MPMI Vol. 25, No. 3, 2012, pp. 279–293. http://dx.doi.org/10.1094/MPMI-09-11-0238. © 2012 The American Phytopathological Society *e*-Xtra*

# A Comprehensive Analysis of Genes Encoding Small Secreted Proteins Identifies Candidate Effectors in *Melampsora larici-populina* (Poplar Leaf Rust)

**Stéphane Hacquard,[1] David L. Joly,[2] Yao-Cheng Lin,[3] Emilie Tisserant,[1] Nicolas Feau,[2] Christine Delaruelle,[1] Valérie Legué,[1] Annegret Kohler,[1] Philippe Tanguay,[2] Benjamin Petre,[1] Pascal Frey,[1] Yves Van de Peer,[3] Pierre Rouzé,[3] Francis Martin,[1] Richard C. Hamelin,[2,4] and Sébastien Duplessis[1]**

[1]Unité Mixte de Recherche 1136 Institut National de la Recherche Agronomique-Nancy Université, Interactions Arbres/Microorganismes, INRA Nancy, 54280 Champenoux, France; [2]Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du PEPS, P.O. Box 10380, Stn. Sainte-Foy, Québec, QC, G1V 4C7, Canada; [3]Department of Plant Systems Biology, VIB, Ghent University, 9052 Ghent, Belgium; [4]Department of Forest Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

Submitted 15 September 2011. Accepted 26 October 2011.

**The obligate biotrophic rust fungus *Melampsora larici-populina* is the most devastating and widespread pathogen of poplars. Studies over recent years have identified various small secreted proteins (SSP) from plant biotrophic filamentous pathogens and have highlighted their role as effectors in host–pathogen interactions. The recent analysis of the *M. larici-populina* genome sequence has revealed the presence of 1,184 SSP-encoding genes in this rust fungus. In the present study, the expression and evolutionary dynamics of these SSP were investigated to pinpoint the arsenal of putative effectors that could be involved in the interaction between the rust fungus and poplar. Similarity with effectors previously described in *Melampsora* spp., richness in cysteines, and organization in large families were extensively detailed and discussed. Positive selection analyses conducted over clusters of paralogous genes revealed fast-evolving candidate effectors. Transcript profiling of selected *M. larici-populina* SSP showed a timely coordinated expression during leaf infection, and the accumulation of four candidate effectors in distinct rust infection structures was demonstrated by immunolocalization. This integrated and multifaceted approach helps to prioritize candidate effector genes for functional studies.**

Worldwide, *Melampsora* spp. (Basidiomycota, Pucciniales) are the most devastating pathogens of poplars (Steenackers et al. 1996), and *Melampsora larici-populina* is a major threat in European poplar plantations (Pinon and Frey 2005). Like many other rust fungi, *M. larici-populina* has a complex life cycle which includes five different spore types and requires two phylogenetically distinct host plants; here, *Populus* and *Larix* spp. It is also an emerging model organism in forest pathology because this is one of the first rust fungi to have its genome sequenced, and one of the rare pathosystems for which both host and pathogen genomes are available (Duplessis et al. 2009; Feau et al. 2007; Hacquard et al. 2011a). Moreover, *M. larici-populina* is also a close relative of *Melampsora lini*, the causal agent of flax rust, which is often considered to be the model rust species. After genetic studies demonstrating that single pairs of allelic genes determine the avirulence or virulence phenotype on host lines with particular resistance genes (i.e., the "gene-for-gene" hypothesis) (Flor 1955), avirulence factors were identified in *M. lini* and their direct interaction with flax resistance (R) proteins was demonstrated (Ravensdale et al. 2011).

In order to manipulate host defenses and enable parasitic colonization, many prokaryotic and eukaryotic biotrophic plant pathogens have evolved highly advanced strategies to deliver suites of effector proteins into host cells during infection (Dodds and Rathjen 2010). In turn, their hosts have evolved diverse families of R proteins, which confer resistance via effector-triggered immunity (ETI) after recognition of specific effector proteins, dubbed avirulence (Avr) proteins (Dodds et al. 2009; Jones and Dangl 2006). Unlike the extensively characterized bacterial type III secretion system (Zhou and Chai 2008), little is known about the cellular machineries responsible for translocation of filamentous pathogen effectors into host cells. However, recent studies demonstrated that some small regions within the N-terminal part of oomycete and fungal effectors are required for translocation into host cells in a pathogen-independent manner, suggesting that a host-encoded process is responsible for effector internalization (Kale et al. 2010; Rafiqi et al. 2010). The exact function of conserved N-terminal host cell entry motifs such as RXLR in oomycetes, as well as their presence in fungal effectors, remains unclear and needs to be clarified (Ellis and Dodds 2011; Rafiqi et al. 2010; Yaeno et al. 2011). Many biotrophic fungi and oomycetes are known to share a common infection process involving the formation of haustoria, which invaginate and engage intimate contact with the plasma membrane of

S. Hacquard and D. L. Joly contributed equally to this work.

Current address for S. Hacquard: Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany.

Current address for D. L. Joly: Agriculture and Agri-Food Canada, Pacific Agri-Food Research Centre, Summerland, BC, V0H 1Z0, Canada.

Current address for N. Feau: Department of Forest Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.

Corresponding author: S. Duplessis; Telephone: +33 383 39 40 13; Fax: +33 383 39 40 69; E-mail: duplessi@nancy.inra.fr

*The *e*-Xtra logo stands for "electronic extra" and indicates that 11 supplementary figures and four supplementary tables are published online. Figures 2 and 3 also appear in color online.

host cells (Dodds et al. 2009). Haustoria have been studied for their role in nutrient acquisition and metabolism (Voegele and Mendgen 2003; Voegele et al. 2009), and there is now evidence to suggest that these structures play crucial roles in the delivery of virulence effectors (Catanzariti et al. 2006; Dodds et al. 2009; Panstruga and Dodds 2009).

