed a higher Cq at 24 hpi inconsistent with the ITS Cq (Fig. 3A). GeNorm [27] was used to define the qPCR expression stability of the four genes. *Mlp-aTub* and *Mlp-ELF1a* showed the lowest expression stability values (M) which correspond to the most stable expressed genes (Fig. 3B). Based on the qPCR profiles and the GeNorm assay, we picked-up the *Mlp-aTub* and the *Mlp-ELF1a* as the best candidate reference genes for RT-qPCR normalization. To further validate the use of those two reference genes, a total of five distinct technical replicates of *Mlp-aTUB* and *Mlp-ELF1a* RT-qPCR amplification (three biological replicates in each assay) were performed by two different laboratories in different laboratory conditions such as different laboratories, different days, different iQ SYBR Green Supermixes (Bio-Rad) and different pools of cDNA synthesized from the same RNA (technical replicates 1–5 in Supplemental Fig. 2).

At early stages of infection, *Mlp-aTUB* and *Mlp-ELF1a* could be detected with a Cq of 30 and 29 cycles respectively (Fig. 3A), although the *in planta* fungal biomass was rather low compared to later time-points. During the infection process, the Cq of those genes showed similar decreases to 17 and 16 cycles respectively, at 168 hpi. Cq profiles of reference genes and ITS were almost comparable during the overall time-course infection. The curves

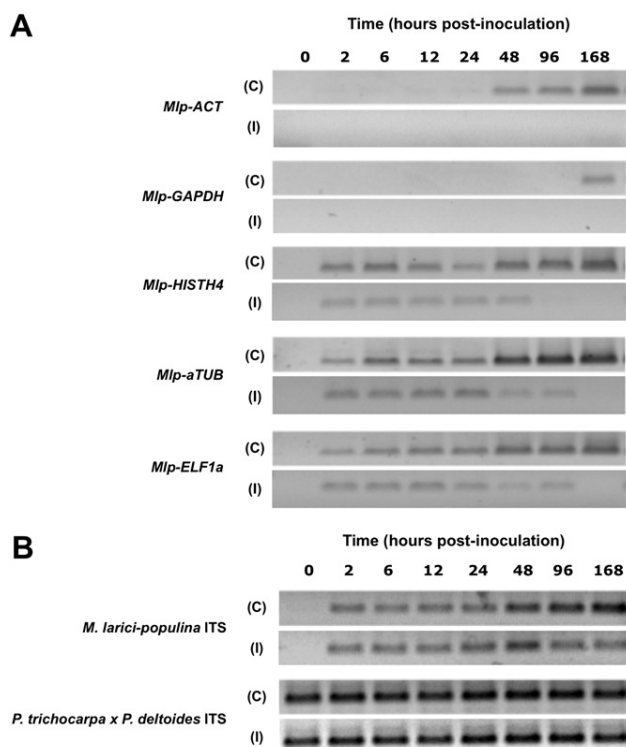


Fig. 2. Semi-quantitative RT-PCR expression profiles for *M. larici-populina* transcripts encoding actin (*Mlp-ACT*), alpha-tubulin (*Mlp-aTUB*), elongation factor-1-alpha (*Mlp-ELF1a*), glyceraldehyde-3-phosphate-dehydrogenase (*Mlp-GAPDH*), histone H4 (*Mlp-HISTH4*) during time-course infection of *P. trichocarpa x P. deltooides* 'Beaupré' (compatible interaction, C) and 'Hoogvorst' (incompatible interaction, I) leaves at 2, 6, 12, 24, 48, 96 and 168 h post-inoculation (hpi) and non inoculated leaves (0 h). A, total RNA of 'Beaupré' or 'Hoogvorst' leaves inoculated with *M. larici-populina* strain 98AG31 was isolated and aliquots of 1 µg were used for first-strand cDNA synthesis. PCR was performed using 1 µl of first-strand cDNA. A control with no RT in the first-strand cDNA synthesis reaction mix was included to control for the lack of genomic DNA. B, PCR amplification of poplar and fungal ITS fragments from DNA samples are indicated as a reference of *in planta* fungal growth progression. In Fig. 2A and B, amplified DNA was sampled after a total of 35 cycles of PCR.

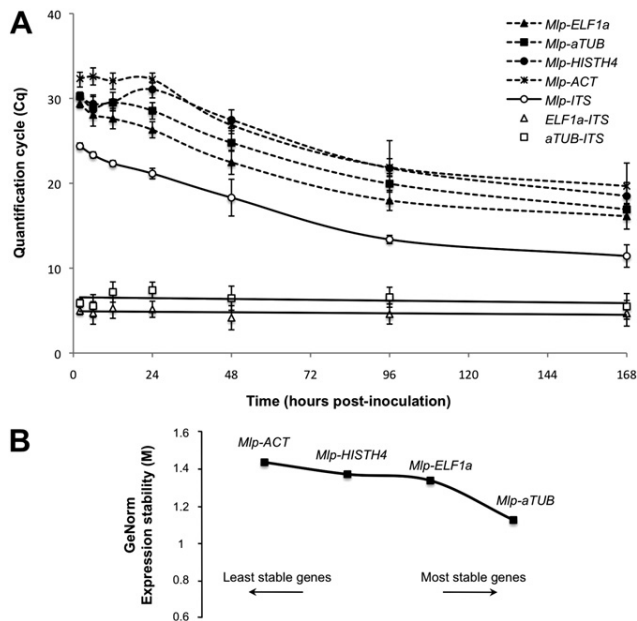


Fig. 3. Comparison of quantitative PCR (qPCR) quantification cycle Cq between fungal ITS fragments and fungal transcripts encoding an actin (*Mlp-ACT*), an alpha-tubulin (*Mlp-aTUB*), an elongation factor-1-alpha (*Mlp-ELF1a*) and an histone H4 (*Mlp-HISTH4*) amplified from *P. trichocarpa* × *P. deltooides* 'Beaupré' infected leaves during time-course infection with *M. larici-populina* strain 98AG31. A, *M. larici-populina* ITS Cq were measured by qPCR from DNA isolated from inoculated leaves at 2, 6, 12, 24, 48, 96, 168 hpi. *Mlp-ACT*, *Mlp-aTUB*, *Mlp-ELF1a* and *Mlp-HISTH4* transcripts Cq were obtained by RT-qPCR from total RNA isolated from the same tissues at the same time-points. Direct subtraction of ITS Cq from *Mlp-aTUB* and *Mlp-ELF1a* transcripts Cq shows a linear correlation between Cq, respectively $y = 0.004x + 6.58$ ($R^2 = 0.11$) and $y = 0.002x + 4.92$ ($R^2 = 0.16$), during time-course infection of poplar leaves. B, Expression stability values (M) of *Mlp-ACT*, *Mlp-aTUB*, *Mlp-ELF1a* and *Mlp-HISTH4* candidate reference genes calculated with GeNorm. Lowest M value indicates most stable expressed genes.

obtained with *Mlp-aTUB* and *Mlp-ELF1a* genes were parallel to that obtained with ITS, suggesting a strong correlation between reference genes expression and fungal biomass content in plant tissue during time-course infection. Subtraction of the ITS Cq from *Mlp-aTUB* and *Mlp-ELF1a* transcripts Cq shows parallel slopes (Fig. 3A), respectively of $y = 0.004x + 6.58$ ($R^2 = 0.11$) and $y = 0.002x + 4.92$ ($R^2 = 0.16$), indicating a good correlation between reference genes expression and fungal biomass content in plant tissue during time-course infection.

3.5. RT-qPCR expression assays

RT-qPCR analysis was used to investigate the relative expression of *Mlp-THI1* and *Mlp-THI2*, two homologs of thiamine biosynthesis genes highly expressed in haustoria as previously reported in *U. fabae* [9,16]. cDNA templates derived from *M. larici-populina* ungerminated urediniospores and infected 'Beaupré' leaves (2, 6, 12, 24, 48, 96, 168 hpi) were used to follow *in planta* transcript profiles during compatible poplar-rust interaction. *Mlp-aTUB* and *Mlp-ELF1a* were used as constitutively expressed endogenous controls for normalization of gene expression. Mean Cq of the two genes was considered at each time-points for $2^{-\Delta Cq}$ calculations. Calculations performed separately with either *Mlp-aTUB* or *Mlp-ELF1a* as a unique reference gene, did not show significant differences (data not shown). Thus, the mean Cq of the two reference genes was used rather than comparison to a single reference gene. *Mlp-THI1* and *Mlp-THI2* were both expressed in *M. larici-populina* during plant infection after 48 hpi and 24 hpi respectively (Fig. 4).

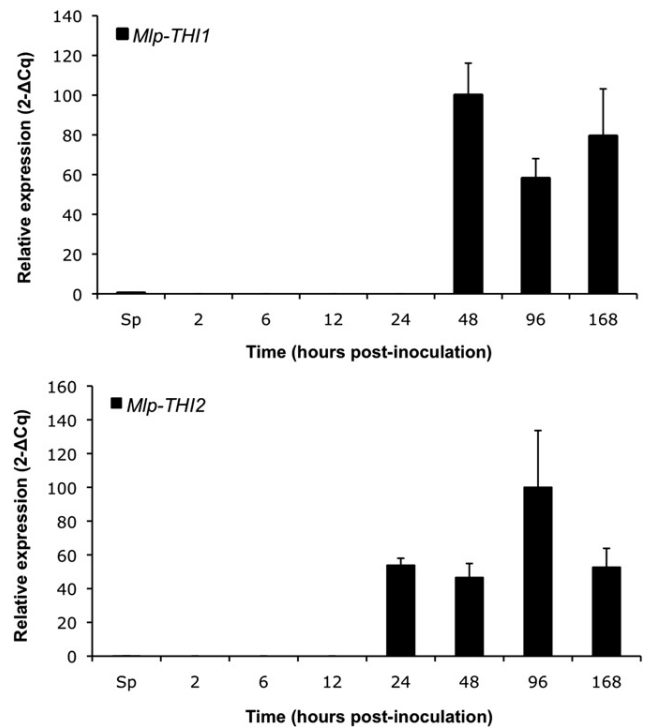


Fig. 4. Expression patterns of *M. larici-populina* thiamine biosynthesis transcripts *Mlp-THI1* and *Mlp-THI2*. Total RNA of ungerminated urediniospores and inoculated *P. trichocarpa* × *P. deltooides* 'Beaupré' leaves with *M. larici-populina* strain 98AG31 at 2, 6, 12, 24, 48, 96 and 168 h post-inoculation were isolated and aliquots of 1 µg were used for first-strand cDNA synthesis. PCR was performed using 2 µl of one-tenth diluted first-strand cDNA. A control with no RT in the first-strand cDNA synthesis reaction mix was included to control the lack of genomic DNA. Both *Mlp-aTUB* and *Mlp-ELF1a* transcripts were considered as reference transcripts for the calculation of *Mlp-THI1* and *Mlp-THI2* regulation levels ($2^{-\Delta Cq}$). For each transcript, expression regulation levels are presented as percentages of the highest regulation level measured (100%).

Expression was not or barely detectable at early stages of 'Beaupré' leaves infection or in ungerminated urediniospores. Both transcripts showed at least a 200-fold induction in infected leaves compared to ungerminated urediniospores suggesting their important role during the biotrophic growth phase.

4. Discussion

Expression profiling is an approach of choice to have a global view and a better understanding of the biological processes that control plant–pathogen interactions. Attempts to study *in planta* development of the barley powdery mildew fungus *Blumeria graminis* f. sp. *hordei* had been made possible through the use of transcriptome profiling [30]. Similar attempts had been also conducted in hemi-biotroph and biotroph fungal pathogens including *Magnaporthe oryzae* [31,32], *Ustilago maydis* [33] and *U. fabae* [34] although only few studies addressed gene expression *in planta*.

Most of our knowledge on expressed genes in rust fungi mostly derived from the sequencing of expressed sequence tags (ESTs) from infected leaves or isolated haustoria [5,9,10]. In the bean rust *U. fabae*, such approaches led to the identification of *in planta* induced genes (PIGs) expressed after the host plant invasion [9,34]. Transcript profiling were analysed in isolated haustoria, in infected leaves at late time-points of interaction, and in fungal structures grown *in vitro* in a synthetic culture system which permits to produce *ex planta* early infection structures similar to those produced in the plant by the rust fungus. This useful system allows

to harvest synchronized structures such as germ tubes, substomatal vesicles or infection hyphae [9]. Northern blot analyses of leaf infection by *U. fabae* helped in identifying transcripts highly or specifically expressed in the haustorial structure such as the secreted rust transferred protein 1 [17], the hexose transporter *HXT1* [15] or genes involved in thiamine biosynthesis such as *THI1* and *THI2* [9,16]. Specific *U. fabae* gene expression was further validated *in planta* after haustorium formation using cDNA-array [34], unfortunately, very little is known about *in planta* expression of rust genes during the early stages of the plant–fungus interaction.

The qPCR technique is very sensitive and had been successfully applied to pathogen detection in field samples for various pathosystems [35,36,37]. Particularly, this technique has allowed the detection and quantification of rust fungi belonging to *Melampsora* spp. in poplar leaves [7,19]. The RT-qPCR technique is now commonly used in molecular studies replacing Northern blotting as a routine tool to quantify transcripts. This approach should be an excellent mean to assess gene expression in biotrophic pathogens infecting their hosts. To date, only few studies reported the application of real time RT-PCR to study the relative quantification of plant pathogen gene expression *in planta* [38,39,40,41]. However, there are serious pitfalls to avoid when conducting such quantification, and it is now required to comply with simple but necessary guidelines prior using the qPCR technique for gene expression [22].

As non-cultivable organisms, biotrophic pathogens such as rust fungi require to set compatible systems with their host plant. Thus, a major concern to perform gene expression study by RT-qPCR *in planta* is the increase of fungal biomass during the time-course of interaction with the host. Considering this limitation, identification of genes showing a constant expression related to fungal growth is a crucial point in order to have a common reference to derive expression profiles and set comparison between genes at different time-points. To validate the use of RT-qPCR for *in planta* gene expression of the poplar-rust *M. larici-populina*, five candidate reference genes typically used in plant pathogen expression studies [23,42] were selected and primers were defined to perform their amplification. Time-course infections have been established between *M. larici-populina* strain 98AG31 and *P. trichocarpa* × *P. deltoides* ‘Beaupré’ (compatible interaction) and ‘Hoogvorst’ (incompatible interaction) leaves. Fungal growth was monitored *in planta* by qPCR using quantification of the ITS of fungal nuclear rDNA [19]. Fungal ITS amplification by qPCR during time-course infection of *P. trichocarpa* × *P. deltoides* ‘Beaupré’ leaves was considered as an *in planta* growth curve reference for the rust fungus (Fig. 1). The assessment of the fungal growth during the compatible poplar–rust interaction is consistent with our knowledge of the progression of rust infection structures in poplar leaves as described using histology and microscopy approaches [7,41,43,44]. Fungal growth was limited to a basal level during the time-course incompatible interaction, which is consistent with the early and strong defense response mounted by poplar to constrain rust progression in the mesophyll [3,7]. RT Semi-qPCR profiles of rust genes during the incompatible interaction indicated lower levels of expression compared to the compatible interaction. However, as the plant express a strong defense reaction during this type of interaction (i.e. hypersensitive response, secreted pathogenesis-related proteins and proteases) [3,7], we considered that potential degradation of fungal nucleic acids might interfere with proper quantification of fungal material. Thus, we focused our analysis to the compatible interaction. The *Mlp-aTUB* and *Mlp-ELF1a* transcripts encoding respectively an alpha-tubulin and an elongation factor-1-alpha of *M. larici-populina*, presented RT-qPCR Cq profiles consistent with ITS qPCR amplification curve, i.e. fungal increase during poplar leaf colonization (Fig. 3), during the compatible interaction. GeNorm is an application designed to

identify the most stable genes within a selection of candidate reference genes across biological samples [27]. Both *Mlp-aTUB* and *Mlp-ELF1a* were considered by GeNorm as more stable than *Mlp-ACT* and *Mlp-HISTH4*. *Mlp-GAPDH* was also tested in our experimental set-up but not detected. Finally, we considered that *Mlp-aTUB* and *Mlp-ELF1a* represent appropriate and reliable candidate reference genes to perform *in planta* rust transcript profiling by RT-qPCR during compatible poplar-rust interaction.

Both *Mlp-aTUB* and *Mlp-ELF1a* transcripts were used to normalize the *in planta* expression profiles of the thiamine biosynthesis genes *THI1* and *THI2* for which a specific expression was previously described in *U. fabae* haustoria [9,16]. In rust infected bean leaves, similar expression patterns were observed for *Uf-THI1* and *Uf-THI2*, with no transcript detection in urediniospores and *in vitro*-differentiated infection structures up to 18 h after germination. A few transcripts were detected in 24 h-old *in vitro* infection structures whereas a high mRNA accumulation was described in haustoria and rust infected leaves at 5 dpi [16]. In poplar leaves infected by *M. larici-populina*, RT-qPCR assays showed *THI* transcript profiles similar to those described for *U. fabae* *THI* genes. In the time-course infection of ‘Beaupré’ leaves, transcripts were barely detected in ungerminated urediniospores or at early stages of infection (i.e., 2, 6 and 12 hpi). *Mlp-THI2* transcripts were detected after 24 hpi and *Mlp-THI1* after 48 hpi (Fig. 4). In *M. larici-populina*-infected poplar leaves, first haustoria have been reported at 17 hpi and their number consistently increase at later stages until uredinia formation [43]. The results described in the present study suggest that thiamine biosynthesis likely occurs in *M. larici-populina* after haustorium formation. These results are in frame with the hypothesis of Sohn et al. [16] that haustorium is an essential structure of the biotrophic rust mycelium not only for nutrient uptake from the plant host cell but also for metabolites biosynthesis such as thiamine. This cofactor, required for the activity of several enzymes of the central carbon metabolism is strongly accumulated in *U. fabae* only during parasitic growth, suggesting that thiamine is growth limiting and not available from the host plant [16]. We can postulate that thiamine biosynthesis also plays an important role for *M. larici-populina* and more generally that it has a key role in rust fungi biology during host infection.

Here we demonstrated that the high sensitivity of real time PCR can help in determining pathogen gene expression *in planta*, even at early stages of infection when the pathogen biomass is very limited in the host. Such approach is crucial to monitor *in planta* transcript profiles of biotrophic fungi. Rust genomic data recently made available to the community with the sequencing of two Pucciniales genomes (*M. larici-populina* at the Joint Genome Institute; *Puccinia graminis* f. sp. *tritici* at the Broad Institute) are valuable resources to design specific primers in order to address expression of rust transcripts during compatible interaction with their respective hosts and to get a better understanding of rust biology.

Acknowledgements

This work was supported by the Région Lorraine and the Institut National de la Recherche Agronomique (Doctoral Scholarship to Stéphane Hacquard and Junior Scientist Research Grants to Sébastien Duplessis). The authors would like to thank Patrice Vion and Bénédicte Fabre for taking care of the poplar nursery and cuttings at INRA Nancy.

Appendix. Supplementary material

Supplementary material associated with this paper can be found, in the online version, at doi:10.1016/j.pmp.2010.10.003.

References

- [1] Pinon J, Frey P. Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In: Pei MH, McCracken AR, editors. Rust diseases of willow and poplar. Wallingford, UK: CAB International; 2005. p. 139–54.
- [2] Barrès B, Halkett F, Dutech C, Andrieux A, Pinon J, Frey P. Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in Europe and recent founder effects overseas. *Infect Genet Evol* 2008;8:577–87.
- [3] Duplessis S, Major I, Martin F, Séguin A. Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit Rev Plant Sci* 2009;28:309–34.
- [4] Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 2006;313:1596–604.
- [5] Feau N, Bergeron M-J, Joly D, Roussel F, Hamelin R. Detection and validation of EST-derived SNPs for poplar leaf rust *Melampsora medusae* f. sp. *deltoides*. *Mol Ecol Notes* 2007;7:1222–8.
- [6] Miranda M, Ralph SG, Mellway R, White R, Heath MC, Bohlmann J, et al. The transcriptional response of hybrid poplar (*Populus trichocarpa* x *P. deltoides*) to infection by *Melampsora medusae* leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. *Mol Plant-Microbe Interact* 2007;20:816–31.
- [7] Rinaldi C, Kohler A, Frey P, Duchaussoy F, Ningre N, Couloux A, et al. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol* 2007;44:347–66.
- [8] Azaiez A, Boyle B, Levée V, Séguin A. Transcriptome profiling in hybrid poplar following interactions with *Melampsora* rust fungi. *Mol Plant-Microbe Interact* 2009;22:190–200.
- [9] Hahn N, Mendgen K. Characterization of *in planta*-induced rust genes isolated from a haustorium-specific cDNA library. *Mol Plant-Microbe Interact* 1997;10:427–37.
- [10] Catanzariti AM, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis JG. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 2006;18:243–56.
- [11] Puthoff DP, Neelam A, Ehrenfried ML, Scheffler BE, Ballard L, Song Q, et al. Analysis of expressed sequence tags from *Uromyces appendiculatus* hyphae and haustoria and their comparison to sequences from other rust fungi. *Phytopathology* 2008;98:1126–35.
- [12] Deising H, Jungblut PR, Mendgen K. Differentiation-related proteins of the broad bean rust fungus *Uromyces viciae-fabae*, as revealed by high resolution two-dimensional polyacrylamide gel electrophoresis. *Arch Microbiol* 1991;155:191–8.
- [13] Wirsel SGR, Voegelé RT, Mendgen KW. Differential regulation of gene expression in the obligate biotrophic interaction of *Uromyces fabae* with its host *Vicia faba*. *Mol Plant-Microbe Interact* 2001;14:1319–26.
- [14] Struck C, Mueller E, Martin H, Lohaus G. The *Uromyces fabae* UFAAT3 gene encodes a general amino acid permease that prefers uptake of *in planta* scarce amino acid. *Mol Plant Pathol* 2004;5:183–9.
- [15] Voegelé RT, Struck C, Hahn M, Mendgen K. The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proc Natl Acad Sci USA* 2001;98:8133–8.
- [16] Sohn J, Voegelé RT, Mendgen K, Hahn M. High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*. *Mol Plant-Microbe Interact* 2000;13:629–36.
- [17] Kemen E, Kemen AC, Raffiqi M, Hempel U, Mendgen K, Hahn M, et al. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol Plant-Microbe Interact* 2005;18:1130–9.
- [18] Gachon C, Saindrenan P. Real-time PCR monitoring of fungal development in *Arabidopsis thaliana* infected by *Alternaria brassicicola* and *Botrytis cinerea*. *Plant Physiol Biochem* 2004;42:367–71.
- [19] Boyle B, Hamelin RC, Séguin A. *In vivo* monitoring of obligate biotrophic pathogen growth by kinetic PCR. *Appl Environ Microb* 2005;71:1546–52.
- [20] Demontis MA, Cacciola SO, Orru M, Balmas V, Chessa V, Maserti BE, et al. Development of real-time PCR systems based on SYBR® Green I and TaqMan® technologies for specific quantitative detection of *Phoma tracheiphila* in infected Citrus. *Eur J Plant Pathol* 2008;120:339–51.
- [21] Stephens AE, Gardiner DM, White RG, Munn AL, Manners JM. Phases of infection and gene expression of *Fusarium graminearum* during crown rot disease of wheat. *Mol Plant-Microbe Interact* 2008;21:1571–81.
- [22] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR Experiments. *Clinic Chem* 2009;55:611–22.
- [23] Fang W, Bidochka MJ. Expression of genes involved in germination, conidiation and pathogenesis in *Metarhizium anisopliae* using quantitative real-time RT-PCR. *Mycol Res* 2006;110:1165–71.
- [24] Bohle K, Jungebloud A, Göcke Y, Dalpiaz A, Cordes C, Horn H, et al. Selection of reference genes for normalisation of specific gene quantification data of *Aspergillus niger*. *J Biotechnol* 2007;132:353–8.
- [25] Sambrook J, Russell DW. Molecular cloning: a laboratory manual. Ed 3, vols. 1–3. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
- [26] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDC(t) method. *Methods* 2001;25:402–8.
- [27] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3:R0034.
- [28] Nolan T, Hands RE, Ogunkolade W, Bustin SA. SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem* 2006;351:308–10.
- [29] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucl Acids Res* 2001;29:e45.
- [30] Both M, Eckert SE, Csukai M, Müller E, Dimopoulos G, Spanu PD. Transcript profiles of *Blumeria graminis* development during infection reveal a cluster of genes that are potential virulence determinants. *Mol Plant-Microbe Interact* 2005;18:125–33.
- [31] Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, et al. The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 2005;434:980–6.
- [32] Oh Y, Donofrio N, Pan H, Coughlan S, Brown DE, Meng S, et al. Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus *Magnaporthe oryzae*. *Genome Biol* 2008;9:R85.
- [33] Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 2006;444:97–101.
- [34] Jakupović M, Heintz M, Reichmann P, Mendgen K, Hahn M. Microarray analysis of expressed sequence tags from haustoria of the rust fungus *Uromyces fabae*. *Fungal Genet Biol* 2006;43:8–19.
- [35] Schena L, Nigro F, Ippolito A, Gallitelli D. Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. *Eur J Plant Pathol* 2004;110:893–908.
- [36] Hayden K, Ivors K, Wilkinson C, Garbelotto M. TaqMan Chemistry for *Phytophthora ramorum* detection and quantification, with a comparison of diagnostic methods. *Phytopathology* 2006;96:846–54.
- [37] Lievens B, Brouwer M, Vanachter ACRC, Cammue BPA, Thomma BPHJ. Real-time PCR for detection and quantification of fungal and oomycete tomato pathogens in plant and soil samples. *Plant Sci* 2006;171:155–65.
- [38] Avrova AO, Venter E, Birch PR, Whisson SC. Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. *Fungal Genet Biol* 2003;40:4–14.
- [39] Desmond OJ, Manners JM, Schenk PM, Maclean DJ, Kazan K. Gene expression analysis of the wheat response to infection by *Fusarium pseudograminearum*. *Physiol Mol Plant Pathol* 2008;70:40–7.
- [40] Zhang Y, Qu Z, Zheng W, Liu B, Wang X, Xue X, et al. Stage-specific gene expression during urediniospore germination in *Puccinia striiformis* f. sp. *tritici*. *BMC Genomics* 2008;9:203.
- [41] Hacquard S, Delaruelle C, Legué V, Tisserant E, Kohler A, Frey P, et al. Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Mol Plant-Microbe Interact* 2010;23:1275–86.
- [42] Yan HZ, Liou RF. Selection of internal control genes for real-time quantitative RT-PCR assays in the oomycete plant pathogen *Phytophthora parasitica*. *Fungal Genet Biol* 2006;43:430–8.
- [43] Laurans F, Pilate G. Histological aspects of a hypersensitive response in Poplar to *Melampsora larici-populina*. *Phytopathology* 1999;89:233–8.
- [44] Boyle B, Levée V, Hamelin R, Nicole MC, Séguin A. Molecular and histochemical characterisation of two distinct poplar *Melampsora* leaf rust pathosystems. *Plant Biol* 2010;12:364–76.

**Chapitre III : Analyse globale du secrétome de
*Melampsora larici-populina***

Chapitre III : Analyse globale du secrétome de *Melampsora larici-populina*

Afin de manipuler les défenses de l'hôte et d'échapper à la reconnaissance de la plante, les pathogènes ont acquis la capacité de sécréter des protéines effectrices dans la cellule hôte (ETS, effector triggered susceptibility) (Jones et Dangl 2006). En délivrant ces effecteurs, le pathogène s'expose au système immunitaire de la plante puisque certains peuvent être reconnus par les produits des gènes de résistance R de l'hôte (ETI, effector triggered immunity). Cette reconnaissance (R-Avr), déclenche une réponse immédiate de la plante appelée réponse d'hypersensibilité qui consiste en une mort cellulaire localisée aux points d'infection pour stopper le développement du pathogène. Au cours de la co-évolution entre le pathogène et son hôte, le pathogène s'adapte en diversifiant ses effecteurs ou en accumulant des mutations au sein de leurs séquences afin d'échapper à la détection par l'hôte (Jiang et al. 2008). Cette pression de sélection conduit généralement à une grande diversification des effecteurs expliquant les faibles niveaux de similarités retrouvés entre leurs séquences. La comparaison des séquences protéiques des effecteurs caractérisés jusqu'à présent a permis d'identifier les caractéristiques propres à chacun ainsi que celles partagées chez certains (Stergiopoulos et de Wit 2009). Ces effecteurs codent, dans la majorité des cas de petits polypeptides (< 30kDa) et possèdent un signal de sécrétion à leur extrémité N-terminale permettant leur adressage au système endomembranaire. Dans de nombreux cas, la richesse en cystéines permet la structuration de la protéine (Rep 2005). Malgré leurs faibles similarités de séquences, certains effecteurs présentent des motifs conservés de quelques acides aminés tel que le motif RxLR identifié chez *Phytophthora infestans* (Whisson et al. 2007). Ce motif, essentiel à l'export de l'effecteur dans la cellule hôte permet ainsi l'identification de nouveaux effecteurs possédant cette séquence dégénérée (Haas et al. 2009). Chez les champignons, de tels motifs conservés entre les effecteurs n'ont pas été identifiés mais certaines études récentes ont rapporté la présence de motifs présentant des structures dégénérées proches également indispensables à l'internalisation de l'effecteur dans la cellule hôte (Rafiqi et al. 2010; Kale et al. 2010). L'accrochage de ces motifs RxLR ou RxLR-like avec les phospholipides (Pi3P et Pi4P) présents à la surface de la membrane de la cellule hôte permet l'internalisation de l'effecteur par exocytose et suggère un mode d'export commun entre les effecteurs d'oomycètes et de champignons (Kale et al. 2010). L'analyse du niveau d'expression des gènes codant pour ces effecteurs a également révélé une accumulation préférentielle de leurs transcrits durant le processus infectieux confirmant leur spécificité *in planta* (Stergiopoulos et de Wit 2009). La récente disponibilité des génomes de champignons phytopathogènes et symbiotiques a permis de dresser le catalogue de

l'ensemble des gènes codant des petites protéines secrétées (secretome) pouvant correspondre à des effecteurs potentiels (Kämper et al. 2006, Martin et al. 2008, Mosquera et al. 2009). C'est dans ce contexte que nous avons entrepris l'analyse du secretome de *M. larici-populina* avec trois objectifs principaux : l'annotation, l'expression et l'évolution des gènes codant pour des SSPs afin d'identifier des effecteurs potentiels chez la rouille du peuplier. Nous avons concentré nos efforts sur l'annotation experte des familles de gènes codant des SSPs comprenant les gènes codant des homologues d'effecteurs déjà décrits chez d'autres rouilles. Ce travail a été réalisé en complète collaboration avec David Joly (doctorant québécois, Centre de Foresterie des Laurentides; Directeur de thèse R. Hamelin) avec qui j'ai partagé l'annotation des 1184 gènes prédits pour coder des SSPs. Dans une deuxième étape, l'analyse de l'évolution des familles de petites protéines secrétées et l'identification de familles sous pression de sélection positive a été réalisée par D. Joly à Québec. J'ai réalisé dans une troisième étape l'analyse des profils d'expression des transcrits codant les SSPs. Pour cela, le séquençage d'ESTs, les puces à oligonucléotides NimbleGen et la RT-qPCR ont été mis en œuvre afin d'identifier des transcrits spécifiques à chaque étape d'infection des feuilles de peuplier par *M. larici-populina*. Enfin, j'ai réalisé l'immunolocalisation *in planta* de plusieurs effecteurs candidats durant le processus infectieux du champignon.

Le manuscrit intitulé : « **Genome-wide analysis of small secreted proteins-encoding genes in *Melampsora larici-populina*, the causal agent of poplar leaf rust** », en attente de soumission à la revue **PloS Pathogens**, est présenté dans ce chapitre de thèse et présente les méthodologies utilisées ainsi que les résultats de l'analyse globale des petites protéines sécrétées de *M. larici-populina*.

Au cours de la réalisation de ces travaux sur le secrétome de *M. larici-populina*, D. Joly a présenté nos résultats conjoints lors de la conférence internationale MPMI au Québec en 2009. A l'issue de cette présentation, nous avons rédigé un acte de congrès qui présente de manière préliminaire notre démarche et nos avancées pour l'annotation et l'analyse des SSPs. Cet acte de congrès intitulé : « **Chasing effectors in the secretome of *Melampsora larici-populina*** » est également inclus à la fin de ce chapitre.

Genome-wide analysis of small secreted protein-encoding genes in *Melampsora larici-populina*, the causal agent of poplar leaf rust

Hacquard S, Joly DL, Tisserant E, Lin YC, Feau N, Delaruelle C, Legué V, Kohler A, Tanguay P, Van de Peer Y, Rouzé P, Martin F, Hamelin RC, Duplessis S.

Publication en attente pour soumission (Journal ciblé: PLoS Pathogens)

Les données supplémentaires de cette publication sont disponibles sur le site web:
http://mycor.nancy.inra.fr/pub/SOM_these_Hacquard

Genome-wide analysis of small secreted protein-encoding genes in *Melampsora larici-populina*, the causal agent of poplar leaf rust

Stéphane Hacquard^{1#}, David L. Joly^{2#}, Émilie Tisserant¹, Yao-Cheng Lin³, Nicolas Feau⁴, Christine Delaruelle¹, Valérie Legué¹, Annegret Kohler¹, Philippe Tanguay², Yves Van de Peer³, Pierre Rouzé³, Francis Martin¹, Richard C. Hamelin^{2,5}, Sébastien Duplessis^{1*}

¹Unité Mixte de Recherche 1136, Institut National de la Recherche Agronomique-Nancy Université, Interactions Arbres/Micro-organismes (IaM), INRA Nancy, F-54280 Champenoux, France.

²Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du PEPS, P.O. Box 10380, Stn. Sainte-Foy, Québec, QC, G1V 4C7, Canada.

³Department of Plant Systems Biology, VIB, Ghent University, 9052 Ghent, Belgium.

⁴Unité Mixte de Recherche 1202, Institut National de la Recherche Agronomique-Université Bordeaux I, Biodiversité, Gènes et Communautés (BioGeCo), INRA Bordeaux-Aquitaine, 33612 Cestas Cedex, France.

⁵Department of Forest Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.

These authors contributed equally to this work.

*E-mail : duplessi@nancy.inra.fr

Abstract

The obligate biotrophic rust *Melampsora larici-populina* is the most devastating and widespread pathogen of poplars. Studies over recent years have identified various effector proteins from plant pathogens and have highlighted their role in host-pathogen interactions. In the wake of the *Populus* genome sequencing, the ~100 Mb genome of *M. larici-populina* has been sequenced. In this study, the expression and evolutionary dynamics of *M. larici-populina* small secreted proteins were investigated at the genomic scale to pinpoint the arsenal of putative effectors that could be involved in setting compatible interaction between the fungus and its host.

Following computational predictions of signal peptide, approximately two thousand secreted proteins were identified and particular attention was given to low molecular weight proteins (300 amino acids arbitrary cut-off), herein called small secreted proteins (SSPs). Similarity with effectors previously described in *Pucciniales* was uncovered, together with an unexpectedly high number of Cys-rich proteins. Interestingly, the poplar rust genome exhibited large families of SSPs, some being organized in gene clusters. Positive selection analyses conducted over clusters of paralogous genes revealed candidate effectors ranging from 58 to 194 amino acids (mean = 120.9), of which more than 85% contained more than 4 cysteine residues. Transcript profiling of *M. larici-populina* SSPs showed timely coordinated expression profiles at different stages of plant tissue infection, most being expressed after the formation of haustoria, a crucial rust infection structure that resides within host cells and where effectors have been shown to be produced. Indeed, immunolocalization validated the accumulation of some candidate effectors in rust infection structures such as haustoria and infection hyphae.

Brought together, our data helped in prioritizing candidate effector genes likely involved in coevolutionary interplay between the pathogen and the plant immune system for functional studies.

Introduction

Worldwide, *M. larici-populina* (Basidiomycota, *Pucciniales*) is the most devastating pathogen of poplars. Like other rust fungi, it has a complex life cycle which includes five different spore types and requires two phylogenetically distinct host plants, here *Populus* and *Larix*. It is also an emerging model organism in forest pathology since this is one of the first rust fungus to have its genome sequenced, and one of the rare pathosystem for which both host and pathogen genomes are available [1,2]. Moreover, *M. larici-populina* is also a close relative to *Melampsora lini*, the

causal agent of flax rust, which has been considered as the model rust species after genetic studies demonstrating that single pairs of allelic genes determine the avirulence/virulence phenotype on host lines with particular resistance genes, i.e. the “gene-for-gene” hypothesis [3]. Poplar leaf rust *M. larici-populina* causes premature defoliation and is thus responsible for growth loss and increased susceptibility to other pests and diseases which has limited the use of poplars for wood production [4]. Therefore, this disease has a strong potential impact on current and future poplar plantations used for carbon sequestration, bioremediation, and wood and biofuel production.

In order to manipulate host defenses and enable parasitic colonization, prokaryotic and eukaryotic plant pathogens have evolved highly advanced strategies to deliver suites of effector proteins directly into host cells [5]. In turn, their hosts have evolved intracellular receptors, called resistance (*R*) genes, which recognize effector molecules on a “gene-for-gene” basis and lead to effector-triggered immunity (ETI) [6]. The recognition of such effectors, dubbed avirulence (*Avr*) proteins, activates plant defense mechanisms [7]. Unlike the extensively characterized bacterial type III secretion system [8] or the RXLR effectors described in oomycetes [9], little is actually known about the cellular machineries responsible for effector translocation in filamentous fungal pathogens. Many biotrophic fungi and oomycetes are known to share a common infection process involving the formation of haustoria, which invaginate host cells and engage in intimate contact with plant membranes [7]. Haustoria have been studied for their role in nutrient acquisition and metabolism [10-12], and there is now evidence to suggest that these structures play crucial roles in the delivery of virulence effectors that alter defense responses [5,7].

The rapid increase in the number of sequenced fungal and oomycete genomes offers the opportunity to predict the whole complement of secreted proteins, i.e. secretomes. Most secreted proteins expressed specifically *in planta* are candidate effectors and there is an expanding effort to define their roles in virulence [13]. The genome sequence of the biotrophic fungus *Ustilago maydis* has revealed the presence of a large set of lineage-specific secreted effectors arranged in clusters [14]. Interestingly, most of these appeared upregulated during biotrophic development and some disruptive mutants were significantly altered in virulence [14-16]. In oomycetes, comparison of three *Phytophthora* genomes showed rapid turnover and extensive expansion of specific families of secreted effector proteins [17,18]. Many of these genes, including the host-translocated RXLR gene family, were shown to be induced during infection and encode proteins with activities predicted to alter host physiology [19]. While most plant *R* proteins and downstream signalling pathways share a conserved nature, the array of

Table 1. Conserved rust effectors in *M. larici-populina* strain 98AG31 (*Mlp*). Homologues of putative and characterized rust effectors were searched in the *Mlp* genome by tblastn using the sequences of the *Melampsora lini* Haustorially Expressed Secreted Proteins (HESPs) and avirulence (Avr) proteins and the *Uromyces fabae* RTP1 protein. Recursive tblastn searches were performed using those *Mlp* putative effector sequences. Percentages of identities and positives were assessed following blastp searches carried out against UNIPROT (E value <1e-5).

Protein ID ^a	Scaffold	Length	Putative effector hit	Species	Identities %	Positives %	Transcript evidence in <i>Mlp</i>		
							Spores	Infected leaves	Haustoria
58459	1-	175	HESP-178	<i>Melampsora lini</i>	29	55			X
33341	3-	146	HESP-178	<i>Melampsora lini</i>	89	91			
117398	37+	207	HESP-178	<i>Melampsora lini</i>	26	46	X	X	
103805	7-	411	HESP-270	<i>Melampsora lini</i>	75	85		X	X
123932	17-	405	HESP-327	<i>Melampsora lini</i>	92	96		X	
123524	23-	216	Rust transferred protein 1	<i>Uromyces fabae</i>	59	76			
123523	31-	238	Rust transferred protein 1	<i>Uromyces fabae</i>	48	63			X
124544	28+	140	HESP-376	<i>Melampsora lini</i>	50	59			
64764	28+	154	HESP-376	<i>Melampsora lini</i>	47	63			X
70587	2+	207	HESP-379	<i>Melampsora lini</i>	27	45			
93569	5-	134	HESP-417	<i>Melampsora lini</i>	86	90	X	X	
71399	11-	132	HESP-417	<i>Melampsora lini</i>	36	52	X	X	X
74496	11+	132	HESP-417	<i>Melampsora lini</i>	29	52	X	X	
71402	11+	135	HESP-417	<i>Melampsora lini</i>	30	50	X	X	
71404	11+	128	HESP-417	<i>Melampsora lini</i>	30	51	X	X	
108269	28-	136	HESP-417	<i>Melampsora lini</i>	76	81	X	X	
35858	17+	234	HESP-570	<i>Melampsora lini</i>	24	44			
53986	69-	226	HESP-570	<i>Melampsora lini</i>	55	65			
42266	6-	85	HESP-767	<i>Melampsora lini</i>	65	76			
110232	42+	142	HESP-897	<i>Melampsora lini</i>	100	100	X	X	
101424	1+	186	HESP-C49 (recursive tblastn)		87	92	X	X	
71126	7+	178	HESP-C49	<i>Melampsora lini</i>	82	86	X	X	
104797	11+	219	HESP-C49 (recursive tblastn)						
85995	16+	492	HESP-C49 (recursive tblastn)					X	
123438	16+	179	HESP-C49 (recursive tblastn)					X	
123437	16+	179	HESP-C49 (recursive tblastn)					X	X
124076	19-	121	HESP-C55	<i>Melampsora lini</i>	72	84			
124415	52+	353	HESP-C61	<i>Melampsora lini</i>	55	70			X
124414	52+	284	HESP-C61	<i>Melampsora lini</i>	35	53		X	
101608	2-	185	HESP-C63	<i>Melampsora lini</i>	27	44			
101611	2+	203	HESP-C63	<i>Melampsora lini</i>	25	46			
63837	22-	175	HESP-C63 (recursive tblastn)					X	
123443	22-	203	HESP-C63	<i>Melampsora lini</i>	24	42		X	
65539	35-	201	HESP-C63	<i>Melampsora lini</i>	51	69			
34793	10+	76	HESP-C66	<i>Melampsora lini</i>	58	69			
34673	10-	74	HESP-C66	<i>Melampsora lini</i>	56	69			
37534	30+	77	HESP-C66	<i>Melampsora lini</i>	52	70		X	
37347	29+	151	AvrL567	<i>Melampsora lini</i>	61	44			
124207	2+	545	AvrM	<i>Melampsora lini</i>	50	33			
124206	12-	568	AvrM	<i>Melampsora lini</i>	53	36			
124204	35+	461	AvrM	<i>Melampsora lini</i>	39	25			
124205	35+	448	AvrM	<i>Melampsora lini</i>	53	36			
124203	65-	518	AvrM	<i>Melampsora lini</i>	54	32			
124202	65-	413	AvrM	<i>Melampsora lini</i>	47	31			X
124208	77-	702	AvrM	<i>Melampsora lini</i>	52	36			
124209	77-	687	AvrM	<i>Melampsora lini</i>	42	28			X
124530	78-	122	AvrP123	<i>Melampsora lini</i>	40	29			
124258	3-	95	AvrP4	<i>Melampsora lini</i>	67	57			
124268	3-	90	AvrP4	<i>Melampsora lini</i>	43	30			
124272	3+	97	AvrP4	<i>Melampsora lini</i>	50	34			
124262	5+	101	AvrP4	<i>Melampsora lini</i>	62	52			
124264	10+	89	AvrP4	<i>Melampsora lini</i>	60	47		X	
124270	15+	89	AvrP4	<i>Melampsora lini</i>	50	38			
124256	18-	88	AvrP4	<i>Melampsora lini</i>	58	44			X
124266	19-	92	AvrP4	<i>Melampsora lini</i>	63	47		X	
124260	35-	92	AvrP4	<i>Melampsora lini</i>	63	47			
124254	39-	90	AvrP4	<i>Melampsora lini</i>	67	52			
124274	51+	97	AvrP4	<i>Melampsora lini</i>	54	34			
112294	64-	105	AvrP4	<i>Melampsora lini</i>	52	30			
124252	167+	96	AvrP4	<i>Melampsora lini</i>	72	51			

^aProtein ID: number of corresponding best predicted gene model in the *M. larici-populina* JGI genome sequence.

structures and functions of Avr proteins is highly diverse, preventing *ab initio* identification from sequence and expression data information alone [13]. In many cases, the only recognizable features are the presence of an N-terminal signal sequence for secretion through the endomembrane pathway, an even number of cysteine residues that may be involved in disulfide bonds, and the absence of homologues in more remotely related species [20,21]. Moreover, consistent with the model of a coevolutionary arms race between actors of the plant immune system and effectors from these pathogens, nucleotide sequences coding for many of these secreted proteins exhibit accelerated evolutionary rates [22-25].