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 (Ellis et al. 2009). 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 (Kämper et al. 2006). Interestingly, most of these appear to be upregulated during biotrophic development, and some disruptive mutants were significantly altered in virulence (Brefort et al. 2009; Doehlemann et al. 2009; Kämper et al. 2006). In oomycetes, comparison of three *Phytophthora* genomes showed rapid turnover and extensive expansion of specific families of genes encoding secreted effector proteins (Haas et al. 2009; Raffaele et al. 2010; Tyler et al. 2006). Many of these genes, including the host-translocated RXLR-containing proteins, were shown to be induced during infection and encode proteins with activities predicted to alter host physiology (Baxter et al. 2010; Morgan and Kamoun 2007). Although most plant R proteins and downstream signaling pathways share a conserved nature, the array of structures and functions of effector proteins is highly diverse, preventing ab initio identification from sequence and expression data information alone (Ellis et al. 2009). In many cases, the only recognizable features are the presence of an N-terminal signal sequence for secretion through the endomembrane pathway and an even number of cysteine residues that may be involved in the formation of disulfide bonds. Effectors frequently have novel sequences and no obvious homologues in more remotely related species (Göhre and Robatzek 2008; Rep 2005), although some, especially among Cys-rich proteins of *Cladosporium fulvum*, are true exceptions to this rule (Bolton et al. 2008; de Jonge et al. 2010; Stergiopoulos et al. 2010). 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 (Brunner et al. 2009; Dodds et al. 2006; Guttman et al. 2006; Win et al. 2007). Another singular—but not exclusive—feature of secreted effector proteins described in fungal plant pathogens is their low molecular weight (Stergiopoulos and de Wit 2009).

The whole-genome draft sequence of *M. larici-populina* 98AG31 consists of 101.1 megabases encoding 16,399 predicted proteins, among which a total of 1,184 small secreted proteins (SSP) (<300 amino acids) was identified (Duplessis et al. 2011a). Although most of the SSP-encoding genes are lineage specific (84%), several present homologies with effectors previously described in the Pucciniales order. Strikingly, 63% of the SSP have more than four Cys residues, including the most expanded SSP-encoding gene family of 111 members (Duplessis et al. 2011a). This large family harbors a conserved YXC motif similar to a motif identified in secreted proteins of the powdery mildew *Blumeria graminis* f. sp. *hordei* and in rust fungi belonging to the genus *Puccinia* (Godfrey et al. 2010; Spanu et al. 2010).

In this study, complementary to the *M. larici-populina* genome analysis (Duplessis et al. 2011a), we report the detailed annotation and genome-wide analysis of *M. larici-populina* SSP-encoding genes. Features such as homology with known effectors, amino acid content, and organization into multigene families are scrutinized. By unraveling the expression and evolutionary dynamics of SSP, fungal genes specifically expressed in planta and exhibiting the hallmarks of positive selection due to coevolutionary arms races with host factors are identified. Moreover, immunolocalization of selected SSP suggests specificity of secretion from different fungal cell types in planta.

## RESULTS

### *M. larici-populina* homologues of rust effectors.

Of the 22 effectors previously described in rust fungi (Catanzariti et al. 2006; Dodds et al. 2004; Kemen et al. 2005), 20 were identified in *M. larici-populina* (Duplessis et al. 2011a). Of these 20 effectors, 12 belong to multigenic families, representing a total of 63 homologues of conserved rust effectors in *M. larici-populina* (Table 1). In all, 8 and 13 divergent homologues of *M. lini AvrM* and *M. lini AvrP4*, respectively, were detected. Six of the *AvrM* paralogs were found to reside in three tandem pairs, with intergenic regions showing similarity to Pfam domains associated with transposable elements (Supplementary Fig. S1). A phylogenetic tree incorporating AvrP4 sequences from multiple *Melampsora* spp. (Van der Merwe et al. 2009) illustrates the high sequence divergence of *M. larici-populina AvrP4* homologues, which lay together in a supported clade (Fig. 1). This suggests duplication and diversification from a unique ancestor of *AvrP4* toward the multigenic family herein presented. As described by Van der Merwe and associates (2009), amino acids at the C-terminal end were highly variable apart from the six Cys residues spaced according to the consensus of a cystine-knot, a structure also found in Avr9 from *C. fulvum* (van den Hooven et al. 2001), as well as in toxic and inhibitory polypeptides (Pallaghy et al. 1994). In contrast, only one homologue of *M. lini AvrP123* and one homologue of *M. lini AvrL567* were identified in the genome of *M. larici-populina*. Strikingly, the 26- to 50-amino acid region containing the translocation signal RFYR identified in *M. lini* AvrL567 (Rafiqi et al. 2010) is not conserved in the *M. larici-populina* homologue. However, the deduced *M. larici-populina* AvrL567 homologue protein threads onto the *M. lini* AvrL567 structure (Wang et al. 2007), indicating that the two proteins might share common structural features (Supplementary Fig. S2). In contrast with the strong diversifying selection reported for the *M. lini AvrL567* loci (Dodds et al. 2004, 2006), a low level of polymorphism as well as an absence of diversifying selection were observed between the six allelic variants identified from 32 *M. larici-populina* isolates (Supplementary Fig. S3). Moreover, no correlation could be established between allelic variants and avirulence phenotypes on a differential set of poplar cultivars with characterized resistance genes (Pinon and Frey 2005). Of the three homologues of the *Uromyces fabae* rust transferred protein (RTP1) identified in *M. larici-populina* (Duplessis et al. 2011a), the N-terminal region of one RTP1 homologue was identical to the haustorially expressed secreted protein (HESP) HESP-327 from *M. lini* (Table 1), which is consistent with the conserved C-terminal region and the more diverse N-terminal region previously reported for RTP1 in other rust fungi (Fernandez et al. 2012; Kemen et al. 2005; Puthoff et al 2008). In silico localization analyses suggested that only the HESP327/RTP1 protein contains a predicted nuclear localization signal. A large proportion of the 63 genes encoding conserved rust effectors were expressed during plant infection (30 genes) compared with urediniospores (10 genes) (Table 1).

### Cys-rich proteins dominate the secretome of *M. larici-populina.*

Of the 1,184 SSP identified in *M. larici-populina*, 746 had a number of Cys residues greater than four (Duplessis et al.

**Table 1.** Conserved rust effectors in *Melampsora larici-populina* 98AG31

| Protein[a] | Scaffold and strand | Length (amino acids) | Putative effector hit[b] | Species[b] | Identities (%) | Positives (%) | Transcript evidence in *M. larici-populina* | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Spores | Infected leaves | Haustoria |
| 58459 | 1– | 175 | HESP-178 | *Melampsora lini* | 29 | 55 | … | X | … |
| 33341 | 3– | 146 | HESP-178 | *M. lini* | 89 | 91 | … | … | … |
| 117398 | 37+ | 207 | HESP-178 | *M. lini* | 26 | 46 | X | X | … |
| 103805 | 7– | 411 | HESP-270 | *M. lini* | 75 | 85 | … | X | X |
| 123932 | 17– | 405 | HESP-327 | *M. lini* | 92 | 96 | … | X | … |
| | | | Rust transferred protein 1 | *Uromyces fabae* | 59 | 76 | … | … | … |
| 123524 | 23– | 216 | Rust transferred protein 1 | *U. fabae* | 48 | 63 | … | X | … |
| 123523 | 31– | 238 | Rust transferred protein 1 | *U. fabae* | 50 | 59 | … | … | … |
| 124544 | 28+ | 140 | HESP-376 | *M. lini* | 47 | 63 | … | X | … |
| 64764 | 28+ | 154 | HESP-376 | *M. lini* | 27 | 45 | … | … | … |
| 70587 | 2+ | 207 | HESP-379 | *M. lini* | 86 | 90 | X | X | … |
| 93569 | 5– | 134 | HESP-417 | *M. lini* | 36 | 52 | X | X | X |
| 71399 | 11– | 132 | HESP-417 | *M. lini* | 29 | 52 | X | X | … |
| 74496 | 11+ | 132 | HESP-417 | *M. lini* | 30 | 50 | X | X | … |
| 71402 | 11+ | 135 | HESP-417 | *M. lini* | 30 | 51 | X | X | … |
| 71404 | 11+ | 128 | HESP-417 | *M. lini* | 76 | 81 | X | X | … |
| 108269 | 28– | 136 | HESP-417 | *M. lini* | 24 | 44 | … | … | … |
| 35858 | 17+ | 234 | HESP-570 | *M. lini* | 55 | 65 | … | … | … |
| 53986 | 69– | 226 | HESP-570 | *M. lini* | 65 | 76 | … | … | … |
| 90824 | 3– | 415 | HESP-735 (recursive tblastn) | … | … | … | … | X | X |
| 141065 | 7– | 223 | HESP-735 | *M. lini* | 90 | 94 | … | … | … |
| 85525 | 14+ | 220 | HESP-735 (recursive tblastn) | … | … | … | … | X | … |
| 42266 | 6– | 85 | HESP-767 | *M. lini* | 100 | 100 | X | X | … |
| 110232 | 42+ | 142 | HESP-897 | *M. lini* | 87 | 92 | X | X | … |
| 101424 | 1+ | 186 | HESP-C49 (recursive tblastn) | … | … | … | … | … | … |
| 71126 | 7+ | 178 | HESP-C49 | *M. lini* | 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 | *M. lini* | 72 | 84 | … | … | … |
| 124415 | 52+ | 353 | HESP-C61 | *M. lini* | 55 | 70 | … | … | X |
| 124414 | 52+ | 284 | HESP-C61 | *M. lini* | 35 | 53 | … | X | … |
| 101608 | 2– | 185 | HESP-C63 | *M. lini* | 27 | 44 | … | … | … |
| 101611 | 2+ | 203 | HESP-C63 | *M. lini* | 25 | 46 | … | … | … |
| 63837 | 22– | 175 | HESP-C63 (recursive tblastn) | … | … | … | … | X | … |
| 123443 | 22– | 203 | HESP-C63 | *M. lini* | 24 | 42 | … | X | … |
| 65539 | 35– | 201 | HESP-C63 | *M. lini* | 51 | 69 | … | … | … |
| 34793 | 10+ | 76 | HESP-C66 | *M. lini* | 58 | 69 | … | … | … |
| 34673 | 10– | 74 | HESP-C66 | *M. lini* | 56 | 69 | … | … | … |
| 37534 | 30+ | 77 | HESP-C66 | *M. lini* | 52 | 70 | … | X | … |
| 37347 | 29+ | 151 | AvrL567 | *M. lini* | 61 | 44 | … | … | … |
| 124207 | 2+ | 545 | AvrM | *M. lini* | 50 | 33 | … | … | … |
| 124206 | 12– | 568 | AvrM | *M. lini* | 53 | 36 | … | … | … |
| 124204 | 35+ | 461 | AvrM | *M. lini* | 39 | 25 | … | … | … |
| 124205 | 35+ | 448 | AvrM | *M. lini* | 53 | 36 | … | … | … |
| 124203 | 65– | 518 | AvrM | *M. lini* | 54 | 32 | … | … | … |
| 124202 | 65– | 413 | AvrM | *M. lini* | 47 | 31 | … | … | X |
| 124208 | 77– | 702 | AvrM | *M. lini* | 52 | 36 | … | … | … |
| 124209 | 77– | 687 | AvrM | *M. lini* | 42 | 28 | … | … | X |
| 124530 | 78– | 122 | AvrP123 | *M. lini* | 40 | 29 | … | … | … |
| 124258 | 3– | 95 | AvrP4 | *M. lini* | 67 | 57 | … | … | … |
| 124268 | 3– | 90 | AvrP4 | *M. lini* | 43 | 30 | … | … | … |
| 124272 | 3+ | 97 | AvrP4 | *M. lini* | 50 | 34 | … | … | … |
| 124262 | 5+ | 101 | AvrP4 | *M. lini* | 62 | 52 | … | … | … |
| 124264 | 10+ | 89 | AvrP4 | *M. lini* | 60 | 47 | … | X | … |
| 124270 | 15+ | 89 | AvrP4 | *M. lini* | 50 | 38 | … | … | … |
| 124256 | 18– | 88 | AvrP4 | *M. lini* | 58 | 44 | … | X | … |
| 124266 | 19– | 92 | AvrP4 | *M. lini* | 63 | 47 | … | X | … |
| 124260 | 35– | 92 | AvrP4 | *M. lini* | 63 | 47 | … | … | … |
| 124254 | 39– | 90 | AvrP4 | *M. lini* | 67 | 52 | … | … | … |
| 124274 | 51+ | 97 | AvrP4 | *M. lini* | 54 | 34 | … | … | … |
| 112294 | 64– | 105 | AvrP4 | *M. lini* | 52 | 30 | … | … | … |
| 124252 | 167+ | 96 | AvrP4 | *M. lini* | 72 | 51 | … | … | … |

[a] Protein ID: number of corresponding best predicted gene model in the *M. larici-populina* Joint Genome Institute genome sequence.