The whole genome draft sequence of *M. larici-populina* strain 98AG31 consists of 101.1 Mb DNA encoding 14,527 predicted proteins (Duplessis et al. in preparation). In this study, small secreted proteins (SSPs, <300 amino acids) were targeted for fine annotation and genome-wide analyses. Features such as homology with known effectors, amino acid content, genomic structure and organization, and duplication into multigene families, were scrutinized. By unravelling the expression and evolutionary dynamics of SSPs, rust genes specifically expressed *in planta* and exhibiting the hallmarks of positive selection due to coevolutionary arms races with host factors were identified. Moreover, immunolocalizations of candidate effectors are presented.

Results

Starting from an initial catalog of 16,694 gene models (JGI *Melampsora* genome annotation, version 1, 2008), 1,848 putative secreted proteins were identified in *M. larici-populina* strain 98AG31 (hereafter *Mlp*), of which 1,016 corresponded to SSP (<300 amino acids). Additional EST data from isolated haustoria (Sanger) [26] and infected leaves (454) (Duplessis et al. in preparation), together with recursive tblastn searches against the *Mlp* genome, allowed the identification of unpredicted genes. Such *de novo* gene discovery approaches extended the catalog of putative SSPs to 1,274. Following expert annotation, 20 and 70 genes were removed because they overlapped transposable elements or had characterized intracellular function (e.g. protein disulfide isomerase), respectively, leaving a total of 1,184 putative SSPs. In total, 857 of the 1,184 SSP genes were supported by EST (390/1,184) and/or identified as expressed above background during plant infection using whole-genome oligoarrays (785/1,184). Expert annotation (i.e. validation of intron/exon boundaries and gene structure using EST support and/or alignment within gene families) allowed the curation of 978 SSP-encoding genes. Among curated genes, 443 were promoted as initially predicted on the JGI website, 365 were corrected, and 170 had no prior prediction and were created *de novo*, as described above. According to

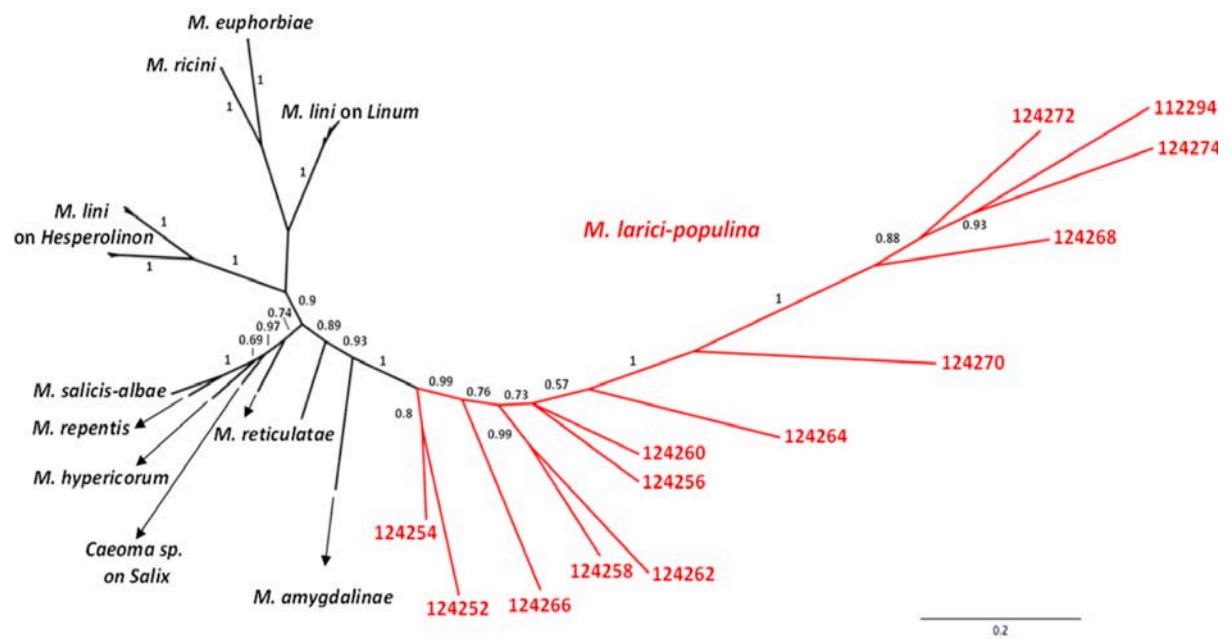


Figure 1. Unrooted bayesian consensus tree of *AvrP4* homologues. *AvrP4* homologues were identified in the genome sequence of *M. larici-populina* strain 98AG31 (*Mlp*, in red) and obtained from diverse *Melampsora* species [30] (in black). The values above branches correspond to posterior probabilities associated with the branch supporting a particular clade. *Mlp* genes encoding Protein ID nos. 124268 and 124272 are located in tandem. The other *AvrP4* homologues are scattered over the *Mlp* genome.

Tribe-MCL analyses, 821/1,184 SSPs were related to at least one other SSP (2 SSPs: 78 families; 3 SSPs: 27 families; 4 SSPs: 18 families; 5 SSPs: 11 families; >5 SSPs: 35 families). Finally, while only 201/1,184 had significant homology in nr, and this number only increased to 251/1,184 when compared to the whole gene set of another rust, *Puccinia graminis* f. sp. *tritici* (E value <1e-6).

***M. larici-populina* genome encodes numerous homologues of rust effectors**

To identify gene loci encoding protein products similar to known effectors, multiple tblastn searches were performed on the Mellp v1.0 genome sequence assembly (November 2007) using available *Pucciniales* putative effector proteins as queries: 17 haustorially expressed secreted proteins (HESPs) and 4 avirulence proteins (*Avr*) from flax rust (*M. lini*) [27,28], and the rust transferred protein (RTP1) from broad bean rust (*Uromyces fabae*) [29]. Gene models were refined on the basis of sequence homology and EST support, and recursive tblastn searches were performed using as queries those *Mlp* putative effector sequences. While the automated gene-prediction algorithms had missed twelve *AvrP4* homologues and the *AvrP123* homologue, the methodology used in this study allowed the identification of 60 loci encoding proteins with similarity to known *Pucciniales* effectors (Table 1). Nine of them were supported by EST data implemented in the JGI genome portal (Sanger sequencing from resting and germinating urediniospores) and 27 had transcript evidence in planta, according to EST databases obtained from infected leaves at 4 and 7 dpi and a partial haustorium-specific library (Table 1). One homologue showing strong structural similarity with *AvrL567* (Figure S1) and one homologue of *AvrP123* were identified in the genome of *Mlp*. Eight highly divergent *AvrM* sequence variants were detected, together with thirteen *AvrP4* homologues. Six of the *AvrM* paralogues were found to reside in three tandem pairs, with intergenic regions showing PFAM homologies associated with transposable elements (Figure S2). A phylogenetic tree incorporating data from Van der Merwe and colleagues [30] demonstrated the high sequence divergence between *Mlp AvrP4* homologues, which lied together in a supported clade (Figure 1). As described in [30], amino acids at the C-terminal end of the alignment were highly variable apart from the six Cys residues. Three different homologues of *RTP1* were identified in the whole genome sequence of *Mlp*. As previously reported by [29] and [31], the C-terminal region of *RTP1* was well conserved, while the N-terminal region was more diverse. Indeed, the N-terminal region of one *RTP1* homologue was identical to HESP-327 from *M. lini*. Among *Melampsora* *RTP1* homologues, Nucleo, NucPred and WoLF PSORT analyses suggested that the mature HESP-327/*RTP1* fusion protein is the only protein with putative nuclear localization.

Table 2. *M. larici-populina* strain 98AG31 small secreted proteins (SSPs) classes with characteristic Cys patterns. Homology searches were carried out against UNIPROT and *Puccinia* Group Database using blastp (*E* value <1e-6).

No. of genes	Length (amino acids)	Cys pattern	Top blast hit
111	107-155	CX ₉₋₁₉ CX ₈₋₁₉ CX ₄ CX ₁₂₋₁₆ CX ₉₋₁₇ CX ₁₁₋₁₈ CX ₈₋₁₄ CX ₈₋₁₃ CX ₁₋₅ C	
39	133-157	CX ₃₋₂₁ CX ₈₋₂₂ CX ₂ CX ₁₂ CX ₁₃₋₁₈ CX ₁₅₋₂₁ CX ₁₀₋₁₄ CX ₈ CX ₂₋₄ C	
8	134-367	CX ₁₂ CX ₅₋₇ CX ₃ CX ₈₋₉ CX ₁₂₋₂₈ CX ₁₃ CX ₈ CX ₉₋₁₁ CX ₂ C	PGTG_20247.2
5	154-159	CX ₁₄₋₁₇ CX ₈₋₁₀ CX ₄ CX ₂₀₋₂₃ CX ₅₁₋₅₄ CX ₉ CX ₄ C	
6	128-149	CX ₁₁₋₁₄ CX ₉₋₁₃ CX ₁₃ CX ₁₁₋₁₄ CX ₁₁₋₁₈ CX ₁₃₋₁₆ CX ₈ C	
6	178-492	CX ₂₁₋₂₂₂ CX ₁₀ CX ₄ CX ₅ CX ₂ CX ₁₃₋₂₃ CX ₉ CX ₃ CX ₅ CX ₂	HESP-C49 (<i>M. lini</i>); PGTG_08705.2
5	142-149	CX ₁₅₋₁₇ CX ₁₂₋₁₃ CX ₃ CX ₇ CX ₁₄₋₁₅ CX ₁₃₋₁₅ CX ₉ CX ₅₋₇ CX ₁₀₋₁₃ CX ₈ CX ₂ C	
4	141-154	CX ₁₃₋₁₄ CX ₂₆₋₂₈ CX ₁₃₋₁₄ CX ₂ CX ₈ CX ₁₅ CX ₁₈₋₁₉ CX ₇₋₈ CX ₈ CX ₂ C	
3	186-507	CX ₂₃ CX ₉ CX ₂ CX ₂₂₋₃₀ CX ₁₆₋₁₇ CX ₁₃₋₁₄ CX ₉ CX ₁₂₋₁₃ CX ₂ C	
3	167-180	CX ₁₆ CX ₉ CX ₄ CX ₂₀₋₂₁ CX ₅₄₋₅₅ CX ₁₂ CX ₂ C	
4	144-153	CX ₁₂₋₁₆ CX ₈₋₁₀ CX ₂ CX ₉₋₁₀ CX ₂₃₋₂₈ CX ₁₀₋₁₁ CX ₁₆ CX ₁₀₋₁₁ CX ₁₀₋₁₁ CX ₂ C	
3	133-139	CX ₆₋₇ CX ₁₃₋₁₅ CX ₁₅₋₂₀ CX ₁₂ CX ₁₇₋₁₈ CX ₁₇₋₁₈ CX ₉ CX ₁₀₋₁₁ C	
2	120-127	CX ₂₁ CX ₄ CX ₂ CX ₁₁₋₁₈ CX ₁₁ CX ₉ CX ₉ CX ₈ CX ₄ C	
1	142	CX ₁₆ CX ₈ CX ₂ CX ₈ CX ₁₈ CX ₁₇ CX ₉ CX ₈ CX ₂ C	HESP-897 (<i>M. lini</i>); PGTG_06517.2
2	203-221	CX ₂₅ CX ₈ CX ₄ CX ₁₃ CX ₈ CX ₅₀₋₅₆ CX ₁₁ CX ₈ CX ₂ C	
2	114-116	CX ₁₈₋₂₁ CX ₁₃₋₁₄ CX ₉₋₁₀ CX ₄ CX ₈ CX ₇ CX ₁₅ CX ₁₀ CX ₇ CX ₄ C	
2	124-130	CX ₅₋₆ CX ₁₃ CX ₉₋₁₄ CX ₄ CX ₈₋₁₀ CX ₁₈ CX ₁₃ CX ₁₁ CX ₇ CX ₂ C	
1	199	CX ₂₂ CX ₃₈ CX ₉ CX ₂ CX ₁₆ CX ₃₁ CX ₂₀ CX ₁₀ CX ₈ CX ₂ C	
2	130-133	CX ₃₋₅ CX ₁₂ CX ₁₀ CX ₄ CX ₉ CX ₁₁₋₁₂ CX ₂₂ CX ₁₀₋₁₂ CX ₁₁ CX ₂ C	
1	157	CX ₅ CX ₁₀ CX ₁₇ CX ₁₀ CX ₂ CX ₉ CX ₂₀ CX ₁₅ CX ₁₀ CX ₂₄ CX ₄ C	
1	189	CX ₁₄ CX ₁₅ CX ₄ CX ₂₂ CX ₆₆ CX ₁₂ CX ₂ C	
1	145	CX ₁₉ CX ₁₃ CX ₁₂ CX ₈ CX ₁₅ CX ₁₆ CX ₇ CX ₇ CX ₃ C	
2	125-136	CX ₁₄ CX ₁₂ CX ₁₃₋₁₄ CX ₁₀₋₁₁ CX ₁₀₋₁₆ CX ₁₃ CX ₈ C	
8	154-164	CX ₁₃₋₁₅ CX ₉ CX ₁₂₋₁₄ CX ₂₁₋₂₃ CX ₁₃₋₁₈ CX ₁₂₋₁₉ CX ₁₂₋₁₄ C	
1	131	CX ₁₀ CX ₈ CX ₁₂ CX ₁₃ CX ₉ CX ₂ CX ₅ CX ₁₀ CX ₁₁ CX ₁₄ CX ₁₁ CX ₂ C	
3	140-143	CX ₁₅ CX ₁₃ CX ₁₄ CX ₉ CX ₁₈ CX ₁₉ CX ₈ C	
1	138	CX ₉ CX ₄ CX ₃ CX ₁₇ CX ₈ CX ₄ CX ₈ CX ₁₂ CX ₉ CX ₈ CX ₂ C	
1	188	CX ₂₀ CX ₂₀ CX ₁₃ CX ₁₁ CX ₈ CX ₄₇ CX ₁₀ CX ₁₁ C	

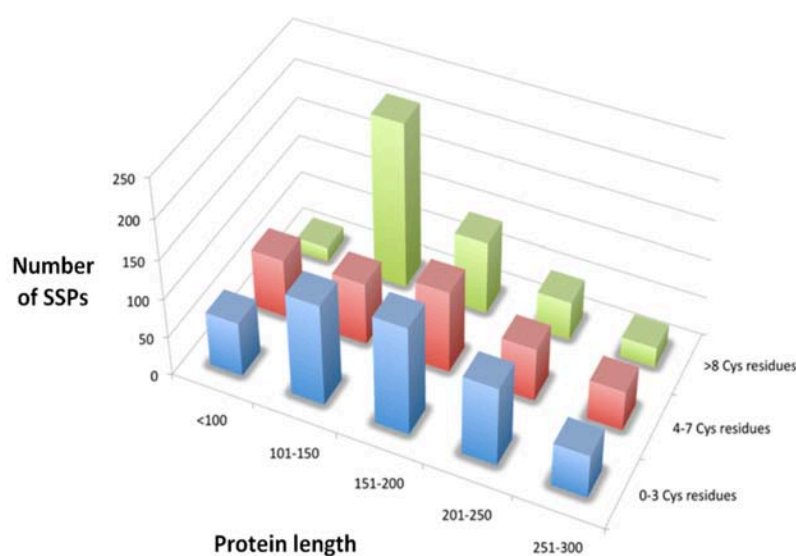


Figure 2. Distribution of *M. larici-populina* strain 98AG31 small secreted proteins (SSPs) according to protein length in amino acids and content in Cys residues

Cysteine-rich proteins dominate the rust secretome

Of the 1,184 small secreted proteins identified in *Mlp*, 746 had a number of Cys residues above 4 (Figure 2). Among those, 228 SSPs were divided in 28 classes according to homology and characteristic spacing of Cys residues (summarized in Table 2; detailed in Table S1). Interestingly, 80% of SSPs grouped in these classes represented 85% of SSPs of 101-150 amino acids with 8 Cys residues or more (Figure 2). Of these 228 Cys-rich SSPs, 99 lied in close vicinity (no more than 5 genes apart) with another member or putative pseudogene of the same class, indicating that they might have arisen from tandem duplication. As already mentioned for the complete set of SSPs, Cys-rich SSPs were under-predicted; 58/228 gene models were created following recursive tblastn against the *Mlp* genome using as queries the sequences from other classes members. Apart from classes VI and XXIV, which contained the HESP-C49 and HESP-897 homologues respectively, only one more class had homology in *Puccinia*. The 28 classes of proteins were characterized by a certain number of features. While significant variability was observed with respect to spacing and number of Cys residues, most of these genes shared a general structure of five exons with the first full codon of exons 2, 4 and 5 encoding a cysteine. Comparison of WebLogos extracted from the largest classes revealed that some protein positions were significantly enriched in particular amino acid residues (Figure 3). In many classes, Tyr/Phe residues were overrepresented at particular positions before conserved Cys #3 and #8 (Y/FXC), and Tyr residues were overrepresented after conserved Cys #6 (CX₂₋₃Y).

Positive selection in candidate effector genes

Using Tribe-MCL clustering analyses as a guide (Duplessis et al. in preparation), we used tblastn searches to identify the extent to which the SSPs have paralogues and to classify them into clusters of paralogous genes (CPGs). We identified a subset of 95 CPGs comprising 536 sequences (509 true SSPs and 27 larger SPs ranging from 301 to 942 amino acids) that contained 3 to 39 genes and could be investigated for adaptive evolution. Eighteen CPGs containing a total of 186 genes showed robust evidence of adaptive evolution at $p < 0.05$ in both M2 vs. M1 and M8 vs. M7 tests, as well as in M8 vs. M8A stringent test of selection (summarized in Table 3; detailed in Table S2). Parameter estimates for each positively selected cluster are shown in Table S3. One of these positively selected CPGs showed homology to the genome of *Puccinia*, and one included *AvrP4* homologues. In addition to the *AvrP4* homologues (CPG5464), five other CPGs consisted of very small (<100 amino acids) Cys-rich proteins (CPG510-1, CPG510-2, CPG2811, CPG4890 and CPGH1). Most genes included in these families presented a conserved LXRR motif (CPG510-1 and CPG510-2) or a region rich in positively charged residues

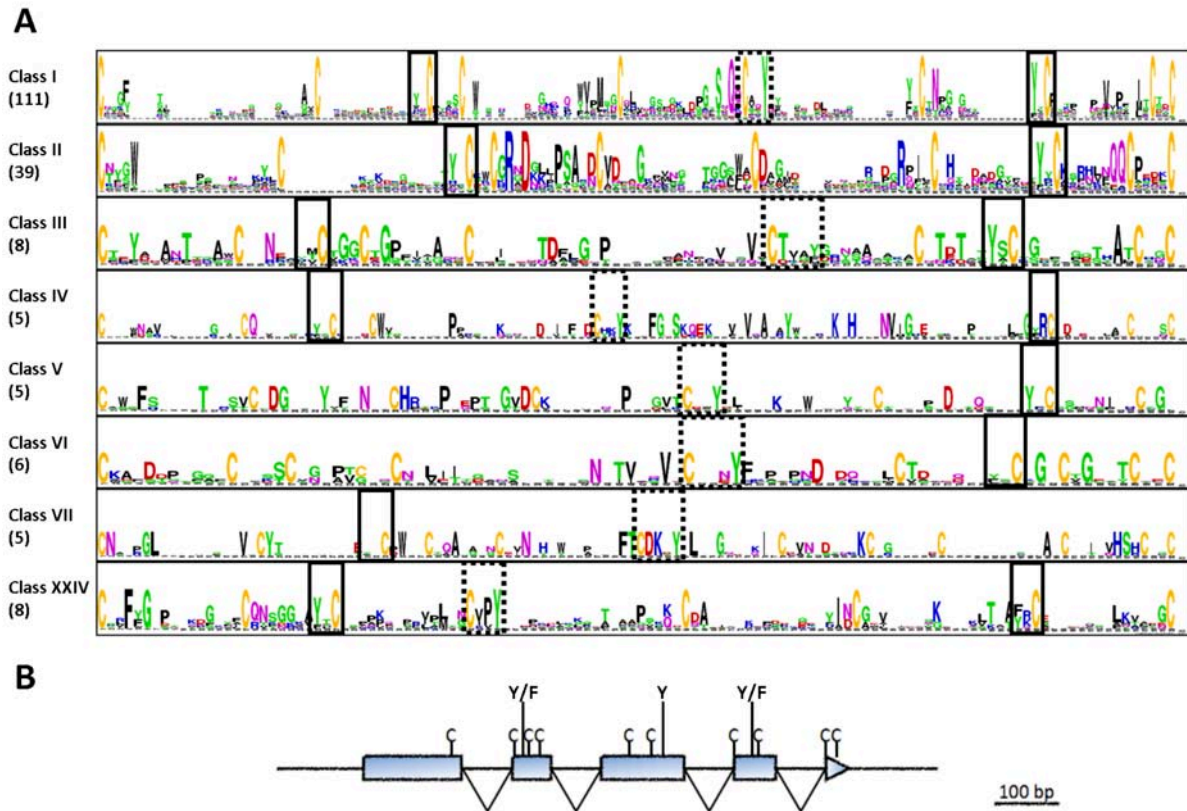


Figure 3. Conservation profiles and gene structure of the eight largest *M. larici-populina* Cys-rich classes. **(A)** Motifs shared by *M. larici-populina* Cys-rich classes. The Y/FXC motif recently identified in powdery mildew [36] (solid lines) is present in both N- and C-terminal (i.e. at the third and eighth conserved Cys, respectively) in 6 and 7 of the 8 classes, respectively. Another motif, CX₂₋₃Y (dashed lines), is present in the central part of the protein (i.e. at the sixth conserved Cys) in 7 of the 8 classes. Conservation profiles were generated with WebLogo. **(B)** Schematic representation of conserved gene structure in *M. larici-populina* Cys-rich classes. The position of conserved Cys residues is shown, together with the position of Tyr/Phe residues found in shared motifs. Introns are represented as “V”s.

CPG	No. of genes	Length (amino acids)	No. of Cys residues	Top blast hit
277	6	181-185	6-7	
335	7	180-187	7-9	
338	13	143-194	4-10	
462-1	5	171-173	10-11	PGTG_17910.2
510-1	16	64-86	4-6	
510-2	11	71-124	6-12	
517-2	3	133-136	8	
653-1	5	150-187	2-3	
1027	3	176-185	8-9	
1252	4	117-119	0	
2811	9	75-96	6	
4438	4	168-172	6-8	
4890	9	69-93	6-10	
5464	13	88-105	6-9	AvrP4 (<i>Melampsora lini</i>)
Class II	39	133-157	9-12	
Class V	5	128-149	8-9	
Class XXVII	3	140-143	8-9	
H1	30	58-84	1-6	

Table 3. *M. larici-populina* strain 98AG31 positively selected clusters of paralogous genes (CPGs) encoding small secreted proteins (SSPs). Homology searches were carried out against UNIPROT and *Puccinia* Group Database using blastp (*E* value <1e-6).

(Arg and Lys) shortly followed by a region rich in acidic residues (Asp and Glu), reminiscent of the RXLR-EER motif from oomycete effectors (Figures 4, S3). Most genes encoding very small proteins from the CPGs under positive selection were absent from the *ab initio* catalog of genes (72/88), exemplifying once again the poor resolution of the automated gene-prediction algorithms. In most of the small Cys-rich secreted proteins with evidence of positive selection, the Bayes Empirical Bayes inference of positively selected amino acid sites revealed that codons with highly significant omega values were concentrated in the C-terminal region (Figure 4).

Melampsora candidate effectors have different expression profiles

The expression level of the SSPs was monitored in spores, germlings and during time course infection of poplar leaves (24, 48, 96 and 168 hpi) using whole genome oligoarrays. Of the SSPs present on the array, 20% displayed expression peaks at the germlings stage, and 35 % reached their highest expression level *in planta* at 96 hpi (Figure S4). In order to validate and extend expression profiles of selected effectors showing homology with previously described rust effectors or presenting contrasted profiles at early and late stages of plant infection on oligoarrays (Tables S4 and S5), RT-qPCR expression profiling was carried out in resting spores, germlings and at 2, 6, 12, 24, 48, 96 and 168 hpi during time-course infection (Figure 5). Coordinated temporal patterns of SSP transcript expression were observed during poplar leaf infection (Figure 5). Comparison of transcript levels measured with distinct profiling approaches (i.e. NimbleGen oligoarrays, EST sequencing and RT-qPCR) revealed a very strong consistency of expression patterns (Figure S5). SSP transcripts encoding homologues of *U. fabae* RTP1 (Protein ID nos. 123932 and 123523) and *M. lini* HESP-417 (Protein ID no. 71404) showed an expression profile related to the first steps of fungal proliferation before the mesophyll invasion with a peak of expression in germlings (Figure 5B). Four of the 9 SSP genes showing transcripts accumulation between 2 and 6 hpi, encodes CFEM domain proteins (Protein ID nos. 72734, 114961, 73660 and 58459), including an homologue of *M. lini* HESP-178, suggestive of a possible role at early stage of rust development such as germ-tube development or appressorium formation. Additionally, two genes identified under positive selection, (Protein ID nos. 123873 and 67604) belonging to the CPG class II and CPG1252, were both preferentially induced at fungal penetration stage. Between 12 and 24 hpi (Figure 5A), fungal development is marked by a shift between compatible and incompatible interactions while formation of the first haustoria occurs [32]. From this pivotal point, several SSP transcripts showed a strong and specific expression during biotrophic growth (Figure 5C). The *Mlp* SSP showing the best homology to *Uf-RTP1* (Protein ID no. 123524) showed a distinct expression profile compared to its two other relatives in the *Mlp-RTP1* gene family (Protein ID nos. 123932 and 123523),

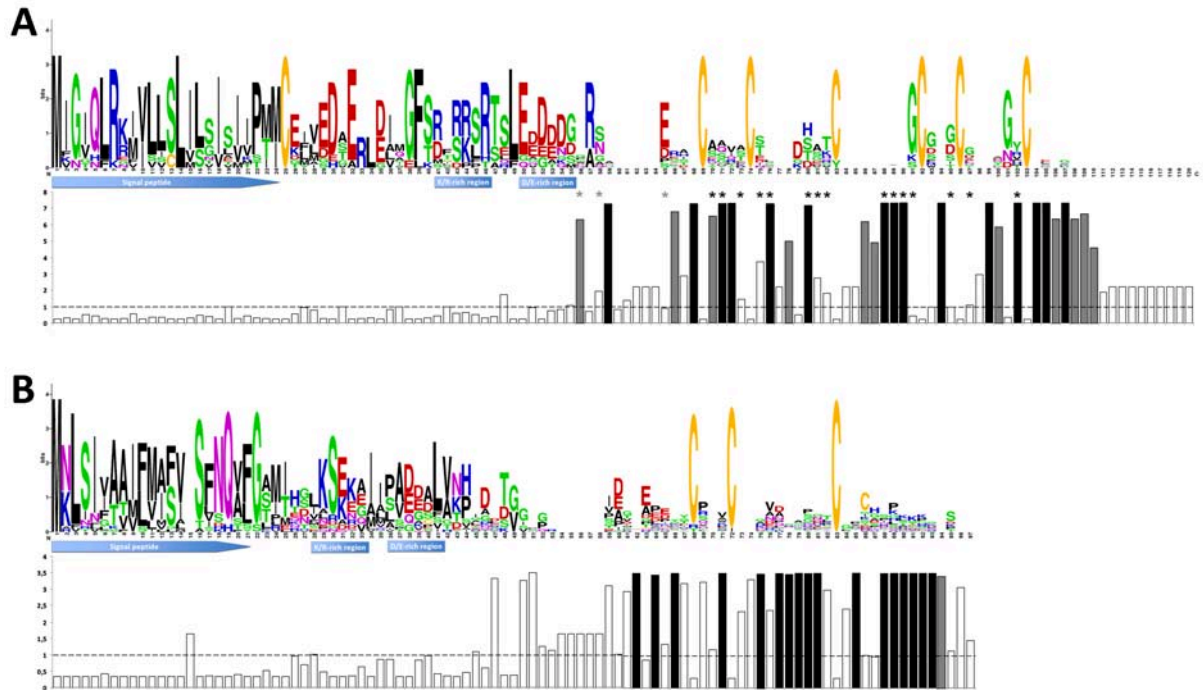


Figure 4. Examples of *M. larici-populina* strain 98AG31 positively selected clusters of paralogous genes (CPGs) encoding small secreted proteins (SSPs). Conservation profiles of **(A)** CPG5464 (*AvrP4* homologues) and **(B)** CPGH1; and associated bar diagram indicating the value of ω calculated with CODEML for each codon under site model M2. The predicted signal peptides, as well as stretches of positively charged (K/R) and acidic (D/E) amino acid residues reminiscent of the RXLR-EER translocation signal, are indicated. The dashed line across the bar diagram indicates an ω value of 1. ω values in black indicate the codons that were found to be under significant positive selection; those in grey indicate codons found to be under selection but without significant ω values. Similarly, asterisks in **(A)** indicate codons found to be under selection in [30]. Conservation profiles were generated with WebLogo.

marked by transcript accumulation concomitant with early haustoria formation. An expression peak was detected for *Mlp-SSP15* (Protein ID no. 123227), an *Mlp*-specific SSP of 15 kDa, at both 12 and 168 hpi corresponding to haustoria and spores formation. After the development of the first haustoria, a consequent colonization of the plant mesophyll is achieved by the rust fungus in the compatible interaction, with a drastic increase of fungal biomass (70-fold) between 48 and 96 hpi (Figure 5A). A large proportion of selected SSP candidates were highly expressed at 48 hpi (Protein ID nos. 37347 and 107091) or 96 hpi (Protein ID nos. 123908, 123532, 112715, 108601, 69798, 123904, 63797, 123530 and 106643) whereas transcripts were not or barely detected at the uredinial stage at 168 hpi which is suggestive of a transcriptional switch between biotrophic and sporulation phase. Strikingly, several of the selected SSP specifically expressed *in planta* appears to be lineage-specific. Of the selected candidates, two (Protein ID nos. 87952 and 85787) were strongly expressed at the uredinial stage, the later stage of rust infection with sporulation occurring (Figure 5D).

Melampsora candidate effectors localize to infection structures

Visualisation of *in planta* rust structures was achieved using PI/Uvitex labelling on infected poplar leaves (Figure 6). Many infection hyphae and haustoria were observed inside both spongy and palisade mesophyll (Figure 6A) during the biotrophic growth phase. The fungal terminal cells differentiated to form haustorial mother cells, and developed haustorial necks inside host cells, where functional haustorial body extended from the neck and invaginated the plant cell plasmalemma. At later stage of rust infection (i.e. 168 hpi), changes in fungal development were observed with aggregates of sporogenous hyphae to produce and release dikaryotic urediniospores (Figure 6B). A late time-frame (120-168 hpi) was chosen for immunological detection to ensure the presence of important fungal biomass and a diversity of fungal tissues (i.e. infection hyphae, haustorial mother cell, haustoria, sporogenous hyphae and newly formed urediniospores). *Mlp-AvrL567*, *Mlp-RTP1sc31* (located on scaffold 31), *Mlp-HESP-327/RTP1* and *Mlp-SSP15* (Protein IDs no. 37347, 123523, 123932 and 123227, respectively) were assayed for localization in rust infected poplar leaves by immunofluorescence microscopy and revealed distinct labelling patterns (Figure 6). Strong labelling of *Mlp-RTP1sc31* was observed in the cytoplasm of infection hyphae (Figure 6C and 6D) whereas no or weak labelling was observed in haustoria, other fungal structures or in the pre-immune serum (Figure S6). In contrast, the HESP-327/RTP1 fusion protein was immunolocalized around haustoria (Figure 6E) and at the periphery of intercellular hyphae (Figure 6F). Specific labelling of *Mlp-AvrL567* was detected at the periphery of fully developed haustoria (Figure 6G and 6H). This SSP is likely present in the extrahaustorial matrix, as well as around the haustorial neck (Figure 6H) whereas

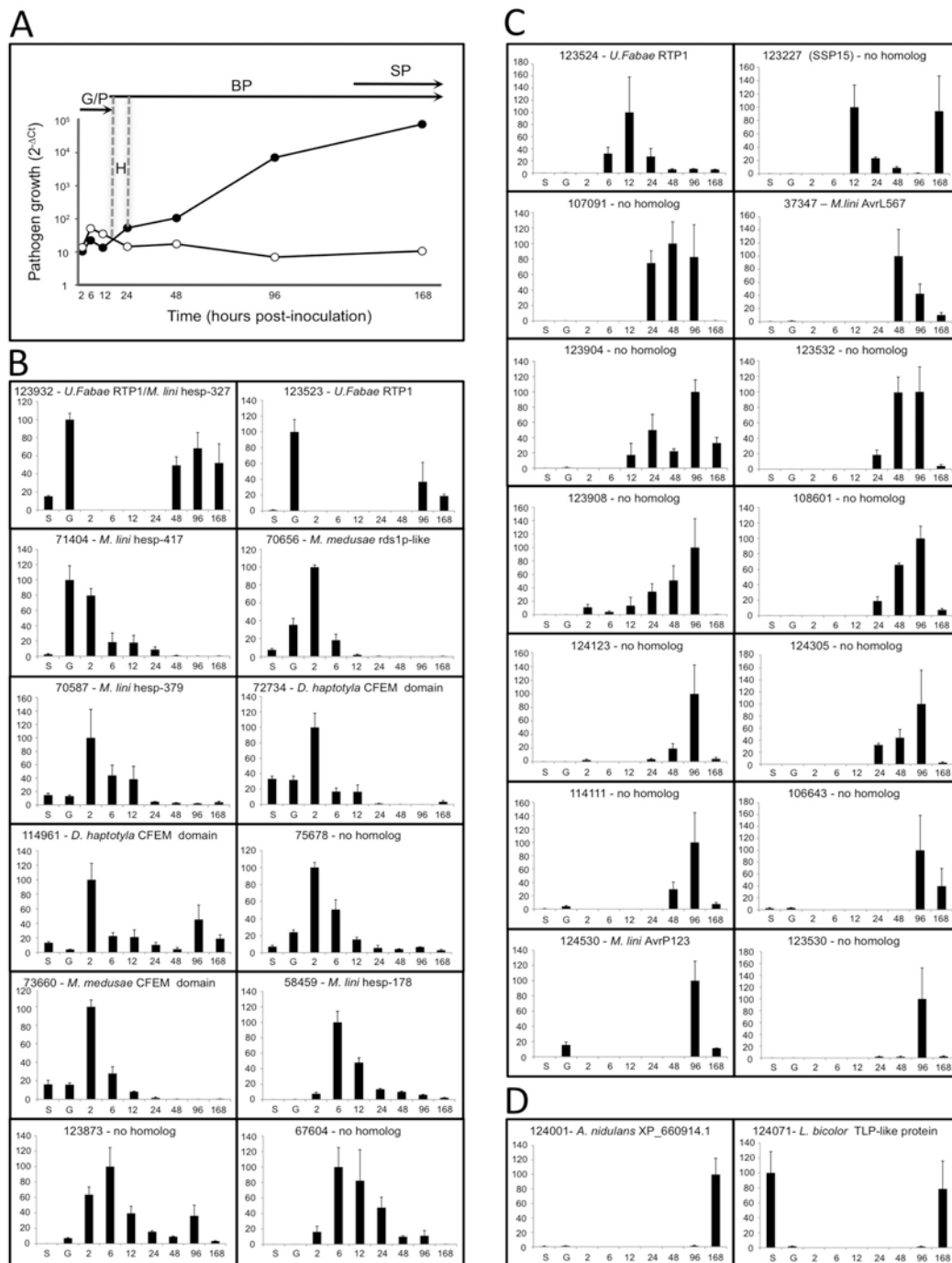


Figure 5. Fungal development of *M. larici-populina* and expression profiles of selected SSP transcripts during poplar infection **(A)** Time-course infection of *M. larici-populina* strain 98AG31 during compatible (black circle) and incompatible (white circle) interactions using respectively the *P. trichocarpa* × *P. deltoides* ‘Beaupré’ and ‘Hoogvorst’ cultivars. Development of the rust strain was monitored by qPCR amplification of the rDNA intergenic *ITS* region from inoculated leaf tissues. Pathogen growth curves correspond to fungal *ITS* amplicons compared to poplar *ITS* amplicons ($2^{-\Delta Ct}$). Note the log-scale. Arrows indicate phases of fungal development. G/P, germination and penetration phases; BP, biotrophic phase, SP: sporulation phase, H: formation of the first haustoria. **(B)** to **(D)**, Expression profiles of selected SSP transcripts in spores (S), germlings (G), and during time course infection of poplar leaves (2-168 hpi, compatible interaction) using RT-qPCR. **(B)** SSP transcripts showing expression peaks during germination or penetration phases (G/A). **(C)** SSP transcripts accumulated *in planta* during biotrophic phase (BP) or **(D)** during sporulation phase (SP) or in newly formed urediniospores. *Mlp-aTUB* and *Mlp-ELF1a* transcripts were used as reference transcripts for the calculation of regulation levels ($2^{-\Delta Ct}$). For each transcript, fold-change expression levels are presented as percentages of the highest regulation level measured (100%).

fluorescence was not detected in haustorial mother cells, infection and sporogenous hyphae or in newly formed urediniospores, suggesting a specific role inside infected host cells. Moreover, labelling of *Mlp-AvrL567* was also detected in young haustoria at earlier stages of host infection (48 hpi, Figure S8), when peak of transcript accumulation was observed by RT-qPCR. *Mlp*-SSP15 showed a very strong labelling around the basal sporogenous hyphae of uredinia (Figure 6I) whereas the newly formed urediniospores remained exempt of any signal. Interestingly, a labelling was also observed around haustoria (Figure 6J) suggesting an accumulation of the protein at the interface between the fungus and the host cell. In contrast, no labelling was detected in infection hyphae. As for the HESP-327/RTP1 fusion protein, localization at multiple sites indicates that these SSPs might play distinct roles in the infection process in different plant and/or fungal compartments. For each immunological detection, controls with pre-immune serum did not show labelling of rust infected tissue (Figures S6, S7, S8, S9).

Discussion

Sequencing genomes of pathogens provides an invaluable resource for biologists, of which the successful mining is a key to new research avenues. Here we have used the data from the *Mlp* genome to scan for genes encoding small secreted proteins. In order to identify putative effectors from these secreted proteins, we designed a multifaceted approach broadly divided in three main themes: 1) annotation, which includes information regarding genomic structure and homology; 2) screening for positive selection; and 3) transcriptomics.

Following a *de novo* gene discovery approach using additional EST data and recursive tblastn searches against the *Mlp* genome, 170 unpredicted SSPs were found, corresponding to more than 14% of total SSPs. Many genes encoding SSPs escaped computational prediction, likely due to a combination of lack of EST support and a relatively small coding size [33]. Small putatively secreted Cys-rich proteins are known to be under-predicted in plants [34], and it seems to also be the case in fungi (Duplessis et al. in preparation; P.J.G.M. de Wit, personal communication). Even with RNA-Seq, transcripts used for genome annotation rarely cover the full exome [35]. In *Mlp*, no cDNA support is available yet for other spore stages, including the aecial stage occurring on *Larix* spp., where additional SSPs probably remain to be discovered.

In *M. larici-populina* strain 98AG31, homologues of known rust effectors exist in multiple copies. In *M. lini*, the *AvrL567* locus presents one to four copies at each haplotype, where intralocus recombination plays a role in generating diversity, while the *AvrM* locus contains multiple

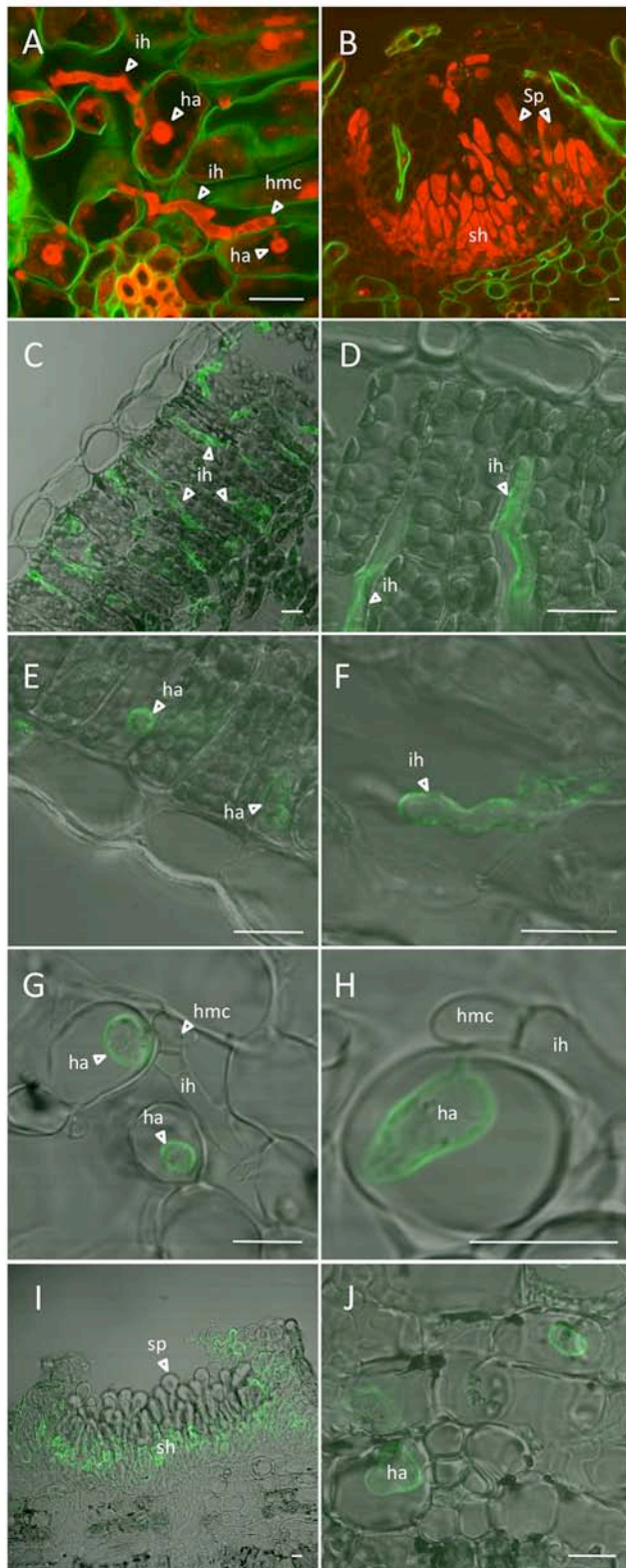


Figure 6. Localization of *M. larici-populina* candidate effector proteins during poplar infection. **(A)** and **(B)**, confocal laser scanning microscopy images of fungal structures *in planta* using Uvitex 2B (green) and propidium iodide (red) to label fungal/plant cell wall and fungal cells respectively. **(A)**, infection hyphae and haustoria colonizing spongy and palisade mesophyll at 96 hpi and **(B)**, uredinium formed after 168 hpi. Pictures were taken using 100X objective and scale bars represent 5 μ m. **(C)** to **(I)**, immunofluorescence localization of rust candidate effectors at late stage of infection. Merged-immunofluorescence and phase contrast images are presented. **(C)** and **(D)**, immunostaining of *Mlp*-RTP1sc31 in infection hyphae at 120 hpi; **(E)** and **(F)**, labelling of *Mlp*-HESP-327/RTP1 around haustoria and infection hyphae at 120 hpi; **(G)** and **(H)** specific *Mlp*-AvrL567 labelling around haustoria at 168 hpi; **(I)** and **(J)**, immunostaining of *Mlp*-SSP15 in sporogenous hyphae and around haustoria at 168 hpi. ih: infection hyphae, ha: haustorium, hmc: haustorial mother cell, sh: sporogenous hyphae, sp: spores. Pictures were taken using 100X objective and scale bars represent 5 μ m.

closely related homologues [22,27,28]. However, apart from these cases, we have no indication that previously characterized rust effectors exist in multigene families in the organisms in which they were first described. Nonetheless, it appears to be the case for most of them in *Mlp*. The occurrence of thirteen *AvrP4* homologues for example, stands in sharp contrast with the rare intraspecific variants reported by [30]. Three RTP1 homologues were found in *Mlp*, while [31] reported 1 or 2 homologues of RTP1 in three *Uromyces* species and found four homologues in *Puccinia graminis* f. sp. *tritici*. Interestingly, *Mlp* and *P. graminis* f. sp. *tritici* are heteroecious, producing spermogonia and aecia on alternate hosts, while *M. lini* and *Uromyces* species are autoecious macrocyclic rusts that complete their life cycles on a single host species. The broad arsenal of effectors in *Mlp* could thus reflect historical associations with phylogenetically diverse host plants and an ability to infect new hosts.