[b] Homologues of putative and characterized rust effectors were searched in the *M. larici-populina* genome by tblastn using the sequences of the *M. lini* haustorially expressed secreted proteins (HESP) and avirulence (Avr) proteins and the *U. fabae* rust transferred protein 1. Recursive tblastn searches were performed using those *M. larici-populina* putative effector sequences. Percentages of identities and positives were assessed following blastp searches carried out against UNIPROT (*E* value < 1e-5).

2011a). Among those, 228 SSP showed a conserved exon and intron structure and a characteristic spacing of Cys residues, as previously reported by Duplessis and associates (2011a) for the largest *M. larici-populina* SSP-encoding gene family. Based on sequence homology, these 228 SSP were grouped into 28 classes (Table 2; Supplementary Table S1). Of those 228 Cys-rich SSP-encoding genes, 99 lay in close vicinity to another member or putative pseudogene of the same class (no more than five genes apart), indicating that they might have arisen by tandem duplication. Apart from classes VI and XIV, which contained the *M. lini* HESP-C49 and HESP-897 homologues, respectively, only one more class had homology to *Puccinia graminis* f. sp. *tritici*. Although significant variability was observed with respect to spacing and number of Cys residues within the 28 classes, most of those SSP-encoding genes shared a general structure of five exons, with the first full codon of exons 2, 4, and 5 encoding a Cys residue, as previously reported for the largest *M. larici-populina* SSP-encoding gene family (Duplessis et al. 2011a). By comparing the conservation profile from the eight largest classes of Cys-rich SSP, specific protein positions were found to be enriched in particular amino acid residues across classes (Fig. 2). In many cases, Tyr residues were found before conserved Cys number 3 and 8 (YXC) and after conserved Cys number 6 ($CX_{2\text{-}3}Y$). In order to assess the importance of these Cys and their surrounding residues, two types of sequence randomization tests were used to verify the significance of the YXC and $CX_{2\text{-}3}Y$ motifs. First, the observed location and frequency of the YXC and $CX_{2\text{-}3}Y$ motifs in the real SSP sequences were compared with randomly shuffled sequences. By plotting all motif locations (excluding the signal peptide) in each class compared with the random data set, both YXC and $CX_{2\text{-}3}Y$ motifs showed strong position preferences (Fig. 2; Supplementary Fig. S4). Furthermore, the observed frequency of YXC and $CX_{2\text{-}3}Y$ motifs in the true SSP classes I, II, III, and XXIV was significantly higher than in the random data set ($P$ value < 0.01). In a second test, hidden Markov models

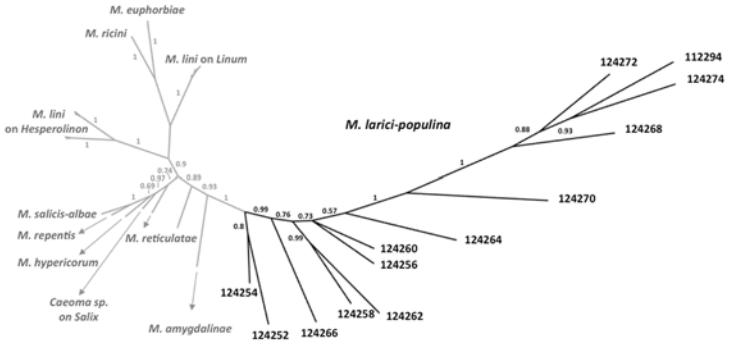


**Fig. 1.** Unrooted Bayesian consensus tree of AvrP4 homologues. AvrP4 homologues were identified in the genome sequence of *Melampsora larici-populina* strain 98AG31 (in black) and obtained from diverse *Melampsora* spp. (Van der Merwe et al. 2009) (in gray). Values above branches correspond to posterior probabilities associated with the branch supporting a particular clade. *M. larici-populina* genes encoding protein ID numbers 124268 and 124272 are located in tandem. The other AvrP4 homologues are scattered over the *M. larici-populina* genome.