A huge number of Cys-rich proteins were unravelled in the *Mlp* genome, of which 228 shared a common gene structure as well as Y/FXC and CX₂₋₃Y conserved motifs. Even though blastp and/or tblastn analyses did not reveal sequence similarities between these classes, the conservation of a particular gene structure including given amino acid residues at particular positions suggests that these classes arose from a single ancestor. Some Cys-rich classes harbouring Y/FXC motifs were found to be expressed in haustoria and two classes were homologous to haustorially expressed secreted proteins (HESPs) from the flax rust fungus, *M. lini*. Recently, [36] identified a new class of candidate effectors with a similar motif (Y/F/WXC) in the barley powdery mildew, *Blumeria graminis* f. sp. *hordei*, as well as in two cereal rusts, *P. graminis* f. sp. *tritici* and *P. triticina*. Consistent with our observation on *Mlp*, overall sequence identity was low and these proteins had a highly similar exon-intron structure [36]. In powdery mildew, these putative effectors were also shown to be highly expressed in haustoria, which suggests that these proteins could serve effector functions essential for haustoria-forming fungi, perhaps in manipulating host defense components [36].

Positive selection was identified in numerous SSP-encoding gene families using computational approaches. In many plant pathogens, secreted proteins show elevated evolutionary rates, consistent with the model of an “arms race” between the plant immune system and the secreted effectors of these pathogens [6]. In *Mlp*, most of positively selected SSP-encoding genes were lineage-specific, Cys-rich, and the majority were expressed *in planta*, suggesting possible interactions with host factors. Cys-poor Avr proteins are suggested to be common in biotrophic species for which the host plant-pathogen interface primarily consists of the closely associated haustorial and plant cell membranes [37,38]. Uncovering such a large content of Cys-rich proteins in the genome of a rust fungus was thus surprising. This feature is more characteristic

of fungal pathogens that colonize the intercellular space of the plant tissue, in which disulfide bridges could enhance stability in the plant apoplast [20,39]. Rather than providing resistance to apoplastic proteases, the presence of disulfide bonds in rust secreted proteins could be important for their structure and function [40]. In several *Mlp* SSP families under positive selection, Cys patterns were conserved despite low sequence identity. While Cys patterns probably form the network of disulfide bridges necessary for the maintenance of the protein structure, variable regions could provide the sequences required for the specific function of the proteins. In such cases, significant changes in amino acid sequences (except at the Cys residues) could occur without altering the overall fold topology, making Cys-rich SSPs ideal for performing roles in recognition and specificity [41,42]. A high content in Cys residues could thus play a disproportionately important role in the evolution of virulence effectors, leading to a very rapid diversification and contributing to the emergence of new virulences. In this study, 18/95 CPGs showed robust evidence of adaptive evolution supporting the assumption that paralogous gene expansions often form the substrate for adaptive change. Selection events often result in an increased gene number in the pathogen to promote virulence while, on the other side, the need to evade host recognition by resistance genes continues to expand [43]. In such cases, rapid sequence divergence in newly formed effector paralogues allows the number of effector genes in the pathogen to increase while minimizing the likelihood of host recognition [43]. In most of the small Cys-rich secreted proteins with evidence of positive selection, codons with highly significant omega values were concentrated in the C-terminal region (Figure 4), suggesting that this region interacts with the host defense system, and that this interaction drives the diversification of the gene family. In rust and oomycete pathogens, diversifying selection was demonstrated as being confined to the C-terminal region in many effector gene families [24,30,44-46]. These findings are consistent with the view that effectors are modular proteins, with the N terminus involved in secretion and host translocation and the C-terminal domain dedicated to modulating or interacting with host defenses inside plant cells [24,47].

Positively selected CPGs encoding different families of small Cys-rich SSPs were shown to possess various N-terminal motifs reminiscent of the RXLR-EER translocation signal. Archetypal oomycete effectors harbour the bipartite RXLR-EER amino acid motif that is dispensable for exocytosis but required for plant access [9,48]. The RXLR domain appears to mediate entry of effectors into host cells independently of any pathogen-derived machinery [48]. Some fungal effectors can also enter plant cells in the absence of the pathogen, and while they lack an RXLR domain and do not share obvious conserved peptide motifs, sequences loosely related to the RXLR motif are often found [28,49,50]. The RXLR motif is overrepresented and positionally constrained in the secretome of *Phytophthora* and *Hyaloperonospora* relative to other

eukaryotes, which seems to be the result of the RXLR effector reservoir being dominated by a single highly successful and rapidly evolving superfamily [24,43]. A systematic search to uncover such overrepresented motifs in other eukaryotes has revealed the overrepresentation of LXRR in *Puccinia* (E. Kemen and J.D.G. Jones, unpublished), and our data suggest that this motif is abundantly represented in *Melampsora*.

Transcript profiling of *Mlp* during time-course infection of poplar leaves revealed temporal and tissue-specific expression of fungal SSPs. In several plant-microbe interactions such as *Magnaporthe oryzae*/rice [51], *U. maydis*/maize [14] and *Laccaria bicolor*/Populus [52], secretome prediction along with transcript profiling led to the identification of candidate fungal effectors specifically expressed during fungal colonization of plant tissues. Using various expression profiling methods we showed that more than 70% of SSP genes are expressed in the biological conditions tested. Time course infection of poplar leaves by *Mlp* was marked by highly coordinated temporal pattern of SSP expression (Figure 5 and Figure S4). This result highlights the intricate processes of cellular control developed by plant pathogens during host colonization through delivery of signal molecules as previously reported for *Phytophthora infestans* RXLR effectors [9,18,47]. Most SSP transcripts showed specific accumulation during or after haustoria formation and many encode homologues of rust effectors such as *Uf*-RTP1 [29] and *M. lini* avirulence factors [13,27]. Interestingly, several SSP transcripts induced at 96hpi *in planta* (Figure 5 and Figure S4) showed lower expression levels at 168 hpi. These results are consistent with the SSP transcriptional switch observed at the uredinial stage of *Mlp* between biotrophy and sporulation related-areas isolated using laser capture microdissection [53].

In *Mlp*, several putative rust effectors immunolocalized around haustorial structures during poplar leaves colonization (Figure 6). In order to manipulate host cell structure and function, effector proteins must be delivered inside host cells during plant tissues invasion [7]. While indirect translocation of fungal effectors across the host membrane into cytoplasm has been suggested due to the recognition in the plant cytoplasm between effectors with avirulence activity and their cognate resistance genes [22,54], only few examples reported direct evidence of effector translocation inside plant cell [29,50,55] or into plant apoplast [16]. In rust fungi, immunolocalization of RTP1 from *U. fabae* and AvrM from *M. lini* showed effector accumulation in the extrahaustorial matrix but also inside plant cell suggesting that haustorial structure might be involved in effector secretion [29,50]. Interestingly, accumulation of candidate effectors (*Mlp*-HESP-327/RTP1, *Mlp*-AvrL567 and *Mlp*-SSP15) was observed around haustorial structure of *Mlp* (Figure 6). The absence of labelling inside host cell could suggest an absence of protein translocation, nevertheless, we cannot exclude that low quantity of fungal effector could be

released inside host cells at undetectable levels by immunodetection [50]. The candidate effectors *Mlp*-SSP15 and *Mlp*-HESP-327/RTP1 displayed localization in different fungal tissues, around haustoria and sporogenous hyphae, and in haustoria and infection hyphae, respectively (Figure 6). Such a secretion at distinct sites was previously showed for *M. lini* AvrM, which localizes both at hyphal tips and haustoria during flax infection [50]. In spite of the cytoplasmic and nuclear localization reported for *U. fabae* RTP1 [29], the localization of *Mlp*-HESP-327/RTP1 was restricted to haustoria and infection hyphae and was not detected beyond the host plasmalemma. In contrast, *Mlp*-AvrL567 and *Mlp*-RTP1sc31 were specifically detected in a single fungal infection structure. During *M. lini*/*Linum* interaction, intracellular recognition between AvrL567 and the cognate resistance genes suggest effector translocation from haustoria toward infected host cell [22,27]. Based on the strong secondary structure similarities between *M. lini* and *Mlp* AvrL567 homologues (Figure S1) and the specific localization of *Mlp*-AvrL567 around haustoria (Figure 6), it is tempting to speculate on a conserved function between those effectors.

Our analysis describes genome-wide patterns of SSP expression and variation and lead to the identification of a large reservoir of candidate rust effectors. Hopefully, the successful stable genetic transformation system developed for *M. lini* [56] should be transferable to other rust pathosystems in order to validate the function of these candidates. A better understanding of effectors evolution will allow the rational design of experiments to assess functions of positively evolving gene families. Information regarding the function of *Mlp* effectors, combined with accurate data on the *Populus* immune system, will be vital for elucidating mechanisms underlying processes such as virulence and effector-triggered immunity and is likely to lead to new breeding strategies. More rust genomes are currently sequenced (Broad Institute and Joint Genome Institute), which will clearly enhance our capacity to discern patterns in related effector repertoires. Effector identification in rust pathogens and their functional characterization should facilitate the comprehension of obligate biotrophy.

Materials and methods

Genomic resources

Genome sequence data of the *M. larici-populina* strain 98AG31 (i.e., sequences of assembled scaffolds, sets of predicted genes and deduced protein sequences, ESTs) were downloaded from the Joint Genome Institute (<http://genome.jgi-psf.org/Mellp1/Mellp1.home.html>) and additional sequence data (i.e., *in planta* *M. larici-populina* ESTs, basidiomycete genomic data for comparative analyses) were available on the *Melampsora* Genome Consortium website

(<http://mycor.nancy.inra.fr/IMGC/MelampsoraGenome/index.php>). Comparisons to the *P. graminis* f. sp. *tritici* genome sequence and nucleic/proteic datasets were performed through the Puccinia Group database webportal at the Broad Institute (http://www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.html).

Sequence analysis

Similarity searches for full length sequences and conserved domains were performed using a combination of standard bioinformatics programs and customized Python scripts, the main search program being BLAST [57]. Multiple alignments were conducted using the programs ClustalW [58] and MUSCLE [59] and adjusted manually. Sequence alignments were submitted to the WebLogo server (<http://weblogo.berkeley.edu>) to generate graphical consensus displays. Phylogenetic analyses of the AvrP4 family and of the Cys-rich classes were performed using Bayesian inference and Markov chain Monte Carlo simulations (B/MCMC) implemented in MrBayes version 3.0 [60]. The models for nucleotide substitutions were selected prior to the MCMC using the likelihood ratio test implemented in the Modeltest 3.7 program [61]. One cold and seven incrementally heated chains were run for 1 million generations, with a random starting tree. Trees were sampled every 100 generations giving 10,000 trees, and the first 10% of the trees (Burn-in period set to 1,000) were excluded to compute the majority rule consensus tree of posterior probabilities.

Signal peptide prediction

In silico predictions of secreted proteins were carried out using combinations of SignalP 3.0, TargetP 1.1 and TMHMM 2.0 [62] as previously described [52] (Duplessis et al. in preparation). Briefly, the output files were manipulated to select signal peptides containing sequences based on different criteria for SignalP-HMM Sprob score, SignalP-NN Smax and D scores, and TargetP signal peptide prediction [63]. Deduced proteins with a single transmembrane domain overlapping the predicted signal peptide were also considered as secreted proteins. Finally, SSPs were selected based on an arbitrary cut-off of 300 amino acids [52]. In order to produce a large catalog of SSP for expert curation, different sets of predictions (lose to stringent) were produced by applying different weights to the prediction program parameters.

Positive selection analyses

M. larici-populina multigene families were generated using Tribe-MCL tools [64] using default settings. SSP families (i.e. gene families containing at least one SSP) were enlarged using tblastn searches. Using this approach, 536 *M. larici-populina* genes encoding secreted proteins were grouped into clusters of paralogous genes (CPGs) that contained a minimum of three sequences. CPGs were manually edited to remove either poorly aligned regions or sequences for which similarity was not found throughout the majority of the coding sequence. The resulting 95 CPGs were then submitted to positive selection analyses using a suite of programs regrouped in a single Python script, as previously described [26].

Plant material and fungal inoculation procedure

All experiments were performed using strain 98AG31 (pathotype 3-4-7) of *M. larici-populina* on *Populus trichocarpa* × *Populus deltoides* ‘Beaupré’ leaves (compatible interaction). The cultivar *P. trichocarpa* × *P. deltoides* ‘Hoogvorst’ was used to monitor fungal growth during the incompatible interaction. Urediniospores of *M. larici-populina* were propagated on detached leaves of susceptible *P. deltoides* × *Populus nigra* ‘Robusta’ as previously reported [65]. Germlings were obtained from 1 mg of urediniospores grown on water agar medium (2%) in Petri dishes for 3 hours at 19±1°C. Plant inoculation procedures were performed as previously described using the same inoculum dose of 100,000 urediniospores/ml and strictly identical culture conditions [65]. The samples harvested at different time points in the different treatments were immediately fixed in 4% (w/v) paraformaldehyde for microscopy analyses or snap-frozen in liquid nitrogen and kept at -80°C for further nucleic acid isolation.

DNA and RNA isolation

For fungal growth monitoring experiment, genomic DNA was isolated from 100 mg of infected leaf tissues (compatible and incompatible interactions from 2 to 168 hpi) using the DNeasy Plant Mini kit (Qiagen, Courtaboeuf, France). RNA was removed by the addition of ribonuclease A. Isolation of total RNA was performed with the RNeasy Plant Mini kit (Qiagen) from 1 mg of ungerminated and germinated spores, and from 100 mg of infected leaf tissues during compatible interaction (2-168 hpi), including a DNase I (Qiagen) treatment according to the manufacturer’s instructions to eliminate traces of genomic DNA. Electrophoretic RNA profiles were assessed with an Experion analyzer using the Experion RNA Standard-sens analysis kit (Bio-Rad, Marnes la Coquette, France).

Fungal ITS amplification

Kinetics of *M. larici-populina* strain 98AG31 growth in poplar leaves (compatible and incompatible interactions) were obtained by qPCR amplification of the fungal rDNA ITS. Amplification was performed as previously described [66]. The pathogen growth corresponded to the relative difference between fungal ITS amplicons and poplar ITS amplicons ($2^{-\Delta Ct}$) calculated in compatible and incompatible interactions. Each point of the curve represented the mean of two biological replicates and two technical qPCR replicates.

Expression profiling during time-course infection

Expression profiling of SSP genes was monitored using NimbleGen oligoarrays containing whole *M. larici-populina* gene models as previously described (Duplessis et al. in preparation) and RT-qPCR. For NimbleGen exon oligoarrays, RNA amplification, double-stranded cDNA synthesized, hybridization and data normalisation were strictly similar as previously described (Duplessis et al. submitted). For RT-qPCR, specific primers amplifying fragments ranging from 111 to 307 nucleotides were designed with the Primer 3 and Amplify 3X programs. Homology searches against the *M. larici-populina* and *P. trichocarpa* genome sequences were performed using the blastn algorithm to verify absence of cross annealing with other *M. larici-populina* transcripts as well as with host plant transcripts. First-strand cDNA synthesized and PCR amplifications were performed as previously described [53]. Expression of the *M. larici-populina* candidate genes was determined using the $2^{-\Delta Ct}$ calculation and calibrated to the highest level of expression observed [67]. The α -tubulin (*Mlp- α TUB*) and elongation factor (*Mlp-ELF1a*) genes were used as previously described to normalize the expression of fungal candidate genes [66].

Laser scanning confocal microscopy

Infection structures were visualized from infected leaves harvested at 96 and 168 hpi. Leaves were cut and fixed for 3 hours at 4°C in 4% (w/v) paraformaldehyde prepared in PBS, pH 7, embedded in 6% agarose (w/v), and cut into 15 and 20 μ m sections using a vibratome VT1000S (Leica, Nanterre, France). Propidium Iodide (PI) and Uvitex staining were performed on 20 μ m sections as previously described [53]. For indirect immunofluorescent localization, peptides were synthesized for AvrL567, SSP15, RTP1sc31 and HESP-327/RTP1 fusion protein and were used as antigen for the generation of antibodies in rabbits according to the manufacturer's procedure (Eurogentec, Seraing, Belgium). The anti *Mlp*-AvrL567, anti *Mlp*-SSP15, anti *Mlp*-RTP1sc31 and anti *Mlp*-HESP-327/RTP1 IgG fractions as well as pre-immune sera were purified

and desalted using MabTrap kit (GE Healthcare, Orsay, France) according to the manufacturer's recommendations. Transversal sections of 15 µm from Beaupré infected leaves were fixed and embedded in agarose as described above. Immunolocalization was performed essentially as previously described [52] with the following modifications. Sections were digested for 10 min in PBS buffer supplemented with 0,1% (w/v) cellulase, 0,01% (w/v) pectolyase, and 0,1% (w/v) BSA. After digestion, sections were washed five times for 5 min each with PBS buffer and then incubated in PBS containing 1% (w/v) BSA. The BSA was removed and the sections were incubated overnight at 4°C with purified anti-*Mlp*-AvrL567 IgG, anti-*Mlp*-SSP15 IgG, anti-*Mlp*-RTP1sc31 IgG or anti-*Mlp*-HESP-327/RTP1 IgG at a final concentration of 0.034 mg/ml, 0.023 mg/ml, 0.022 mg/ml and 0.022 mg/ml, respectively in PBS buffer with 1% BSA (w/v). Sections were incubated with IgG purified from pre-immune sera at the same concentrations for control observation. After 5 PBS washes, sections were incubated in PBS containing of the goat anti-rabbit IgG Alexa fluor 488 conjugate (1:400) used as secondary antibody (Molecular Probes, Carlsbad, CA). After five more PBS washes, sections were mounted in antifade reagent (Molecular Probes) and observed with a Radiance 2100 AGR3Q-BLD Rainbow microscope (Bio-Rad) using X60 and X100 objectives.

Acknowledgments

We thank the Joint Genome Institute for allowing early access to the *M. larici-populina* genome sequence. JGI sequencing is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Author contributions

Conceived and designed the experiments: SH DLJ FM RCH SD. Performed the experiments: SH DLJ ET YCL NF CD VL AK PT. Analyzed the data: SH DLJ SD. Contributed reagents/materials/analysis tools: YVdP PR. Wrote the paper: SH DLJ SD.

References

1. Feau N, Joly DL, Hamelin RC (2007) Poplar leaf rusts: model pathogens for a model tree. *Can J Bot* 85: 1127-1135.
2. Duplessis S, Major I, Martin F, Séguin A (2009) Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit Rev Plant Sci* 28: 309-334.

3. Flor HH (1955) Host-parasite interaction in flax rust – its genetics and other implications. *Phytopathology* 45: 680-685.
4. Steenackers J, Steenackers M, Steenackers V, Stevens M (1996) Poplar diseases, consequences on growth and wood quality. *Biomass Bioenergy* 10: 267-274.
5. Panstruga R, Dodds PN (2009) Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens. *Science* 324: 748-750.
6. Jones JDG, Dangl JL (2006) The plant immune system. *Nature* 444: 323-329.
7. Dodds PN, Rafiqi M, Gan PHP, Hardham AR, Jones DA, et al. (2009) Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytol* 183: 993-1000.
8. Zhou JM, Chai J (2008) Plant pathogenic bacterial type III effectors subdue host responses. *Curr Opin Microbiol* 11: 179-185.
9. Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, et al. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450: 115-118.
10. Hahn M, Mendgen K (2001) Signal and nutrient exchange at biotrophic plant–fungus interfaces. *Curr Opin Plant Biol* 4: 322–327.
11. Voegelé RT, Mendgen K (2003) Rust haustoria: nutrient uptake and beyond. *New Phytol* 159: 93–100.
12. Voegelé RT, Hahn M, Mendgen K (2009) The Uredinales: cytology, biochemistry, and molecular biology. In: Deising HB, editor. *The Mycota, volume 5: plant relationships*. Berlin: Springer. pp. 69-98.
13. Ellis JG, Rafiqi M, Gan P, Chakrabarti A, Dodds PN (2009) Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. *Curr Opin Plant Biol* 12: 399-405.
14. Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, et al. (2006) Insights from the genome of the biotrophical fungal plant pathogen *Ustilago maydis*. *Nature* 444: 97-101.
15. Brefort T, Doehlemann G, Mendoza-Mendoza A, Reissmann S, Djamei A, et al. (2009) *Ustilago maydis* as a Pathogen. *Annu Rev Phytopathol* 47:423-45.
16. Doehlemann G, van der Linde K, Assmann D, Schwammbach D, Hof A, et al. (2009) Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLoS Pathog* 5: e1000290.
17. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RHY, et al. (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313: 1261-1266.

18. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, et al. (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461: 393-398.
19. Morgan W, Kamoun S (2007) RXLR effectors of plant pathogenic oomycetes. *Curr Opin Microbiol* 10: 332-338.
20. Rep M (2005) Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiol Lett* 253: 19-27.
21. Göhre V, Robatzek S (2008) Breaking the barriers: microbial effector molecules subvert plant immunity. *Annu Rev Phytopathol* 46: 189-215.
22. Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang CIA, et al. (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Natl Acad Sci U S A* 103: 8888-8893.
23. Guttman DS, Gropp SJ, Morgan RL, Wang PW (2006) Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae*. *Mol Biol Evol* 23: 2342-2354.
24. Win J, Morgan W, Bos J, Krasileva KV, Cano LM, et al. (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 19: 349-2369.
25. Brunner PC, Keller N, McDonald BA (2009) Wheat domestication accelerated evolution and triggered positive selection in the β -xylosidase enzyme of *Mycosphaerella graminicola*. *PLoS ONE* 4: e7884.
26. Joly DL, Feau N, Tanguay P, Hamelin RC (2010) Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (*Melampsora* spp.). *BMC Genomics* 11: 422.
27. Dodds PN, Lawrence GJ, Catanzariti A-M, Ayliffe MA, Ellis JG (2004) The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* 16: 755-768.
28. Catanzariti A-M, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis JG (2006) Haustorially-expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18: 243-256.
29. Kemen E, Kemen AC, Rafiqi M, Hempel U, Mendgen K, et al. (2005) Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol Plant Microbe Interact* 18: 1130-1139.
30. Van der Merwe MM, Kinnear MW, Barrett LG, Dodds PN, Ericson L, et al. (2009) Positive selection in AvrP4 avirulence gene homologues across the genus *Melampsora*. *Proc R Soc B* 276: 2913-2922.

31. Puthoff DP, Neelam A, Ehrenfried ML, Scheffler BE, Ballard L, et al. (2008) Analysis of expressed sequence tags from *Uromyces appendiculatus* hyphae and haustoria and their comparison to sequences from other rust fungi. *Phytopathology* 98: 1126-1135.
32. Laurans F, Pilate G (1999) Histological aspects of a hypersensitive response in poplar to *Melampsora larici-populina*. *Phytopathology* 89: 233-238.
33. Lease K, Walker JC (2006) The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. *Plant Physiol* 142: 831-838.
34. Silverstein KA, Moskal WA, Wu HC, Underwood BA, Graham MA, et al. (2007) Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J* 51: 262-280.
35. Brent MA (2008) Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nature Rev Genet* 9: 62-73.
36. Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, et al. (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics* 11: 317.
37. Ellis J, Catanzariti A-M, Dodds P (2006) The problem of how fungal and oomycete avirulence proteins enter plant cells. *Trends Plant Sci* 11: 61-63.
38. Fudal I, Ross S, Gout L, Blaise F, Kuhn ML, et al. (2007) Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of AvrLm6. *Mol Plant Microbe Interact* 20: 459-470.
39. Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu Rev Phytopathol* 44: 41-60.
40. Catanzariti A-M, Dodds PN, Ellis JG (2007) Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiol Lett* 269: 181-188.
41. Templeton MD, Rikkerink EHA, Beever RE (1994) Small, cysteine-rich proteins and recognition in fungal-plant interactions. *Mol Plant Microbe Interact* 7: 320-325.
42. Povolotskaya IS, Kondrashov FA (2010) Sequence space and the ongoing expansion of the protein universe. *Nature* 465: 922-926.
43. Jiang RHY, Tripathy S, Govers F, Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc Natl Acad Sci U S A* 105: 4874-4879.
44. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, et al. (2004) Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science* 306: 1957-60.
45. Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, et al. (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell* 17: 1839-1850.

46. Barrett LG, Thrall PH, Dodds PN, van der Merwe M, Linde CC, et al. (2009) Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. *Mol Biol Evol* 26: 2499-2513.
47. Schornack S, Huitema E, Cano LM, Bozkurt TO, Oliva R, et al. (2009) The things to know about oomycete effectors. *Mol Plant Pathol* 10: 795-803.
48. Dou D, Kale SD, Wang X, Jiang RHY, Bruce NA, et al. (2008) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell* 20: 1930-1947.
49. Manning VA, Ciuffetti LM (2005) Localization of Ptr ToxA produced by *Pyrenophora tritici-repentis* reveals protein import into wheat mesophyll cells. *Plant Cell* 17: 3203-3212.
50. Rafiqi M, Gan PH, Ravensdale M, Lawrence GJ, Ellis JG, et al. (2010) Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell* 22: 2017-2032.
51. Mosquera G, Giraldo MC, Hyun Khang C, Coughlan S, Valent B (2009) Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy-associated secreted proteins in rice blast disease. *Plant Cell* 21: 1273-1290.
52. Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, et al. (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452: 88-92.
53. Hacquard S, Delaruelle C, Legué V, Tisserant E, Kohler A, et al. (2010) Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Mol Plant Microbe Interact*. In press.
54. Houterman PM, Ma L, van Ooijen G, de Vroomen MJ, Cornelissen BJ, et al. (2009) The effector protein Avr2 of the xylem-colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. *Plant J* 58: 970-978.
55. Khang CH, Berruyer R, Giraldo MC, Kankanala P, Park SY, et al. (2010) Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. *Plant Cell* 22: 1388-1403.
56. Lawrence GJ, Dodds PN, Ellis JG (2010) Transformation of the flax rust fungus, *Melampsora lini*: selection via silencing of an avirulence gene. *Plant J* 61: 364-369.
57. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.

58. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
59. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792-1797.
60. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
61. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
62. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953-971.
63. Klee EW, Ellis LBM (2005) Evaluating eukaryotic secreted protein prediction. *BMC Bioinformatics* 6: 256.
64. Enright AJ, Van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30: 1575-1584.
65. Rinaldi C, Kohler A, Frey P, Duchaussoy F, Ningre N, et al. (2007) Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol* 144: 347-366.
66. Hacquard S, Frey P, Martin F, Duplessis S. Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection. Submitted to *Physiol Mol Plant Pathol*.
67. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) method. *Methods* 25: 402-408.

**Chasing effectors in the secretome of
*Melampsora larici-populina***

Joly DL, Hacquard S, Feau N, Tanguay P, Martin F, Hamelin RC,
Duplessis S.

Acte de congrès publié suite à la conférence internationale MPMI à Québec (2010)

Paper 77 in: Biology of Plant-Microbe Interactions, volume 7.

Chasing effectors in the secretome of *Melampsora larici-populina*

David L. Joly¹, Stéphane Hacquard², Nicolas Feau³, Philippe Tanguay¹, Francis Martin², Richard C. Hamelin¹, and Sébastien Duplessis²

¹Laurentian Forestry Centre, Canadian Forest Service, Natural Resources Canada, 1055 du PEPS, PO Box 10380, Stn. Sainte-Foy, Québec, QC, Canada, G1V 4C7. ²INRA Nancy, UMR 1136, Interactions Arbres-Microorganismes, 54280 Champenoux, France. ³INRA Bordeaux-Aquitaine, UMR 1202, Biodiversité, Gènes et Communautés, 33883 Villenave d'Ornon Cedex, France.

rhamelin@nrcan.gc.ca / duplessi@nancy.inra.fr

The basidiomycete *Melampsora larici-populina* Kleb. (*MLP*), causal agent of poplar leaf rust, is an obligate macrocyclic biotroph which has its sexual stage on larch (*Larix* sp.) and its asexual stage on poplar (*Populus* species from the sections *Aigeiros* and *Tacamahaca*, and their hybrids). In the last decades, the distribution range of *M. larici-populina* has expanded significantly worldwide. Even though it rarely causes mortality, this rust is responsible for growth loss, premature defoliation, and predisposition to other pests and diseases. Although breeders developed several cultivars with qualitative resistance to *M. larici-populina*, new strains of the pathogen were able to overcome host resistances and new virulent genotypes produced by recombination were amplified by a strong selection pressure (Frey et al. 2005). Deployment of sustainable resistance requires a better understanding of the molecular mechanisms underlying *Populus-Melampsora* interactions (Feau et al. 2007).

Like most rust fungi, *M. larici-populina* urediniospore germination results in a germ tube that differentiates into a series of infection structures in the intercellular space (Rinaldi et al. 2007). The haustorium invaginates host cells and makes near-direct contact with the host plasma membrane where it plays a crucial role in nutrient acquisition (Voegelé et al. 2009). This structure is also the core of cellular communication involved in the redirection of host metabolism and the suppression of host defenses (Voegelé et al. 2009), which

allows the pathogen to deliver a suite of effector proteins to modulate plant innate immunity and enable parasitic colonization (Dodds et al. 2009).

With the advent of genomics and bioinformatics, the identification of putative effectors was facilitated by tracking secreted proteins exported through the eukaryotic general secretory pathway via N-terminal amino acid sequences known as signal peptides (Torto et al. 2003). While signal peptides cannot be directly identified using molecular methods, computational tools can assign signal peptide prediction scores with a high degree of accuracy (Nielsen et al. 1997). Such bioinformatic approaches have been applied to rust fungi cDNA libraries, resulting in the identification of a number of effectors with avirulence functions (Catanzariti et al. 2006). Moreover, the availability of filamentous pathogen genome sequences enabled genome-wide survey of secreted proteins, allowing the identification of putative effectors (Soanes et al. 2007).

In the wake of the *Populus* genome sequencing, the ~100 Megabases genome of *M. larici-populina* has been sequenced by the Joint Genome Institute (<http://genome.jgi-psf.org/>), generating a great opportunity to identify and characterize *Melampsora* effectors likely involved in molecular interactions with its host plant (Duplessis et al. 2009). The analysis of this genomic sequence by an international consortium (<http://mycor.nancy.inra.fr/>) has revealed that this fungus is equipped with an arsenal of approximately two thousand secreted proteins, with so far 1,058 Small Secreted Proteins (<300 amino acids), herein termed SSPs. In this report, we briefly outline our current approach to identify poplar leaf rust candidate effectors in the secretome of *M. larici-populina*.

An integrated approach for the identification of novel rust effectors

Since effectors act during infection and outside the fungal cell, their corresponding genes can be steadily identified on the basis of *in planta* expression and by the presence of a signal peptide for secretion. Some of these proteins face the daunting challenge of carrying out virulence functions, while also potentially exposing the pathogen to recognition by resistance proteins, leading to an activation of host defense systems (Dodds et al. 2009). Strong selective pressures are imposed by these competing roles, and the subsequent genetic changes leave their footprints in the extant genomic variation (Ma and Guttman 2008). Such information are thus of critical importance for a comprehensive identification of candidate effectors. In order to identify effector genes in the *M. larici-populina* genome sequence, we

developed a multifaceted approach broadly divided into three main themes: annotation of SSPs, which includes information regarding genomic organization, structure and homologies; gene expression profiling in infected plant tissues; and screening for positive selection.

Amino acid sequence comparisons frequently yield the first clues on possible functions of novel proteins. However, database searches with sequences of SSPs from fungi commonly do not yield recognizable homologues (Rep 2005). Indeed, among the 1,058 *M. larici-populina* SSP, less than 25% have homologs in public sequence databases. Surprisingly, only one third of *Melampsora* SSPs have identifiable homologs in the genome sequence of *Puccinia graminis* f. sp. *tritici* (PGT), the only other *Pucciniales* sequenced so far (<http://www.broad.mit.edu>). Apart from the presence of the signal peptide, the only recognizable feature of these SSPs is often a large content in cysteine residues. These cysteines are presumed to play an important role in the stability of secreted proteins, and are typical features of some fungal and oomycete effectors (Kamoun 2006; Stergiopoulos and de Wit 2009). For instance, one fourth of *M. larici-populina* predicted SSPs have been grouped in 30 classes according to characteristic spacings of cysteine residues, the largest family encompassing more than one hundred members. Similarities with effectors previously described in *Pucciniales* were uncovered, such as homologs of the Rust Transferred Protein (RTP1) described in *Uromyces fabae* (Kemen et al. 2005), or of the Haustorially Expressed Secreted Proteins (HESPs) from *Melampsora lini* (Catanzariti et al. 2006). Several of these HESPs are elicitors of the hypersensitive response in flax and co-segregate with known avirulence genes (Catanzariti et al. 2006). While only one homolog could be found for two of them (*AvrL567* and *AvrP123*), eight and twelve homologs of *AvrM* and *AvrP4*, respectively, were recovered in the *M. larici-populina* genome.

Expression profiling of *M. larici-populina* transcripts during time-course infection of poplar leaves using exon oligonucleotide arrays (NimbleGen Systems) allowed to derive groups of genes expressed at specific stages of the interaction with the host. The number of expressed SSP genes increased all along the infection process and the expression of 56 genes peaked at 96 hpi during the biotrophic phase (Fig. 1). Pyrosequencing of transcripts (Roche 454) isolated from infected poplar leaf tissues confirmed the accumulation of several SSP transcripts at specific stages of the infection. Monitoring of candidate effector transcripts previously described in *Melampsora* spp. (Feau et al. 2007) by RT-qPCR also confirmed their specific temporal expression during plant colonization and biotrophic development.

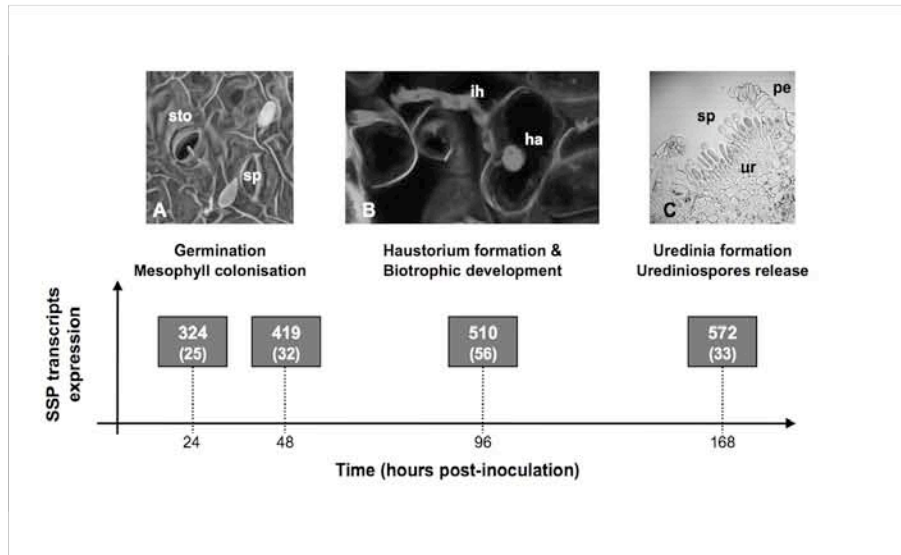


Fig. 1. Expression of *Melampsora larici-populina* Small Secreted Protein (SSP) transcripts during time-course infection of poplar leaves. Urediniospores of the virulent 98AG31 strain were inoculated on compatible poplar leaves as previously described (Rinaldi et al., 2007). Transcripts isolated from infected plant tissues were used to probe expression of the ~17,000 *M. larici-populina* predicted genes using oligonucleotide arrays. The number of genes expressed above the background level is given. Number of genes highly expressed is given in brackets. Microscopy pictures (© INRA, Hacquard & Duplessis) illustrate major steps of poplar leaf infection by *M. larici-populina*: **A**, spore (sp) germination and penetration through stoma (sto); **B**, infection hyphae (ih) and haustorium (ha) in the plant mesophyll; **C**, uredinia (ur) formation and release of newly formed urediniospores (sp) through the plant epidermis (pe).

Since about one third of the SSPs belong to gene families, intra- and inter-specific comparative genomics using closely related species with different host specificities appeared as a complementary strategy to target candidate effector genes. Clusters of paralogous or orthologous genes were generated to uncover the role of positive selection in the evolution of genes encoding secreted proteins. Maximum likelihood approaches were used to estimate nonsynonymous nucleotide substitution rate (d_N) and synonymous substitution rate (d_S) across all possible pairwise sequence comparisons within each cluster and to contrast codon substitution models (Yang et al. 2000). Our preliminary analyses suggest that most SSP-encoding genes specifically and/or highly expressed *in planta* with evidence of positive selection have an even number of cysteine residues, no homology in public databases, and a high rate of tandem duplication (Table 1). These results are

consistent with other studies conducted in filamentous pathogens. Further characterization of the rust candidate effectors reported here is currently ongoing to precise their exact roles in pathogenicity.

Table 1. Examples of SSPs expressed *in planta* with evidence of positive selection

SSP family	# of <i>MLP</i>	# in tandem	# of <i>PGT</i>	BLAST hits	# of Cys
CPG 510	4	2	0	-	4
CPG 518	3	3	0	-	8
CPG 1193+	5	2, 2	0	<i>M. lini</i> HESP-C63	8
COG 747	1	-	12	-	4

Acknowledgements

The authors would like to thank Igor Grigoriev, Susan Lucas, Harris Shapiro (JGI) and Jane Grimwood (SHGC) for their outstanding efforts in sequencing, assembling and annotating the *Melampsora* genome. Emilie Tisserant (INRA Nancy) and Yao-Cheng Lin (VIB Ghent) are acknowledged for their support in bioinformatics. Investigations carried out in R. C. Hamelin laboratory were supported by the Canadian Forest Service Genomics Research and Development Initiative. Investigations at UMR 1136 were supported by grants from INRA (AIP Bioressources to F. Martin, innovative grant to S. Duplessis) and the Région Lorraine.

Literature cited

- Catanzariti, A.-M., Dodds, P.N., Lawrence, G.J., Ayliffe, M.A., and Ellis, J.G. 2006. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18: 243-256.
- Dodds, P.N., Rafiqi, M., Gan, P.H.P., Hardham, A.H., Jones, D.A., and Ellis, J.G. 2009. Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytol.* 183: 993-1000.
- Duplessis, S., Major, I., Martin, F., Séguin, A. 2009. Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit. Rev. Plant Sci.* (in press)
- Feau, N., Joly, D.L., and Hamelin, R.C. 2007. Poplar leaf rusts: model pathogens for a model tree. *Can. J. Bot.* 85: 1127-1135.
- Frey, P., Gérard, P., Feau, N., Husson, C., and Pinon, J. 2005. Variability and population biology of *Melampsora* rusts on poplars. In: *Rust Diseases of Willow and Poplar*. Ed. Pei, M.H., and McCracken, A.R. CABI Publishing, Wallingford, UK. pp. 63-72.

- Kamoun, S. 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.* 44: 41-60.
- Kemen, E., Kemen, A.C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., and Voegelé, R.T. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol. Plant-Microbe Interact.* 18: 1130-1139.
- Ma, W., and Guttman, D.S. 2008. Evolution of prokaryotic and eukaryotic virulence effectors. *Curr. Opin. Plant Biol.* 11: 412-419.
- Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Prot. Engineer.* 10: 1-6.
- Rep, M. 2005. Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiol. Lett.* 253: 19-27.
- Rinaldi, C., Kohler, A., Frey, P., Duchaussoy, F., Ningre, N., Couloux, A., Wincker, P., Le Thiec, D., Fluch, S., Martin, F., Duplessis, S. 2007. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant. Physiol.* 144: 347-36.
- Soanes, D.M., Richards, T.A., Talbot, N.J. 2007. Insights from sequencing fungal and oomycete genomes: What can we learn about plant disease and the evolution of pathogenicity? *Plant Cell* 19: 3318-3326.
- Stergiopoulos, I., and de Wit, P.J.G.M. 2009. Fungal effector proteins. *Annu. Rev. Phytopathol.* 47: 233-263.
- Torto, T.A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N.A.R., van West, P., and Kamoun, S. 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res.* 13: 1675-1685.
- Voegelé, R.T., Hahn, M., Mendgen, K. 2009 The Uredinales: cytology, biochemistry, and molecular biology. In: *The Mycota* (Vol. V, 2nd edition). Ed. Deising, H.B. Springer-Verlag, Berlin-Heidelberg. pp 69-98.
- Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.-M.K. 2000. Codon-substitution models for heterogeneous selection pressures at amino acid sites. *Genetics* 155: 431-449.

**Chapitre IV : La microdissection à capture laser,
un outil puissant pour l'analyse du
transcriptome de tissus complexes: le cas du
stade urédien chez *Melampsora larici-populina***

Chapitre IV : La microdissection à capture laser, un outil puissant pour l'analyse du transcriptome de tissus complexes: le cas du stade urédien chez *Melampsora larici-populina*

Comme la plupart des agents de rouilles, *Melampsora larici-populina* se distingue par la complexité de son cycle biologique. Celui-ci est dit hétéroïque en référence aux deux hôtes (mélèze et peuplier) et macrocyclique en référence aux cinq types de spores produites au cours du cycle biologique. Alors que la reproduction sexuée a lieu sur le mélèze au printemps (hôte alternatif), la reproduction asexuée se déroule durant tout l'été sur les feuilles de peuplier (hôte principal). La phase importante dans la maladie est le stade urédien au cours duquel se développe une structure sporifère appelée urédie. Des quantités extrêmement importantes d'urédospores dicaryotiques de couleur orangée sont libérées à partir de cette structure et sont disséminées par les vents sur de larges distances et permettent l'infection de nouveaux hôtes peuplier compatibles et ainsi de suite durant tout l'été (Barrès et al. 2008). Le caractère cyclique de l'infection aboutit à de sévères épidémies limitant la photosynthèse de l'arbre et induisant des pertes de biomasse pouvant atteindre 60% (Miot et al. 1999). L'infection est rapide puisqu'il s'écoule une semaine entre la germination de l'urédospore à la surface de la feuille de peuplier et la formation de l'urédie (Rinaldi et al. 2007). Au cours de ce processus infectieux, le champignon développe différentes cellules dans les tissus foliaires tels que les hyphes d'infection, les haustoria, les hyphes sporifères ainsi que des spores nouvellement formées. Cette complexité tissulaire est l'une des principales limitations à l'analyse moléculaire des processus de biotrophie et de sporulation.

La microdissection à capture laser est un outil qui permet d'isoler, au sein de tissus complexes des cellules d'intérêt afin de les analyser. Cette technique a récemment été utilisée dans les interactions plantes/microorganismes afin d'étudier les niveaux de transcrits au sein de cellules ou tissus infectés (Tang et al. 2006; Chandran et al. 2010). L'application de cette technique à l'étude des structures d'infection formées *in planta* par les organismes phytopathogènes biotrophes ouvre ainsi des perspectives prometteuses dans la compréhension des interactions plantes/pathogènes (Balestrini et al. 2009).

L'acquisition récente d'un microscope à dissection laser au sein de l'UMR 1136 IAM nous a permis d'envisager l'utilisation de la microdissection laser pour séparer différents tissus

fongiques au moment de la formation de l'urédie. En effet, on note à cette étape du stade urédien la présence concomitante d'hyphes d'infection et d'haustoria plutôt en lien avec le mode de vie biotrophe du champignon et des hyphes sporifères ainsi que des spores nouvellement formées dans l'urédie plutôt en lien avec l'étape de sporulation et de reproduction asexuée. Dans un premier temps nous nous sommes attachés à décrire la formation de l'urédie et à identifier les différentes structures fongiques la composant en utilisant la microscopie confocale. Dans un second temps nous avons utilisé la microscopie à dissection laser afin d'isoler des zones spécifiquement au niveau des urédies contenant des tissus fongiques en lien soit avec la biotrophie, soit avec la sporulation dans le but d'établir par une approche transcriptomique les programmes génétiques exprimés dans ces différentes zones. Avant de pouvoir procéder à l'établissement des profils d'expression génique, il était nécessaire d'établir une procédure adaptée à l'extraction des acides ribonucléiques de qualité. Nous avons donc apporté une attention particulière à la préparation des tissus en aval de la microdissection. Nous avons alors pu analyser le niveau d'expression de l'ensemble des transcrits de *M. larici-populina* pour les zones microdisséquées en utilisant les puces à oligonucléotides.

Le manuscrit intitulé : « **Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation** », publié dans la revue **Molecular Plant-Microbe Interactions**, est présenté dans ce chapitre de thèse et présente les méthodologies utilisées ainsi que les résultats obtenus au cours de cette étude.

**Laser capture microdissection of uredinia formed by
Melampsora larici-populina revealed a transcriptional switch
between biotrophy and sporulation**

Hacquard S, Delaruelle C, Legué V, Tisserant E, Kohler A,
Frey P, Martin F, Duplessis S.

Article publié dans la revue Molecular Plant-Microbe Interactions (Octobre 2010)

Laser Capture Microdissection of Uredinia Formed by *Melampsora larici-populina* Revealed a Transcriptional Switch Between Biotrophy and Sporulation

Stéphane Hacquard, Christine Delaruelle, Valérie Legué, Emilie Tisserant, Annegret Kohler, Pascal Frey, Francis Martin, and Sébastien Duplessis

Unité Mixte de Recherche 1136 INRA/Nancy Université 'Interactions Arbres/Micro-organismes', INRA Nancy, F-54280 Champenoux, France

Submitted 17 May 2010. Accepted 8 June 2010.

The foliar rust caused by the basidiomycete *Melampsora larici-populina* is the main disease affecting poplar plantations in Europe. The biotrophic status of rust fungi is a major limitation to study gene expression of cell or tissue types during host infection. At the uredinial stage, infected poplar leaves contain distinct rust tissues such as haustoria, infection hyphae, and uredinia with sporogenous hyphae and newly formed asexual urediniospores. Laser capture microdissection (LCM) was used to isolate three areas corresponding to uredinia and subjacent zones in the host mesophyll for expression analysis with *M. larici-populina* whole-genome exon oligoarrays. Optimization of tissue preparation prior to LCM allowed isolation of RNA of good integrity for genome-wide expression profiling. Our results indicate that the poplar rust uredinial stage is marked by distinct genetic programs related to biotrophy in the host palisade mesophyll and to sporulation in the uredinium. A strong induction of transcripts encoding small secreted proteins, likely containing rust effectors, is observed in the mesophyll, suggesting a late maintenance of suppression of host defense in the tissue containing haustoria and infection hyphae. On the other hand, cell cycle and cell defense rescue transcripts are strongly accumulated in the sporulation area. This combined LCM-transcriptomic approach brings new insights on the molecular mechanisms underlying urediniospore formation in rust fungi.

The foliar rust caused by the basidiomycete fungus *Melampsora larici-populina* is the main disease affecting poplar in Europe, causing severe damages in plantations (Gérard et al. 2006). At present, almost all planted poplar cultivars are susceptible to *M. larici-populina*, and sustainability of newly selected resistances requires a better understanding of the molecular mechanisms underlying *Populus*–*Melampsora* spp. interaction. The recent sequencing of the *M. larici-populina* genome is a great opportunity to identify loci coding for the arsenal developed by the rust fungus to penetrate and exploit its host (Duplessis et al. 2009). *M. larici-populina* has a complex bio-

logical lifecycle, performing sexual reproduction on an alternate host (*Larix* spp.), whereas polycyclic production of asexual urediniospores occurs during summer on poplar trees, leading to severe epidemics (Barrès et al. 2008). The fungus penetrates through stomata on the abaxial epidermis of poplar leaves between 2 and 6 h after a spore lands on a leaf surface (Rinaldi et al. 2007). A drastic increase in fungal biomass is observed at 96 h postinfection (hpi) inside parenchyma cells after the formation of haustoria (Rinaldi et al. 2007). This infection structure, involved in the redirection of host metabolism and the suppression of host defenses, allows the establishment of a successful biotrophic relationship (Heath 1997; Voegelé et al. 2009). By 168 hpi, orange pustules corresponding to uredinia release urediniospores on the leaf surface to achieve further infection of poplar leaves.

The obligate biotroph status of rust fungi is a limitation to the study of gene expression in structures involved in biotrophic growth or sporulation stage because rust cannot be cultivated outside of their hosts. Most of the knowledge gained on rust gene expression in planta derived from sequencing of expressed sequence tags (EST) from haustorial structures or infected plant tissues (Hahn and Mendgen 1997; Thara et al. 2003; Broeker et al. 2006; Catanzariti et al. 2006; Feau et al. 2007; Puthoff et al. 2008; Yin et al. 2009). Identification of genes specifically expressed during the biotrophic phase was reported in the bean rust *Uromyces fabae* such as the thiamine biosynthesis genes *THI1* and *THI2* (Sohn et al. 2000), the amino-acid transporter *AAT2* and the sugar transporter *HXT1* (Hahn et al. 1997; Voegelé et al. 2001), or the rust-transferred protein 1 *RTP1* (Kemen et al. 2005). In the flax rust *Melampsora lini*, haustorially expressed secreted proteins (HESP) were reported as highly enriched for avirulence elicitors (Dodds et al. 2004; Catanzariti et al. 2006, 2010). In *Uromyces appendiculatus*, a large number of EST confirmed the importance of secreted proteins, primary metabolism, and nutrient uptake in haustorial structures (Puthoff et al. 2008). In contrast to the knowledge of molecular events related to rust invasion in plant hosts, nothing is known about gene networks underlying spore formation in rust fungi. The complexity of fungal structures at late infection stages (i.e., host tissues containing haustoria, infection hyphae, sporogenous hyphae, and newly formed spores) is a major limitation to identifying genes specifically involved in sporulation.

Laser capture microdissection (LCM) is a powerful tool to isolate populations of cells from heterogeneous tissues and

Corresponding author: S. Duplessis; Telephone: +33 383 39 40 13; E-mail: duplessi@nancy.inra.fr

*The e-Xtra logo stands for “electronic extra” and indicates that one supplementary table and one supplementary figure are published online.

offers new possibilities to explore transcript expression analysis on specific cells or tissues. LCM has been successfully used to study gene expression in different plant tissues such as epidermal cells, shoot meristem tissues, and root cap tissues

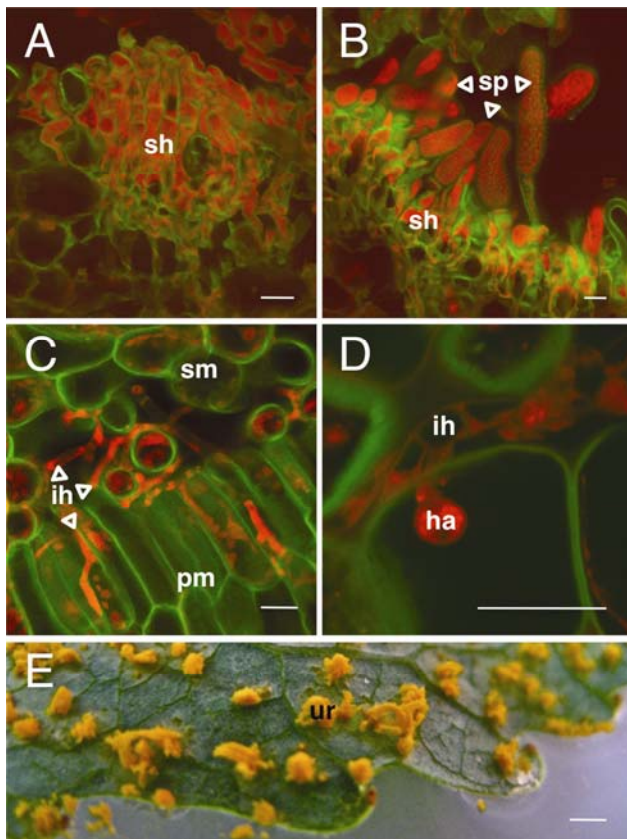


Fig. 1. Development of uredinia from *Melampsora larici-populina* 98AG31 in the susceptible *Populus trichocarpa* × *P. deltoides* cv. Beaupré. **A** through **D**, Confocal laser scanning microscopy images of fungal structures in planta at uredinial stage using Uvitex 2B staining plant and fungal cell walls (green) and propidium iodide staining fungal material (red). **A**, Early and **B**, mature uredinium formed after 96 and 168 h postinfection (hpi). **C**, Infection hyphae and **D**, haustorium colonizing spongy and palisade mesophyll at 168 hpi. Pictures were taken using a ×100 objective lens; scale bar represents 5 μm. **E**, Macroscopic view of mature uredinia. Bar represents 1 mm; sh = sporogenous hyphae, sp = spores, ih = infection hyphae, ha = haustorium, sm = spongy mesophyll, pm = palisade mesophyll, ur = uredinium.

(Nakazono et al. 2003; Jiang et al. 2006; Brooks et al. 2009). So far, only a few studies have reported the use of LCM to study plant–microbe interactions (Balestrini et al. 2009) such as those between plants and nematodes (Ramsay et al. 2004; Klink et al. 2007; Ithal et al. 2007), arbuscular mycorrhizal fungi (Balestrini et al. 2007; Gomez and Harrison 2009), and pathogenic fungi (Tang et al. 2006; Tremblay et al. 2009; Chandran et al. 2010). An LCM-based approach allowed collecting maize stalk cells associated with *Colletotrichum graminicola* hyphae at early stage of infection and microarray analysis led to the identification of several transcripts highly accumulated and potentially involved in pathogenesis (Tang et al. 2006). In a study conducted on soybean infected by the rust fungus *Phakopsora pachyrhizi*, RNA isolation through LCM from uredinia allowed cDNA library construction and access to a few rust unisquences (Tremblay et al. 2009). More recently, microdissection of infection sites from *Arabidopsis* leaves challenged by the powdery mildew *Golovinomyces orontii* helped in identifying downstream elements of the plant defense response not previously detected through whole-leaf transcriptome analysis (Chandran et al. 2010).

In this study, we successfully used LCM to dissect three distinct areas from poplar leaves infected by *M. larici-populina* at the uredinial stage. Using this strategy, we were able to separate fungal sporulation and fungal infection tissues. High-quality total RNA were isolated and amplified from these areas to assess rust transcript profiles through transcriptome analysis using *M. larici-populina* whole-genome exon oligoarrays and reverse-transcription quantitative polymerase chain reaction (RT-qPCR).

RESULTS

Development of *M. larici-populina* uredinia in poplar leaves.

M. larici-populina uredinia development was visualized in planta at 96 and 168 hpi using laser-scanning confocal microscopy. Primordia of uredinia containing initial sporogenous hyphae were observed at 96 hpi inside spongy mesophyll (Fig. 1A) whereas mature uredinia, from which newly formed dikaryotic urediniospores are released, were observed at 168 hpi (Fig. 1B). In addition, these stages were marked by a massive accumulation of fungal infection hyphae (Fig. 1C), and the presence of many haustoria was also noticeable (Fig. 1D) in both spongy and palisade mesophylls. At 168 hpi, uredinia emerged as yellow-orange pustules on the abaxial leaf surface

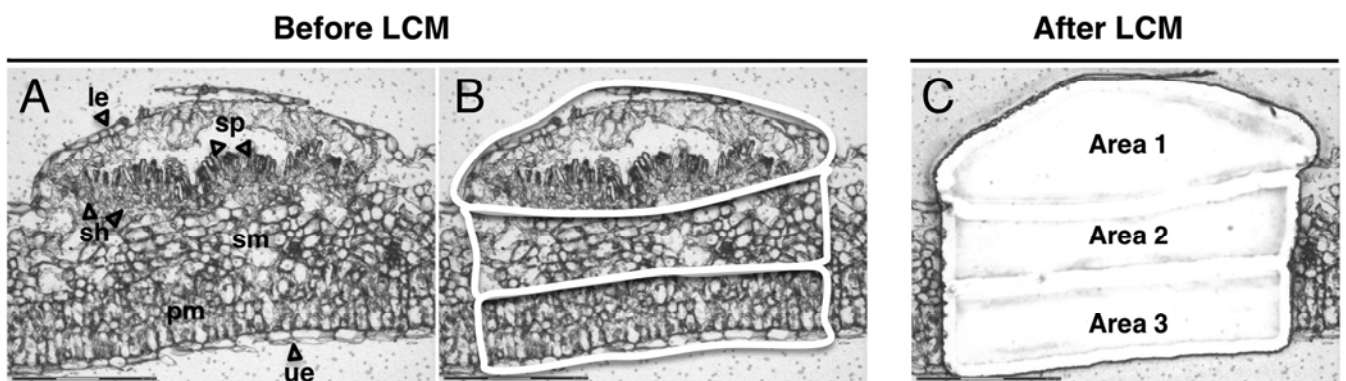


Fig. 2. Laser capture microdissection (LCM) of an uredinium formed by *Melampsora larici-populina* strain 98AG31 on the susceptible poplar cv. Beaupré. **A**, Transversal section of an uredinium; **B**, selection of areas for LCM; and **C**, section after laser capture of the three selected areas. Area 1, subepidermal *M. larici-populina* uredinium below poplar abaxial surface containing urediniospores and sporogenous hyphae; area 2, poplar leaf spongy mesophyll containing infection hyphae and haustoria; and area 3, poplar leaf palisade mesophyll containing *M. larici-populina* infection hyphae and haustoria. Pictures were taken using a ×40 objective lens; scale bar represents 150 μm; sh = sporogenous hyphae, sp = spores, sm = spongy mesophyll, pm = palisade mesophyll, le = lower leaf epidermis, ue = upper leaf epidermis.

of poplar and led to a massive production of asexual urediniospores (Fig. 1E).

Optimization of RNA isolation from LCM uredinia.

Ethanol-acetic acid (EAA) fixation has been reported to maintain histological integrity while enabling good RNA yield recovery (Kerk et al. 2003) and has been successfully used for LCM and microarray analysis of plant tissues (Nakazono et al. 2003; Casson et al. 2008). EAA-fixed and sectioned uredinia are shown before (Fig. 2A and B) and after (Fig. 2C) laser capture of area 1, 2, and 3 (i.e., spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll, respectively) (Fig. 2B). Good preservation of plant and fungal cell structures were obtained from EAA-fixed, paraffin-embedded sections but initial efforts to isolate intact RNA from laser-captured areas were unsuccessful. Using the conventional method of slide mounting, where paraffin ribbons were floated onto a bead of diethylpyrocarbonate (DEPC) water and mounted onto a slide, strong RNA degradation was observed for the different areas as revealed by the absence of ribosomal (r)RNA peaks on RNA electrophoresis profiles (Fig. 3A). When paraffin ribbons were directly transferred onto a microscope slide without water and slide drying, a drastic increase of RNA integrity was observed (Fig. 3B). A moderate degradation of RNA was still observed as revealed by the attenuated signal of the 28S ribosomal subunits of rRNA. In order to improve RNA integrity, we used a 37°C melting point paraffin instead of conventional paraffin at 58°C to prevent exposition of RNA molecules to high temperatures during paraffin impregnation of samples. Using this low-melting-point paraffin coupled with a direct transfer of paraffin ribbons onto microscope slides, two distinct peaks of rRNA were observed for each sampled area on electrophoresis profile, indicating improved RNA integrity after LCM (Fig. 3C; Supplementary Fig. 1A). Less than 5 ng of total RNA were isolated from the three LCM uredinia areas and more than 4 µg of amplified RNA (aRNA) were produced for all samples through linear amplification (Nakazono et al. 2003; Chandran et al. 2010), enabling proper synthesis of cDNA amounts to perform labeling and array hybridization.

Gene expression profiling of LCM fungal tissues.

M. larici-populina whole-genome exon oligoarrays were used to measure transcripts expression in the three areas isolated from uredinia. Among the 15,388 genes analyzed on the oligoarrays, 9,650 (63%) were expressed above background level in at least one of the different areas. In detail, 8,145, 7,786, and 7,288 transcripts were detected in area 1, 2, and 3, respectively. Few differentially expressed transcripts (*t* test, *P* value < 0.05, fold-change > 3) were detected between area 2 and 1 or area 2 and 3 (0 and 56 transcripts respectively; *data not shown*). In contrast, a strong proportion of transcripts were differentially accumulated between area 1 and 3, suggesting a continuum of transcript expression between the three sampled areas in uredinia. In total, 542 transcripts were induced in area 3 compared with area 1, whereas 762 were induced in area 1 compared with area 3. The comparison of transcript profiles between these two contrasted areas at the uredinial stage allowed us to gain insights into the genes associated with biotrophy (area 3) and sporulation (area 1).

Poplar rust biotrophy-related transcripts.

Among the 542 *M. larici-populina* transcripts upregulated in area 3 compared with area 1, 117 shared similarities with known proteins and were assigned to functional categories (Table 1) based on the Functional Catalogue (FunCat) scheme (Ruepp et al. 2004) to decipher molecular mechanisms under-

lying biotrophic growth. Six transcripts encoding *M. lini* HESP homologs previously identified in haustoria during flax infection (Catanzariti et al. 2006) (JGI [Joint Genome Institute] protein ID nos. 64764, 58916, 103805, 110232, 71126, and 93569) were gradually expressed in spongy mesophyll toward palisade mesophyll and showed inductions ranking between seven- and 90-fold compared with the sporulation area (Table 1). Two of them were identified among the 25 most highly induced *M. larici-populina* transcripts (>80-fold) in area 3 compared with area 1 (Table 2). Interestingly, 19 of these 25 most highly upregulated transcripts encoded secreted proteins of less than 300 amino acids (small secreted proteins [SSP]) and most of them were lineage specific but of unknown function. The most highly induced transcript (JGI protein ID no. 106078; 400-fold induction) encoded a 137-amino-acid cysteine-rich SSP with one ortholog in the wheat stem rust *Puccinia graminis* f. sp. *tritici* genome and no ortholog in other fungi (Table 2). Induction of expression of transcripts related to lipid or polysaccharide degradation such as cutinase (JGI protein ID no. 123735), lipases (JGI protein ID nos. 110949 and 106095), and β-glycosidase (JGI protein ID no. 91623) was observed in the biotrophic area (Table 1), indicating that the fungus might still degrade the host cell wall at the uredinial stage. Several enzymes involved in carbohydrate metabolism were also strongly induced in the biotrophic area (>10-fold) such as an invertase (JGI protein ID no. 44167), able to cleave sucrose from the host plant cell in glucose and fructose. Transcripts encoding sugar transporters (JGI protein ID nos. 38418 and 85306), an amino acid permease (JGI protein ID no. 34045), or an oligopeptide transporter (JGI protein ID no. 105630) were also highly induced in the biotrophic area (more than sixfold), highlighting the importance of nutrients uptake in the host mesophyll. Interestingly, the two most induced transporter transcripts (JGI protein ID nos. 38418 and 34045), with 31.1- and 12.1-fold changes, respectively, are homologs of the *U. fabae* hexose transporter HXT1 and amino acid permease AAT2 previously described for their expression restricted to haustoria (Hahn et al. 1997; Voegele et al. 2001). In addition, exocytosis- and cellular-trafficking-related transcripts (JGI protein ID nos. 95032, 47066, and 103049) were also

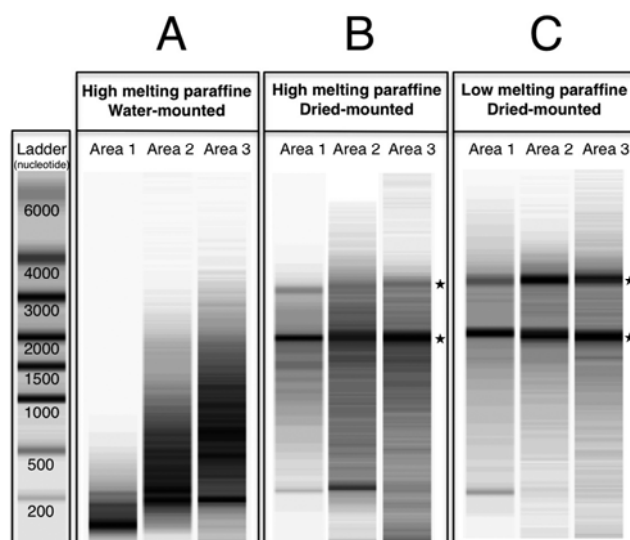


Fig. 3. Electrophoretic profiles of total RNA isolated from laser microdissected areas 1, 2, and 3 using different sample preparation methods. **A**, Sample impregnation using high-melting point paraffin and water mounted on the slide; **B**, sample impregnation using high-melting point paraffin and dry mounted on the slide; and **C**, sample impregnation using low-melting point paraffin and dry mounted on the slide. Stars indicate ribosomal RNA subunits.

upregulated, consistent with an intense cellular transport activity in the biotrophic area. Several transcripts involved in protein degradation were also differentially expressed in area 3 compared with area 1, such as an ubiquitin-conjugating enzyme (JGI protein ID no. 41986) and a RING-finger-containing ubiquitin ligase (JGI protein ID no. 118355) with 5.5- and fivefold regulation, respectively. Moreover, two proteases (JGI protein ID nos. 69042 and 66970) acting in peptide bonds hydrolysis were upregulated 13.5- and 9.1-fold, respectively, suggesting a strong activation of fungal proteasome and proteolysis inside palisade mesophyll at late uredinial stage.

Poplar rust sporulation-related transcripts.

Among the 762 transcripts upregulated in area 1 compared with area 3, 381 showed homologies with known proteins and were assigned to functional cellular categories in order to identify gene networks associated with urediniospore formation (Table 3). An important proportion of transcripts induced in the sporulation area encode detoxification enzymes such as cytosolic Cu/Zn superoxide dismutases (JGI protein ID nos. 72564 and

72563), peroxidases (JGI protein ID nos. 87334 and 94172), and a cytochrome P450 (JGI protein ID no. 76380), suggesting cellular protection against oxygen-derived free radicals during sporogenesis. Nine transcripts related to cell cycle control and encoding a centrin 3, septin, M-phase inducer phosphatase, and DNA replication initiation factor *cdc45* were also upregulated in area 1, highlighting the high rate of cell division in the sporulation area compared with the biotrophic area. In addition, DNA replication-, recombination-, and repair-related transcripts were strongly upregulated, consistent with cell division activity. Interestingly, most of the genes involved in cell division and detoxification were among the 25 most highly upregulated transcripts (>90-fold) in the sporulation area compared with the biotrophic area (Table 4). Transcripts encoding lipid metabolism enzymes, such as a squalene synthase (86.7-fold induction, JGI protein ID no. 94543) involved in sterol biosynthesis pathway, and three enzymes showing lipase activity, such as a phospholipase (60.4-fold induction, JGI protein ID no. 59333), an epoxide hydrolase (39.2-fold induction, JGI protein ID no. 123933), and a lipase (31.1-fold induction, JGI protein ID no. 63115), were observed

Table 1. Selected *Melampsora larici-populina* genes presenting transcript induction in the palisade mesophyll (area 3) versus sporogenous (area 1) microdissected structures, assessed by whole-genome expression oligoarrays and categorized according to their cellular category (Functional Catalogue)

Protein ID ^b	Putative function or feature	Uredinia laser capture microdissection ^a				
		Expression level			Expression ratio	
		Area1	Area2	Area3	Area3/Area1	P
	Subcellular localization					
64764	Small secreted protein (hesp-376 homology)	15	185	1,363	90.0	0.012
58916	Small secreted protein (hesp-C63 homology)	15	246	1,247	85.7	0.010
103805	Small secreted protein (hesp-270 homology)	109	2,994	4,509	41.5	0.010
110232	Small secreted protein (hesp-897 homology)	53	300	1,512	28.3	0.042
71126	Small secreted protein (hesp-C49 homology)	355	4,389	8,691	24.4	0.023
93569	Small secreted protein (hesp-417 homology)	5,665	18,108	40,009	7.0	0.033
	Metabolism					
	Lipid					
123735	Cutinase	88	2,311	6,193	70.2	0.011
110949	Lipase B	53	649	2,338	44.0	0.020
106095	Lysophospholipase A	88	1,161	2,659	30.0	0.025
87140	Polyprenyl synthetase	22	36	349	16.0	0.046
72337	α/β hydrolase	553	1,663	2,822	5.1	0.039
	Carbon					
53785	Ribokinase	24	263	1,020	42.7	0.018
44167	β -Fructosidase (invertase) (glycosyl hydrolase family 32)	417	1,377	4,421	10.6	0.023
29139	Phosphoribosylglycinamide formyltransferase	470	1,537	4,720	10.0	0.018
91623	β -Glycosidase (glycosyl hydrolase family 5)	30	54	205	6.8	0.034
	Vitamin					
79837	Riboflavin biosynthesis protein	155	490	874	5.6	0.044
	Redox					
86038	Oxidoreductase, 2OG-Fe(II) oxygenase family	408	4,197	10,964	26.8	0.039
90083	Indoleamine 2,3-dioxygenase	142	1,324	3,357	23.6	0.049
63255	NAD dependent epimerase/dehydratase	72	666	1,502	20.8	0.031
114173	Aldo/keto reductase	794	9,026	15,541	19.5	0.037
93258	Oxidoreductase 2OG-Fe(II) oxygenase family	131	837	1,273	9.7	0.047
33839	FAD dependent oxidoreductase	686	4,556	5,160	7.5	0.030
109018	Glutaredoxin	207	653	1,551	7.4	0.036
84159	NAD dependent epimerase/dehydratase	204	644	1,433	7.0	0.013
33377	Short-chain dehydrogenase/reductase	1,197	6,383	7,263	6.0	0.041
103191	Oxidoreductase, 2OG-Fe(II) oxygenase family	1,028	3,604	5,722	5.5	0.032
25325	D-arabinono-1,4-lactone oxidase	11	16	57	5.1	0.020
	Cellular transport, transport facilitation					
38418	HXT1p Sugar transporter	29	329	901	31.1	0.014
34045	Amino acid permease AAT2	34	90	415	12.1	0.039
85306	Sugar transporter-MFS monosaccharide transporter	213	313	2,315	10.8	0.016
95032	Exocyst complex component, sec3 subunit	13	17	112	8.6	0.036
47066	Clathrin assembly protein	178	168	1,179	6.6	0.029
105630	Oligopeptide transporter	121	458	792	6.5	0.040
103049	Got1-like vesicle transport protein	42	113	237	5.6	0.013

(continued on next page)

^a Normalized transcript levels in area 1, 2, and 3 (i.e., spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll, respectively), transcript expression ratio (area 3/area 1), and associated *t* test *P* value (<0.05).

^b Protein ID number of corresponding best gene model in the *M. larici-populina* JGI genome sequence.

in the sporulation area (Table 3). Other enzymes involved in carbohydrate degradation, such as α -glycosidase (JGI protein ID nos. 92210), β -glycosidases (JGI protein ID nos. 116147 and 55983), and α -L-arabinofuranosidase (JGI protein ID no. 107307), also showed transcript induction during sporogenesis. In contrast with the biotrophic area, genes involved in nucleotide catabolism, such as ribonucleases (JGI protein ID nos. 117154 and 54248) and an endonuclease (JGI protein ID no. 26899), were strongly induced in the sporulation area, indicating a strong turnover of nucleic acids. Several transporters were specifically detected during urediniospores formation. The most induced encoded a GDP-fucose transporter (JGI protein ID no. 49361) that showed a 93-fold induction. MFS permeases (JGI protein ID nos. 94938 and 25355), an amino acid permease (JGI protein ID no. 49500), and an oligopeptide transporter (JGI protein ID no. 94692) were among the highly induced transcripts (>30-fold), indicating strong metabolite trafficking concomitant with urediniospore synthesis. In contrast with the strong proportion of SSP identified among the most highly induced transcripts in the biotrophic area (Tables 1 and 2), one *M. lini* HESP homo-

log was induced in the sporulation tissues (Table 3) and only five SSP were identified among the 25 most highly induced transcripts (>80-fold) in the sporulation area (Table 4), suggesting the preferential role of SSP during the biotrophic growth phase.

Validation of differential gene expression profiles.

The expression level of 29 transcripts encoding *M. laricipopulina* SSP and known rust protein homologs or that showed a preferential expression pattern in area 1 or 3 was assessed by RT-qPCR. A linear correlation ($R^2 = 0.96$) was observed between expression profiles measured by RT-qPCR and oligoarrays validating the transcriptomic approach (Fig. 4). The slope of the regression line was close to unity (0.86), indicating a minimal variation in expression data and consistency between both approaches.

DISCUSSION

As noncultivable organisms, obligate biotrophic fungi are dependent on their host to achieve their growth and develop-

Table 1. (continued from preceding page)

Protein ID ^b	Putative function or feature	Uredinia laser capture microdissection ^a				
		Expression level			Expression ratio	
		Area1	Area2	Area3	Area3/Area1	P
	Protein fate (folding, modification, destination)					
69042	Pepsin	741	9,008	10,207	13.5	0.011
66970	Cysteine proteases	24	64	228	9.1	0.048
41986	Ubiquitin-conjugating enzyme E2	990	2,606	5,531	5.5	0.035
118355	C3HC4 type RING finger containing ubiquitin ligase	1,193	3,193	6,042	5.0	0.013
	Transcription					
108564	GCN5-related N-acetyltransferase	20	73	575	28.7	0.025
107345	C2H2-type Zn finger protein	175	2,448	2,377	13.5	0.021
74108	Transcriptional regulator Prz1	1,147	4,880	11,866	10.3	0.016
94759	C2H2-type Zn finger protein 7	282	3,356	2,585	9.1	0.024
39532	Transcription elongation factor	627	1,579	3,293	5.2	0.027
	Cellular communication or signal transduction mechanism					
106239	NUDIX hydrolase	44	392	991	22.3	0.010
94327	Protein kinase AGC family	23	30	165	7.3	0.039
43665	Response regulator receiver	682	1,908	4,450	6.5	0.048
	Cell rescue, defense, and virulence					
40544	Survival protein SurE	844	1,810	13,302	15.7	0.010
106535	Heat shock protein	304	755	1,554	5.1	0.018
	Cell cycle and DNA processing					
60363	DNA topoisomerase I	27	69	349	13.1	0.034
47226	G1 Cyclin	472	2,140	3,863	8.1	0.028
	Protein synthesis					
30347	Ribosomal protein S8	30	62	418	13.8	0.038
37443	Diphthine synthase	618	531	3,279	5.3	0.020
	Biogenesis of cellular components					
27318	Cell wall protein	23	78	248	10.6	0.030
95757	Proteophosphoglycan 5	35	62	277	7.9	0.040
64466	Proteophosphoglycan ppg4	20	42	123	6.1	0.038
	Transposable elements, viral and plasmid proteins					
88937	Transposase	42	110	514	12.2	0.048
94469	Transposable element	14	22	75	5.2	0.048
	Energy					
93294	Malate dehydrogenase	485	986	3,429	7.0	0.042
95852	Malate dehydrogenase, NAD-dependent	40	73	235	5.9	0.014
65530	Electron transport complex	12	45	72	5.8	0.031
	Protein with binding function or cofactor requirement					
105154	Phosphatidylethanolamine-binding protein	200	2,567	13,666	68.4	0.013
66416	Zinc finger SWIM domain-containing protein	21	37	184	8.9	0.049
62481	Glycine-rich RNA binding protein	1,682	8,050	12,280	7.3	0.021
38287	Phosphatidylethanolamine-binding protein	34	44	234	6.8	0.046
	Unclassified proteins					
28849	Histidine-rich protein	21	49	280	13.5	0.039
93926	Sporozoite surface protein	22	62	294	13.2	0.043
61950	Glycoprotein	20	34	162	8.0	0.042
85862	CRKRS protein	14	19	84	6.0	0.027
87097	EPS15 homology (EH) domain-containing protein	20	31	105	5.3	0.018

ment. This is a major limitation to studying the biology of rust fungi and, thus far, there is a limited knowledge of molecular mechanisms at stake during rust proliferation in planta and, particularly, at the late stage of infection when sporulation occurs. In order to show that physiological heterogeneity and metabolic zonation occurred within fungal tissues (Cairney and Burke 1996), we used an LCM-based approach to study gene expression at the uredinial stage of rust infection in plant tissue (i.e., production of asexual urediniospores) targeting different leaf areas below uredinial structures. Using this strategy, we were able to separate two contrasted fungal tissues from poplar leaf tissues infected by *M. larici-populina*: a biotrophic area in the plant mesophyll and a sporulation area below the abaxial epidermis where the uredinial dome is formed and from which newly formed urediniospores are released.

Transcriptome analyses strongly depend on RNA quality (Copojs et al. 2007) and the use of LCM-collected samples to accurately determine integrity of RNA isolated from microdissected tissues. Low-volume automated electrophoresis bioanalyzers now allow assessment of RNA profiles at very low concentrations, down to 100 pg/μl. In microdissection studies, RNA quality is often based on RT-PCR transcript amplification (Kerk et al. 2003; Inada and Wildermuth 2005; Yu et al. 2007) and not on total RNA profiles. Tissue preparation prior to LCM and fixation methods, for instance, can strongly impact quality of RNA (Kerk et al. 2003). Farmer's fixation has been reported to maintain good histological and RNA integrity (Kerk et al. 2003; Ivashikina et al. 2003; Nakazono et al. 2003; Ramsay et al. 2004; Nakada et al. 2006). Loss of RNA integrity during slide mounting was previously pointed out and resolved by the

use of a tape transfer system (Cai and Lashbrook 2006). Here, we report that other factors impact RNA quality during tissue preparation, such as sample impregnation and slide mounting. Low-melting paraffin coupled with a direct transfer of paraffin ribbons onto microscope slides without water and slide drying led to a drastic increase of RNA integrity (Fig. 3).

Transcriptome analysis of LCM-sampled tissues revealed gene networks of *M. larici-populina* associated with particular biological features of this rust fungus such as biotrophy and sporulation. Many transcripts were strongly induced in the biotrophic tissues compared with the sporulation tissues, suggesting that the corresponding proteins might be specifically involved in the parasitic growth (Tables 1 and 2). A striking result is the massive induction of genes encoding SSP in the palisade mesophyll, including several homologs of *M. lini* haustorially expressed secreted protein (Catanzariti et al. 2006). These SSP represent a reservoir of putative effectors that might be required to alter the host defense system (Ellis et al. 2007). Thus, it is tempting to link the induction of *M. larici-populina* SSP to the maintenance of the biotrophic interaction (i.e. through suppression of host defense) at the uredinial stage. Concomitantly with this massive induction of transcripts encoding SSP, the *sec3* subunit of the exocyst complex, involved in the tethering, docking, and fusion of post-Golgi vesicles with the plasma membrane (Wu et al. 2008) as well as the Got1-like vesicle transport protein, showed consistent inductions in the fungal tissue (Table 1), suggesting that the exocytosis mechanism, one of the suggested routes for effector secretion (Voegelé et al. 2009), might be required for SSP release by rust infection structures. Several transcripts related to car-

Table 2. Most highly upregulated *Melampsora larici-populina* transcripts in the palisade mesophyll (area 3) versus sporogenous (area 1) microdissected structures identified using whole-genome expression oligoarrays and ranked by decreasing fold changes

Protein ID ^b	SP ^c	Length	Cysteine ^d	Putative function or features	Blastp ^e	Uredinia laser capture microdissection ^a				
						Expression level			Expression ratio	P
						Area1	Area2	Area3	Area3/Area1	
106078	Y	137	11	Rust-specific SSP	PGTG_15874	23	4,804	9,025	397.5	0.013
Mlp0002_0042	Y	76	1	Mlp-specific SSP	No significant hit	11	517	3,802	340.4	0.012
106150	Y	118	1	Mlp-specific SSP	No hits found	28	3,457	7,441	262.1	0.012
104758	–	136	0	Mlp-specific protein	No hits found	19	804	4,909	255.8	0.013
123218	Y	210	6	SSP	PGTG_02151	68	7,025	17,338	254.9	0.011
105547	Y	206	1	Mlp-specific SSP	No hits found	28	1,349	6,440	231.8	0.013
108588	–	109	4	Mlp-specific protein	No hits found	23	3,575	5,265	225.6	0.014
109323	–	142	3	Mlp-specific protein	No hits found	32	2,668	6,840	215.0	0.013
123731	Y	184	2	Mlp-specific SSP	No significant hit	27	1,417	5,769	212.8	0.011
Mlp0016_0150	Y	98	7	Mlp-specific SSP	No significant hit	18	820	3,662	202.1	0.011
112715	Y	142	9	Mlp-specific SSP	No hits found	33	2,788	6,355	190.9	0.021
56132	–	155	4	Rust-specific protein	PGTG_05628	27	130	5,186	189.5	0.018
123531	Y	103	9	Mlp-specific SSP	No significant hit	41	2,966	7,457	179.8	0.040
60448	Y	156	3	Mlp-specific SSP	No hits found	48	978	8,533	179.4	0.013
72983	Y	220	8	SSP	PGTG_06969	102	10,925	16,923	166.1	0.027
102036	Y	107	1	Mlp-specific SSP	No hits found	18	927	2,928	165.3	0.011
86274	Y	244	10	SSP	PGTG_04481	16	763	2,523	162.3	0.007
106925	–	105	10	Mlp-specific protein	No hits found	21	1,668	3,022	144.1	0.010
Mlp0032_0018	Y	76	2	Mlp-specific SSP	No significant hit	52	3,818	7,088	136.8	0.006
66715	Y	459	2	Mlp-specific secreted protein	No hits found	14	521	1,718	119.8	0.014
123438	Y	180	11	Rust-specific SSP	PGTG_08705	223	15,297	25,016	112.1	0.021
109567	Y	156	2	Mlp-specific SSP	No hits found	48	976	5,287	109.1	0.046
64764	Y	154	1	SSP (hesp-376 homology)	No hits found	15	185	1,363	90.0	0.012
68912	Y	144	2	Mlp-specific SSP	No hits found	38	1,255	3,390	89.7	0.014
58916	Y	219	4	SSP (hesp-C63 homology)	No hits found	15	246	1,247	85.7	0.010

^a Normalized transcript levels in area 1, 2, and 3 (i.e., spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll, respectively), transcript expression ratio (area 3/area 1), and associated *t* test *P* value (<0.05).

^b Protein ID number of corresponding best predicted gene model in the *M. larici-populina* JGI genome sequence; Mlp0002_0042, Mlp0016_0150, and Mlp0032_0018 correspond to additional predicted gene models not described on the *M. larici-populina* genome portal.

^c Signal peptide (SP) prediction. Presence (Y) or absence (–) of a signal peptide is indicated.

^d Number of cysteines.

SSP = small secreted protein.

^e Best blast hit against the genome sequence of *Puccinia graminis* f. sp. *tritici* deposited at the Broad Institute genome portal.

bohydrate transport and metabolism and expressed in haustoria of *U. fabae* were also detected in the biotrophic area of *M. larici-populina*. The strong expression of a secreted invertase, a β -glycosidase, and an hexose transporter (*U. fabae* HXT1 homolog) is consistent with the hypothesis that fungal enzymes should be present in the haustorial matrix to provide the substrate for the hexose transporter (Voegelé et al. 2001, 2006, 2009). In addition, transcripts encoding an amino acid permease showing homology with the haustoria-specific Uf-AAT2 permease (Hahn et al. 1997) were also strongly induced, suggesting a conserved role for rust transporters in nutrient uptake from infected host cells.

Most recent biochemical and molecular studies of rust fungi focused on the parasitic growth and, more precisely, the haustorial structure. In addition, the molecular basis underlying the sporulation process in rust is poorly known. This step is essential for rust development and leads to huge production of asexual urediniospores disseminated by winds over very large distances (Nagarajan and Singh 1990; Barrès et al. 2008). In ascomycetes, molecular genetic approaches have been successfully applied to dissect the genetic mechanisms controlling conidia morphogenesis in model fungal species (Adams et al. 1998). In rust fungi, LCM and EST sequencing of uredinia formed by *Phakopsora pachyrhizi* identified a limited number of unisquences displaying similarities to known proteins in databases (Tremblay et al. 2009). Among them, transcripts encoding ribosomal protein and serine/threonine-protein phosphatase were also strongly expressed in *M. larici-populina* uredinia. Our LCM-based transcriptome study allowed the unraveling of essential functions of rust sporulation in the basidiomycete *M. larici-populina* (Tables 3 and 4). Asexual urediniospores are the result of cell division occurring via mitosis. Several genes preferentially expressed in the sporulation area are involved in the cell cycle, such as the M phase inducer phosphatase *cdc25*, as well as *cdc45* required for initiation of DNA replication in *Saccharomyces cerevisiae* (Saha et al. 1998). This mitotic activity during sporogenesis is also supported by transcript induction of a septin involved in cytokinesis. Although cytokinesis mechanisms have been well described during yeast division and sporulation (Wolfe and Gould 2005; Diamond et al. 2009) and hyphae septation (Walther and Wendland 2003), their role in sporogenesis of basidiomycetes remains unclear. The transport mechanisms used to translocate metabolites from haustoria to urediniospores also remains to be determined. We report the

strong induction of several transporters that likely play an important role in metabolite trafficking toward the newly formed urediniospores. Peroxidases and superoxide dismutases are involved in the prevention of oxidative damages (Alscher et al. 2002). At later stages of compatible interaction between poplar leaves and *Melampsora* rust fungi, previous studies have shown the late expression of plant defense-related proteins and the accumulation of reactive oxygen species (Duplessis et al. 2009; Boyle et al. 2010). High levels of detoxification enzyme-related transcripts in the sporulation area could indicate that the fungus is setting protective mechanisms against late plant host defenses. In contrast with the biotrophic area, several transcripts involved in the degradation of plant cell wall polysaccharides were expressed in the sporulation area. The release of a cocktail of degrading enzymes by *M. larici-populina* during sporulation might be required to alter the host cell wall in order to break the epidermis cell layer and release the urediniospores.

In summary, this LCM-transcriptomic-coupled study allowed us to distinguish the metabolic and developmental functions of two contrasted rust tissues at the uredinial stage: a biotrophy-related tissue in the host palisade mesophyll and a sporulation-related tissue below the abaxial leaf epidermis. Optimization of LCM technical conditions allowed us to isolate RNA of good integrity to perform an array-based expression analysis. This approach provides insights on distinct genetic programs activated by rust at the uredinial stage. Particularly, a massive induction of transcripts encoding putative effectors, such as SSP, support a late maintenance of biotrophy at the late stage of rust infection, when urediniospores are released. In contrast, cell cycle and cell defense rescue are important components of the rust sporulation process.

MATERIALS AND METHODS

Growth conditions and inoculation procedures.

M. larici-populina isolate 98AG31 (pathotype 3-4-7) was used in this study. Urediniospore multiplication and plant inoculation procedures were performed as previously described using the same inoculum of 100,000 urediniospores/ml, leaf plastochrony index for detached *Populus trichocarpa* \times *P. deltoides* 'Beaupré' leaves, and strictly identical culture conditions (Rinaldi et al. 2007). For LCM analysis, leaves were harvested after 168 hpi when uredinial pustules were visible on the leaf abaxial epidermis.