**Table 2.** *Melampsora larici-populina* small secreted proteins classes with characteristic Cys patterns

| Class | Number of genes | Length (amino acids) | Cys pattern | Top blast hit[a] |
|---|---|---|---|---|
| I[b] | 111 | 107 to 155 | $CX_{9\text{-}19}CX_{8\text{-}19}CX_4CX_{12\text{-}16}CX_{9\text{-}17}CX_{11\text{-}18}CX_{8\text{-}14}CX_{8\text{-}13}CX_{1\text{-}5}C$ | … |
| II | 39 | 133 to 157 | $CX_{3\text{-}21}CX_{8\text{-}22}CX_2CX_{12}CX_{13\text{-}18}CX_{15\text{-}21}CX_{10\text{-}14}CX_8CX_{2\text{-}4}C$ | … |
| III | 8 | 134 to 367 | $CX_{12}CX_{5\text{-}7}CX_3CX_{8\text{-}9}CX_{12\text{-}28}CX_{13}CX_8CX_{9\text{-}11}CX_2C$ | PGTG_20247.2 |
| IV | 5 | 154 to 159 | $CX_{14\text{-}17}CX_{8\text{-}10}CX_4CX_{20\text{-}23}CX_{51\text{-}54}CX_9CX_4C$ | … |
| V | 6 | 128 to 149 | $CX_{11\text{-}14}CX_{9\text{-}13}CX_{13}CX_{11\text{-}14}CX_{11\text{-}18}CX_{13\text{-}16}CX_8C$ | … |
| VI | 6 | 178 to 492 | $CX_{21\text{-}222}CX_{10}CX_4CX_5CX_2CX_{13\text{-}23}CX_9CX_3CX_5CX_2$ | HESP-C49 (*M. lini*); PGTG_08705.2 |
| VII | 5 | 142 to 149 | $CX_{15\text{-}17}CX_{12\text{-}13}CX_4CX_7CX_{14\text{-}15}CX_{13\text{-}15}CX_9CX_{5\text{-}7}CX_{10\text{-}13}CX_8CX_2C$ | … |
| VIII | 4 | 141 to 154 | $CX_{13\text{-}14}CX_{26\text{-}28}CX_{13\text{-}14}CX_2CX_8CX_{15}CX_{18\text{-}19}CX_{7\text{-}8}CX_8CX_2C$ | … |
| IX | 3 | 186 to 507 | $CX_{23}CX_9CX_2CX_{22\text{-}30}CX_{16\text{-}17}CX_{13\text{-}14}CX_9CX_{12\text{-}13}CX_2C$ | … |
| X | 3 | 167 to 180 | $CX_{16}CX_9CX_4CX_{20\text{-}21}CX_{54\text{-}55}CX_{12}CX_2C$ | … |
| XI | 4 | 144 to 153 | $CX_{12\text{-}16}CX_{8\text{-}10}CX_2CX_{9\text{-}10}CX_{23\text{-}28}CX_{10\text{-}11}CX_{16}CX_{10\text{-}11}CX_{10\text{-}11}CX_2C$ | … |
| XII | 3 | 133 to 139 | $CX_{6\text{-}7}CX_{13\text{-}15}CX_{15\text{-}20}CX_{12}CX_{17\text{-}18}CX_{17\text{-}18}CX_9CX_{10\text{-}11}C$ | … |
| XIII | 2 | 120 to 127 | $CX_{21}CX_4CX_2CX_{11\text{-}18}CX_{11}CX_9CX_9CX_8CX_4C$ | … |
| XIV | 1 | 142 | $CX_{16}CX_8CX_2CX_8CX_{18}CX_{17}CX_9CX_8CX_2C$ | HESP-897 (*M. lini*); PGTG_06517.2 |
| XV | 2 | 203 to 221 | $CX_{25}CX_8CX_4CX_{13}CX_8CX_{50\text{-}56}CX_{11}CX_8CX_2C$ | … |
| XVI | 2 | 114 to 116 | $CX_{18\text{-}21}CX_{13\text{-}14}CX_{9\text{-}10}CX_4CX_8CX_7CX_{15}CX_{10}CX_7CX_4C$ | … |
| XVII | 2 | 124 to 130 | $CX_{5\text{-}6}CX_{13}CX_{9\text{-}14}CX_4CX_{8\text{-}10}CX_{18}CX_{13}CX_{11}CX_7CX_2C$ | … |
| XVIII | 1 | 199 | $CX_{22}CX_{38}CX_9CX_2CX_{16}CX_{31}CX_{20}CX_{10}CX_8CX_2C$ | … |
| XIX | 2 | 130 to 133 | $CX_{3\text{-}5}CX_{12}CX_{10}CX_4CX_9CX_{11\text{-}12}CX_{22}CX_{10\text{-}12}CX_{11}CX_2C$ | … |
| XX | 1 | 157 | $CX_5CX_{10}CX_{17}CX_{10}CX_2CX_9CX_{20}CX_{15}CX_{10}CX_{24}CX_4C$ | … |
| XXI | 1 | 189 | $CX_{14}CX_{15}CX_4CX_{22}CX_{66}CX_{12}CX_2C$ | … |
| XXII | 1 | 145 | $CX_{19}CX_{13}CX_{12}CX_8CX_{15}CX_{16}CX_7CX_7CX_5C$ | … |
| XXIII | 2 | 125 to 136 | $CX_{14}CX_{12}CX_{13\text{-}14}CX_{10\text{-}11}CX_{10\text{-}16}CX_{13}CX_8C$ | … |
| XXIV | 8 | 154 to 164 | $CX_{13\text{-}15}CX_9CX_{12\text{-}14}CX_{21\text{-}23}CX_{13\text{-}18}CX_{12\text{-}19}CX_{12\text{-}14}C$ | … |
| XXVI | 1 | 131 | $CX_{10}CX_8CX_{12}CX_{13}CX_9CX_2CX_8CX_{10}CX_{11}CX_{14}CX_{11}CX_2C$ | … |
| XXVII | 3 | 140 to 143 | $CX_{15}CX_{13}CX_{14}CX_9CX_{18}CX_{19}CX_8C$ | … |
| XXVIII | 1 | 138 | $CX_9CX_4CX_3CX_{17}CX_8CX_4CX_8CX_{12}CX_9CX_8CX_2C$ | … |
| XXIX | 1 | 188 | $CX_{20}CX_{20}CX_{13}CX_{11}CX_8CX_{47}CX_{10}CX_{11}C$ | … |

[a] Homology searches were carried out against UNIPROT and *Puccinia* Group Database using blastp ($E$ value < 1e-6).
[b] Class I has been reported previously by Duplessis and associates (2011a). … = no hit.

(HMM) scores confirmed positional constraints on the YXC and $CX_{2\text{-}3}Y$ motifs and their surrounding sequences in classes I, II, III, and XXIV (discussed below) (Supplementary Fig. S5).

**Positive selection in candidate effector genes.**

Homology searches were used to refine the sequence clustering of SSP paralogous families and classify them into clusters of paralogous genes (CPG). We identified a subset of 95 CPG comprising 536 sequences (509 true SSP and 27 larger secreted proteins ranging from 301 to 942 amino acids) that contained three to 39 genes that could be investigated by positive selection analysis (codeml program of the PAML package) (Yang 2007). Eighteen CPG containing a total of 185 genes showed robust evidence of diversifying selection at $P < 0.05$ in both codon sub-