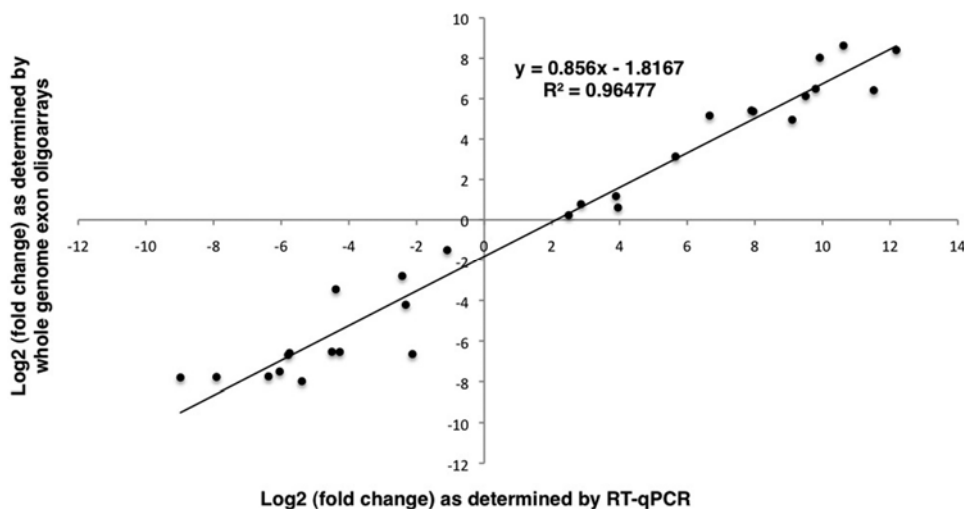


Fig. 4. Validation of *Melampsora larici-populina* whole genome exon oligoarrays data using reverse-transcription quantitative polymerase chain reaction (RT-qPCR). Log₂ transcript fold changes measured on NimbleGen oligoarrays and by RT-qPCR are plotted for a set of 29 differentially-expressed genes between area 3 and area 1.

Laser-scanning confocal microscopy.

Infected leaves showing the first fungal decoloration and mature uredinia were harvested at 96 and 168 hpi, respectively, and were directly cut and fixed in 4% (wt/vol) paraformaldehyde in phosphate-buffered saline (PBS, pH 7) for 3 h at 4°C. Fixed tissues were washed in PBS buffer (1×) and embedded in 6% agarose (wt/vol). Sections of 20 µm were prepared using a VT1000S Leica vibratome (Leica Microsystems, Nanterre, France). Sections were dual stained with 1% (wt/vol) Uvitex 2B (Polysciences, Eppelheim, Germany) in PBS for 2 min and washed with propidium iodide (PI; 10 µg/ml) (Invitrogen, Cergy Pontoise, France). PI was used to stain fungal infection structures (Koh et al. 2005) and Uvitex was used to stain plant and fungal cell walls. Transversal sections were observed using a NIKON-BIORAD Radiance 2100 Rainbow confocal scanning laser microscope. Sections were excited with blue laser beams at 405 nm, scanned with a filter setting at 500 to 560 nm for Uvitex, excited with argon laser beams at 515 nm, and scanned with filter settings at 550 to 650 nm for PI. Merged images

were obtained using the LaserSharp 2000 program (Bio-Rad, Marnes la Coquette, France).

Tissue preparation for LCM.

Ten leaf pieces of 0.4 cm², each containing a single uredinium, were randomly sampled on the overall leaf surface and immediately fixed on ice in Farmer's solution (EAA) containing 75% (vol/vol) ethanol and 25% (vol/vol) acetic acid (Kerk et al. 2003). This fixative was first vacuum infiltrated into the section (400 mm of Hg) for 15 min and then left in contact with leaf pieces overnight at 4°C. Farmer's solution was removed and samples were dehydrated in a graded series of ethanol (45 min each on ice [vol/vol] at 75, 80, and 90% and three times in absolute ethanol), followed by an ethanol:otix+ (Microm Microtech, Francheville, France) series (45 min each on ice [vol/vol] at 75:25, 50:50, and 25:75% and three times in 100% Otix+). After Otix+ removal, two paraffins were used for sample impregnation. The first was a diawax paraffin (Diapath, Martinengo, Italy) with a 58°C melting point and the sec-

Table 3. Selected *Melampsora larici-populina* genes presenting transcript induction in sporogenous (area 1) versus palisade mesophyll (area 3) micro-dissected structures, assessed by whole-genome expression oligoarrays and categorized according to their cellular category (Functional Catalogue)

Protein ID ^b	Putative function or features	Uredinia laser capture microdissection ^a				
		Expression level			Expression ratio	
		Area1	Area2	Area3	Area1/Area3	P
	Cell rescue, defense and virulence					
72564	Cytosolic Cu/Zn superoxide dismutase	6,548	173	33	196.8	0.020
87334	Peroxidase, family 2	7,075	487	39	181.6	0.034
72563	Cytosolic Cu/Zn superoxide dismutase	5,662	1,102	59	96.4	0.046
94172	Peroxidase	6,961	3,624	81	86.2	0.046
76380	Cytochrome P450	401	126	11	37.3	0.013
	Metabolism					
	Lipid					
94543	Squalene synthase	985	348	11	86.7	0.022
59333	Phospholipase/lecithinase/hemolysin	2,393	645	40	60.4	0.045
123933	Epoxide hydrolase	3,162	2,553	81	39.2	0.016
123555	Coesterase/carboxylesterase	646	472	14	45.7	0.007
35173	Sphingomyelin phosphodiesterase	554	372	17	32.3	0.025
63115	Lipase	608	120	20	31.1	0.014
	Carbon					
92210	α-Glycosidase (glycoside hydrolase family 71)	700	101	13	55.7	0.019
116147	β-Glycosidase (glycoside hydrolase family 5)	2,919	1,197	55	52.8	0.044
55983	β-Glycosidase (glycoside hydrolase family 16)	2,246	717	48	47.1	0.046
107307	α-L-arabinofuranosidases (glycoside hydrolase family 51)	1,197	1,125	35	33.9	0.021
87698	Chitin deacetylase (carbohydrate esterase family 4)	907	112	29	31.1	0.010
	Nucleotide					
117154	Ribonuclease P protein subunit	2,665	1,312	12	214.4	0.010
88931	Adenosine deaminase-like protein	629	23	13	49.1	0.010
26899	Endonuclease III	393	105	9	45.9	0.009
54248	Ribonuclease	831	189	19	44.4	0.032
93485	Spermidine synthase	471	467	11	41.4	0.013
	Redox					
38530	Oxidoreductase, short chain dehydrogenase	3,257	336	13	252.9	0.010
	Cell cycle and DNA processing					
40080	Centrin 3	2,308	1,741	22	103.1	0.023
87088	Septin	3,232	119	35	92.5	0.013
110193	6-O-methylguanine-DNA methyltransferase	2,253	623	25	89.1	0.024
91152	20S cyclosome subunit	857	168	11	75.5	0.009
76389	Meiotic chromosome segregation protein	1,067	284	23	46.1	0.014
71714	Nuclear transcription factor	512	640	14	35.5	0.022
85106	Cell division control protein 45	312	98	9	33.8	0.026
30176	M-phase inducer phosphatase	638	42	19	33.0	0.047
89448	Transcription-associated recombination protein	397	91	12	31.8	0.038
	Cellular transport, transport facilitation					
49361	GDP-fucose transporter	1,119	21	12	93.0	0.024
79553	ER-Golgi vesicle-tethering protein	2,865	527	33	86.9	0.021
39514	Synaptobrevin	625	73	10	64.4	0.013

(continued on next page)

^a Normalized transcript levels in area 1, 2, and 3 (i.e., spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll, respectively), transcript expression ratio (area 1/area 3), and associated *t* test *P* value (<0.05).

^b Protein ID number of corresponding best gene model in the *M. larici-populina* JGI genome sequence.

ond had a 37°C melting point (VWR, Fontenay sous bois, France). Samples were transferred to an oven at 37 or 58°C and approximately 10 ml of liquefied paraffin (VWR) was added to each vial and left at 37 or 58°C overnight. The paraffin was replaced five times at 1.5-h intervals. Then, samples were embedded using the conventional diawax paraffin with the 58°C melting point. This high-melting-point paraffin was retained for tissue embedding in order to realize proper sections. Sections (10 µm each) were cut on a rotary microtome and transferred to microscope membrane slide (Zeiss, Bernried, Germany) with or without water bath impregnation of paraffin ribbons. In order to improve adhesion between microscope slide and ribbons, slides were kept at 45°C during 15 min to spread out ribbons and stored at -20°C in darkness under dehydrating conditions. Just before LCM, slides were deparaffinized twice in OtiX+ for 10 min and then in otiX+:ethanol series (3 min each [vol/vol] at 67:33, 50:50, and 33:67% and three times in absolute ethanol).

LCM.

For each biological replicate, 100 sites of uredinium formation were isolated after sectioning of 10 uredinia randomly sampled from a single infected poplar leaf (10 serial sections per uredinium). For each of the 100 uredinia sections, three distinct areas were collected by LCM at each uredinial site corresponding to spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll (Fig. 2). A PALM Robot-Microbeam system (Carl Zeiss MicroImaging, Munich, Germany) was used to microdissect cells from deparaffinized and dried tissue sections. Laser parameters for dissection performed with the ×40 objective lens were energy, 52; delta, 12; and focus, 68 using the RoboLPC cutting and catapulting function. The catapulted material was collected in an adhesive cap of a 500-µl tube (Zeiss). Three biological replicates were analyzed in total (i.e., three leaves from different poplar plants independently infected by rust urediniospores) and three areas were LCM collected from each replicate.

Table 3. (continued from preceding page)

Protein ID ^b	Putative function or features	Uredinia laser capture microdissection ^a				
		Expression level			Expression ratio	
		Area1	Area2	Area3	Area1/Area3	P
	Cellular transport, transport facilitation (<i>continued</i>)					
71469	Aquaporin-1	8,125	1,313	143	56.6	0.049
94938	Permease-MFS transporter	710	461	15	48.2	0.015
46095	Endosomal membrane protein	811	113	20	40.6	0.044
32963	Vacuolar assembly/sorting protein	1,578	865	39	40.2	0.023
33903	Translocation protein	1,263	273	34	37.3	0.027
49500	Amino acid permease	585	129	16	35.7	0.031
94692	Oligopeptide transporter 4	312	28	10	31.8	0.041
25355	Permease-MFS transporter	410	34	13	31.2	0.020
	Transcription					
32431	Methyltransferase	1,211	774	16	74.7	0.011
48934	Ribosomal RNA assembly protein	2,092	1,298	30	70.5	0.010
50186	Elongation factor-type GTP-binding protein	3,359	2,787	51	65.7	0.021
25365	Glutamyl-tRNA amidotransferase	728	393	14	52.0	0.030
117808	tRNA methyltransferase	1,222	177	29	41.7	0.014
27169	tRNA methyltransferase complex subunit Cpd1	612	762	17	35.2	0.008
38981	Transcription factor C2H2 type	1,059	35	32	33.3	0.028
75483	ATP-dependent RNA helicase	1,441	209	48	30.3	0.013
	Cell fate					
73052	Expansin/rare lipoprotein A domain protein	2,384	166	41	58.1	0.023
55582	Riboflavin aldehyde-forming enzyme/Rare lipoprotein A domain protein	3,761	308	67	56.5	0.037
70937	Expansin/rare lipoprotein A domain protein	3,015	1,735	81	37.4	0.035
76387	Expansin/rare lipoprotein A domain protein	6,249	518	207	30.1	0.019
	Protein fate (folding, modification, destination)					
123605	Carboxypeptidase S1	2,270	698	31	73.5	0.029
92176	COP9 signalosome complex subunit	1,003	165	21	48.7	0.036
76645	TPR repeat-containing protein	1,327	324	41	31.3	0.038
	Protein synthesis					
71776	Ribosomal protein	1,042	98	12	84.0	0.010
96447	Capsular associated protein	1,538	796	34	45.2	0.040
	Energy					
110853	NADH-ubiquinone oxidoreductase	5,248	1,767	74	70.7	0.023
38407	FAD monooxygenase	384	340	12	31.9	0.027
	Subcellular localization					
33341	Small secreted protein containing CFEM domain (hesp-178 homology)	2,558	328	58	44.3	0.039
117611	GPI transamidase subunit	319	369	9	36.5	0.010
	Cellular communication/signal transduction mechanism					
117612	Ankyrin repeat protein	1,579	521	38	41.0	0.026
40575	Ras-related GTPase	551	350	18	31.3	0.011
	Transposable elements, viral and plasmid proteins					
72990	Transposase	1,280	705	32	40.4	0.012
	Protein with binding function or cofactor requirement					
104398	Cold-shock DNA/RNA-binding protein	881	625	23	38.6	0.040
77832	Nuclear cap-binding protein subunit	477	241	15	31.7	0.020
	Unclassified proteins					
117548	fmHP	1,950	575	24	80.3	0.018
85991	Pentatricopeptide (PPR) repeat-containing protein	2,097	109	35	60.4	0.016
95993	Major allergen TRI14	954	197	20	47.5	0.032

RNA extraction and amplification.

For the LCM experiment, total RNA (<5 ng) was isolated from all samples using the PicoPure RNA isolation kit (Arcturus, Mountain View, CA, U.S.A.). The RNase-Free DNase Set kit (Qiagen, Courtaboeuf, France) was used during the RNA isolation procedure according to the manufacturer's instructions to remove traces of genomic DNA. Elution was performed with 12 µl of prewarmed DEPC water (65°C). Quality and quantity of RNA samples were assessed using the Bio-Rad Experion analyzer and Experion RNA high-sense analysis kit (Bio-Rad). Total RNA from each sample (<5 ng) was subjected to two rounds of amplification using the MessageAmp II aRNA amplification kit (Ambion, Austin, TX, U.S.A.) following the manufacturer's instructions. RNA amplification generated more than 4 µg of aRNA for each sample, and aRNA profiles were verified using the Experion analyzer and Experion RNA standard-sense analysis kit (Bio-Rad).

Exon oligoarray analyses.

The *M. larici-populina* custom-exon expression array (4 × 72 K) manufactured by Roche NimbleGen Systems Limited (Madison, WI, U.S.A.) contained four independent, nonidentical, 60-mer probes per gene model coding sequence. In total, 17,556 predicted genes from *M. larici-populina* gene catalog (JGI Melampsora portal) and 1,063 random 60-mer control probes and labeling controls were included on the array (GEO platform no. GPL10350). Double-stranded cDNA were synthesized from 2.5 µg of aRNA using the Superscript Double-stranded cDNA synthesis kit (Invitrogen) according to the NimbleGen user protocol. Single-dye labeling of samples, hybridization procedures, and data acquisition were performed at the NimbleGen facilities (NimbleGen Systems, Reykjavik,

Iceland) following their standard protocol. Microarray probe intensities were quantile normalized across chips and average expression levels were calculated for each gene from the independent probes on the array and were used for further analyses. A transcript was deemed expressed when its signal intensity was threefold higher than the mean signal-to-noise threshold (cut-off value) of 1,063 random oligonucleotide probes present on the array. A Student *t* test with false discovery rate (Benjamini-Hochberg) multiple testing correction was applied to the data using the ARRAYSTAR software (DNASTAR, Inc. Madison, WI, U.S.A.). Transcripts presenting a threefold change in transcript level (*P* value < 0.05) were considered to be significantly differentially expressed. Data filtering, normalization process, and complete expression datasets are available at the Gene Expression Omnibus (National Center for Biotechnology Information) as series no. GSE21487.

RT-qPCR.

In order to assess transcript levels by RT-qPCR, specific primers of 29 genes were designed with the Primer 3 and Amplify 3X programs based on the *M. larici-populina* genome sequence (Supplementary Table 1). Amplified fragments showed a length ranging between 121 and 262 nucleotides. Homology searches against *M. larici-populina* and *P. trichocarpa* genome sequences were performed using the blastn algorithm to verify absence of cross annealing with other *M. larici-populina* transcripts as well as with host plant transcripts. First-strand cDNA was synthesized from 150 ng of aRNA using SuperScript II reverse transcriptase (Bio-Rad) with random primers in a final volume of 20 µl. cDNA derived from aRNA (1.5 ng) was amplified by qPCR in iQ SYBR Green supermix (Bio-Rad) with 0.3 µM specific 5' and 3' primers. Amplifications were

Table 4. Most highly upregulated *Melampsora larici-populina* transcripts in sporogenous (area 1) versus palisade mesophyll (area 3) microdissected structures using whole-genome expression oligoarrays and ranked by decreasing fold changes

Protein ID ^b	SP ^c	Length	Cysteine ^d	Putative function or features	Blastp ^e	Uredinia laser capture microdissection ^a				
						Expression level			Expression ratio	<i>P</i>
						Area1	Area2	Area3	Area1/Area3	
38530	–	240	4	Oxidoreductase, short chain dehydrogenase	PGTG_01366	3,257	336	13	252.9	0.010
85149	–	131	4	Mlp-specific protein	No hits found	5,230	6,754	23	222.6	0.032
112748	–	245	4	Mlp-specific protein	No hits found	5,637	2,493	26	218.9	0.022
117154	–	422	5	Ribonuclease P protein subunit	PGTG_13270	2,665	1,312	12	214.4	0.010
72564	–	122	4	Cytosolic Cu/Zn superoxide dismutase	PGTG_05675	6,548	173	33	196.8	0.020
87334	–	323	5	Peroxidase, family 2	PGTG_00594	7,075	487	39	181.6	0.034
114474	–	371	11	Rust-specific protein	PGTG_07899	1,568	1,170	11	149.2	0.009
73728	–	81	1	Mlp-specific protein	No hits found	1,557	347	11	145.5	0.009
76897	Y	185	10	Mlp-specific protein	No hits found	5,523	3,451	39	140.0	0.030
104937	Y	183	9	Mlp-specific protein	No hits found	2,650	1,282	21	128.9	0.011
112967	Y	185	6	Mlp-specific protein	No hits found	4,569	2,810	44	103.4	0.044
40080	–	166	0	Centrin 3	PGTG_06324	2,308	1,741	22	103.1	0.023
73760	–	92	1	Mlp-specific protein	No hits found	1,154	712	11	102.0	0.010
55089	–	74	0	Mlp-specific protein	No hits found	4,608	478	46	100.8	0.022
105220	Y	160	4	Mlp-specific protein	No hits found	1,105	22	11	100.0	0.026
34854	–	105	4	Hypothetical protein	PGTG_12203	4,042	363	40	100.0	0.031
72563	Y	237	5	Cytosolic Cu-Zn superoxide dismutase	PGTG_05675	5,662	1,102	59	96.4	0.046
63262	Y	572	3	Mlp-specific protein	No hits found	1,041	982	11	93.8	0.007
103398	–	116	3	Mlp-specific protein	No hits found	939	39	10	93.3	0.030
71121	–	113	3	Mlp-specific protein	No hits found	1,177	1,358	13	93.2	0.014
67049	–	376	7	Rust-specific protein	PGTG_15359	995	57	11	93.0	0.007
49361	–	486	6	GDP-fucose transporter	PGTG_17642	1,119	21	12	93.0	0.024
87088	–	369	2	Septin	PGTG_13103	3,232	119	35	92.5	0.013
110193	–	183	1	6-O-methylguanine-DNA methyltransferase	No hits found	2,253	623	25	89.1	0.024
57035	–	69	1	Mlp-specific protein	No hits found	1,082	2,251	12	89.1	0.008

^a Normalized transcript levels in area 1, 2 and 3 (i.e., spores and sporogenous hyphae, fungal infection tissues in spongy mesophyll, and fungal infection tissues in palisade mesophyll, respectively), transcript expression ratio (area 1/area 3), and associated *t* test *P* value (<0.05).

^b Protein ID number of corresponding best gene model in the *M. larici-populina* JGI genome sequence.

^c Signal peptide (SP) prediction. Presence (Y) or absence (–) of a signal peptide is indicated.

^d Number of cysteines.

^e Best blast hit against the genome sequence of *Puccinia graminis* f. sp. *tritici* deposited at the Broad Institute genome portal.

carried out in triplicate with an MJ-opticon2 DNA engine (Bio-Rad). After 3 min of denaturation, 40 cycles (95°C for 15 s and 60°C for 1 min) were performed and a melting curve was determined to verify absence of multiple amplifications. *M. larici-populina* transcript expression was normalized to the transcript level of α -tubulin (*Mlp- α TUB*, JGI protein ID no. 73137) and elongation factor (*Mlp-ELF1a*, JGI protein ID no. 73378) reference genes using the Pfaffl calculation (Pfaffl 2001).

ACKNOWLEDGMENTS

We thank D. Le Thiec and D. Cohen for their valuable advice on LCM, B. Pêtre for providing macroscopic pictures of uredinia, and A. Rigal for her help in setting sample impregnation. This work was supported by the Région Lorraine and the Institut National de la Recherche Agronomique (Doctoral Scholarship to S. Hacquard and Junior Scientist Research Grants to S. Duplessis) and the Fabelor and IFR 110 Ecogenomics facilities. The Zeiss P.A.L.M. LCM device was co-funded by the Institut National de la Recherche Agronomique, Région Lorraine, and the Conseil Général 54. The Joint Genome Institute (JGI) is acknowledged for the early access made to the *M. larici-populina* genome sequence. JGI sequencing is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

LITERATURE CITED

- Adams, T. H., Wieser, J. K., and Yu, J. H. 1998. Asexual sporulation in *Aspergillus nidulans*. *Microbiol. Mol. Biol. Rev.* 62:35-54.
- Alscher, R. G., Erturk, N., and Heath, L. S. 2002. Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. *J. Exp. Bot.* 53:1331-13441.
- Balestrini, R., Gomez-Ariza, J., Lanfranco, L., and Bonfante, P. 2007. Laser microdissection reveals that transcripts for five plant and one fungal phosphate transporter genes are contemporaneously present in arbusculated cells. *Mol. Plant-Microbe Interact.* 20:1055-1062.
- Balestrini, R., Gomez-Ariza, J., Klink, V. P., and Bonfante, P. 2009. Application of laser microdissection to plant pathogenic and symbiotic interactions. *J. Plant Interact.* 4:81-92.
- Barrès, B., Halkett, F., Dutech, C., Andrieux, A., Pinon, J., and Frey, P. 2008. Genetic structure of the poplar rust fungus *Melampsora larici-populina*: Evidence for isolation by distance in Europe and recent founder effects overseas. *Infect. Genet. Evol.* 8:577-587.
- Boyle, B., Levée, V., Hamel, L. P., Nicole M. P., and Séguin, A. 2010. Molecular and histochemical characterisation of two distinct poplar *Melampsora* leaf rust pathosystems. *Plant Biol.* 12:364-376.
- Broeker, K., Bernard, F., and Moerschbacher, B. M. 2006. An EST library from *Puccinia graminis* f. sp. *tritici* reveals genes potentially involved in fungal differentiation. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett.* 256:273-281.
- Brooks, L., Strable, J., Zhang, X., Ohtsu, K., Zhou, R., Sarkar, A., Hargreaves, S., Elshire, R. J., Eudy, D., Pawlowska, T., Ware, D., Janick-Buckner, D., Buckner, B., Timmermans, M. C., Schnable, P. S., Nettleton, D., and Scanlon, M. J. 2009. Microdissection of shoot meristem functional domains. *PLoS Genet.* 5:e1000476. Published online.
- Cai, S., and Lashbrook, C. C. 2006. Laser capture microdissection of plant cells from tape-transferred paraffin sections promotes recovery of structurally intact RNA for global gene profiling. *Plant J.* 48:628-637.
- Cairney, J. W. G., and Burke, R. M. 1996. Physiological heterogeneity within fungal mycelia: An important concept for a functional understanding of the ectomycorrhizal symbiosis. *New Phytol.* 134:685-695.
- Casson, S. A., Spencer, M. W., and Lindsey, K. 2008. Laser-capture microdissection to study global transcriptional changes during plant embryogenesis. *Methods Mol. Biol.* 427:111-120.
- Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A., and Ellis, J. G. 2006. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18:243-256.
- Catanzariti, A. M., Dodds, P. N., Ve, T., Kobe, B., Ellis, J. G., and Staskawicz, B. J. 2010. The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein. *Mol. Plant-Microbe Interact.* 23:49-57.
- Chandran, D., Inada, N., Hather, G., Kleindt, C. K., and Wildermuth, M. C. 2010. Laser microdissection of *Arabidopsis* cells at the powdery mildew infection site reveals site-specific processes and regulators. *Proc. Natl. Acad. Sci. U.S.A.* 107:460-465.
- Copois, V., Bibeau, F., Bascou-Mollevi, C., Salvétat, N., Chalbos, P., Baireil, C., Candeil, L., Fraslou, C., Conseiller, E., Granci, V., Mazière, P., Kramar, A., Ychou, M., Pau, B., Martineau, P., Molina, F., and Del Rio, M. 2007. Impact of RNA degradation on gene expression profiles: Assessment of different methods to reliably determine RNA quality. *J. Biotechnol.* 127:549-559.
- Diamond, A. E., Park, J. S., Inoue, I., Tachikawa, H., and Neiman, A. M. 2009. The anaphase promoting complex targeting subunit Ama1 links meiotic exit to cytokinesis during sporulation in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 20:134-145.
- Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Ayliffe, M. A., and Ellis, J. G. 2004. The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* 16:755-768.
- Duplessis, S., Major, I. T., Martin, F., and Séguin, A. 2009. Poplar and pathogen interactions: Insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Crit. Rev. Plant. Sci.* 28:309-334.
- Ellis, J. G., Dodds, P. N., and Lawrence, G. J. 2007. The role of secreted proteins in diseases of plants caused by rust, powdery mildew and smut fungi. *Curr. Opin. Microbiol.* 10:326-331.
- Feau, N., Bergeron, M. J., Joly, D., Roussel, F., and Hamelin, R. 2007. Detection and validation of EST-derived SNPs for poplar leaf rust *Melampsora medusae* f. sp. *deltoidae*. *Mol. Ecol. Notes* 7:1222-1228.
- Gérard, P. R., Husson, C., Pinon, J., and Frey, P. 2006. Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathology* 96:1027-1036.
- Gomez, S. K., and Harrison, M. J. 2009. Laser microdissection and its application to analyze gene expression in arbuscular mycorrhizal symbiosis. *Pest. Manage. Sci.* 65:504-511.
- Hahn, M., and Mendgen, K. 1997. Characterization of in planta-induced rust genes isolated from a haustorium-specific cDNA library. *Mol. Plant-Microbe Interact.* 10:427-437.
- Hahn, M., Neef, U., Struck, C., Göttfert, M., and Mendgen, K. 1997. A putative amino acid transporter is specifically expressed in haustoria of the rust fungus *Uromyces fabae*. *Mol. Plant-Microbe Interact.* 10:438-445.
- Heath, M. C. 1997. Signalling between pathogenic rust fungi and resistant or susceptible host plants. *Ann. Bot.* 80:713-720.
- Inada, N., and Wildermuth, M. C. 2005. Novel tissue preparation method and cell-specific marker for laser microdissection of *Arabidopsis* mature leaf. *Planta* 221:9-16.
- Ithal, N., Recknor, J., Nettleton, D., Maier, T., Baum, T. J., and Mitchum, M. G. 2007. Developmental transcript profiling of cyst nematode feeding cells in soybean roots. *Mol. Plant-Microbe Interact.* 20:510-525.
- Ivashikina, N., Deeken, R., Ache, P., Kranz, E., Pommerrenig, B., Sauer, N., and Hedrich, R. 2003. Isolation of AtSUC2 promoter-GFP-marked companion cells for patch-clamp studies and expression profiling. *Plant J.* 36:931-945.
- Jiang, K., Zhang, S., Lee, S., Tsai, G., Kim, K., Huang, H., Chilcott, C., Zhu, T., and Feldman, L. J. 2006. Transcription profile analyses identify genes and pathways central to root cap functions in maize. *Plant Mol. Biol.* 60:343-363.
- Kemen, E., Kemen, A. C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., and Voegelé, R. T. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol. Plant-Microbe Interact.* 18:1130-1139.
- Kerk, N. M., Ceserani, T., Tausta, S. L., Sussex, I. M., and Nelson, T. M. 2003. Laser capture microdissection of cells from plant tissues. *Plant Physiol.* 132:27-35.
- Klink, V. P., Overall, C. C., Alkharouf, N. W., MacDonald, M. H., and Matthews, B. F. 2007. A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection. *Planta* 226:1423-1447.
- Koh, S., André, A., Edwards, H., Ehrhardt, D., and Somerville, S. 2005. *Arabidopsis thaliana* subcellular responses to compatible *Erysiphe cichoracearum* infections. *Plant J.* 44:516-529.
- Nagarajan, S., and Singh, D. V. 1990. Long-distance dispersion of rust pathogens. *Annu. Rev. Plant Pathol.* 28:139-153.
- Nakada, M., Komatsu, M., Ochiai, K., Ohtsu, K., Nakazono, M., Nishizawa, N. K., Nitta, K., Nishiyama, R., Kameya, T., and Kanno, A. 2006. Isolation of MaDEF from *Muscari armeniacum* and analysis of its expression using laser microdissection. *Plant Sci.* 170:143-150.
- Nakazono, M., Qiu, F., Borsuk, L. A., and Schnable, P. S. 2003. Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: Identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* 15:583-596.
- Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45.
- Puthoff, D. P., Neelam, A., Ehrenfried, M. L., Scheffler, B. E., Ballard, L., Song, Q., Campbell, K. B., Cooper, B., and Tucker, M. L. 2008. Analy-

- sis of expressed sequence tags from *Uromyces appendiculatus* hyphae and haustoria and their comparison to sequences from other rust fungi. *Phytopathology* 98:1126-1135.
- Ramsay, K., Wang, Z., and Jones, G. K. 2004. Using laser capture microdissection to study gene expression in early stages of giant cells induced by root-knot nematodes. *Mol. Plant Pathol.* 5:587-592.
- Rinaldi, C., Kohler, A., Frey, P., Duchaussoy, F., Ningre, N., Couloux, A., Wincker, P., Le Thiec, D., Fluch, S., Martin, F., and Duplessis, S. 2007. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol.* 144:347-366.
- Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., and Mewes, H. W. 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* 32:5539-5545.
- Saha, P., Thome, K. C., Yamaguchi, R., Hou, Z., Weremowicz, S., and Dutta, A. J. 1998. The human homolog of *Saccharomyces cerevisiae* CDC45. *J. Biol. Chem.* 273:18205-18209.
- Sohn, J., Voegelé, R. T., Mendgen, K., and Hahn, M. 2000. High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*. *Mol. Plant-Microbe Interact.* 13:629-636.
- Tang, W., Coughlan, S., Crane, E., Beatty, M., and Duvick, J. 2006. The application of laser microdissection to *in planta* gene expression profiling of the maize anthracnose stalk rot fungus *Colletotrichum graminicola*. *Mol. Plant-Microbe Interact.* 19:1240-1250.
- Thara, V. K., Fellers, J. P., and Zhou, J. M. 2003. In planta induced genes of *Puccinia triticina*. *Mol. Plant Pathol.* 4:51-56.
- Tremblay, A., Li, S., Scheffler, B. E., and Matthews, B. F. 2009. Laser capture microdissection and expressed sequence tag analysis of uredinia formed by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Physiol. Mol. Plant Pathol.* 6:163-174.
- Voegelé, R. T., Struck, C., Hahn, M., and Mendgen, K. 2001. The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proc. Natl. Acad. Sci. U.S.A.* 98:8133-8138.
- Voegelé, R. T., Wirsal, S., Möll, U., Lechner, M., and Mendgen, K. 2006. Cloning and characterization of a novel invertase from the obligate biotroph *Uromyces fabae* and analysis of expression patterns of host and pathogen invertases in the course of infection. *Mol. Plant-Microbe Interact.* 19:625-634.
- Voegelé, R. T., Hahn, M., and Mendgen, K. 2009. The Uredinales: Cytology, Biochemistry, and Molecular Biology. Pages 69-98 in: *The Mycota, 5. Plant Relationships Volume*. H. B. Deising, ed. Springer, Berlin.
- Walther, A., Wendland, J. 2003. Septation and cytokinesis in fungi. *Fungal Genet. Biol.* 40:187-196.
- Wolfe, B. A., and Gould, K. L. 2005. Split decisions: Coordinating cytokinesis in yeast. *Trends Cell Biol.* 15:10-18.
- Wu, H., Rossi, G., and Brennwald, P. 2008. The ghost in the machine: Small GTPases as spatial regulators of exocytosis. *Trends Cell Biol.* 18:397-404.
- Yin, C., Chen, X., Wang, X., Han, Q., Kang, Z., and Hulbert, S. H. 2009. Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici*. *BMC Genomics* 10:626.
- Yu, Y., Lashbrook, C. C., and Hannapel, D. J. 2007. Tissue integrity and RNA quality of laser microdissected phloem of potato. *Planta* 226:797-803.

AUTHOR-RECOMMENDED INTERNET RESOURCES

- Broad Institute Puccinia Group database:
www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.html
- Munich Information Center for Protein sequence (MIPS) functional catalogue: mips.helmholtz-muenchen.de/proj/funcatDB
- National Center for Biotechnology Information Gene Expression Omnibus database: www.ncbi.nlm.nih.gov/geo
- NimbleGen website: www.nimblegen.com
- U.S. Department of Energy JGI genome database queries:
genome.jgi-psf.org/Mellp1/Mellp1.home.html and
genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html

Chapitre V : Discussion générale

Chapitre V : Discussion générale

Un vaste projet de recherche visant à identifier des marqueurs moléculaires de la virulence et les fonctions impliquées dans le mode de vie biotrophe du champignon, en cours depuis 2002, est développé dans l'équipe Ecogénomique des Interactions de l'UMR1136. L'exploitation des données génomiques disponibles suite au séquençage du génome de *M. larici-populina* (Souche 98AG31) a ouvert de nouvelles perspectives dans la compréhension des mécanismes moléculaires régissant l'interaction entre ce champignon et son hôte. Mon travail de thèse s'est inséré dans ce projet et a consisté dans un premier temps à améliorer les connaissances sur la biologie du champignon en décrivant sa cinétique de développement, ainsi que les structures d'infection formées *in planta* au cours de l'interaction compatible avec le peuplier (*P. trichocarpa* x *P. deltoides* cv. Beaupré). Dans un deuxième temps, différentes méthodes d'analyse de l'expression génique reposant sur les puces à ADN, le séquençage d'EST et la RT-qPCR, m'ont permis de contribuer à l'étude fine du processus infectieux et de la formation de l'urédie. Un troisième volet de mon projet de thèse, réalisé en collaboration avec le groupe de R. Hamelin (CFC, Québec), a consisté en l'analyse détaillée du sécrétome de *M. larici-populina*. Une attention particulière a été portée à l'annotation, l'expression et l'évolution des gènes codant des SSPs afin d'identifier des effecteurs potentiels du champignon. J'ai notamment utilisé la technique d'immunolocalisation afin de vérifier la localisation de l'expression *in planta* de certaines protéines effectrices candidates sécrétées par le champignon.

I - Analyse du développement de *Melampsora larici-populina in planta*

Afin de détecter et/ou quantifier la présence d'un pathogène *in planta*, la PCR quantitative représente un outil de choix du fait de sa sensibilité et de sa facilité d'utilisation (Schena et al. 2004; Lievens et al. 2006). Cette technique a notamment été utilisée pour amplifier la région ITS de l'ADN ribosomique de *Melampsora* spp. au cours de l'infection de feuilles de peuplier (Boyle et al. 2005; Rinaldi et al. 2007). Par cette même approche, nous avons pu suivre l'accumulation du champignon au cours des interactions compatible et incompatible entre la souche 98AG31 de *M. larici-populina* et les cultivars de peuplier « Beaupré » et « Hoogvorst » à de fortes doses d'inoculum (100,000 spores/ml). Nous avons ainsi pu observer un décrochement de la courbe de croissance entre l'interaction compatible et l'interaction incompatible entre 12 et 24 hpi (Hacquard et al., en préparation pour PLoS Pathogens, chapitre III, publication n°4). De manière intéressante, ce point correspond exactement à la formation des premiers haustoria observés dans le cadre de l'interaction entre *M. larici-populina* et *P. deltoides* x *P. nigra* « Ogy » (Laurans et

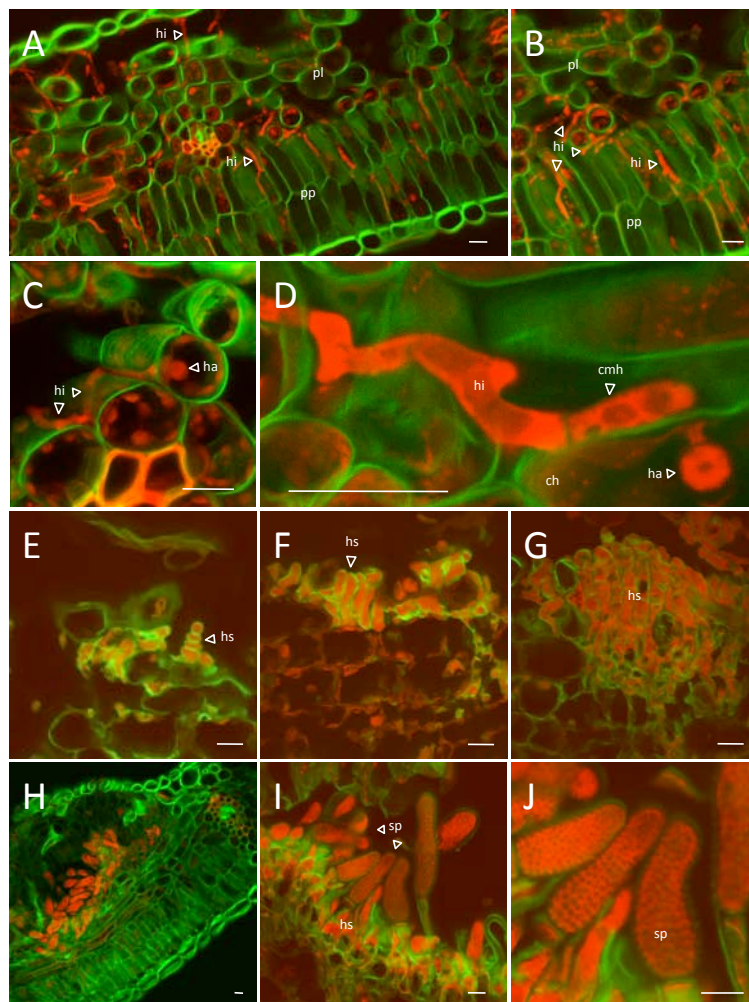


Figure 27: Structures d'infection formées *in planta* par *Melampsora larici-populina* au cours de l'interaction compatible avec *Populus trichocarpa* x *Populus deltoides* « Beupré ». **A,B** Hyphe d'infection colonisant le parenchyme lacuneux et palissadique à des stades tardifs d'infection (168 heures post infection, hpi). **C,D** Haustoria observés au sein de cellules de peuplier entre 96 et 168 hpi. **E-G** Différenciation des premières cellules formant le futur tissu sporifère observées à 96 hpi. **H-J** Urédie mature observée à 168 hpi et à partir de laquelle les urédiniospores dicaryotiques sont libérées. hi, hyphe d'infection; pl, parenchyme lacuneux; pp, parenchyme palissadique; cmh, cellule mère haustoriale; ha, haustoria; hs, hyphe sporifère; sp, urédiniospore. Echelle = 5 μ m .

Pilate 1999), suggérant le rôle clef de l'haustorium dans l'établissement de la biotrophie. Cette structure d'infection caractéristique des agents de rouilles, impliquée dans le détournement des métabolites de la plante et la sécrétion de molécules signal (Dodds et al. 2009) est difficilement observable du fait de sa localisation intracellulaire (Kuck et al. 1981). Des études cytologiques ont néanmoins permis l'identification de ces structures formées *in planta* chez certains agents de rouilles tels que *U. fabae*, *M. lini*, *M. larici-populina*, *H. vastatrix* et *P. graminis* (Kuck et al. 1981; Laurans et Pilate 1999; Ramiro et al. 2009; Rafiqi et al. 2010). Chez *M. larici-populina*, la formation des haustoria au sein des tissus foliaires infectés a été mise en évidence par microscopie à transmission (Laurans et Pilate 1999). Récemment, une description détaillée du développement de *M. larici-populina* durant la colonisation des feuilles du cultivar « Beaupré » a été rendue possible par l'utilisation de la microscopie électronique à balayage mais n'a pas permis la détection de ces structures *in planta* (Rinaldi et al. 2007). Afin de visualiser ces structures discrètes, nous avons utilisé 2 fluorochromes utilisés en microscopie confocale et présentant des longueurs d'ondes d'excitation et d'émission distinctes (Hacquard et al., en préparation pour PLoS Pathogens, chapitre III, publication n°4; Hacquard et al. 2010, chapitre IV, publication n°6). L'iodure de propidium a coloré les structures fongiques (Koh et al. 2005), alors que l'Uvitex 2B a marqué les parois des cellules végétales et fongiques permettant ainsi la discrimination entre les structures de l'hôte et du pathogène (Felten et al. 2009). Cette méthode a ainsi permis de visualiser différentes structures fongiques formées par *M. larici-populina in planta* comme les hyphes d'infection, les cellules mères haustoriales, les haustoria, le tissu sporifère ou encore les urédospores nouvellement formées (**Figure 27**), complétant ainsi les connaissances sur les structures formées par les agents de rouilles décrites jusqu'à présent. Nous avons ainsi pu observer la présence plus importante d'haustoria à 96 et 168 hpi que dans les étapes plus précoces, établissant une bonne corrélation avec les courbes de croissance mesurées par qPCR lors de l'amplification de l'ITS fongique (Hacquard et al. 2011, chapitre II, publication n°3). De manière intéressante, nous avons également pu visualiser les premières étapes de la formation de la future urédie dès 96 hpi au sein du parenchyme lacuneux en détectant les cellules pionnières du tissu sporifère. La présence de ces cellules, bien que très minoritaires en comparaison avec les hyphes d'infection et les haustoria suggère que le déterminisme de formation des urédies est engagé précocement lors de l'interaction avec le peuplier. Cette différenciation rapide de l'urédie semble commune à de nombreux agents de rouilles tels que *U. fabae*, *M. lini*, *P. graminis* ou encore *P. pachyrhizi* chez qui la production d'urédospores a lieu dès 6 à 8 jours post infection (Leonard et Szabo 2005; Lawrence et al. 2007; Tremblay et al. 2009; Voegelé et al. 2009).

II- Analyse du génome de *Melampsora larici-populina*

Le séquençage du génome de *M. larici-populina* a été réalisé au JGI dans le cadre d'un projet visant à étudier les génomes de plusieurs champignons biotrophes interagissant avec le peuplier (Martin et al. 2004). Qu'ils soient symbiotiques, tels que *L. bicolor* et *Glomus intradices* (Martin et al. 2008a, Martin et al. 2008b), ou pathogènes tel que *M. larici-populina* (Duplessis et al. 2009), l'analyse de ces génomes offre l'opportunité d'identifier les déterminants génétiques et les mécanismes moléculaires régissant les interactions entre le peuplier et divers microorganismes qui lui sont associés. L'analyse comparative de ces génomes ouvre également de nouvelles perspectives dans la compréhension de l'évolution de la symbiose ou du caractère pathogène de ces champignons et plus particulièrement dans l'identification des déterminants moléculaires leur permettant de se développer dans les tissus de l'arbre sans être détectés ou combattus. De plus, le séquençage du génome de *M. larici-populina* érige le pathosystème *M. larici-populina*/peuplier au rang de modèle pour l'étude des interactions arbre/pathogène voir plante/pathogène en général puisqu'il fait partie des rares pathosystèmes dont les génomes de l'hôte et du pathogène ont été séquencés (Medina and Sachs 2010).

La plupart des connaissances acquises chez les champignons phytopathogènes ont été obtenues chez les champignons biotrophes non obligatoires tels que *C. fulvum*, (Thomma et al. 2005), *M. oryzae* (Wilson et Talbot 2009), *U. maydis* (Brefort et al. 2009) ou encore *Mycosphaerella graminicola* (Palmer et Skinner 2002). De manière contrastée, les connaissances acquises chez les organismes biotrophes obligatoires restent limitées. Un point non élucidé jusqu'à présent est le caractère obligatoire de leur biotrophie. L'une des hypothèses qui peut être avancée est la perte de certaines voies métaboliques ou de fonctions non essentielles à leur mode de développement et la compensation de cette absence par le détournement de ressources de la plante (Both et al. 2005b). Alors que l'analyse préliminaire des voies de biosynthèses des acides aminés menée chez *B. graminis* ne démontra pas de différences majeures avec les champignons non obligatoires (Giles et al. 2003), l'analyse du génome du champignon a bien révélé une signature de biotrophie (Spanu et al. 2010).

L'analyse des 100 megabases de la séquence génomique de *M. larici-populina* (codante pour 14,527 protéines prédites) a également révélé des signatures caractéristiques reflétant le caractère biotrophe du champignon (Duplessis et al., soumis à PNAS, chapitre II, publication n°1). L'expansion de nombreuses familles multigéniques codant des transporteurs d'acides aminés, d'oligopeptides ou de sucres a ainsi pu être révélée parallèlement à un profil particulier en protéases et notamment l'expansion de familles qu'il serait tentant de mettre en relation avec

celles des oligopeptides; ces 2 classes de gènes étant exprimés *in planta*. De manière intéressante, certaines fonctions de transport ou métaboliques clefs qui avaient été décrites chez l'agent responsable de la rouille de la féverole *U. fabae* (Voegelé et al. 2009) sont ici retrouvées soit en expansion dans le génome de *M. larici-populina*; soit de par leur expression lors de l'infection (voir plus bas). Par ailleurs, les familles codant les enzymes de dégradation des sucres (GH, glycosyl hydrolases, CAZymes) présentent un profil singulier par rapport aux pathogènes héli-biotrophes et necrotrophes avec des contractions ou expansions spécifiques pouvant refléter le mode trophique du champignon. Un résultat marquant de cette analyse génomique concerne l'absence de nitrite réductase au sein du génome de *M. larici-populina*. Cette enzyme clé du cycle de l'azote catalysant la réduction des nitrites en ammonium (au pH cellulaire NH_3 et NH_4^+), forme azotée indispensable à la formation des acides aminés (Marzluf 1997). De manière intéressante, cette déficience est commune à *P. graminis* f. sp. *tritici* (données non publiées), *B. graminis* (Spanu et al. 2010) et *H. arabidopsis* (Baxter et al. 2010) et suggère que, de par leur spécialisation pour la croissance chez un hôte végétal, les agents biotrophes obligatoires auraient perdu la capacité à assimiler les nitrates pour la production d'ammonium.

III - Analyse du transcriptome de *Melampsora larici-populina*

1 - La biotrophie, principale limitation à l'étude de l'expression des gènes de *Melampsora larici-populina*

Les agents de rouilles sont des organismes biotrophes obligatoires dont la croissance et la reproduction sont totalement dépendantes des tissus de la plante hôte (Voegelé et al. 2009). Ceci est la principale limitation à l'étude de l'expression des gènes durant le processus infectieux du champignon. Cette limitation est d'autant plus critique aux premiers stades de colonisation de la plante hôte lorsque la quantité de tissu fongique est extrêmement réduite par rapport à ceux de la plante. Ce problème de représentation de l'hôte versus pathogène rend notamment délicate l'utilisation des méthodes telles que les puces à ADN à des temps précoces d'infection. Jusqu'à présent, le séquençage d'ESTs issus d'haustoria purifiés ou de feuilles infectées à des stades tardifs s'est imposé comme la méthode de choix pour identifier des transcrits fortement accumulés durant l'infection. Cette méthode a conduit à l'identification de nombreux gènes exprimés *in planta* chez plusieurs agents de rouilles tels que *M. lini*, *U. fabae*, *U. appendiculatus*, *M. larici-populina*, *P. triticina* et *P. striiformis* (Hahn and Mendgen 1997a; Catanzariti et al. 2006; Feau et al. 2007; Puthoff et al. 2008; Yin et al. 2009). Des méthodes d'enrichissement en transcrits fongiques par la technique d'hybridation soustractive et suppressive (SSH) ont

également été employées dans le cadre des interactions *P. triticina*/blé et *P. graminis* f. sp. *tritici*/blé (Thara et al. 2003; Broeker et al. 2006).

Dans le cadre de notre analyse de l'infection des tissus foliaires de peuplier par *M. larici-populina*, nous avons mis en place plusieurs approches pour étudier les profils d'expression géniques du champignon à la fois aux stades précoces et tardifs de la cinétique d'infection. Outre le séquençage d'ESTs de feuilles infectées (96 et 168 hpi) et de spores (germées et non germées), nous avons entrepris de couvrir 6 étapes de la colonisation du champignon; 2 représentatives des stades *ex planta* (spores germées et non germées) et 4 représentatives des stades *in planta* depuis la formation des premiers haustoria à la formation des urédies (24, 48, 96 et 168 hpi) par l'utilisation de puces à ADN (Duplessis et al., en préparation pour MPMI, chapitre II, publication n°2). Par cette technique nous avons également entrepris l'analyse exhaustive de l'expression des transcrits de *M. larici-populina* au sein de différentes zones microdisséquées au stade de formation de l'urédie (Hacquard et al. 2010, chapitre IV, publication n°6). Alors que les puces à ADN ont déjà été utilisées pour étudier le niveau d'expression des transcrits chez plusieurs organismes phytopathogènes (Both et al. 2005a; Both et al. 2005b; Kämper et al. 2006; Oh et al. 2008) ou symbiotiques (Duplessis et al. 2005; Martin et al. 2008a; Martin et al. 2010) durant la colonisation des tissus de l'hôte, elle l'ont été de manière anecdotique chez les agents de rouilles et sur un nombre limité de gènes (Jakupović et al. 2005). Du fait de l'accès au génome de *M. larici-populina* et plus particulièrement au 14,527 modèles de gènes prédits par les annotations automatique et manuelle, nous avons entrepris le suivi de l'expression des gènes aux 6 stades choisis pour la quasi-totalité des gènes (13,093 sur 14,527). Les 1,434 gènes manquants ne possédaient pas d'oligonucléotides spécifiques définis pour détecter leur expression (294) ou au contraire ont été éliminés car ils présentaient une trop forte similarité de séquence avec un autre gène parlogue du champignon (859). D'autres ont également été éliminés car ils présentaient une hybridation croisée avec les transcrits de la plante (281). Néanmoins, l'analyse des étapes très précoces d'infection (2, 6 et 12 hpi) n'a pas abouti du fait de la trop faible quantité de champignon dans les tissus foliaires à ces stades. En effet, moins de 500 gènes sont identifiés comme étant exprimés au-dessus du bruit de fond et la plupart correspondent à des protéines ribosomiques.

Afin d'identifier des transcrits accumulés dans les étapes précoces d'infection (2, 6, 12 hpi), nous avons utilisé la technique de RT-PCR quantitative dont la sensibilité de détection des transcrits est meilleure que celle des puces à ADN. Cette technique, bien que très utilisée pour étudier le niveau des transcrits de l'hôte en réponse à son pathogène (McMaugh et Lyon 2003; Rinaldi et al. 2007; Desmond et al. 2008) reste encore assez peu utilisée pour l'étude des profils

d'expression d'organismes phytopathogènes au cours du processus infectieux (Avrova et al. 2003; Kebdani et al. 2010). Chez les agents de rouilles, les techniques de RT-PCR semi-quantitative (Wirsel et al. 2001) et de Northern blot (Hahn et Mendgen 1997a; Sohn et al. 2000; Voegelé et al. 2001; Kemen et al. 2005) ont été utilisées pour analyser le niveau d'expression de gènes candidats. La RT-qPCR est une approche de choix permettant une analyse plus fine de l'évolution des quantités de transcrits ainsi qu'une meilleure détection de ceux-ci durant l'interaction, particulièrement lorsque la biomasse du pathogène au sein des tissus foliaires est faible. Toutefois, afin de pouvoir comparer les niveaux d'expression d'un même gène au cours d'une cinétique d'infection, il est primordial de disposer de gènes de références dont l'expression ne varie pas lors de l'infection des tissus de l'hôte et dont les transcrits sont facilement détectables même aux étapes précoces d'infection. De tels gènes doivent présenter des profils d'expression corrélés à l'augmentation de la biomasse fongique au cours de l'infection. Pour analyser les niveaux d'expression de *M. larici-populina* par RT-qPCR, nous avons d'abord cherché à identifier des gènes de référence dont les profils d'expression étaient corrélés avec la courbe de croissance du champignon (Hacquard et al. 2011, chapitre II, publication n°3). Nous avons sélectionné 5 gènes de référence typiquement utilisés chez les champignons comme gènes de référence et nous avons testé leur expression *in planta* en les comparant systématiquement aux courbes d'amplification de l'ITS fongique. Les gènes codant l'alpha-tubuline et le facteur d'élongation-1-alpha présentaient des profils adéquats de gènes de référence et nous ont permis de normaliser les niveaux d'expression de gènes candidats (codants des SSPs ou d'autres fonctions) mesurés par RT-qPCR au cours du processus infectieux du champignon.

2 - Identification de gènes induits durant le processus infectieux

Le séquençage d'ESTs ainsi que les puces à ADN ont permis l'étude du transcriptome de plusieurs oomycètes (Cheung et al. 2008; Gaulin et al. 2008), champignons pathogènes (Jakupović et al. 2005; Skibbe et al. 2010; Both et al. 2005a; Oh et al. 2008) et symbiotiques (Duplessis et al. 2005; Martin et al. 2008a; Martin et al. 2010) durant la colonisation des tissus de leurs hôtes. La plupart de ces études mettent en avant l'induction de transcrits du pathogène ou du symbiote au cours de l'interaction avec leurs hôtes pour certains stades de développement particuliers. Néanmoins, l'analyse globale des profils d'expression génique au cours de différentes phases du processus de colonisation n'a été rapportée que dans le cadre de quelques interactions. Ainsi, l'identification de gènes impliqués dans les différentes étapes de la formation de l'appressorium chez *M. oryzae* (Oh et al. 2008), la détection de gènes marqueurs des différents stades de formation de l'ectomycorrhize dans le cadre de l'interaction *Pisolithus*

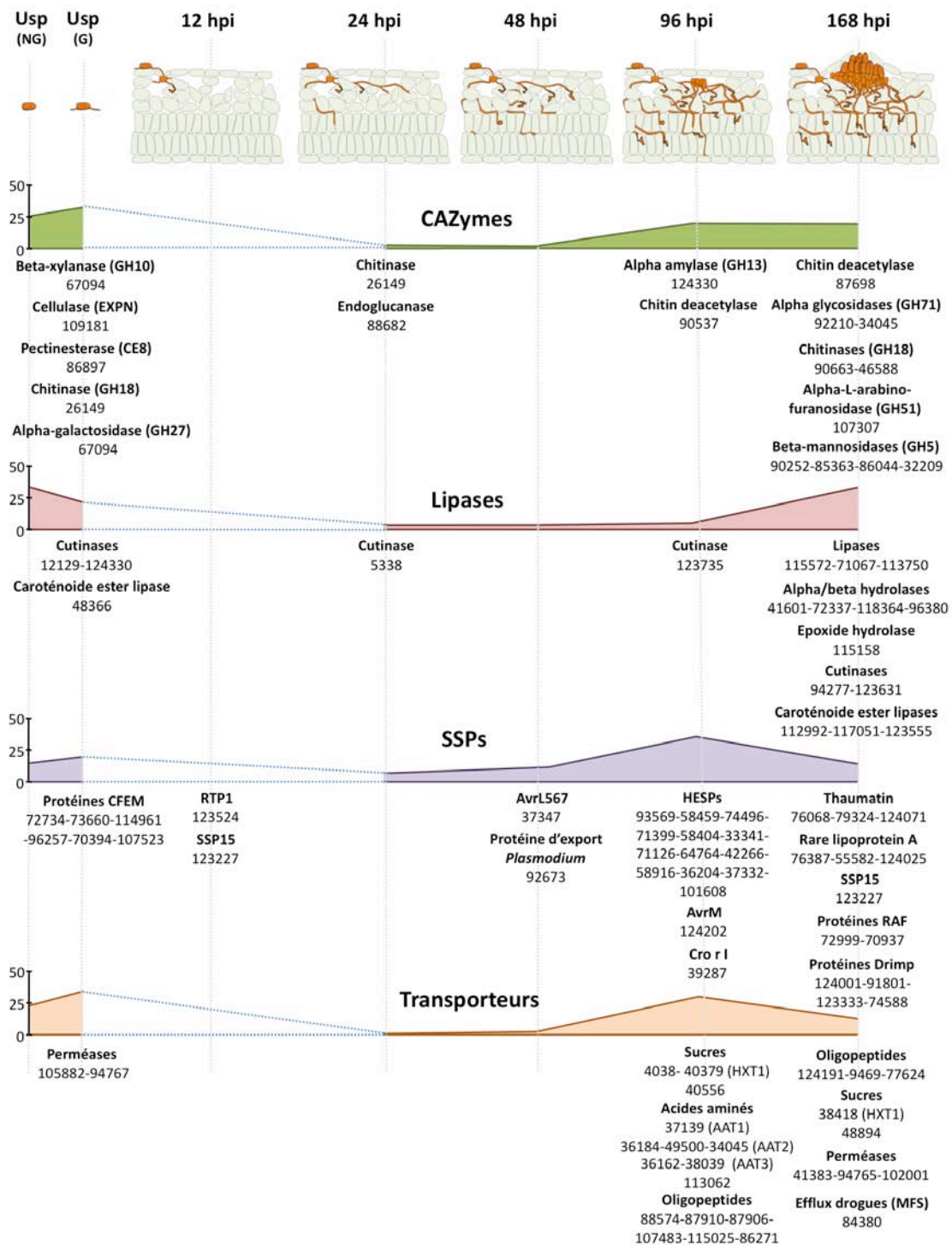


Figure 28: Représentation des pics d'expression (exprimés en pourcentage) de l'ensemble des transcrits codant pour des CAZymes, des lipases, des petites protéines sécrétées (SSPs) et des transporteurs mesurés par les puces à oligonucléotides NimbleGen au cours de la cinétique d'infection de *Melampsora larici-populina* (urédiniospores non germées, urédiniospores germées, 24, 48, 96, 168 heures post infection). Pour chaque stade de développement, certains gènes marqueurs fortement accumulés et/ou induits à chacun des stades sont indiqués ainsi que leurs identifiants de séquence. Les gènes présentant un pic d'expression à 12 hpi par RT-qPCR sont également indiqués. na: données non disponibles.

microcarpus/Eucalyptus (Duplessis et al. 2005) ou encore le décryptage des processus activés durant les différentes phases du processus infectieux de *B. graminis* sur orge (Both et al. 2005a) ont été rendus possibles par l'utilisation de puces à ADN.

Chez *M. larici-populina*, ces puces ont également permis l'analyse du niveau d'expression de l'ensemble des gènes du champignon durant son processus infectieux sur feuilles de peuplier (Duplessis et al., en préparation pour MPMI, chapitre II, publication n°2). Cette analyse a permis de révéler des gènes marqueurs de chaque phase de développement du champignon impliqués dans la germination de la spore, la formation des premiers haustoria (24 hpi), la prolifération du champignon *in planta* (48 et 96 hpi) ou encore la formation de l'urédie (168 hpi). Au cours de cette étude nous nous sommes notamment focalisés sur l'ensemble des enzymes impliquées dans la dégradation des polysaccharides (GH, CAZymes) et des lipides (lipases) ainsi que sur les SSPs et les transporteurs du champignon. L'analyse des pics d'accumulation des transcrits pour ces différentes catégories cellulaires a permis de révéler des profils contrastés (**Figure 28**). La majorité des cazymes et lipases présentent des pics d'expression dans les spores, les spores germées ainsi que lors de la formation de l'urédie suggérant leur rôle préférentiel lors de ces étapes de l'infection. L'analyse transcriptomique du développement de *B. graminis* au cours de l'interaction avec l'orge a également révélé l'expression massive de nombreuses lipases dans les conidies ainsi que dans les stades précoces de germination suivie par une diminution importante de leurs niveaux d'expression durant les stades plus tardifs d'infection (Both et al. 2005a). Chez *M. oryzae*, le métabolisme lipidique est également fortement stimulé lors de la formation de l'appressorium (Oh et al. 2008) et l'augmentation de l'activité lipase a également été détectée au niveau de cette structure entraînant l'accumulation de glycérol et l'augmentation de la pression de turgescence au sein de cette structure (Thines et al. 2000). Au cours de cette étude, nous avons également observé l'induction de nombreux gènes codant des chitinases, cellulases, xylanases ou pectinesterase au niveau des spores germées qui pourraient être impliquées dans la croissance du tube germinatif ou/et dans le remodelage des parois végétales pour faciliter la pénétration du champignon *in planta*.

Afin d'identifier des gènes fortement induits durant la phase de croissance biotrophe du champignon, nous avons entrepris la comparaison entre le niveau d'expression des transcrits mesuré dans les spores avec celui mesuré à 96 hpi (Duplessis et al., en préparation pour MPMI, chapitre II, publication n°1). De plus la microdissection laser de feuilles infectées à 168 hpi a également permis de mettre à jour des transcrits induits dans l'aire de biotrophie par rapport à l'aire de sporulation pouvant correspondre à des déterminants clés du processus infectieux (Hacquard et al. 2010, chapitre IV, publication n°6). La comparaison entre ces deux approches

Protein ID	Function	Best blast hit Puccinia ID	NR ID	NimbleGen oligoarrays			EST sequencing		RT-qPCR	
				96 hpi/Usp FC	P-value	area3/area1 FC	P-value	Usp		96-168 hpi
Small Secreted Proteins										
123532	small secreted protein			1037.7	3.68 E-05	ns	ns	nd	nd	Y
124305	small secreted protein			918.4	8.20 E-05	190.9	0.021	nd	nd	Y
59518	small secreted protein			913.7	8.20 E-05	ns	0.096	nd	nd	na
123443	small secreted protein			837.5	9.52 E-05	34.1	0.042	nd	Y	na
106150	small secreted protein			831.3	1.94 E-04	262.1	0.012	nd	Y	na
63414	small secreted protein			823.4	8.15 E-05	15.0	0.011	nd	Y	na
68912	small secreted protein			820.3	2.19 E-04	89.7	0.014	nd	Y	na
123530	small secreted protein			806.7	4.91 E-05	ns	ns	nd	Y	na
124017	small secreted protein			798.7	4.78 E-05	74.1	0.018	nd	Y	na
105045	small secreted protein			779.6	1.74 E-04	26.2	0.034	nd	Y	na
39287	small secreted protein, Cronartium Cro r I homolog	AAF87492		263.9	0.026	ns	ns	nd	Y	na
64764	small secreted protein, M. lini hesp-376 homolog			217.1	1.26 E-03	90.0	0.012	nd	nd	na
101611	small secreted protein, M. lini hesp-C63 homolog	ABB96282		159.1	7.55 E-05	85.7	0.01	nd	nd	na
124202	secreted protein, M. lini AvrM-B homolog	ABB96259		139.5	4.12 E-04	ns	ns	nd	nd	na
101608	small secreted protein, M. lini hesp-C63 homolog			68.5	2.13 E-04	ns	ns	nd	nd	na
55212	small secreted protein, M. lini hesp-735 homolog	ABB96276		67.4	5.20 E-04	ns	ns	nd	nd	na
37347	small secreted protein, M. lini AvrL567-F homolog	ABF56218		49.0	2.45 E-04	ns	ns	nd	nd	Y
103805	secreted protein, M. lini hesp-270 homolog	ABB96269		47.7	1.16 E-04	41.5	0.01	nd	Y	na
124206	secreted protein, M. lini AvrM-B homolog	ABB96259		10.7	1.44 E-03	ns	ns	nd	nd	na
Transporters										
88574	oligopeptide transporter, OPT	PGTG_17016	XP_001394363	440.1	1.40 E-04	ns	ns	Y	Y	na
33238	transporter, MFS (Major Facilitator Superfamily)	PGTG_15200	XP_001833110	146.8	1.45 E-03	ns	ns	nd	Y	na
32692	Anion-Cation Symporter, MFS (Major Facilitator Superfamily)	PGTG_10920	XP_002560051	58.2	1.70 E-04	ns	ns	nd	Y	na
38418	sugar transporter HXT1, MFS (Major Facilitator Superfamily)	PGTG_15147	CAC41332	53.6	2.83 E-04	31.1	0.014	nd	Y	na
94985	transporter, MFS (Major Facilitator Superfamily)	PGTG_08573	XP_001885160	28.3	4.49 E-03	ns	ns	nd	Y	na
92950	Anion-Cation Symporter, MFS (Major Facilitator Superfamily)	PGTG_12524	XP_001831769	26.7	5.59 E-04	ns	ns	nd	nd	na
36162	amino acid transporter AAT3, YAT (Yeast Amino Acid Transporter)	PGTG_15547	XP_001873273	21.5	3.62 E-03	ns	ns	nd	nd	na
115025	oligopeptide transporter (OPT)	PGTG_17016	EEU39757	10.9	4.36 E-03	ns	ns	Y	nd	na
34045	amino acid transporter AAT2, YAT (Yeast Amino Acid Transporter)	PGTG_10749	AAB39866	10.9	5.74 E-03	12.1	0.039	nd	nd	na
86448	transporter, AEC (Auxin Efflux Carrier) family	PGTG_06747	XP_759229	428.2	1.33 E-04	ns	ns	nd	Y	na
63226	transporter, AEC (Auxin Efflux Carrier) family	PGTG_06747	XP_759229	235.9	7.94 E-04	ns	ns	nd	nd	na
Lipid metabolism										
110949	lipase, secreted	PGTG_15782	XP_001486627	713.8	1.26 E-04	44.0	0.02	nd	Y	na
48366	carotenoid ester lipase, secreted	PGTG_13346	XP_001875752	123.1	1.26 E-03	ns	ns	Y	Y	na
94890	lipase, secreted	PGTG_15782	XP_001273241	108.1	1.32 E-03	3.9	0.044	nd	Y	na
85985	lysophospholipase	PGTG_09202	XP_384051	125.9	1.65 E-04	ns	ns	nd	nd	na
87140	geranylgeranyl pyrophosphate synthase/Polyprenyl synthetase	PGTG_19831	XP_001889705	64.1	2.79 E-04	16.0	0.046	nd	Y	na
Redox metabolism										
106552	peroxidase	PGTG_12961	XP_001840251	368.4	7.59 E-05	83.0	0.015	nd	Y	na
63255	oxidoreductase	PGTG_15129	XP_571921	233.2	1.41 E-04	20.8	0.031	nd	Y	na
93258	oxidoreductase	PGTG_19479	XP_001829618	133.4	5.20 E-04	9.7	0.047	nd	nd	na
34897	catalase	PGTG_10338	XP_002475416	73.4	2.78 E-03	ns	ns	nd	nd	na
38712	catalase	PGTG_01221	XP_002475416	41.3	3.46 E-04	ns	ns	nd	Y	na
Cazymes										
123735	cutinase, secreted, carboxylesterase 5, CE5	PGTG_13478	XP_383480	47.7	2.44 E-03	70.2	0.011	nd	N	na
91623	b-glycosidase, glycosyl hydrolase 5, GH5	PGTG_07933	YP_003009446	16.8	5.77 E-04	6.8	0.034	nd	Y	na
DNA processing										
91139	SNF2 family DNA-dependent ATPase	PGTG_07740	XP_001873970	51.6	2.39 E-03	ns	ns	nd	nd	na
95635	SNF2-related helicase	PGTG_07740	XP_001873970	33.9	0.024	ns	ns	nd	nd	na
Protein fate										
65069	tyrosinase, secreted	PGTG_11655	XP_001885552	63.3	1.13 E-04	ns	ns	nd	nd	na
92040	OTU-like cysteine protease, protease C85 family	PGTG_16664	XP_001416751	39.0	3.42 E-03	ns	ns	nd	Y	na
95405	aspartic peptidase A1, secreted	PGTG_10570	1WKR	31.4	0.016	13.7	0.011	nd	nd	na
Transcription										
39714	GATA factor, cutinase gene palindrome-binding protein	PGTG_06212	XP_661040	459.4	5.39 E-05	ns	ns	nd	nd	na
Carbon metabolism										
112995	trehalose phosphatase	PGTG_13607	XP_002173099	115.1	1.90 E-03	ns	ns	nd	nd	na
Vitamin biosynthesis										
67013	thiamin biosynthesis enzyme Thi1p (PIG1)	PGTG_10151	ABK96768	130.8	1.51 E-04	4.1	0.013	nd	Y	Y

Tableau 5: Identification de gènes fortement induits au cours de la phase biotrophe et communs entre les deux études transcriptomiques (96 heures post infection versus urediniospores et aire de biotrophie versus aire de sporulation durant la formation de l'urédie). Les 10 gènes les plus régulés codant des petites protéines secrétées ainsi que les gènes les plus fortement induits et présentant des homologues dans la base de donnée nr ont été sélectionnés et regroupés par catégories cellulaires. Le niveau de régulation (Fold-change, FC) est indiqué pour les deux études ainsi que la *p-value*. Les gènes soulignés sont spécifiquement détectés *in planta*. La présence d'ESTs identifiés au sein de banques d'urediniospores (Usp) et de feuilles infectées (96-168 heures post infection) est également présentée ainsi que la validation de l'expression *in planta* par RT-qPCR. na: données non disponibles; nd: non détecté; ns: gènes régulés plus de trois fois mais pas de manière significative.

(**Tableau 5**) nous permet de mettre à jour un jeu de gènes fortement induits durant le processus infectieux. Ces deux études ont tout d'abord révélé l'induction conséquente de gènes codant des SSPs au cours de la phase biotrophe (**Figure 28**). Ainsi, près de 70% des gènes identifiés parmi le top 25 des transcrits les plus régulés à 96 hpi dans l'analyse globale du transcriptome ou par l'approche de microdissection laser codent des SSPs et une majorité sont retrouvés dans ces deux études. Outre la présence importante de gènes homologues des HESPs de *M. lini* parmi les gènes les plus régulés, certains codent également des effecteurs potentiels pour lesquels un caractère d'avirulence a été mis en évidence chez *M. lini* (Dodds et al. 2004; Catanzariti et al. 2006) (**Figure 28**). Ce résultat marquant reflète l'abondance et le rôle important de ces protéines au cours du processus infectieux et plus particulièrement après que les haustoria sont formés. De manière intéressante, parmi les transcrits codant des SSPs fortement accumulés lors de la croissance biotrophe et qui présentent une homologie avec des HESPs de *M. lini*, plusieurs ne partagent pas d'homologie avec des gènes de *P. graminis* f. sp. *tritici* (**Tableau 5**), ce qui indiquerait que des effecteurs putatifs parmi les SSPs seraient partagés entre ces deux espèces proches (*Melampsoraceae*) mais pas avec des espèces plus éloignées (*Pucciniaceae*). De plus, leur expression à 168 hpi dans le parenchyme palissadique alors que le champignon produit par ailleurs des urédies suggère un rôle dans le maintien de l'interaction biotrophe aux stades tardifs d'infection. Notre analyse montre également l'accumulation chez *M. larici-populina*, de transcrits précédemment identifiés chez des agents de rouilles comme étant fortement exprimés au niveau de la structure haustoriale. C'est notamment le cas d'homologues de certains transporteurs d'acides aminés ou de sucres identifiés chez *U. fabae* (*Uf-AAT2*, Hahn et al. 1997b; *Uf-HXT1*, Voegelé et al. 2001; *Uf-AAT3*, Struck et al. 2004). La forte induction *in planta* de ces transporteurs chez *M. larici-populina* suggère également leur implication dans le détournement des réserves carbonées et azotées de la plante hôte (**Figure 28**). Nous avons également identifié parmi les gènes très fortement exprimés *in planta* des homologues de *Uf-THI1* et *Uf-THI2* (Hahn et Mendgen 1997a; Sohn et al. 2000), tous deux impliqués dans la biosynthèse de la thiamine (Vitamine B1). Les profils d'expression que nous avons établis par RT-qPCR pour *Mlp-THI1* et *Mlp-THI2* sont en adéquation avec les niveaux d'expression rapportés pour *U. fabae* (Hacquard et al. 2011, chapitre II, publication n°3). Chez cet agent de rouille, les transcrits *THI* étaient également parmi les plus fortement représentés dans la banque d'ESTs issus d'haustoria purifiés ce qui reflète l'importance de la vitamine B1 dans le métabolisme du champignon *in planta* et plus généralement chez les agents de rouilles.

L'analyse transcriptomique des étapes tardives d'infection (168 hpi), que ce soit lors de l'analyse globale du transcriptome ou par l'approche de microdissection à capture laser a également permis l'identification de gènes marqueurs pouvant jouer un rôle clef dans la sporulation de *M.*

larici-populina. La technique de microdissection à capture laser appliquée à la formation de l'urédie a notamment permis de discriminer les structures fongiques associées à la biotrophie tels que les haustoria et les hyphes d'infection des structures associées à la sporulation tels que les tissus sporifères et les urédospores nouvellement formées (Hacquard et al. 2010, chapitre IV, publication n°6). Nous avons ainsi pu montrer que la formation des urédospores au sein de la zone de sporulation est accompagnée à la fois par l'induction de nombreux transcrits en relation avec le cycle cellulaire, supportant l'intense division observée mais également par l'accumulation de nombreux transcrits impliqués dans les mécanismes de détoxification suggérant la sollicitation active des systèmes de protection cellulaire du champignon durant cette phase. Ces systèmes de protection pourraient être activés en réponse aux réactions de défense tardives mises en place par le peuplier dans le cadre de l'interaction compatible avec *Melampsora* spp. (Duplessis et al. 2009). Les deux approches transcriptomiques (Analyse globale du transcriptome à 168 hpi et microdissection laser de l'urédie ont permis de mettre en évidence l'induction d'un panel de gènes codant des enzymes de dégradation telles que des alpha/beta glycosidases ou une alpha-L-arabinofuranosidase qui pourraient être requises pour dégrader l'épiderme de la plante hôte et permettre le relargage des spores (**Figure 28**).

IV - Analyse du secretome de *Melampsora larici-populina*

La disponibilité croissante de génomes fongiques a permis de réaliser la prédiction des gènes codant des protéines sécrétées chez les champignons pathogènes ou symbiotiques des plantes (Kämper et al. 2006; Martin et al. 2008a). Récemment, une base de données dédiée à la prédiction et l'analyse des sécrétomes fongiques y compris non pathogènes a été développée afin d'améliorer les paramètres de prédiction et l'accessibilité des données (Choi et al. 2010). La majorité des effecteurs fongiques identifiés jusqu'à présent correspondent, à de rares exceptions près, à des protéines sécrétées de petite taille (Stergiopoulos et de Wit 2010). Ainsi, la prédiction à grande échelle de l'ensemble des protéines sécrétées des micro-organismes pathogènes a ouvert de nouvelles perspectives dans l'identification plus systématique d'effecteurs fongiques candidats potentiellement impliqués dans le dialogue moléculaire avec l'hôte (Dodds et al. 2009). Dans le cadre de l'analyse du secretome de *M. larici-populina*, nous nous sommes focalisés sur l'ensemble des modèles de gènes prédits pour coder des protéines sécrétées de moins de 300 aa (Cut-off arbitraire proche de 30kDa) pouvant correspondre à des effecteurs potentiels. De manière intéressante, un large arsenal de 1,184 gènes codant des SSPs a été identifié chez *M. larici-populina* ce qui représente plus de 8% des gènes prédits du champignon. La plupart de ces gènes codant des SSPs n'ont pas de fonctions connues (> 60%) et seulement 25% d'entre eux présentent une homologie de séquence avec des gènes prédits chez l'agent

responsable de la rouille du blé *P. graminis* suggérant une évolution et une diversification rapide entre les secrétomes des ces deux pathogènes.

1 - Identification de petites protéines sécrétées présentant des homologies avec des effecteurs précédemment caractérisés.

Malgré les faibles niveaux d'homologie de séquence observés jusqu'à présent entre les effecteurs de différents pathogènes, la proximité taxonomique entre *M. larici-populina* et les agents responsables de la rouille du lin *M. lini* et de la fêverole *U. fabae* a permis de révéler la présence de SSPs dont les séquences sont plus ou moins conservées (de 24% à 100%) chez *M. larici-populina* (Hacquard et al., en préparation pour PLoS Pathogen, chapitre III, publication n°4). De manière intéressante, les gènes codant quatre facteurs d'avirulence et 14 des 17 HESPs identifiés chez *M. lini* (Dodds et al. 2004, Catanzariti et al. 2006) ainsi que RTP1 identifiée chez *U. fabae* (Kemen et al. 2005) ont été retrouvés chez *M. larici-populina* indiquant la présence de SSPs conservées chez ces agents de rouilles.

Chez *M. lini*, *AvrM* appartient à une famille multigénique sous pression de sélection positive (Catanzariti et al. 2006; Catanzariti et al. 2010). Un excès de substitutions non synonymes comparé aux substitutions synonymes a été identifié entre *AvrM* et les autres membres de la famille chez *M. lini* (Catanzariti et al. 2006). De la même manière, 8 séquences très divergentes ont été identifiées au sein de la séquence génomique de *M. larici-populina*. L'organisation en tandem de certains de ces gènes et la présence d'éléments transposables entre les régions intergéniques suggèrent, comme chez l'agent de la rouille du lin, une diversification rapide de cette famille chez *M. larici-populina*.

Une analyse de la diversité de séquence de gènes codant des homologues du gène d'avirulence *AvrP4* de *M. lini* a été entreprise au sein du genre *Melampsora* (Van der Merwe et al. 2009). Cette étude a révélé une forte pression de sélection positive jouant un rôle majeur dans l'évolution rapide de cette famille. Chez *M. larici-populina*, 13 modèles de gènes homologues d'*AvrP4* ont été identifiés et l'analyse du taux de substitutions dans notre étude révèle également une pression de sélection positive s'exerçant sur cette famille (Hacquard et al., en préparation pour PLoS Pathogen, chapitre III, publication n°4). De manière intéressante, cette analyse a également démontré le contraste entre la parfaite conservation des cystéines et l'intense diversification des autres acides aminés au niveau de la région C-terminale de la protéine suggérant le potentiel rôle de ce domaine dans l'interaction avec les protéines de défense de l'hôte. L'analyse de cette famille multigénique a également permis l'identification d'une région riche en acides aminés basiques (Arg, Lys) suivie par une région riche en acides aminés acides (Asp, Glu) dans la partie

N-terminal de la protéine reminiscence du domaine RxLR dEER identifié chez *Phytophthora* spp. (Rehmany et al. 2005; Tyler et al. 2006; Jiang et al. 2007). Ce motif pourrait permettre l'ancrage de l'effecteur à l'inositol 3 et/ou 4 phosphate présents au sein de la membrane plasmique de l'hôte pour assurer ensuite son internalisation dans la cellule de la plante hôte par endocytose (Kale et al. 2010).

De manière contrastée, un seul homologue d'AvrL567 a été identifié au sein du génome de *M. larici-populina* alors que différents variants (1 à 4 copies pour chaque haplotype) ont été identifiés chez *M. lini* (Dodds et al. 2004). De nombreuses similarités ont été observées entre les deux orthologues que ce soit dans la structure du gène et de la protéine, dans l'accumulation spécifique des transcrits *in planta* ainsi que dans la localisation de la protéine au niveau de l'haustorium (Hacquard et al., en préparation pour PLoS Pathogens, chapitre III, publication n°4). Bien que la localisation de la protéine n'ait pas été réalisée directement *in planta* chez *M. lini*, la reconnaissance entre le produit du gène d'avirulence *AvrL567* du champignon et les protéines cytoplasmiques codées par les gènes de résistance correspondants *L5*, *L6*, *L7* chez le lin est une démonstration indirecte de la translocation de l'effecteur de l'haustorium dans la cellule hôte (Dodds et al. 2004; Dodds et al. 2006). Du fait des similarités de séquences entre les deux espèces de *Melampsora*, il est tentant de spéculer sur une fonction commune conservée entre ces effecteurs. Chez *M. lini*, cette famille a été identifiée comme étant sous pression de sélection positive et il apparaissait important d'étudier le niveau de polymorphisme de ce gène chez *M. larici-populina*. *AvrL567* étant un gène orphelin au sein du génome de *M. larici-populina*, nous avons entrepris le séquençage de sa partie codante chez 32 souches de l'espèce *M. larici-populina* collectées sur le terrain sur des peupliers infectés en Lorraine, en France ou ailleurs dans le monde (**Figure 29**). A partir des 64 allèles séquencés (32 souches dicaryotiques), seuls 5 variants alléliques ont été mis en évidence et leur fréquence au sein de l'espèce *M. larici-populina* oscille entre 3% et 56%. Contrairement à la forte variabilité nucléotidique observée entre les séquences d'*AvrL567* de *M. lini* (27 mutations non synonymes et 3 mutations synonymes, Dodds et al. 2004), cette étude a révélé un très faible niveau de polymorphisme entre les différentes souches analysées ainsi que l'absence de sélection positive pour ce gène de *M. larici-populina* (6 mutations non synonymes et 5 mutations synonymes). La forte conservation de séquence d'*AvrL567* et l'absence de zones de polymorphisme communes à certains pathotypes de *M. larici-populina* ne permettent pas d'affirmer qu' *AvrL567* est un facteur d'avirulence conservé au sein des deux pathosystèmes Lin/*M. lini* et Peuplier/*M. larici-populina*.

L'analyse de la séquence génomique de *M. larici-populina* a également révélé la présence de 3 gènes homologues de RTP1 dont un correspond à une protéine plus grande dont la partie N-terminale est homologue d'un HESP de *M. lini*. Ce gène, précédemment identifié chez *U. fabae* à partir d'une banque d'ESTs d'haustoria enrichie, est spécifique du processus infectieux (Hahn et Mendgen 1997). La sécrétion de la protéine au niveau de l'haustorium vers le cytoplasme ainsi que son adressage au noyau de l'hôte ont été validés expérimentalement par immunolocalisation dans le cadre de l'interaction *U. fabae*/féverole (Kemen et al. 2005). Chez *M. larici-populina*, les trois membres de la famille présentent des profils d'expression différents par RT-qPCR suggérant des rôles ou des fenêtres d'action différents durant le processus infectieux. Alors que chez *M. larici-populina*, l'immunolocalisation du plus proche homologue de Uf-RTP1 n'a pas abouti du fait de l'absence de marquage des structures fongiques et végétales, la localisation des deux autres membres de la famille a révélé une localisation distincte *in planta*. Au sein de cette famille, seule la protéine fusion entre HESP-327 et RTP1 présentait un domaine d'adressage nucléaire potentiel. Néanmoins, l'immunolocalisation de la protéine, bien que détectée au niveau de l'haustorium, n'a pas été visualisée au-delà dans le cytoplasme ou le noyau des cellules de la plante hôte.

2 - Identification d'effecteurs candidats spécifiques à *Melampsora larici-populina* riches en cystéine et sous pression de sélection positive

Chez les eucaryotes, de nombreux effecteurs possèdent un nombre important de cystéines au sein de leurs séquences protéiques. C'est notamment le cas de certains effecteurs identifiés chez *C. fulvum* (Avr2, Avr4, Avr4E, Avr9, Ecp1, Ecp2, Ecp4, Ecp6 et Ecp7), *L. maculans* (AvrLm6 et AvrLm4-7), *M. oryzae* (Avr-Pita), *M. lini* (AvrP123 et AvrP4), *R. secalis* (Nip1, Nip2, Nip3) ou encore chez *P. infestans* (elicitines) (Stergiopoulos et de Wit 2010). Cette forte concentration de résidus cystéine reflète généralement la formation de ponts disulfures qui vont assurer la stabilité structurale de la protéine. Un nombre important de ponts disulfures au sein d'une protéine permettra de pouvoir faire face aux protéases contenues dans le milieu apoplastique de la plante hôte (Stergiopoulos et de Wit 2010). De manière intéressante, nous avons pu observer une forte proportion de gènes codant des protéines riches en cystéines parmi les SSPs de *M. larici-populina* (60% possèdent au moins 4 cystéines). Au sein de ce SSPs riches en cystéines, certaines, réparties en 28 classes, partagent une structure génique commune (Hacquard et al., en préparation pour PLoS Pathogens, chapitre III, publication n°4). Malgré la faible identité de séquence au sein de chacune des classes, ces protéines possèdent des résidus cystéines quasi systématiquement conservés ce qui suggère une structure tridimensionnelle probablement conservée au sein de ces familles. Alors que les ponts disulfures seraient nécessaires au maintien

de la structure de ces protéines, les régions variables pourraient être essentielles à leurs fonctions et leurs spécificités (Povolotskaya et Kondrashov 2010). De manière intéressante, certaines de ces familles partagent un motif Y/FxC similaire au motif Y/F/WxC récemment décrit chez une nouvelle classe d'effecteurs candidats, fortement exprimés au niveau de l'haustorium chez *B. graminis* et *P. graminis* (Godfrey et al. 2010). Bien que nous ne disposions pas de données d'expression pour certains de ces gènes du fait de l'absence de prédiction bioinformatique au moment de la fabrication des puces à oligonucléotides NimbleGen, plusieurs ont été identifiés au sein des banques d'haustoria et de feuilles infectées confirmant un rôle probable de ces protéines au niveau de l'haustorium. C'est notamment le cas pour les classes VI et XIV dont les membres présentent des homologues avec les HESPs identifiés chez *M. lini* (HESP-C49 et HESP-897) et pour lesquelles certains gènes ont été identifiés comme étant spécifiquement exprimés *in planta* chez *M. larici-populina*. Cette spécificité d'expression *in planta* a été confirmée par la microdissection laser de l'urédie révélant l'induction de plus de 20 fois de leurs transcrits dans la zone de biotrophie comparé à la zone de sporulation du champignon (Hacquard et al. 2010, chapitre IV, publication n°6).

Chez de nombreux pathogènes, les protéines sécrétées présentent des niveaux d'évolution élevés en accord avec le modèle de « course aux armements » entre le système immunitaire de l'hôte et les effecteurs sécrétés par le pathogène (Dangl et Jones 2006). Cette pression de sélection positive, essentielle au contournement des mécanismes de défense de l'hôte a été observée pour de nombreux effecteurs (Dodds et al. 2006; Guttman et al. 2006; Win et al. 2007; Brunner et al. 2009). Afin de tester si certains gènes codant des SSPs pourraient être soumis à de telles pressions de sélection, nous avons identifié 95 clusters de groupes de gènes paralogues (de 3 à 39 gènes) au sein du génome de *M. larici-populina* et représentant au total un sous ensemble de 509 SSPs. Parmi ceux-ci, nous avons vérifié que 18 CPGs (186 gènes) présentaient des niveaux d'évolution accélérés (Hacquard et al., en préparation pour PloS pathogens, chapitre III, publication n°4). De manière troublante, l'ensemble des gènes identifiés sous pression de sélection codent des protéines de petite taille (<194 acides aminés) parmi lesquelles 85% sont riches en cystéines (> 4 cystéines). Ces SSPs pourraient présenter de tels taux d'évolution afin d'échapper à la reconnaissance de l'hôte et représentent des facteurs d'avirulence potentiels susceptibles d'être reconnus par certains cultivars de peuplier (résistance R1 à R7). Il serait intéressant de vérifier cette hypothèse en séquençant ces ensembles de gènes chez plusieurs souches aux fonds génétiques différents pour les virulences afin de tester leur variabilité au niveau intraspécifique. Comme il l'a été démontré chez de nombreux effecteurs, la plupart des petites protéines sécrétées riches en cystéines identifiées sous pression de sélection chez *M. larici-populina* présentent des variations importantes de leurs acides aminés dans la partie C-

Tableau 6: Cluster de gènes paralogues (CPG) codant des petites protéines sécrétées identifiés sous pression de sélection positive au sein du génome de *Melampsora larici-populina* et présentant un motif réminiscent du motif RxLR. Pour chacun des gènes, la présence d'ESTs (en rouge) dans les différentes banques: urediniospores, 96/168 heures post infection et haustoria (Joly et al. 2010) est également présentée.