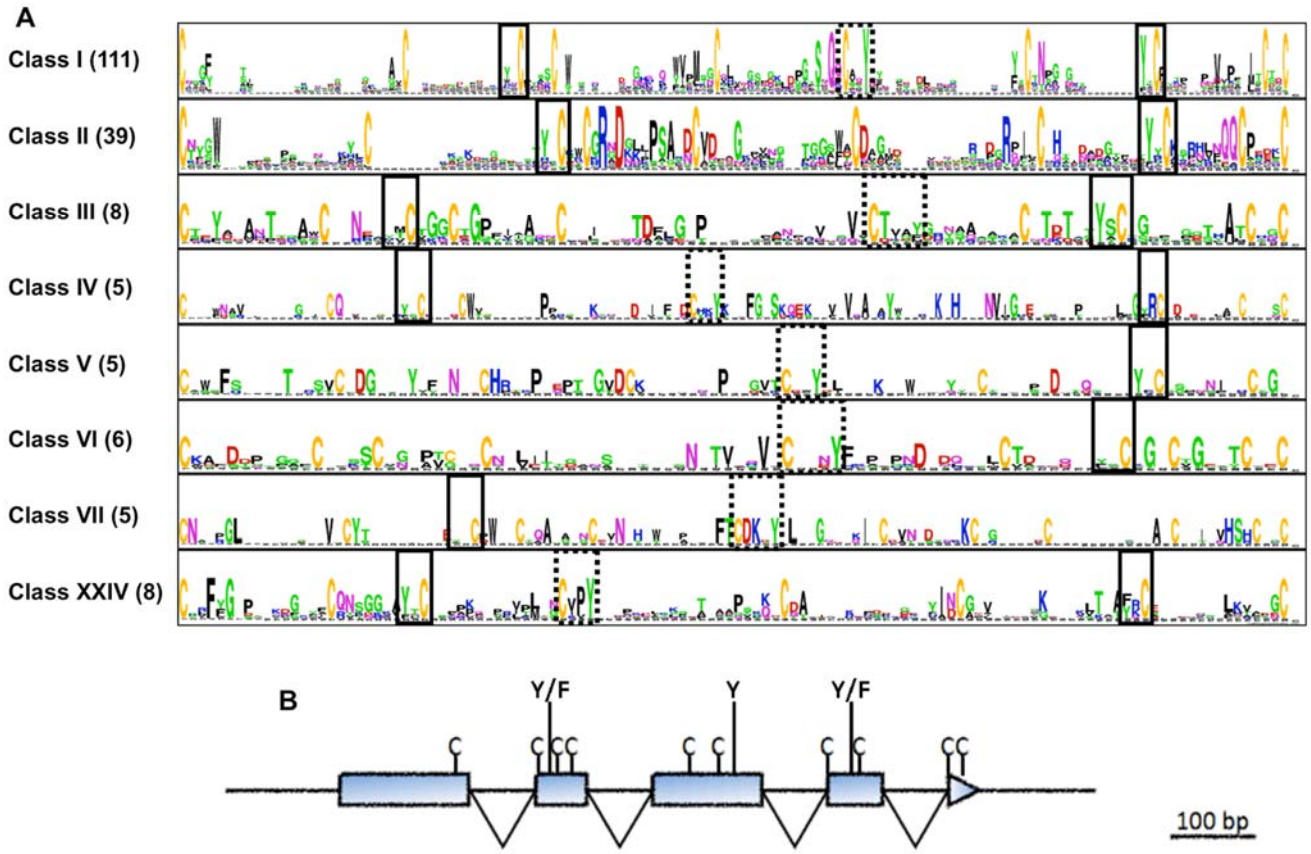


**Fig. 2.** Conservation profiles and gene structure of the eight largest *Melampsora larici-populina* Cys-rich classes. **A,** Motifs shared by *M. larici-populina* Cys-rich classes. The Y/FXC motif identified in powdery mildew (Godfrey et al. 2010) (solid lines) is present in both the N and C termini (i.e., at the third and eighth conserved Cys, respectively) in six and seven of the eight classes, respectively. Another motif, $CX_{2\text{-}3}Y$ (dashed lines), is present in the central part of the protein (i.e., at the sixth conserved Cys) in seven of the eight classes. Conservation profiles were generated with WebLogo. The number of proteins in each class is given in parentheses. Conservation profile of class I as in Duplessis and associates (2011a). **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 residues found in shared motifs. Introns are represented as "V"s.

**Table 3.** *Melampsora larici-populina* positively selected clusters of paralogous genes (CPG) encoding small secreted proteins

| CPG | Number of genes | Length (amino acids) | Number of Cys residues | Top blast hit[a] |
|---|---|---|---|---|
| 277 | 6 | 181 to 185 | 6 to 7 | … |
| 335 | 7 | 180 to 187 | 7 to 9 | … |
| 338 | 13 | 143 to 194 | 4 to 10 | … |
| 462-1 | 5 | 171 to 173 | 10 to 11 | PGTG_17910.2 |
| 510-1 | 16 | 64 to 86 | 4 to 6 | … |
| 510-2 | 11 | 71 to 124 | 6 to 12 | … |
| 517-2 | 3 | 133 to 136 | 8 | … |
| 653-1 | 5 | 150 to 187 | 2 to 3 | … |
| 1027 | 3 | 176 to 185 | 8 to 9 | … |
| 1252 | 4 | 117 to 119 | 0 | … |
| 2811 | 9 | 75 to 96 | 6 | … |
| 4438 | 4 | 168 to 172 | 6 to 8 | … |
| 4890 | 9 | 69 to 93 | 6 to 10 | … |
| 5464 | 13 | 88 to 105 | 6 to 9 | AvrP4 (*M. lini*) |
| Class II | 39 | 133 to 157 | 9 to 12 | … |
| Class V | 5 | 128 to 149 | 8 to 9 | … |
| Class XXVII | 3 | 140 to 143 | 8 to 9 | … |
| H1 | 30 | 58 to 84 | 1 to 6 | … |

[a] Homology searches were carried out against UNIPROT and *Puccinia* Group Database using blastp (*E* value < 1e-6). … = no hit.

stitution model M2 versus M1 and M8 versus M7 tests, as well as in the M8 versus M8A stringent test of selection (Table 3; Supplementary Table S2). Parameter estimates for each positively selected cluster are shown in Supplementary Table S3. One of these positively selected CPG shows homology to *P. graminis* f. sp. *tritici* (CPG462-1) and one is composed of *M. lini AvrP4* homologues (CPG5464). In addition to the *AvrP4* homologues, five other CPG encode very small (<100 amino acids) Cys-rich proteins (CPG510-1, CPG510-2, CPG2811, CPG4890, and CPGH1). Most genes included in the latter families present a conserved LXRR motif (CPG510-1 and CPG510-2) or a region rich in positively charged residues (Arg and Lys) shortly followed by a region rich in acidic residues (Asp and Glu), reminiscent of the RXLR-EER motif from oomycete effectors (Fig. 3; Supplementary Fig. S6). Most genes encoding very small proteins from the CPG under positive selection were absent from the ab initio gene prediction of the *M. larici-populina* genome (72 of 88), exemplifying the importance of expert annotation beyond automated gene prediction. In most of the small Cys-rich secreted proteins with evidence of positive selection, the Cys residues are conserved and the Bayes Empirical Bayes inference of positively selected amino acid sites revealed that codons with highly significant ω values were concentrated in the C-terminal region (Fig. 3).