CPG	Protéine ID	motif RXLR	Taille protéine	Cysteines	EST Usp	EST 96-168 hpi	EST Haustoria
CPGH1	124505		68	4	-	-	-
	124509		68	3	-	Oui	-
	124516		68	4	-	Oui	-
	124517		77	3	-	-	-
	124497		77	4	-	Oui	Oui
	124498		73	3	-	-	Oui
	124499		72	3	-	Oui	Oui
	124511		66	3	-	-	-
	124525		71	3	-	Oui	Oui
	124526		76	3	-	Oui	Oui
	124514		70	4	-	-	-
	124515		73	4	-	-	-
	124522		82	2	-	Oui	-
	124527		84	5	-	Oui	Oui
	124528		84	5	-	Oui	Oui
	124529	Région basique (K/R)	83	6	-	Oui	-
	124502	suivie par une	73	3	-	-	Oui
	124503	Région acide (D/E)	81	3	-	-	-
	124506		58	2	-	-	-
	124507		77	5	-	Oui	Oui
	124510		74	4	-	Oui	-
	124520		77	4	-	-	-
	124513		75	4	-	-	Oui
	124521		76	4	-	Oui	-
124500		95	4	-	-	-	
124518		76	3	-	Oui	Oui	
124501		116	2	-	-	-	
124512		82	5	-	-	-	
124504		73	3	-	-	-	
124519		80	3	-	-	-	
124523		75	5	-	Oui	-	
124524		70	1	-	-	-	
CPG5464 (AvrP4)	112294		105	6	-	-	-
	124252		96	7	-	-	-
	124254		90	7	-	-	-
	124256		88	7	-	Oui	-
	124258		95	7	-	-	-
	124260	Région basique (K/R)	92	7	-	-	-
	124262	suivie par une	101	7	-	-	-
	124264	Région acide (D/E)	89	9	-	Oui	-
	124266		92	8	-	Oui	-
	124268		90	7	-	-	-
124270		89	8	-	-	-	
124272		97	7	-	-	-	
124274		97	7	-	-	-	
CPG510-1	86947		86	5	-	-	-
	89503		78	4	-	-	-
	89506		78	4	-	-	-
	95971		77	4	-	-	-
	124531		70	4	-	Oui	-
	124532		70	4	-	-	-
	124533		70	4	-	-	-
	124534		69	4	-	-	-
	124535	LXRR	64	4	-	-	-
	124536		79	4	-	-	-
	124537		68	4	-	Oui	-
	124538		77	4	-	Oui	-
	124539		77	4	-	Oui	-
124541		76	6	-	-	-	
124542		86	4	-	Oui	-	
124543		83	4	-	-	Oui	
124540		21	4	-	-	-	
CPG510-2	74800		79	7	Oui	Oui	-
	87935		74	6	-	-	-
	95419		74	6	-	-	-
	124387		76	6	-	-	-
	124388		78	6	-	-	-
	124391	LXRR	73	6	-	-	-
	124392		80	6	-	-	-
	124393		73	6	-	-	-
	124394		124	12	-	-	-
	124395		73	6	-	-	-
124396		71	6	-	-	-	

terminale de la protéine suggérant que c'est cette région qui interagit avec le système de défense de l'hôte (Allen et al. 2004; Rehmany et al. 2005; Win et al. 2007; Barrett et al. 2009; Van der Merwe et al. 2009). De manière contrastée, la partie N-terminale de ces protéines, où résident les signaux de sécrétion et probablement de translocation de l'effecteur semble mieux conservée. En effet, tout comme chez AvrP4 (CPG5464), certaines familles sous pression de sélection (CPGH1, CPG510-1 et CPG510-2) présentent des domaines réminiscent du motif RXLR qui pourraient être impliqués dans la translocation de l'effecteur dans la cellule hôte (Rafiqi et al. 2010; Kale et al. 2010). Le séquençage d'ESTs a permis de révéler la présence de transcrits pour 28 gènes appartenant à ces quatre familles (CPG5464, CPGH1, CPG510-1 et CPG510-2) et a révélé l'accumulation spécifique de leurs transcrits *in planta* (haustoria ou feuilles infectées) pour 27 d'entre eux confirmant leur statut d'effecteur potentiel (**Tableau 6**).

3 - Identification d'effecteurs candidats induits *in planta*

Bien que de nombreux effecteurs d'organismes phytopathogènes mis à jour jusqu'à présent soient sous pression de sélection positive, ceci n'est cependant pas une caractéristique partagée par tous. Par exemple, les effecteurs extracellulaires (Ecp) identifiés chez *C. fulvum* ne sont pas sous pression de sélection positive. De plus, pour une large majorité d'effecteurs, ce paramètre évolutif n'est pas disponible (Stergiopoulos et de Wit 2010). Chez *M. larici-populina*, nous n'avons pu tester ce paramètre évolutif que pour les clusters de gènes paralogues à un niveau intra-souche et ceux-ci ne représentaient que 509 SSPs. Afin d'identifier des effecteurs potentiels au sein du sécrétome de *M. larici-populina* il apparaissait essentiel de ne pas se focaliser uniquement sur cette approche évolutive. Conjointement au séquençage des génomes de champignons et d'oomycètes interagissant avec les plantes, l'analyse de leurs transcriptomes a permis l'identification de gènes spécifiquement induits durant la colonisation des tissus végétaux. Ces approches ont par exemple conduit à l'identification d'effecteurs potentiels induits durant la formation de la symbiose ectomycorhize (MISSPs, mycorrhiza-induced SSPs) chez le couple peuplier/*L. bicolor* (Martin et al. 2008a; Plett et al., non publié). De la même manière chez *M. oryzae*, l'identification de gènes codant des protéines sécrétées induits durant l'invasion des cellules de riz a confirmé la présence de nombreux facteurs d'avirulence (Pwl2, Avr-Pita1) et a permis l'identification d'effecteurs candidats (BAS, Biotrophy-Associated Secreted proteins) spécifiquement associés à la phase biotrophe de l'agent de la pyriculariose du riz (Mosquera et al. 2009; Khang et al. 2010). Chez *M. larici-populina*, l'analyse globale des transcrits codant ces petites protéines sécrétées a révélé leur accumulation préférentielle *in planta* (96 hpi) suggérant l'importance de ces protéines dans la biotrophie du champignon (Duplessis et al., en préparation pour science, chapitre II, publication n°1; Hacquard et al. en préparation pour PLoS Pathogens,

A

Souches	Vir	Avr	Hôte	Echantillonnage	Allèles	Position des codons polymorphes						
						15	34	144	160	164	211	299
97A1	0	1_2_3_4_5_6_7_8	<i>P. x euramericana</i>	Maroc	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
97A3	3.4	1_2_5_6_7_8	<i>P. nigra</i>	N-Zélande	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
97J10	3.4	1_2_5_6_7_8	<i>P. x euramericana</i> '1-488'	Afrique du Sud	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
98AE3	2	1_3_4_5_6_7_8	-	Finlande	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
00A19	0	1_2_3_4_5_6_7_8	<i>P. x euramericana</i> '1-488'	Chili	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
98AR1	1_3_4_5_7_8	2_6	<i>P. x interamericana</i> 'B-71085/A1'	Belgique	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
95US1	3.4	1_2_5_6_7_8	<i>P. nigra</i> 'Italica'	USA	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
95XP5	1_3_4_5_6_7	2_8	<i>P. x interamericana</i> 'Beaupré'	Cadillac (33)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
95XW2	1_3_4_5_7	2_6_8	<i>P. x interamericana</i> 'Boelare'	Guémené-Penfao (44)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
02AO33	4	1_2_3_5_6_7_8	<i>P. nigra</i>	La Brillanne (04)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991006	3.4	1_2_5_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991056	1.4_5	2_3_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991061	1_3_4_5_6	2_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991108	4	1_2_3_5_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991192	4.7	1_2_3_5_6_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
97C3	1_3_4_5_6_7	2_8	<i>P. x interamericana</i> 'Beaupré'	Royaume Uni	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
01L31	2_3_7	1_4_5_6_8	<i>P. x euramericana</i> 'AAA'	Nogent-sur-Vernisson (45)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
01L66	1_2_3_4_5_7	6_8	<i>P. x euramericana</i> 'Laxo 7'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991009	2.4	1_3_5_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991036	3_4_7	1_2_5_6_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991046	2_3_4	1_5_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
991094	1_2_3_4_5_7	6_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
95AAGH12	1_3_4_5_6_7	2_8	<i>P. x interamericana</i> 'Beaupré'	Raincourt (70)	2	CGT	ATC	TCC	ACC	TAT	TTC	AGT
99D1	3.4	1_2_5_6_7_8	<i>P. trichocarpa</i>	Islande	1	CGC	GTC	TCC	ACC	TAT	CTC	AGT
94Z25	1_3_4_5_7	2_6_8	<i>P. x euramericana</i> 'Ghoy'	Chereng (59)	2	CGT	ATC	TCC	ACC	TAT	CTC	AGT
98A02	3_4_7	1_2_5_6_8	<i>P. x interamericana</i> 'Beaupré'	Bayonne (64)	1	CGC	GTC	TCC	ACC	TAT	CTC	AGT
991179	3_4_5	1_2_6_7_8	<i>P. x euramericana</i> 'Robusta'	Champenoux (54)	2	CGT	ATC	TCC	ACC	TAT	CTC	AGT
98AG31	3_4_7	1_2_5_6_8	<i>P. x interamericana</i> 'Beaupré'	Moy (02)	1	CGT	ATC	TCC	ACC	TAT	TTC	AGT
00M10	3_4_7_8	1_2_5_6	<i>P. x interamericana</i> '69038-1'	Charrey-sur-Saône (21)	2	CGC	GTC	TCC	ACC	TAT	CTC	AGT
97EA2	4	1_2_3_5_6_7_8	-	Chine	1	CGC	GTC	TCT	AAC	TAC	CTC	AAT
97EA1	4	1_2_3_5_6_7_8	-	Chine	2	CGC	GTC	TCT	AAC	TAC	CTC	AAT
97EA3	4	1_2_3_5_6_7_8	-	Chine	2	CGC	GTC	TCT	AAC	TAC	CTC	AAT

- ★ Substitutions non synonymes
- ★ Substitutions synonymes

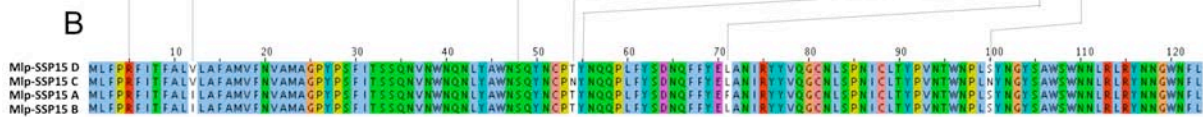


Figure 30: Analyse du polymorphisme de séquence de *Mlp-SSP15*. **A**, 64 séquences ont été séquencées à partir de 32 souches (dicaryotiques) appartenant à l'espèce *Melampsora larici-populina* et présentant des combinaisons virulence (Vir) /avirulence (Avr) différentes. Pour chacune des souches, les deux allèles ont été séquencés et seuls les codons présentant un polymorphisme sont présentés. Les nucléotides divergeants entre les différentes souches sont représentés en rouge. Les différents acides aminés ont été déduits à partir des différents codons et sont présentés en gras sous le tableau **B**, Alignement des séquences protéiques déduites à partir des séquences nucléotidiques de chaque gène. ★ mutations synonymes ou/et ★ non synonymes observées

chapitre III, publication n°4) (**Figure 28**). Un cinquième des gènes codant pour ces SSPs, incluant de nombreux homologues d'HESPs et de facteurs d'avirulence déjà identifiés chez *M. lini*, est spécifiquement exprimé au cours de la phase biotrophe (entre 24 et 168 hpi). L'analyse comparative des interactions peuplier/*Melampsora* spp. a permis de mettre en évidence une manipulation très précoce des systèmes de défense de l'hôte (Duplessis et al. 2009). Il apparaissait donc essentiel de vérifier le niveau d'expression des gènes lors d'étapes très précoces d'infection (entre 2 et 24 hpi) pour identifier des transcrits qu'il ne nous était pas possible de détecter avec les puces à ADN pour des raisons techniques (dilution des transcrits fongiques dans les feuilles infectées). Nous avons ainsi pu montrer grâce à la technique de RT-PCR quantitative une diversité de profils d'expression pour les gènes codant des SSPs sur la totalité du processus infectieux et notamment l'accumulation très précoce de certains transcrits à 2, 6 et 12 hpi. Parmi ces transcrits, plusieurs codent des SSPs à domaine CFEM identifiées aussi parmi les déterminants des étapes précoces d'infection chez *M. oryzae* (DeZwaan et al. 1999).

Le séquençage d'ESTs nous a permis d'identifier 221 gènes dont les transcrits sont accumulés à 96-168 hpi alors qu'ils n'ont pas été identifiés dans les spores (germées et non germées) suggérant leur spécificité dans la phase biotrophe. Le transcrit le plus fortement représenté parmi la banque d'ESTs de feuilles infectées code une petite protéine sécrétée de 15kDa nommée SSP15. De manière intéressante, cette spécificité d'expression *in planta* a été confirmée par l'utilisation des puces à ADN ainsi que par l'analyse par RT-qPCR. Cette dernière étude a permis de révéler l'absence de transcrit dans les spores et les spores germées ainsi qu'une induction biphasique du gène à 12 hpi au moment de la mise en place des premiers haustoria et à 168 hpi lors de la formation de l'urédie (Hacquard et al., en préparation pour PLoS Pathogens, chapitre III, publication n°4). De manière très intéressante, nous avons pu déterminer par immunolocalisation à un stade tardif d'infection (168 hpi) sa présence spécifique dans les tissus foliaires à deux sites distincts. L'accumulation de la protéine au niveau du tissu sporifère ainsi qu'à la périphérie de l'haustorium et en accord avec les résultats de RT-qPCR et suggère des rôles distincts de la protéine durant le processus infectieux. Il serait intéressant de déterminer si cette protéine est également accumulée au niveau de l'haustorium dans les étapes plus précoces d'infection (stades 12, 24 et 48 hpi) pour confirmer sa fonction d'effecteur potentiel. Afin d'étayer cette hypothèse, nous avons entrepris le séquençage de ce gène chez un panel de 32 souches de l'espèce *M. larici-populina* afin de mettre à jour sa variabilité nucléotidique (**Figure 30**). A partir des 64 séquences analysées, seuls quatre variants dont les fréquences varient entre 9,5 et 62% ont été identifiés. Entre ces différents variants, 4 mutations non synonymes et 3 mutations synonymes ont été identifiées reflétant l'absence de sélection positive pour ce gène chez *M. larici-populina*. Bien que la fonction de ce gène soit difficile à établir du fait de l'absence

d'homologie dans les bases de données internationales ainsi qu'avec *P. graminis* f. sp. *tritici*, l'accumulation spécifique et importante des ses transcrits *in planta* suggère son rôle prépondérant dans la biotrophie et la sporulation du champignon.

Au cours de cette étude, nous avons mis à jour un nombre important de gènes codant des SSPs au sein du génome de *M. larici-populina*. Par nos approches d'analyse d'expression et d'évolution des SSPs couplés à l'immunolocalisation de certains candidats, nos travaux apportent une contribution essentielle pour la recherche d'effecteurs chez les agents de rouilles et pour la compréhension des mécanismes d'interaction entre *M. larici-populina* et son hôte. Ce travail unique d'analyse fine des effecteurs potentiels à l'échelle génomique va grandement contribuer à l'amélioration des connaissances sur la biologie du champignon ainsi qu'à l'identification de moyens de lutte efficaces contre la maladie.

Conclusions et perspectives

Conclusions et Perspectives

Le séquençage du génome de *M. larici-populina* a permis de révéler des signatures caractéristiques pouvant expliquer le statut biotrophe de cet agent de rouille. Outre l'absence de nitrite réductase, la faible proportion d'enzymes de dégradation et l'expansion de certaines familles de transporteurs de sucres ou d'oligopeptides, l'analyse du génome de la rouille du peuplier a également révélé un large répertoire de gènes codant des SSPs dont une bonne part lui est propre. L'analyse du transcriptome du champignon, basée sur l'utilisation de différentes approches transcriptomiques, a permis de mettre à jour des gènes marqueurs des différentes phases d'infection du champignon allant de la spore dormante à la formation de l'urédie. Alors que l'induction de nombreuses lipases et enzymes de dégradation des sucres a été observée lors de la germination et la sporulation (168 hpi) du champignon, la majorité des transcrits codant des transporteurs et des SSPs a été détectée durant le processus infectieux (96 hpi) suggérant le rôle prépondérant de ces protéines lors de la croissance biotrophe du champignon. De manière intéressante cette induction importante des transcrits codant des SSPs, incluant de nombreux HESPs se poursuit à 168 hpi au niveau du parenchyme palissadique lors du stade urédien, suggérant que certaines SSPs pourraient intervenir dans le maintien de l'interaction biotrophe aux stades tardifs d'infection. Nos travaux, en plus de confirmer l'induction spécifique *in planta* de nombreux gènes précédemment identifiés chez les agents de rouilles comme accumulés au niveau de l'haustorium (Catanzariti et al. 2006; Voegelé et al. 2009), ont également permis de mettre à jour de nouveaux gènes marqueurs de la biotrophie.

L'analyse fine du sécrétome de *M. larici-populina*, basée sur l'annotation, l'analyse de l'évolution et de l'expression des gènes codant les SSPs ainsi que l'immunolocalisation de certains gènes candidats a permis l'identification d'effecteurs potentiels au sein du génome de *M. larici-populina*. Les différentes approches transcriptomiques utilisées au cours de cette étude ont permis de mettre à jour une diversité importante des profils d'expression des gènes codant ces SSPs au cours des stades précoces et tardifs de la cinétique d'infection et ont révélé l'induction spécifique et importante de nombreux transcrits pendant ou après la formation des premiers haustoria. Parmi eux, Mlp-SSP15 présente un profil d'expression biphasique par RT-qPCR coïncidant avec la formation des premiers haustoria (12 hpi) et la sporulation du champignon (168 hpi). L'immunolocalisation de la protéine au niveau du tissu sporifère ainsi que des haustoria pourrait indiquer une implication dans le dialogue moléculaire avec la plante ou encore un rôle dans la protection des différentes structures fongiques formées dans les tissus de l'hôte. D'autres effecteurs candidats ont été identifiés au sein du génome de *M. larici-populina*

sur la base de leur homologie avec des effecteurs précédemment identifiés chez *M. lini* et *U. fabae*. La fonction de ces effecteurs pourrait être conservée entre les différents agents de rouilles et l'immunolocalisation de Mlp-AvrL567 et de Mlp-HESP327/RTP1 a notamment confirmé la présence de ces protéines au niveau de l'haustorium chez *M. larici-populina* suggérant leur implication potentielle dans le dialogue moléculaire avec la plante hôte. Au sein du sécrétome de *M. larici-populina*, nous avons également identifié de nombreuses classes de SSPs présentant des cystéines conservées au sein de leurs séquences protéiques et arborant un domaine Y/FxC proche du domaine Y/F/WxC récemment identifié chez une nouvelle classe d'effecteurs candidats exprimés au niveau de l'haustorium chez *B. graminis* et retrouvés chez *P. graminis* (Godfrey et al. 2010). Enfin, de nombreuses familles de gènes codant des SSPs sont sous pression de sélection positive chez *M. larici-populina*. De manière très intéressante, les gènes présents au sein de quatre familles identifiées sous pression de sélection positive sont spécifiquement exprimés *in planta* et possèdent à l'extrémité N-terminale de leurs séquences protéiques un motif réminiscent du motif RxLR identifié chez les oomycètes (Whisson et al. 2007). L'évolution accélérée de ces gènes, leur expression spécifique *in planta*, la présence d'un motif potentiel de translocation ainsi que l'identification au sein de l'une de ces familles d'homologues du facteur d'avirulence AvrP4 identifié chez *M. lini* font de ces SSPs d'excellents effecteurs candidats.

Par cette analyse détaillée du sécrétome de *M. larici-populina*, de nombreux effecteurs candidats ont ainsi pu être identifiés et les perspectives à ce travail sont multiples :

- **Identifier de nouveaux effecteurs au sein du génome de *M. larici-populina*.**

Au cours de cette étude, nous avons pu mettre en évidence que de nombreux gènes codant des SSPs n'avaient pas été prédits au cours de l'annotation automatique du génome et ont été identifiés à partir de séquences exprimées *in planta*. Il se peut donc que certains effecteurs ne soient pas présents dans le catalogue de SSPs actuel. Un moyen de les détecter pourrait être l'utilisation des nouvelles techniques de séquençage (454, Illumina) afin d'identifier un large nombre de séquences spécifiques par exemple des étapes précoces d'infection ou lors de la formation des premiers haustoria. Dans le cadre de ma thèse, j'ai produit des ADNc des interactions compatible et incompatible (Beaupré/*M. larici-populina*). Ces ADNc ont été séquencés par 454 titanium et leur analyse, actuellement en cours, permettra de révéler la présence ou non de nouveaux gènes codant des effecteurs candidats chez *M. larici-populina*

- Identifier des effecteurs candidats exprimés au niveau de l'haustorium grâce à la technique de microdissection à capture laser

Au cours de ce projet de thèse, la technique de microdissection à capture laser a été utilisée afin de microdisséquer de larges zones en relation avec la biotrophie et la sporulation du champignon au stade urédien. La préparation des échantillons ainsi que la technique de microdissection étant maintenant au point, il est envisageable d'utiliser cette méthode pour microdisséquer des structures haustoriales du champignon formées durant l'infection de l'hôte à 96 hpi par exemple, point de la cinétique où l'induction des SSPs est le plus important. Cette étude pourrait être une alternative à la purification d'haustoria à partir de feuilles infectées et représente une technique de choix permettant l'analyse de l'expression des gènes au niveau de cette structure *in planta*. Une telle analyse pourrait conduire à l'identification de nouveaux effecteurs candidats et permettrait de valider l'expression des candidats identifiés au cours de notre étude au niveau de cette structure.

- Confirmer la sécrétion des SSPs par le système YSST

Les SSPs ont été identifiées au sein du génome de *M. larici-populina* sur la base de la prédiction bioinformatique d'un signal de sécrétion présent à l'extrémité N-terminal de leurs séquences protéiques et il apparaît important de valider leur sécrétion grâce au système YSST (Yeast Signal Sequence Trap). Au cours de ce travail de thèse, j'ai préparé une banque d'ADNc YSST enrichie en 5' à partir d'ARN issus de spores, spores germées et tissus infectés (96 hpi et 168 hpi). Cette banque pourra ensuite être utilisée afin de réaliser un criblage sur levure. Pour cela, une souche de *Saccharomyces cerevisiae* *SUC2* déficiente, incapable de sécréter une invertase et de pousser sur un milieu saccharose sera utilisée. Un vecteur comprenant le gène *SUC2* sans son codon start et son signal de sécrétion est requis pour compléter la levure. Seuls les ADNc tronqués possédant un codon start et un signal de sécrétion intégrés dans le vecteur permettront de restaurer la fonction de l'invertase et le développement de la levure sur un milieu saccharose. Cette approche permettra ainsi de confirmer la sécrétion de nombreuses SSPs et d'identifier peut-être de nouveaux candidats pour lesquels le signal de sécrétion n'avait pas été prédit bioinformatiquement.

- Localiser les effecteurs candidats dans la cellule de l'hôte

Au cours de cette étude nous avons utilisé la technique d'immunofluorescence afin de localiser certains effecteurs au sein de feuilles infectées. Par cette approche, nous avons notamment pu

localiser certaines protéines au niveau de la périphérie de l'haustorium. L'absence de marquage à l'intérieur de la cellule hôte pourrait être liée à la trop faible concentration de la protéine dans la cellule de la plante. Il semble essentiel d'utiliser la technique d'immunogold, permettant une localisation plus précise de la protéine à un niveau sub-cellulaire, afin de vérifier si ces SSPs sont exportées ou non à l'intérieur de la cellule hôte où elles pourraient y accomplir leurs fonctions.

- Identifier les interactants des effecteurs candidats chez la plante par le système Y2H

L'identification d'effecteurs candidats au sein du sécrétome de *M. larici-populina* offre maintenant la possibilité de rechercher leurs interactants chez la plante. Au cours de ma thèse, j'ai également préparé une banque d'ADNc Y2H (Yeast Two Hybrid) enrichie en fragments d'ADNc pleine longueur issus de spores de *M. larici-populina* et de feuilles infectées. Pour réaliser cette étude, les levures seront dans un premier temps transformées avec la banque d'ADNc Y2H puis avec le vecteur contenant le gène d'intérêt. Un criblage est ensuite réalisé afin d'identifier les colonies où l'interaction protéine-protéine a eu lieu. Cette interaction se traduit chez la levure par l'expression d'un gène rapporteur permettant la visualisation des levures exprimant cette interaction. Cette analyse peut potentiellement conduire à l'identification des cibles végétales où d'interactants du côté fongique ou encore déboucher sur l'identification de facteurs d'avirulence si un interactant de l'hôte (Gène de résistance) est utilisé à la place de l'effecteur.

- Analyser le polymorphisme de séquence des effecteurs candidats au sein de l'espèce *M. larici-populina*

Au cours de cette étude, nous avons identifié des effecteurs candidats sous pression de sélection en nous basant sur des groupes de gènes paralogues. Il semble indispensable d'analyser le polymorphisme de séquence au niveau intraspécifique en utilisant un grand nombre de souches de *M. larici-populina* comme nous l'avons réalisé pour *Mlp-AvrL567* et *Mlp-SSP15*. Comme nous disposons au sein de l'UMR1136 (Pascal Frey, équipe Ecologie des Champignons Pathogènes Forestiers) d'une large collection de nombreuses souches présentant des combinaisons de virulences différentes (pathotypes différents), il serait possible de mettre en relation un polymorphisme observé sur la séquence nucléotidique avec les virulences du champignon. Par exemple, si toutes les souches avirulentes 1 « Avr1 » ont toutes un polymorphisme à une position donnée qui n'est pas partagé par les autres souches possédant la virulence 1 « Vir1 », l'effecteur peut potentiellement représenter le facteur d'avirulence Avr1. En plus de pouvoir

conduire à l'identification de facteurs d'avirulence potentiels, cette approche peut également confirmer une pression de sélection positive des effecteurs candidats au niveau inter-souches. Cette approche peut être réalisée en définissant des amorces spécifiques pour un effecteur donné. Une approche globale pourrait également être envisagée avec le séquençage du génome de plusieurs souches dicaryotiques de *M. larici-populina*. Ainsi, le polymorphisme de séquence pourrait être établi à une plus large échelle (tous les gènes codant des SSPs) mais aussi les événements de type délétion/insertion ou encore de duplication pouvant eux aussi contribuer à l'émergence de nouveaux effecteurs. Cette approche est envisagée dans le cadre d'un projet ANR PopRust (S. Duplessis) qui débutera en janvier 2011.

- **Transformer génétiquement *M. larici-populina***

Le système de transformation stable développé chez *M. lini* (Lawrence et al. 2010), basé sur l'inhibition génique post transcriptionnelle d'un gène (gene silencing), pourrait être applicable à *M. larici-populina*. Elle pourrait notamment permettre de valider la fonction des effecteurs candidats identifiés au cours de cette étude et conduire à l'identification de facteurs d'avirulence. Des collaborateurs québécois du *Melampsora* Genome Consortium cherchent actuellement à réaliser cette transformation (P. Tanguay, CFS).

- **Valider l'implication des motifs RxLR-like dans la translocation des effecteurs candidats de *M. larici-populina* chez l'hôte**

Certains effecteurs candidats identifiés au cours de cette étude possèdent un domaine réminiscent du domaine RxLR qui pourrait permettre leur translocation à l'intérieur de la cellule hôte. Afin de tester cette hypothèse, des protéines de synthèse ou recombinantes visualisables en microscopie confocale et possédant ou non ce motif, pourraient être mises en contact avec des cultures cellulaires de peuplier afin de valider ou non l'export de l'effecteur dans la cellule hôte. En plus de permettre de visualiser le compartiment végétal ciblé par les effecteurs (noyau, chloroplaste, mitochondrie ...), l'absence de translocation chez les protéines mutées pour ce domaine pourrait confirmer la fonction de tels motifs dans l'export des effecteurs de *M. larici-populina*. Le même type d'approche chez des cellules ou tissus de plantes non-hôtes pourrait permettre de révéler des fonctions cellulaires ciblées par les effecteurs d'agent de rouilles qui seraient conservées chez les plantes.

Références bibliographiques

Références bibliographiques

- Abramovitch RB, Anderson JC, Martin GB** (2006) Bacterial elicitation and evasion of plant innate immunity. *Nature Reviews Molecular Cell Biology* **7**: 601-611.
- AFOCEL** (2006) Synergie et concurrence entre utilisation énergétique du bois et autres usages : état des lieux, évolution et propositions d'amélioration. In Rapport Provisoire.
- AFOCEL** (2006) Biomasse Normandie, CTBA. 60p.
- Ahmadian A, Ehn M, Hober S** (2006) Pyrosequencing: history, biochemistry and future. *Clinica Chimica Acta* **363**: 83-94.
- Aime MC** (2006) Toward resolving family level relationships in rust fungi (*Uredinales*). *Mycoscience* **47**: 112-122.
- Alexopoulos CJ, Mims CW, Blackwell M** (1996) Introductory Mycology, 4th edn. New York: Wiley.
- Alfano JR, Collmer A** (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annual Review of Phytopathology* **42**: 385-414.
- Alfano JR** (2009) Roadmap for future research on plant pathogen effectors. *Molecular Plant Pathology* **10**: 805-813.
- Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, Beynon JL** (2004) Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. *Science* **306**: 1957-1960.
- Angot A, Peeters N, Lechner E, Vaillau F, Baud C, Gentzbittel L, Sartorel E, Genschik P, Boucher C, Genin S** (2006) *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. *Proceedings of the National Academy of Sciences, USA* **103**: 14620-14625.
- Armstrong MR, Whisson SC, Pritchard L, Bos JI, Venter E, Avrova AO, Rehmany AP, Böhme U, Brooks K, Cherevach I, Hamlin N, White B, Fraser A, Lord A, Quail MA, Churcher C, Hall N, Berriman M, Huang S, Kamoun S, Beynon JL, Birch PR** (2005) An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognised in the host cytoplasm. *Proceedings of the National Academy of Sciences, USA* **102**: 7766-7771.
- Avrova AO, Venter E, Birch PR, Whisson SC** (2003) Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. *Fungal Genetics and Biology* **40**: 4-14.
- Azaiez A, Boyle B, Levée V, Séguin A** (2009) Transcriptome profiling in hybrid poplar following interactions with *Melampsora* rust fungi. *Molecular Plant-Microbe Interactions* **22**: 190-200.

- Balestrini R, Gomez-Ariza J, Klink VP, Bonfante P** (2009) Application of Laser Microdissection to plant pathogenic and symbiotic interactions. *Journal of Plant Interactions* **4**: 81-92.
- Barrès B, Dutech C, Andrieux A, Caron H, Pinon J, Frey P** (2006) Isolation and characterization of 15 microsatellite loci in the poplar rust fungus, *Melampsora larici-populina*, and cross-amplification in related species. *Molecular Ecology Notes* **6**: 60-64.
- Barrès B, Halkett F, Dutech C, Andrieux A, Pinon J, Frey P** (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in Europe and recent founder effects overseas. *Infection, Genetics and Evolution* **8**: 577-587.
- Barrett LG, Thrall PH, Dodds PN, van der Merwe M, Linde CC, Lawrence GJ, Burdon JJ** (2009) Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. *Molecular Biology and Evolution* **26**: 2499-2513.
- Bauer R, Begerow D, Sampaio JP, Weiß M, Oberwinkler F** (2006) The simple-septate basidiomycetes: a synopsis. *Mycological Progress* **5**: 41-66.
- Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, Thines M, Ah-Fong A, Anderson R, Badejoko W, Bittner-Eddy P, Boore JL, Chibucos MC, Coates M, Dehal P, Delehaunty K, Dong S, Downton P, Dumas B, Fabro G, Fronick C, Fuerstenberg SI, Fulton L, Gaulin E, Govers F, Hughes L, Humphray S, Jiang RH, Judelson H, Kamoun S, Kyung K, Meijer H, Minx P, Morris P, Nelson J, Phuntumart V, Qutob D, Rehmany A, Rougon-Cardoso A, Ryden P, Torto-Alalibo T, Studholme D, Wang Y, Win J, Wood J, Clifton SW, Rogers J, Van den Ackerveken G, Jones JD, McDowell JM, Beynon J, Tyler BM** (2010) Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* **330**: 1549-1551.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S** (2004) Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**: 783-795.
- Bent AF, Mackey D** (2007) Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology* **45**: 399-436.
- Berthelot A, Chantre G** (2004) Poplar stands in France - An essential link in the forestry industry. *Biofutur* **247**: 20-23.
- Bhattacharjee S, Hiller NL, Liolios K, Win J, Kanneganti TD, Young C, Kamoun S, Haldar K** (2006) The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PLoS Pathogens* **2**: e50.
- Birch PR, Boevink PC, Gilroy EM, Hein I, Pritchard L, Whisson SC** (2008) Oomycete RXLR effectors: delivery, functional redundancy and durable disease resistance. *Current Opinion in Plant Biology* **11**: 373-379.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U**

(2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509-12.

Boddey JA, Hodder AN, Günther S, Gilson PR, Patsiouras H, Kapp EA, Pearce JA, de Koning-Ward TF, Simpson RJ, Crabb BS, Cowman AF (2010) An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463**: 627-31.

Böhnert HU, Fudal I, Dioh W, Tharreau D, Notteghem JL, Lebrun MH (2004) A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *Plant Cell* **16**: 2499-2513.

Boller T, Felix GA (2009) Renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**: 379-406.

Bonas U, Stall RE, Staskawicz B (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular Genetics and Genomics* **218**: 127-136.

Bos JI, Kanneganti TD, Young C, Cakir C, Huitema E, Win J, Armstrong MR, Birch PR, Kamoun S (2006) The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant Journal* **48**: 165-76.

Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, Taylor RM, Zhendong T, Engelhardt S, Vetukuri RR, Harrower B, Dixelius C, Bryan G, Sadanandom A, Whisson SC, Kamoun S, Birch PR (2010) *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences, USA* **107**: 9909-9914.

Both M, Csukai M, Stumpf MP, Spanu PD (2005a) Gene expression profiles of *Blumeria graminis* indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. *Plant Cell* **17**: 2107-2122.

Both M, Eckert SE, Csukai M, Mueller E, Dimopoulos G, Spanu PD (2005b). Transcript profiles of *Blumeria graminis* development during infection reveal a cluster of genes that are potential virulence determinants. *Molecular Plant-Microbe Interactions* **18**: 125-133.

Boyle B, Hamelin RC, Seguin A (2005) *In vivo* monitoring of obligate biotrophic pathogen growth by kinetic PCR. *Applied and Environmental Microbiology* **71**: 1546-1552.

Boyle B, Levée V, Hamel LP, Nicole MC, Séguin A (2010) Molecular and histochemical characterisation of two distinct poplar *Melampsora* leaf rust pathosystems. *Plant Biology* **12**: 364-376.

Brefort T, Doehlemann G, Mendoza-Mendoza A, Reissmann S, Djamei A, Kahmann R (2009) *Ustilago maydis* as a Pathogen. *Annual Review of Phytopathology* **47**: 423-445.

- Bretz JR, Mock NM, Charity JC, Zeyad S, Baker CJ, Hutcheson SW** (2003) A translocated protein tyrosine phosphatase of *Pseudomonas syringae* pv. *tomato* DC3000 modulates plant defence response to infection. *Molecular Microbiology* **49**: 389-400.
- Broeker K, Bernard F, Moerschbacher BM** (2006) An EST library from *Puccinia graminis* f. sp. *tritici* reveals genes potentially involved in fungal differentiation. *FEMS Microbiology Letters* **256**: 273-281.
- Brunner PC, Keller N, McDonald BA** (2009) Wheat domestication accelerated evolution and triggered positive selection in the β -xylosidase enzyme of *Mycosphaerella graminicola*. *PLoS ONE* **4**: e7884.
- Burdon JJ, Thrall PH** (2009) Coevolution of plants and their pathogens in natural habitats. *Science* **324**: 755-756.
- Cai DG, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, KleinLankhorst RM, Salentijn EMJ, Lange W, Stiekema WJ, Wyss U, Grundler FMW, Jung C** (1997) Positional cloning of a gene for nematode resistance in sugar beet. *Science* **275**: 832-834
- Cantrill LC, Deverall BJ** (1993) Isolation of haustoria from wheat leaves infected by the leaf rust fungus. *Physiological and Molecular Plant Pathology* **42** :337-341
- Catanzariti A-M, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis JG** (2006) Haustorially-expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* **18**: 243-256.
- Catanzariti AM, Dodds PN, Ellis JG** (2007) Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiology Letters* **269**: 181-188.
- Catanzariti AM, Dodds PN, Ve T, Kobe B, Ellis JG, Staskawicz BJ** (2010) The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein. *Molecular Plant-Microbe Interactions* **23**: 49-57.
- Chandran D, Inada N, Hather G, Kleindt CK, Wildermuth MC** (2010) Laser microdissection of Arabidopsis cells at the powdery mildew infection site reveals site-specific processes and regulators. *Proceedings of the National Academy of Sciences, USA* **107**: 460-465.
- Chardenon J** (1982) Le peuplier : aujourd'hui et demain. Ed. Institut du développement Forestier, Limoges, France, 279p.
- Cheung F, Win J, Lang JM, Hamilton J, Vuong H, Leach JE, Kamoun S, André Lévesque C, Tisserat N, Buell CR** (2008) Analysis of the *Pythium ultimum* transcriptome using Sanger and Pyrosequencing approaches. *BMC Genomics* **9**: 542.
- Choi J, Park J, Kim D, Jung K, Kang S, Lee YH** (2010) Fungal secretome database: integrated platform for annotation of fungal secretomes. *BMC Genomics* **11**: 105.
- Colcombet J, Hirt H** (2008) Arabidopsis MAPKs: a complex signalling network involved in multiple biological processes. *Biochemical Journal* **413**: 217-26.

- Cooper JI** (1979) Virus diseases of trees and shrubs. Ed. Institute of Terrestrial Ecology, Cambridge, UK, 46-47.
- Cornelis GR** (2006) The type III secretion injectisome. *Nature Reviews Microbiology* **4**: 811-825.
- Crute IR, Holub EB, Burdon JJ** (1997) The gene-for-gene relationship in plant-parasite interactions. New York: CAB International; pp. 359-378.
- Cuomo CA, Güldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Münsterkötter M, Nelson D, O'donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC** (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* **317**: 1400-1402.
- Dangl JL, Jones JDL** (2001) Plant pathogens and integrated defence responses to infection. *Nature* **411**: 826-833.
- Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeier C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW** (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* **434**: 980-986.
- De Bary** (1863) Recherches sur le développement des quelques champignons phytoparasites. *Annales des Sciences Naturelles, Botanique* **20**: 3-148.
- De Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MH, Thomma BP** (2010) Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science* **329**: 953-955.
- Desmond OJ, Manners JM, Schenk PM, Maclean DJ, Kazan K** (2008) Gene expression analysis of the wheat response to infection by *Fusarium pseudograminearum*. *Physiological and Molecular Plant Pathology* **70**: 40-47.
- DeZwaan TM, Carroll AM, Valent B, Sweigard JA** (1999) *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. *Plant Cell* **11**: 2013-2030.
- Dietel P** (1928) *Uredinales*. In: Engler A, Prantl K (eds) Die Natürlichen Pflanzenfamilien, vol 2. Engelmann, Leipzig, pp 24-98.
- Hiratsuka Y, Cummins GB** (1963) Morphology of the spermogonia of the rust fungi. *Mycologia* **55**: 487-507.
- Diguistini S, Liao NY, Platt D, Robertson G, Seidel M, Chan SK, Docking TR, Birol I, Holt RA,**

- Hirst M, Mardis E, Marra MA, Hamelin RC, Bohlmann J, Breuil C, Jones SJ** (2009) De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biology* **10**: R94.
- Cummins GB, Hiratsuka Y** (2003) Illustrated genera of rust fungi, 3rd edn. American Phytopathological Society, St. Paul, MN.
- Dodds PN, Lawrence GJ, Catanzariti A-M, Ayliffe MA, Ellis JG** (2004) The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* **16**: 755-768.
- Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CI, Ayliffe MA, Kobe B, Ellis JG** (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences, USA* **103**: 8888-8893.
- Dodds PN, Rafiqi M, Gan PHP, Hardham AR, Jones DA, Ellis JG** (2009) Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytologist* **183**: 993-1000.
- Doehlemann G, van der Linde K, Assmann D, Schwammbach D, Hof A, Mohanty A, Jackson D, Kahmann R** (2009) Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLoS Pathogens* **5**: e1000290.
- Dou D, Kale SD, Wang X, Chen Y, Wang Q, Wang Y, Jiang RH, Arredondo FD, Anderson RG, Thakur PB** (2008a) Carboxy-terminal motifs common to many oomycete RXLR effectors are required for avirulence and suppression of BAX-mediated programmed cell death by *Phytophthora sojae* effector Avr1b. *Plant Cell* **20**: 1118-1133.
- Dou D, Kale SD, Wang X, Jiang RH, Bruce NA, Arredondo FD, Zhang X, Tyler BM** (2008b) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell* **20**: 1930-1947.
- Dowkiw A** (2003) Analyse génétique de la résistance et de la tolérance de peupliers hybrides *Populus deltoides* x *Populus trichocarpa* aux rouilles foliaires à *Melampsora larici-populina*. Thèse de doctorat, Université d'Orléans.
- Duplessis S, Courty PE, Tagu D, Martin F** (2005) Transcript patterns associated with ectomycorrhiza development in *Eucalyptus globulus* and *Pisolithus microcarpus*. *New Phytologist* **165**: 599-611.
- Duplessis S, Major IT, Martin F, Séguin A** (2009) Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. *Critical Reviews in Plant Sciences* **28**: 309-334.
- Eckenwalder JE** (1996) Systematics and evolution of *Populus*. dans *Biology of Populus*. Stettler RF, Bradshaw HD, Heilman PE, et Hinckley TM (Eds) NRC Research Press, p. 7-32.

- Eitas TK, Dangl JL** (2010) NB-LRR proteins: Pairs, pieces, perception, partners and pathways. *Current Opinion in Plant Biology* **13**: 1-6.
- Ellis JG, Lawrence GJ, Dodds PN** (2007) Further analysis of gene-for-gene disease resistance specificity in flax. *Molecular Plant Pathology* **8**: 103-109.
- Ellis JG, Rafiqi M, Gan P, Chakrabarti A, Dodds PN** (2009) Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. *Current Opinion in Plant Biology* **12**: 399-405.
- Espinosa A, Alfano JR** (2004) Disabling surveillance: bacterial type III secretion system effectors that suppress innate immunity. *Cellular Microbiology* **6**: 1027-1040.
- FAO** (1981) Les maladies des peupliers. Commission Internationale du peuplier. AFOCEL, France, 193 p.
- Farman ML, Leong SA** (1998) Chromosome walking to the *AVR1-CO39* avirulence gene of *Magnaporthe grisea*: discrepancy between the physical and genetic maps. *Genetics* **150**: 1049-1058.
- Farman ML, Eto Y, Nakao T, Tosa Y, Nakayashiki H, Mayama S, Leong SA** (2002) Analysis of the structure of the *AVR1-CO39* avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Molecular Plant-Microbe Interactions* **15**: 6-16.
- Feau N, Bergeron MJ, Joly D, Roussel F, Hamelin R** (2007) Detection and validation of EST-derived SNPs for poplar leaf rust *Melampsora medusae* f. sp. *deltoidea*. *Molecular Ecology Notes* **7**: 1222-1228.
- Felten J, Kohler A, Morin E, Bhalerao RP, Palme K, Martin F, Ditengou FA, Legué V** (2009) The ectomycorrhizal fungus *Laccaria bicolor* stimulates lateral root formation in poplar and *Arabidopsis* through auxin transport and signaling. *Plant Physiology* **151**: 1991-2005.
- Flor HH** (1947) Inheritance of réaction to rust in flax. *Journal of Agricultural Research* **74**: 241-262.
- Flor HH** (1956) The complementarity genetic systems in flax and flax rust. *Advances in genetics* **8**: 29-54.
- Flor HH** (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**: 275-296.
- Frey P, Pinon J** (2004) La rouille du peuplier: un pathosystème modèle. *Biofutur* **247**: 28-32.
- Frey P, Gérard P, Feau N, Husson C, Pinon J** (2005) Variability and population of *Melampsora* rusts on poplars. In *Rust diseases on willow and poplar*. Pei, M. and McCracken, A. R., eds. Wallingford, UK, CAB International: p. 63-72.
- Fu ZQ, Guo M, Jeong BR, Tian F, Elthon TE, Cerny RL, Staiger D, Alfano JR** (2007) A type III effector ADPribosylates RNA-binding proteins and quells plant immunity. *Nature* **447**: 284-288.
- Fudal I, Ross S, Gout L, Blaise F, Kuhn ML, Eckert MR, Cattolico L, Bernard-Samain S,**

- Balesdent MH, Rouxel T** (2007) Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. *Molecular Plant-Microbe Interactions* **20**: 459-470.
- Galán JE, Wolf-Watz H** (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**: 567-573.
- Gaulin E, Madoui MA, Bottin A, Jacquet C, Mathé C, Couloux A, Wincker P, Dumas B** (2008) Transcriptome of *Aphanomyces euteiches*: New Oomycete Putative Pathogenicity Factors and Metabolic Pathways. *PLoS ONE* **3**: e1723.
- Génin S** (2010) Molecular traits controlling host range and adaptation to plants in *Ralstonia solanacearum*. *New Phytologist* **187**:920-928.
- Gérard PR, Husson C, Pinon J, Frey P** (2006) Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathology* **96**: 1027-1036.
- Gibrat JF, Marin A** (2004) Stratégies permettant d'assigner une fonction *in silico* aux séquences protéiques issues de projets de génomique. La génomique en biologie végétale, J-F. Morot-Gaudry, J-F. Briat, coord. Editions INRA: p. 91-113.
- Giles PF, Soanes DM, Talbot NJ** (2003). A relational database for the discovery of genes encoding amino acid biosynthetic enzymes in pathogenic fungi. *Comparative and Functional Genomics* **4**: 4-15.
- Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, Thordal-Christensen H** (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics* **11**: 317.
- Göhre V, Robatzek S** (2008) Breaking the barriers: microbial effector molecules subvert plant immunity. *Annual Review of Phytopathology* **46**: 189-215.
- Gout L, Kuhn ML, Vincenot L, Bernard-Samain S, Cattolico L, Barbetti M, Moreno-Rico O, Balesdent MH, Rouxel T** (2007) Genome structure impacts molecular evolution at the *AvrLm1* avirulence locus of the plant pathogen *Leptosphaeria maculans*. *Environmental Microbiology* **9**: 2978-2992.
- Gout L, Fudal I, Kuhn ML, Blaise F, Eckert M, Cattolico L, Balesdent MH, Rouxel T** (2006) Lost in the middle of nowhere: the *AvrLm1* avirulence gene of the Dothideomycete *Leptosphaeria maculans*. *Molecular Microbiology* **60**: 67-80.
- Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes RW, Dangl JL** (1995) Structure of the Arabidopsis *RPM1* gene enabling dual specificity disease resistance. *Science* **269**: 843-846.