### *M. larici-populina* candidate effector genes display coordinated expression profiles.

A large proportion of SSP-encoding genes (857 of 1,184) were supported by expressed sequence tags (EST) (Joly et al. 2010; E. Tisserant and S. Duplessis, *unpublished*) or identified as expressed during plant infection using whole-genome oligoarrays (Duplessis et al. 2011b). In order to monitor expression profiles of candidate effectors, 34 SSP-encoding genes displaying homology with previously described rust effectors (Table 1), elevated evolution rates (Fig. 3; Table 3), or contrasted expression levels during plant infection using oligoarrays and EST sequencing were selected for reverse-transcription quantitative polymerase chain reaction (RT-qPCR) analysis (Supplementary Table S4). Coordinated temporal patterns of SSP transcript expression were observed during poplar leaf infection (Fig. 4; Supplementary Fig. S7). Four selected SSP-encoding genes, including two *RTP1* and one *AvrP4* homologues (protein ID numbers 123932, 123523, and 124272), showed a biphasic expression profile in spores and at late stages of leaf infection (Fig. 4A). Strikingly, four of the eight SSP-encoding genes showing transcript accumulation between 2 and 6 h postinoculation (hpi) (Fig. 4B) encode CFEM (common in the fungal extracellular membrane) domain proteins (protein ID numbers 72734, 114961, 73660, and 58459), including a homologue of *M. lini* HESP-178, suggestive of a possible role at early stages of rust development such as during germ-tube elongation or appressorium formation. Additionally, two genes identified as evolving under positive selection and belonging to CPG class II and CPG1252 (protein ID numbers 123873 and 67604) were both preferentially induced at the fungal penetration stage (6 to 12 hpi). Between 12 and 24 hpi, fungal development is marked by the formation of the first haustoria (Laurans and Pilate 1999). From this point, several SSP transcripts showed a

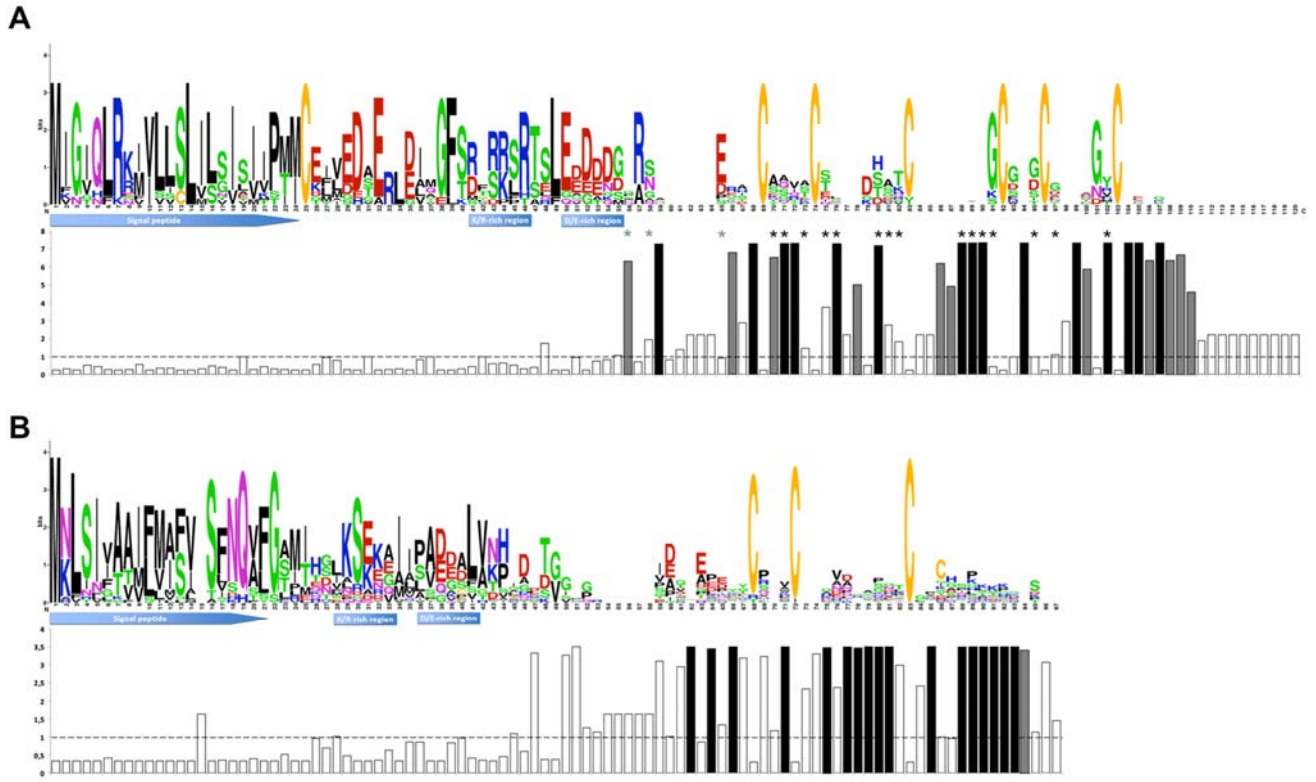


**Fig. 3.** Examples of positively selected clusters of paralogous genes encoding small secreted proteins. Conservation profiles of *Melampsora larici-populina* clusters of paralogous genes (CPG) **A,** CPG5464 (AvrP4 homologues) and **B,** CPGH1; and associated bar diagram indicating the value of ω calculated with codeml for each codon under site model M2. 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 gray indicate codons found to be under selection but without significant ω values. Asterisks in A indicate codons found to be under selection by Van der Merwe and associates (2009). Conservation profiles were generated with WebLogo.