- Guttman DS, Gropp SJ, Morgan RL, Wang PW** (2006) Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae*. *Molecular Biology and Evolution* **23**: 2342-2354.
- Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AM, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JI, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, Fischbach MA, Fugelstad J, Gilroy EM, Gnerre S, Green PJ, Grenville-Briggs LJ, Griffith J, Grünwald NJ, Horn K, Horner NR, Hu CH, Huitema E, Jeong DH, Jones AM, Jones JD, Jones RW, Karlsson EK, Kunjeti SG, Lamour K, Liu Z, Ma L, Maclean D, Chibucos MC, McDonald H, McWalters J, Meijer HJ, Morgan W, Morris PF, Munro CA, O'Neill K, Ospina-Giraldo M, Pinzón A, Pritchard L, Ramsahoye B, Ren Q, Restrepo S, Roy S, Sadanandom A, Savidor A, Schornack S, Schwartz DC, Schumann UD, Schwessinger B, Seyer L, Sharpe T, Silvar C, Song J, Studholme DJ, Sykes S, Thines M, van de Vondervoort PJ, Phuntumart V, Wawra S, Weide R, Win J, Young C, Zhou S, Fry W, Meyers BC, van West P, Ristaino J, Govers F, Birch PR, Whisson SC, Judelson HS, Nusbaum C** (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**: 393-398.
- Hacquard S, Delaruelle C, Legué V, Tisserant E, Kohler A, Frey P, Martin F, Duplessis S** (2010) Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Molecular Plant-Microbe Interactions* **23**: 1275-1286.
- Hacquard S, Veneault-Fourrey C, Delaruelle C, Frey P, Martin F, Duplessis S** (2011) Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection of poplar leaves *Physiological and Molecular Plant Pathology* doi:10.1016/j.pmpp.2010.10.003.
- Hahn M, Mendgen K** (1992) Isolation by ConA binding of haustoria from different rust fungi and comparison of their surface qualities. *Protoplasma* **170**: 95-103.
- Hahn M, Jungling S, Knogge W** (1993) Cultivar-specific elicitation of barley defense reactions by the phytotoxic peptide NIP1 from *Rhynchosporium secalis*. *Molecular Plant-Microbe Interactions*. **6**: 745-754.
- Hahn M, Mendgen K** (1997a). Characterization of *in planta*-induced rust genes isolated from a haustorium-specific cDNA library. *Molecular Plant-Microbe Interactions* **10**: 427-437.
- Hahn M, Neef U, Struck C, Göttfert M, Mendgen K** (1997b) A putative amino acid transporter is specifically expressed in haustoria of the rust fungus *Uromyces fabae*. *Molecular Plant-Microbe Interactions* **10**: 438-445.

Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, Thorsten Lumbsch H, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer R, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai YC, Gams W, Geiser DM, Griffith GW, Gueidan C, Hawksworth DL, Hestmark G, Hosaka K, Humber RA, Hyde KD, Ironside JE, Kõljalg U, Kurtzman CP, Larsson KH, Lichtwardt R, Longcore J, Miadlikowska J, Miller A, Moncalvo JM, Mozley-Standridge S, Oberwinkler F, Parmasto E, Reeb V, Rogers JD, Roux C, Ryvarden L, Sampaio JP, Schüssler A, Sugiyama J, Thorn RG, Tibell L, Untereiner WA, Walker C, Wang Z, Weir A, Weiss M, White MM, Winka K, Yao YJ, Zhang N (2007) A higher-level phylogenetic classification of the Fungi. *Mycological Research* **111**: 509-547.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez- Estraño C, Haldar K (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**: 1934-1937.

Hiratsuka Y, Hiratsuka N (1980) Morphology of spermogonia and taxonomy of rust fungi. *Reports of the Tottori Mycological Institute* **18**: 257-268.

Hotson A, Mudgett MB (2004) Cysteine proteases in phytopathogenic bacteria: identification of plant targets and activation of innate immunity. *Current Opinion in Plant Biology* **7**: 384-390.

Houterman PM, Speijer D, Dekker HL, De Koster CG, Cornelissen BJC, Rep M (2007). The mixed xylem sap proteome of *Fusarium oxysporum*-infected tomato plants. *Molecular Plant Pathology* **8**: 215-221.

Houterman PM, Cornelissen BJ, Rep M (2008) Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathogens* **4**: e1000061.

Houterman P, Ma LS, van Ooijen G, de Vroomen M, Cornelissen B, Takken F, Rep M (2009) The effector protein Avr2 of the xylem colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. *Plant Journal* **58**: 970-978.

Janjusevic R, Abramovitch RB, Martin GB, Stebbins CE (2006) A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* **311**: 222-226.

Jansson S, Douglas CJ (2007) *Populus*: a model system for plant biology. *Annual Review of Plant Biology* **58**: 435-458.

Jha G, Rajeshwari R, Sonti RV (2007) Functional interplay between two *Xanthomonas oryzae* pv., *oryzae* secretion systems in modulating virulence on rice. *Molecular Plant-Microbe Interactions* **20**: 31-40.

Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO Journal* **19**: 4004-4014.

Jiang RH, Tripathy S, Govers F, Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members.

Proceedings of the National Academy of Sciences, U S A **105**: 4874-4879.

Joly DL, Feau N, Tanguay P, Hamelin RC (2010) Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (*Melampsora* spp.). *BMC Genomics* **11**: 422.

Jones JDG, Dangl JL (2006) The plant immune system. *Nature* **444**: 323-329.

Joosten MHAJ, de Wit PJGM (1999) The tomato-*Cladosporium fulvum* interaction: a versatile experimental system to study plant-pathogen interactions. *Annual Review of Phytopathology* **37**: 335-367.

Jakupović M, Heintz M, Reichmann P, Mendgen K, Hahn M (2006) Microarray analysis of expressed sequence tags from haustoria of the rust fungus *Uromyces fabae*. *Fungal Genetics and Biology* **43**: 8-19.

Jülich W (1981) Higher taxa of Basidiomycetes. *Bibliotheca Mycologia* **85**: 1-483

Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Annual Review of Phytopathology* **44**: 41-60.

Kale SD, Gu B, Capelluto DG, Dou D, Feldman E, Rumore A, Arredondo FD, Hanlon R, Fudal I, Rouxel T, Lawrence CB, Shan W, Tyler BM (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* **142**: 284-295.

Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O, Perlin MH, Wösten HAB, de Vries R, Ruiz-Herrera J, Reynaga-Peña CG, Snetselaar K, McCann M, Pérez-Martín J, Feldbrügge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, González-Prieto JM, Kennel JC, Molina L, Schirawski J, Mendoza-Mendoza A, Greilinger D, Münch K, Rössel N, Scherer M, Vraneš M, Ladendorf O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho ECH, Cahill MJ, Boyce KJ, Weixi Li J, Sanchez-Alonso P, Schreier PH, Häuser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schlüter T, Margolis J, Platt D, Swimmer C, Gnrirke A, Chen F, Vysotskaia V, Mannhaupt G, Güldener U, Münsterkötter M, Haase D, Oesterheld M, Mewes H-W, Mauceli EW, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan JE, Birren BW (2006) Insights from the genome of the biotrophical fungal plant pathogen *Ustilago maydis*. *Nature* **444**: 97-101.

Kang S, Sweigard JA, Valent B (1995) The *PWL* host specificity gene family in the blast fungus *Magnaporthe grisea*. *Molecular Plant-Microbe Interactions* **8**: 939-948.

Kawchuk LM, Hachey J, Lynch DR, Kulcsar F, Van Rooijen G, Waterer DR, Robertson A, Kokko E, Byers R, Howard RJ, Fischer R, Pruffer D (2001) Tomato *Ve* disease resistance genes encode cell surface-like receptors. *Proceedings of the National Academy of Sciences, USA* **98**: 6511-6515.

Kay S, Hahn S, Marois E, Hause G, Bonas U (2007) A bacterial effector acts as a plant

transcription factor and induces a cell size regulator. *Science* **318**: 648-651.

Kebdani N, Pieuchot L, Deleury E, Panabières F, Le Berre JY, Gourgues M (2010) Cellular and molecular characterization of *Phytophthora parasitica* appressorium-mediated penetration. *New Phytologist* **185**: 248-257.

Kemen E, Kemen AC, Rafiqi M, Hempel U, Mendgen K, Hahn M, Voegelé RT (2005) Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Molecular Plant-Microbe Interactions* **18**: 1130-1139.

Khang CH, Park SY, Lee YH, Valent B, Kang S (2008) Genome organization and evolution of the AVR-pita avirulence gene family in the *Magnaporthe grisea* species complex. *Molecular Plant-Microbe Interactions* **21**: 658-670.

Khang CH, Berruyer R, Giraldo MC, Kankanala P, Park SY, Czymmek K, Kang S, Valent B (2010) Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. *Plant Cell* **22**: 1388-403.

Koh S, André A, Edwards H, Ehrhardt D, Somerville S (2005) *Arabidopsis thaliana* subcellular responses to compatible *Erysiphe cichoracearum* infections. *Plant Journal* **44**: 516-529.

Kohler A, Rinaldi C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, Meyers BC, Boerjan W, Martin F (2008) Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Molecular Biology* **66**: 619-636.

Kolmer JA, Ordonez ME, Groth JV (2009) The Rust Fungi. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0021264.

Kuck KH, Tiburzy R, Hänßler G, Reisener HJ (1981) Visualization of rust haustoria in wheat leaves by using fluorochromes. *Physiological and Molecular Plant Pathology* **19**: 439-441.

Lamoril J, Ameziane N, Daybach JC, Bouizegarène P, Bogard M (2008) Les techniques de séquençage de l'ADN: Une révolution en marche. Première partie. *Immuno-analyse et biologie spécialisée* **23**: 260-279.

Lamour KH, Win J, Kamoun S (2007) Oomycete genomics: new insights and future directions. *FEMS Microbiology Letters* **274**: 1-8.

Lanier L, Joly P, Bondoux P, Bellemère A (1976) Mycologie et pathologie forestières. II Pathologie forestière. Masson, Paris, France, 332-377.

Laugé R, Goodwin PH, De Wit PJGM, Joosten MHAJ (2000) Specific HR-associated recognition of secreted proteins from *Cladosporium fulvum* occurs in both host and non-host plants. *Plant Journal* **23**: 735-745.

Laurans F, Pilate G (1999) Histological aspects of a hypersensitive response in poplar to *Melampsora larici-populina*. *Phytopathology* **89**: 233-238.

Lawrence GJ, Mayo GME, Shepherd KW (1981) Interactions between genes controlling pathogenicity in the flax rust fungus. *Phytopathology* **71**: 12-19.

- Lawrence GJ, Finnegan EJ, Ayliffe MA, Ellis JG** (1995) The *L6* Gene for Flax Rust Resistance Is Related to the *Arabidopsis* Bacterial-Resistance Gene *Rps2* and the Tobacco Viral Resistance Gene-*N*. *Plant Cell* **7**: 1195-1206.
- Lawrence GJ, Dodds PN, Ellis JG** (2007) Rust of flax and linseed caused by *Melampsora lini*. *Molecular Plant Pathology* **8**: 349-364.
- Lawrence GJ, Dodds PN, Ellis JG** (2010) Transformation of the flax rust fungus, *Melampsora lini*: selection via silencing of an avirulence gene. *Plant Journal* **61**: 364-369.
- Leister RT, Ausubel FM, Katagiri F** (1996) Molecular recognition of pathogen attack occurs inside of plant cells in plant disease resistance specified by the *Arabidopsis* genes RPS2 and RPM1. *Proceedings of the National Academy of Sciences, USA* **93**: 15497-15502.
- Leonard KJ, Szabo LJ** (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Molecular Plant Pathology* **6**: 99-111.
- Lévesque CA, Brouwer H, Cano L, Hamilton JP, Holt C, Huitema E, Raffaele S, Robideau GP, Thines M, Win J, Zerillo MM, Beakes GW, Boore JL, Busam D, Dumas B, Ferriera S, Fuerstenberg SI, Gachon CM, Gaulin E, Govers F, Grenville-Briggs L, Horner N, Hostetler J, Jiang RH, Johnson J, Krajaejun T, Lin H, Meijer HJ, Moore B, Morris P, Phuntmart V, Puiu D, Shetty J, Stajich JE, Tripathy S, Wawra S, van West P, Whitty BR, Coutinho PM, Henrissat B, Martin F, Thomas PD, Tyler BM, De Vries RP, Kamoun S, Yandell M, Tisserat N, Buell CR** (2010) Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology* **11**: R73.
- Libault M, Wan J, Czechowski T, Udvardi M, Stacey G** (2007) Identification of 118 *Arabidopsis* transcription factor and 30 ubiquitin-ligase genes responding to chitin, a plant defense elicitor. *Molecular Plant-Microbe Interactions* **20**: 900-911.
- Lievens B, Brouwer M, Vanachter A, Cammue BPA, Thomma B** (2006) Real-time PCR for detection and quantification of fungal and oomycete tomato pathogens in plant and soil samples. *Plant Science* **171**: 155-165.
- Liu Z, Bos JI, Armstrong M, Whisson SC, da Cunha L, Torto-Alalibo T, Win J, Avrova AO, Wright F, Birch PR, Kamoun S** (2005) Patterns of diversifying selection in the phytotoxin-like *scr74* gene family of *Phytophthora infestans*. *Molecular Biology and Evolution* **22**: 659-672.
- Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie X, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RA, Chapman S, Coulson R, Coutinho PM, Danchin EG, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC,**

Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**: 367-73.

Ma W, Guttman DS (2008) Evolution of prokaryotic and eukaryotic virulence effectors. *Current Opinion in Plant Biology* **11**: 412-419.

Mackey D, Holt BF, Wiig A, Dangl JL (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**: 743-754.

McDonald BA (2004) Population Genetics of Plant Pathogens. The Plant Health Instructor. DOI:10.1094/PHI-A-0524-01.

McDowell JM, Woffenden BJ (2003) Plant disease resistance genes: recent insights and potential applications. *Trends Biotechnology* **21**: 178-183.

McMaugh SJ, Lyon BR (2003) Real-time quantitative RT-PCR assay of gene expression in plant roots during fungal pathogenesis. *BioTechniques* **34**: 982-986.

Maier W, Begerox D, WeiB M, Oberwinkler F (2003) Phylogeny of rust fungi: An approach using nuclear large subunit ribosomal DNA sequences. *Canadian Journal of Botany* **81**: 12-23.

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**: 1930-1933.

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganai MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**: 1432-1436.

Martin F, Tuskan GA, DiFazio SP, Lammers P, Newcombe G, Podila GK (2004) Symbiotic sequencing for the *Populus* mesocosm. *New Phytologist* **161**: 330-335.

Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rahashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kües U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008a) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**: 88-92.

Martin F, Gianinazzi-Pearson V, Hijri M, Lammers P, Requena N, Sanders IR, Shachar-Hill Y, Shapiro H, Tuskan GA, Young JPW (2008b). The long hard road to a completed *Glomus intraradices* genome. *New Phytologist* **180**: 747-750.

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R, Porcel B, Rubini A, Amicucci A, Amselem J, Anthouard V, Arcioni S, Artiguenave F, Aury JM, Ballario P, Bolchi A, Brenna A, Brun A, Buée M, Cantarel B, Chevalier G, Couloux A, Da Silva C, Denoeud F, Duplessis S, Ghignone S, Hilselberger B, Iotti M, Marçais B, Mello A, Miranda M, Pacioni G, Quesneville H, Riccioni C, Ruotolo R, Splivallo R, Stocchi V, Tisserant E, Viscomi AR, Zambonelli A, Zampieri E, Henrissat B, Lebrun MH, Paolocci F, Bonfante P, Ottonello S, Wincker P (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**: 1033-1038.

Marzluf GA 1997 Genetic regulation of nitrogen metabolism in the fungi. *Microbiology and Molecular Biology Reviews* **61**: 17-32.

Medina M, Sachs JL (2010) Symbiont genomics, our new tangled bank. *Genomics* **95**: 129-137.

Melotto M, Underwood W, Koczan J, Nomura K, He SY (2006) Plant stomata function in innate immunity against bacterial invasion. *Cell* **126**: 969-980.

Mendgen K, Struck C, Voegelé RT, Hahn M (2000) Biotrophy and rust haustoria. *Physiological and Molecular Plant Pathology* **56**: 141-145.

Mendgen K, Hahn M (2002) Plant infection and the establishment of fungal biotrophy. *Trends in Plant Science* **7**: 352-356.

Metzker ML (2010) Sequencing technologies - the next generation. *Nature Reviews Genetics* **11**: 31-46.

Mims CW, Rodriguez-Lothar C, Richardson EA (2002) Ultrastructure of the host-pathogen interface in daylily leaves infected by the rust fungus *Puccinia hemerocallidis*. *Protoplasma* **219**: 221-226.

Miot S, Frey P, Pinon J (1999) Varietal mixture of poplar clones: Effects on infection by *Melampsora larici-populina* and on plant growth. *European Journal of Forest Pathology* **29**: 411-423.

Miranda M, Ralph SG, Mellway R, White R, Heath MC, Bohlmann J, Constabel CP (2007) The transcriptional response of hybrid poplar (*Populus trichocarpa* x *P. deltoides*) to infection by *Melampsora medusae* leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. *Molecular Plant Microbe Interactions* **20**: 816-831.

Mosquera G, Giraldo MC, Hyun Khang C, Coughlan S, Valent B (2009) Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy-associated secreted proteins in rice blast disease. *Plant Cell* **21**: 1273-1290.

Mueller O, Kahmann R, Aguilar G, Trejo-Aguilar B, Wu A, de Vries RP (2008) The secretome of the maize pathogen *Ustilago maydis*. *Fungal Genetics and Biology* Suppl 1:S63-70.

Nagarajan S, Singh DV (1990) Long-distance dispersion of rust pathogens. *Annual Review of Plant Pathology* **28**: 139-153.

- Nicolas P, Chiapello H** (2004) Prédiction de gènes. dans La génomique en biologie végétale, J-F. Morot-Gaudry, J-F. Briat, coord. Editions INRA: p. 77-89.
- Nicastro G, Orsomando G, Ferrari E, Manconi L, Desario F, Amici A, Naso A, Carpaneto A, **Pertinhez TA, Ruggieri S, Spisni A** (2009) Solution structure of the phytotoxic protein PcF: the first characterized member of the *Phytophthora* PcF toxin family. *Protein Science* **18**: 1786-1791.
- Nicole MC, Hamel LP, Morency MJ, Beaudoin N, Ellis BE, Séguin A** (2006) MAP-ping genomic organization and organ-specific expression profiles of poplar MAP kinases and MAP kinase kinases. *BMC Genomics* **7**: 223.
- Oh Y, Donofrio N, Pan H, Coughlan S, Brown DE, Meng S, Mitchell T, Dean RA** (2008) Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus *Magnaporthe oryzae*. *Genome Biology* **9**:R85.
- Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B** (2000) Atelomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. *Plant Cell* **12**:2019–2032.
- Palmer CL, Skinner W** (2002) *Mycosphaerella graminicola*: Latent infection, crop devastation and genomics. *Molecular Plant Pathology* **3**: 63-70.
- Parlange F, Daverdin G, Fudal I, Kuhn ML, Balesdent MH, Blaise F, Grezes-Besset B, Rouxel T** (2009) *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. *Molecular Microbiology* **71**: 851-863.
- Pedley KF, Martin GB** (2005) Role of mitogen-activated protein kinases in plant immunity. *Current Opinion in Plant Biology* **8**: 541-547.
- Pinon J** (1995) Présence en France d'une nouvelle race de *Melampsora larici-populina*, agent de la rouille foliaire des peupliers cultivés. *Revue Forestière Française* **47**: 230-234.
- Pinon J, Frey P** (2005) Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In Pei M. H., McCracken A. R., eds. Rust diseases of Willow and Poplar. CAB International, Wallingford, UK, 139-154.
- Povolotskaya IS, Kondrashov FA** (2010) Sequence space and the ongoing expansion of the protein universe. *Nature* **465**: 922-926.
- Puthoff DP, Neelam A, Ehrenfried ML, Scheffler BE, Ballard L, Song Q, Campbell KB, Cooper B, Tucker ML** (2008) Analysis of expressed sequence tags from *Uromyces appendiculatus* hyphae and haustoria and their comparison to sequences from other rust fungi. *Phytopathology* **98**: 1126-1135.
- Qutob D, Tedman-Jones J, Dong S, Kuflu K, Pham H, Wang Y, Dou D, Kale SD, Arredondo FD, Tyler BM, Gijzen M** (2009) Copy number variation and transcriptional polymorphisms of *Phytophthora sojae* RXLR effector genes *Avr1a* and *Avr3a*. *PLoS ONE* **4**: e5066.
- Rafiqi M, Gan PH, Ravensdale M, Lawrence GJ, Ellis JG, Jones DA, Hardham AR, Dodds PN**

(2010) Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell* **22**: 2017-2032.

Ramiro DA, Escoute J, Petitot AS, Nicole M, Maluf MP, Fernandez D (2009) Biphasic haustorial differentiation of coffee rust (*Hemileia vastatrix* race II) associated with defence responses in resistant and susceptible coffee cultivars. *Plant pathology* **58**: 944-955.

Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, Whisson SC, Kamoun S, Tyler BM, Birch PRJ, Beynon JL (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. *Plant Cell* **17**: 1839-1850.

Rep M, van der Does HC, Meijer M, van Wijk R, Houterman PM, Dekker HL, de Koster CG, Cornelissen BJ (2004) A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology* **53**: 1373-1383.

Rep M (2005) Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiology Letters* **253**: 19-27.

Ridé M (1995) Evolution des maladies et programmes européens d'amélioration du peuplier. *Comptes-rendus de l'Académie d'agriculture de France* **81**: 91-97.

Ridout CJ, Skamnioti P, Porritt O, Sacristan S, Jones JDG, Brown JKM (2006) Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *Plant Cell* **18**:2402-2414.

Rinaldi C, Kohler A, Frey P, Duchaussoy F, Ningre N, Couloux A, Wincker P, Le Thiec D, Fluch S, Martin F, Duplessis S (2007) Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiology* **144**: 347-366.

Rinaldi C (2007) Caractérisation des déterminant géniques impliqués dans l'interaction entre un agent phytopathogène, *Melampsora larici-populina*, et le peuplier. Thèse de doctorat, Université de Nancy.

Rivas S, Thomas CM (2005) Molecular interactions between tomato and the leaf mold pathogen *Cladosporium fulvum*. *Annual Review of Phytopathology* **43**: 395-436.

Ronaghi M, Uhlén M, Nyrén PA (1998) Sequencing method based on real-time pyrophosphate, *Science* **281**: 363-365.

Rooney HC, Van't Klooster JW, van der Hoorn RA, Joosten MH, Jones JD, de Wit PJ (2005) *Cladosporium Avr2* inhibits tomato *Rcr3* protease required for Cf-2-dependent disease resistance. *Science* **308**: 1783-1786.

Rosebrock TR, Zeng L, Brady JJ, Abramovitch RB, Xiao F, Martin GB (2007) A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature* **448**: 370-374.

- Rouhier N, Gelhaye E, Gualberto JM, Jordy MN, De Fay E, Hirasawa M, Duplessis S, Lemaire SD, Frey P, Martin F, Manieri W, Knaff DB, Jacquot JP** (2004) Poplar peroxiredoxin Q. A thioredoxin-linked chloroplast antioxidant functional in pathogen defense. *Plant Physiology* **134**: 1027-1038.
- Rouxel T, Duplessis S, Viaud M, Amselem J, Lebrun MH** (2009) La génomique fongique. *Biofutur* **296**: 31-35.
- Rubin EM** (2008) Genomics of cellulosic biofuels. *Nature* **454**: 841-5.
- Russo I, Babbitt S, Muralidharan V, Butler T, Oksman A, Goldberg DE** (2010) Plasmeppin V licenses Plasmodium proteins for export into the host erythrocyte. *Nature* **463**: 632-6.
- Sacristán S, Vigouroux M, Pedersen C, Skamnioti P, Thordal-Christensen H, Micali C, Brown JK, Ridout CJ** (2009) Coevolution between a family of parasite virulence effectors and a class of LINE-1 retrotransposons. *PLoS One* **4**: e7463.
- Savile DBO** (1953) Short-season adaptation in the rust fungi. *Mycologia* **45**: 75-87.
- Schena L, Nigro F, Ippolito A, Gallitelli D** (2004). Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. *European Journal of Plant Pathology* **110**: 893-908.
- Schornack S, van Damme M, O. Bozkurt T, Cano LM, Smoker M, Thines M, Gaulin E, Kamoun S, Huitema E** (2010) Ancient class of translocated oomycete effectors targets the host nucleus. *Proceedings of the National Academy of Sciences USA*, doi:10.1073/pnas.1008491107.
- Shan W, Cao M, Leung D, Tyler BM** (2004) The Avr1b locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps1b. *Molecular Plant-Microbe Interactions* **17**: 394-403.
- Shan L, He P, Sheen J** (2007) Intercepting host MAPK signaling cascades by bacterial type III effectors. *Cell Host Microbe* **1**: 167-174.
- Shattock RC, Preece TF** (2000) Tranzschel revisited: modern studies of the relatedness of different rust fungi confirm his Law. *Mycologist* **14**: 113-117.
- Skibbe DS, Doehlemann G, Fernandes J, Walbot V** (2010) Maize tumors caused by *Ustilago maydis* require organ-specific genes in host and pathogen. *Science* **328**: 89-92.
- Soanes DM, Talbot NJ** (2008) Moving targets: rapid evolution of oomycete effectors. *Trends in Microbiology* **16**: 507-510.
- Sohn J, Voegelé RT, Mendgen K, Hahn M** (2000) High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*. *Molecular Plant-Microbe Interactions* **13**: 629-636.
- Sohn KH, Lei R, Nemri A, Jones JD** (2007) The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*. *Plant Cell* **19**: 4077-4090.
- Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC** (1997) Evolution of the rice Xa21

disease resistance gene family. *Plant Cell* **9**: 1279-1287.

Spanu P, Kämper J (2010) Genomics of biotrophy in fungi and oomycètes - emerging patterns. *Current Opinion in Plant Biology* **13**: 409-414.

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Ver Loren van Themaat E, Brown JK, Butcher SA, Gurr SJ, Lebrun MH, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjdia M, Bidzinski P, Bindschedler LV, Both M, Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, López-Ruiz FJ, Lu X, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H, Reinhardt R, Rott M, Sacristán S, Schmidt SM, Schön M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M, Wessling R, Wicker T, Panstruga R (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**: 1543-1546.

Staples RC, Macko V (1984) Germination of uredospores and differentiation of infection structures. In: The Cereal Rusts. Vol. 1. Bushnell, W.R. and Roelfs, A.J. eds. Academic Press, New York, pp. 255-289.

Staples RC (2001) Nutrients for a rust fungus: the role of haustoria. *Trends in Plant Science* **6**: 496-498.

Staskawicz BJ, Dahlbeck D, Keen NT (1984) Cloned avirulence gene of *Pseudomonas syringae* pv. *Glycinea* determines race-specific incompatibility on glycine max (L.) Merr. *Proceedings of the National Academy of Sciences, USA* **81**: 6024-6028.

Stépien V (1992) Contribution à l'étude des protéines de réserve végétatives du peuplier (*Populus x euramericana*). Thèse de biologie forestière. Université Henri Poincaré - Nancy I., pp. 145.

Stergiopoulos I, de Wit PJ (2009) Fungal effector proteins. *Annual Review of Phytopathology* **47**: 233-263.

Struck C, Hahn M, Mendgen K (1996) Plasma Membrane H⁺-ATPase Activity in Spores, Germ Tubes, and Haustoria of the Rust Fungus *Uromyces viciae-fabae*; *Fungal Genetics and Biology* **20**: 30-35.

Struck C, Siebels C, Rommel O, Wernitz M, Hahn M (1998) The plasma membrane H⁽⁺⁾-ATPase from the biotrophic rust fungus *Uromyces fabae*: molecular characterization of the gene (*PMA1*) and functional expression of the enzyme in yeast. *Molecular Plant-Microbe Interactions* **11**: 458-465.

Struck C, Ernst M, Hahn M (2002) Characterization of a developmentally regulated amino acid transporter (AAT1p) of the rust fungus *Uromyces fabae*. *Molecular Plant Pathology* **3**: 23-30.

- Struck C, Mueller E, Martin H, Lohaus G** (2004) The *Uromyces fabae* UfAAT3 gene encodes a general amino acid permease that prefers uptake of in planta scarce amino acids. *Molecular Plant Pathology* **5**: 183-189.
- Sweigard JA, Carroll AM, Kang S, Farrall L, Chumley FG, Valent B** 1995. Identification, cloning, and characterization of *PWL2*, a gene for host species specificity in the rice blast fungus. *Plant Cell* **7**: 1221-1233.
- Szurek B, Rossier O, Hause G, Bonas U** (2002) Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Molecular Microbiology* **46**: 13-23.
- Taylor G** (2007) Linking physiology, molecular genetics and genomics in *Populus* to understand and improve yield and quality for biomass and timber production in Europe. In: Final Report of the POPYOMICS European Project. (QLK5-CT-2002-00953) European Commission.
- Tang W, Coughlan S, Crane E, Beatty M, Duvick J** (2006) The application of laser microdissection to *in planta* gene expression profiling of the maize anthracnose stalk rot fungus *Colletotrichum graminicola*. *Molecular Plant-Microbe Interactions* **19**:1240-1250.
- Thara VK, Fellers JP, Zhou JM** (2003) *In planta* induced genes of *Puccinia triticina*. *Molecular Plant Pathology* **4**: 51-56.
- Thines E, Weber RW, Talbot NJ** (2000) MAP kinase and protein kinase A-dependent mobilization of triacylglycerol and glycogen during appressorium turgor generation by *Magnaporthe grisea*. *Plant Cell* **12**: 1703-1718.
- Thomma BPHJ, van Esse HP, Crous PW, De Wit PJGM** (2005) *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Molecular Plant Pathology* **6**: 379-393.
- Thompson JN, Burdon JJ** (1992) Gene-for-gene coevolution between plants and parasites. *Nature* **360**: 121-125.
- Tian M, Huitema E, da Cunha L, Torto-Alalibo T, Kamoun S** (2004) A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. *The Journal of Biological Chemistry* **279**: 26370-26377.
- Tiburzy R, Martins EMF, Reisener HJ** (1992) Isolation of haustoria of *Puccinia graminis* f.sp. *tritici* from wheatleaves. *Experimental Mycology* **16**: 324-328.
- Tremblay A, Li S, Scheffler BE, Matthews BF** (2009) Laser capture microdissection and expressed sequence tag analysis of uredinia formed by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Physiological and Molecular Plant Pathology* **6**: 163-174.
- Tuori RP, Wolpert T.J. and Ciuffetti LM** (2000) Heterologous expression of functional Ptr ToxA. *Molecular Plant-Microbe Interactions* **13**: 456-464.
- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D,**

Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596-1604.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee MK, McDonald WH, Medina M, Meijer HJ, Nordberg EK, Maclean DJ, Ospina-Giraldo MD, Morris PF, Phuntumart V, Putnam NH, Rash S, Rose JK, Sakihama Y, Salamov AA, Savidor A, Scheuring CF, Smith BM, Sobral BW, Terry A, Torto-Alalibo TA, Win J, Xu Z, Zhang H, Grigoriev IV, Rokhsar DS, Boore JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313**: 1261-1266.

Tyler BM (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. *Cell Microbiology* **11**: 13-20.

Underwood W, Zhang S, He SY (2007) The *Pseudomonas syringae* type III effector tyrosine phosphatase HopA01 suppresses innate immunity in *Arabidopsis thaliana*. *Plant Journal* **52**: 658-672.

Valent B, Farrall L, Chumley FG (1991) *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. *Genetics* **127**: 87-101.

Van den Burg HA, Harrison SJ, Joosten MH, Vervoort J, de Wit PJ (2006) *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interactions* **19**: 1420-1430.

Van den Burg HA, Westerink N, Francoijs KJ, Roth R, Woestenenk E, Boeren S, de Wit PJ, Joosten MH, Vervoort J (2003) Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin binding ability. *The Journal of Biological Chemistry* **278**: 27340-

27346.

Van den Hooven HW, van den Burg HA, Vossen P, Boeren S, de Wit PJ, Vervoort J (2001) Disulide bond structure of the AVR9 elicitor of the fungal tomato pathogen *Cladosporium fulvum*: evidence for a cystine knot. *Biochemistry* **40**: 3458–3466.

Van der Hoorn RA, Kamoun S (2008) From guard to decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**:2009-2017.

Van der Merwe MM, Kinnear MW, Barrett LG, Dodds PN, Ericson L, Thrall PH, Burdon JJ (2009) Positive selection in *AvrP4* avirulence gene homologues across the genus *Melampsora*. *Proceedings of the Royal Society B* **276**: 2913-2922.

Van Esse HP, Bolton MD, Stergiopoulos I, de Wit PJ, Thomma BP (2007) The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor. *Molecular Plant-Microbe Interactions* **20**: 1092-1101.

Van Kan JAL, Van Den Ackerveken GFJM, de Wit PJGM (1991) Cloning and characterization of cDNA of avirulence gene *avr9* of the fungal pathogen *Cladosporium fulvum*, causal agent of tomato leaf mold. *Molecular Plant-Microbe Interactions* **4**: 52-59.

Van Poppel PMJ, Guo J, van der Vondervoort PJI, Jung MWM, Birch PRJ, Whisson SC, Govers F (2008) The *Phytophthora infestans* avirulence gene Avr4 encodes an RXLR-dEER effector. *Molecular Plant-Microbe Interactions* **21**: 1460-1470.

Van't Slot KA, Van den Burg HA, Kloks CP, Hilbers CW, Knogge W, Papavoine CH (2005) Solution structure of the plant disease resistance-triggering protein NIP1 from the fungus *Rhynchosporium secalis*. *The Journal of Biological Chemistry* **278**: 45730-45736.

Van't Slot KAE, Gierlich A, Knogge W (2007) A single binding site mediates resistance- and disease associated activities of the effector protein NIP1 from the barley pathogen *Rhynchosporium secalis*. *Plant Physiology* **144**: 1654-1666.

Vervoort J, van den Hooven HW, Berg A, Vossen P, Vogelsang R, Joosten MHAJ, de Wit, PJGM (1997) The race-specific elicitor AVR9 of the tomato pathogen *Cladosporium fulvum*: a cystine knot protein. Sequence-specific ¹H NMR assignments, secondary structure and global fold of the protein. *FEBS Letters* **404**: 153-158.

Viard M (1999) Un siècle de populiculture en France. Les changements intervenus de 1947 à 1997 et leurs conséquences. *La Forêt privée* **246**: 27-42.

Vleeshouwers VG, Rietman H, Krenek P, Champouret N, Young C, Oh SK, Wang M, Bouwmeester K, Vosman B, Visser RG, Jacobsen E, Govers F, Kamoun S, Van der Vossen EA (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS ONE* **3**: e2875.

Voegelé RT, Struck C, Hahn M, Mendgen K (2001) The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proceedings of the National Academy*

of Sciences, USA **98**: 8133-8138.

Voegelé RT, Mendgen K (2003) Rust haustoria: nutrient uptake and beyond. *New Phytologist* **159**: 93-100.

Voegelé RT, Wirsal S, Möll U, Lechner M, Mendgen K (2006) Cloning and characterization of a novel invertase from the obligate biotroph *Uromyces fabae* and analysis of expression patterns of host and pathogen invertases in the course of infection. *Molecular Plant-Microbe Interactions* **19**: 625-634.

Voegelé RT, Hahn M, Mendgen K (2009) The Uredinales: Cytology, Biochemistry, and Molecular Biology. First publ. in: *The Mycota, 5. Plant relationships / Vol. ed.*: H. B. Deising. Berlin: Springer, 2. ed., pp. 69-98.

Wang CI, Guncar G, Forwood JK, Teh T, Catanzariti AM, Lawrence GJ, Loughlin FE, Mackay JP, Schirra HJ, Anderson PA, Ellis JG, Dodds PN, Kobe B (2007) Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *Plant Cell* **19**: 2898-2912.

Wevelslep L, Kogel KH, Knogge W (1991) Purification and characterization of peptides from *Rhynchosporium secalis* inducing necrosis in barley. *Physiological and Molecular Plant Pathology* **39**: 471-482.

Wevelslep L, Ruppig E, Knogge W (1993) Stimulation of barley plasmalemma H⁺-ATPase by phytotoxic peptides from the fungal pathogen *Rhynchosporium secalis*. *Plant Physiology* **101**: 297-301.

Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, Gilroy EM, Armstrong MR, Grouffaud S, van West P, Chapman S, Hein I, Toth IK, Pritchard L, Birch PR (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* **405**: 115-119.

White FF, Yang B, Johnson LB (2000) Prospects for understanding avirulence gene function. *Current Opinion in Plant Biology* **3**: 291-298.

Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B (1994) The product of the tobacco mosaic virus resistance gene *N* : similarity to Toll and the interleukin-1 receptor. *Cell* **78**: 1-20.

Wilson RA, Talbot NJ (2009) Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. *Nature Reviews Microbiology* **7**: 185-195.

Win J, Morgan W, Bos J, Krasileva KV, Cano LM, Chaparro-Garcia A, Ammar R, Staskawicz BJ, Kamoun S (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* **19**: 2349-2369.

- Wirsel SG, Voegelé RT, Mendgen KW** (2001) Differential regulation of gene expression in the obligate biotrophic interaction of *Uromyces fabae* with its host *Vicia faba*. *Molecular Plant-Microbe Interactions* **14**: 1319-1326.
- Wynn WK and Staples RC** (1981) Tropisms of fungi in host recognition. In: Staples RC and Toenniessen GA (eds) *Plant Disease Control: Resistance and Susceptibility*. New York: Wiley.
- Xhaard C, Andrieux A, Halkett F, Frey P** (2009) Characterization of 41 microsatellite loci developed from the genome sequence of the poplar rust fungus, *Melampsora larici-populina*. *Conservation Genetics Resources* **1**: 21-25.
- Xiang T, Zong N, Zou Y, Wu Y, Zhang J, Xing W, Li Y, Tang X, Zhu L, Chai J, Zhou JM** (2008) *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Current Biology* **18**: 74-80.
- Yang S, Zhang X, Yue JX, Tian D, Chen JQ** (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. *Molecular Genetics and Genomics* **280**: 187-198.
- Yin C, Chen X, Wang X, Han Q, Kang Z, Hulbert SH** (2009) Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus *Puccinia striiformis* f. sp. *Tritici*. *BMC Genomics* **10**: 626.
- Zhang J, Shao F, Li Y, Cui H, Chen L, Li H, Zou Y, Long C, Lan L, Chai J, Chen S, Tang X, Zhou JM** (2007) A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host Microbe* **1**: 175-185.
- Zipfel C** (2009) Early molecular events in PAMP-triggered immunity. *Current Opinion in Plant Biology* **12**: 414-420.

La maladie de la rouille foliaire du peuplier, causée par le basidiomycète *Melampsora larici-populina* (*Mlp*) cause des dégâts importants dans les peupleraies européennes. Le séquençage du génome de la souche 98AG31 de *Mlp* a ouvert de nouvelles perspectives pour l'identification de déterminants géniques impliqués dans le processus infectieux du champignon et notamment ceux codant des effecteurs fongiques capables de manipuler la structure et le fonctionnement de la cellule hôte pour assurer le succès de l'infection. L'analyse du transcriptome du champignon au cours des différentes phases du processus infectieux, basée sur l'utilisation de puces à oligonucléotides NimbleGen ou le séquençage massifs d'ESTs, a permis d'identifier des gènes marqueurs de la germination, de la phase de croissance biotrophe et de la sporulation du champignon. Nous avons notamment pu montrer l'induction importante de nombreux gènes codant des petites protéines sécrétées (SSPs) au cours de la phase biotrophe à 96 hpi heures post-inoculation (hpi) ainsi qu'au sein du parenchyme lacuneux à 168 hpi par microdissection à capture laser. L'analyse fine du sécrétome de *Mlp*, basée sur l'annotation, l'évolution et l'expression des gènes codant des SSPs a permis de mettre à jour des effecteurs candidats. Certains, spécifiquement exprimés *in planta* ou présentant des homologies de séquence avec des effecteurs de rouilles ont été localisés au niveau de l'haustorium. De manière intéressante, d'autres gènes candidats appartenant à des familles multigéniques sous pression de sélection positive, sont riches en cystéines, spécifiquement exprimés *in planta* et possèdent un motif de translocation potentiellement impliqué dans l'export de l'effecteur dans la cellule hôte. Ce travail d'analyse fine des effecteurs potentiels d'un agent de rouille à l'échelle génomique va contribuer à l'amélioration des connaissances sur la biologie de ces champignons biotrophes et contribuera à faciliter la recherche de nouvelles méthodes de lutte contre la maladie.

Mots clefs : rouille, biotrophe, peuplier, génomique, transcriptome, protéines sécrétées, effecteurs.

The leaf rust disease caused by *Melampsora larici-populina* (*Mlp*) is the main disease affecting poplar plantations in Europe with severe economic losses. The recent sequencing of the genome of *Mlp* (strain 98AG31) opens new perspectives to identify key genes involved in the fungal infection process and particularly those encoding fungal effectors that could manipulate host cell structure and function to facilitate host colonization. Analysis of the rust transcriptome during time course infection of poplar leaves, based on NimbleGen oligoarrays and massive EST sequencing led to the identification of genes related to fungal germination, biotrophy and sporulation. A consistent induction of genes encoding small-secreted proteins (SSPs) was observed during the biotrophic growth at 96 hours post-inoculation (hpi) but also at 168 hpi in the palisade mesophyll using laser capture microdissection. *Mlp* Secretome analysis, based on annotation, evolution and expression of genes encoding SSPs helped in identifying candidate poplar rust effectors. Some, specifically expressed *in planta* or showing homologies with known rust effectors were localized around the haustorium. Interestingly, other candidate genes, belonging to multigenic families under diversifying selection are cysteine-rich, specifically expressed *in planta* and harbour a translocation signal potentially involved in effector export inside host cell. This genome-wide analysis of putative fungal effectors will contribute to the general knowledge of rust biology and will help to set new approaches to prevent and control the disease.

Keywords : rust fungi, biotroph, poplar, genomics, transcriptome, secreted proteins, effectors.