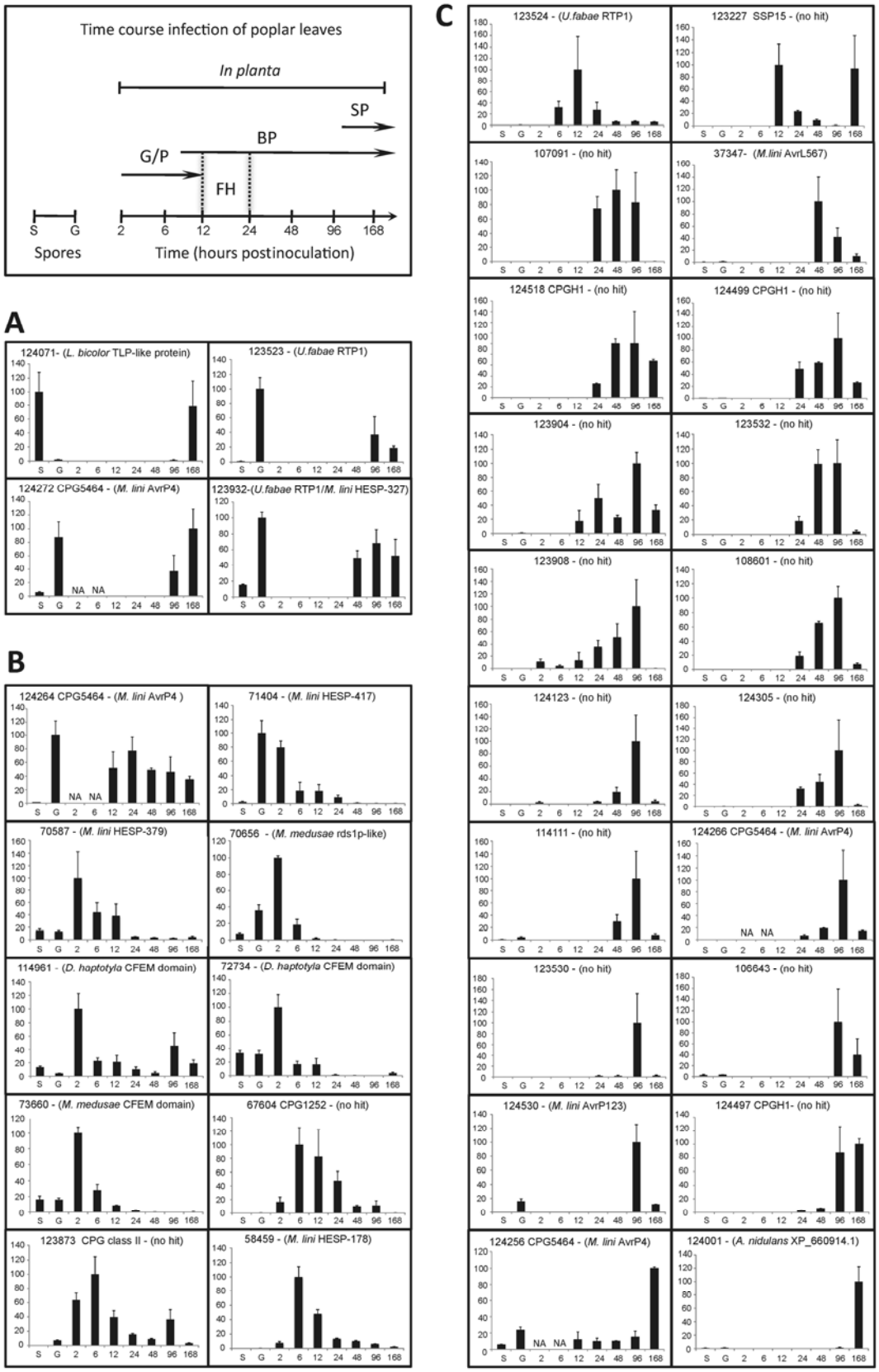


**Fig. 4.** Expression profiles of *Melampsora larici-populina* small secreted protein (SSP) transcripts during poplar leaf infection. Urediniospores and *M. larici-populina* time-course infection stages on poplar leaves (2 to 168 h postinoculation [hpi]) surveyed in the analysis are detailed in the insert upper-left. S, resting urediniospores; G, germinating urediniospores; G/P, germination/penetration; BP, biotrophic phase; SP, sporulation; FH, first haustoria. Reverse-transcription quantitative polymerase chain reaction expression levels were normalized with *Mlp-aTub* and *Mlp-ElF1α* reference genes (Hacquard et al. 2011b). For each transcript, expression regulation levels are presented as percentages according to the highest level. **A,** SSP transcripts with biphasic expression profile in spores and during leaf infection. **B,** SSP transcripts with a preferential accumulation in urediniospores (S, G) and during germination or penetration (G/P). **C,** SSP transcripts preferentially accumulated in planta after haustoria formation (BP) or sporulation (SP). NA, not assessed. Error bars represent standard errors.

strong and specific expression associated with biotrophic growth (Fig. 4C). Expression peaks concordant with both early haustoria formation (12 hpi) and sporulation (168 hpi) were detected for a gene encoding an *M. larici-populina*-specific SSP of 15 kDa (protein ID number 123227). The third *Mlp-RTP1* homologue (protein ID number 123524) showed a distinct expression profile compared with its two other relatives in the *Mlp-RTP1* gene family, marked by transcript accumulation concomitant with haustoria formation. Numerous candidate effectors preferentially expressed after the differentiation of the first haustoria appeared to be lineage specific or showed homology with characterized avirulence genes (protein ID numbers 37347, 124266, 124256, and 124530). Interestingly, candidate effectors belonging to the fast-evolving CPGH1 (Fig. 3) were also specifically detected in planta after haustoria formation (protein ID numbers 124518, 124499, and 124497).

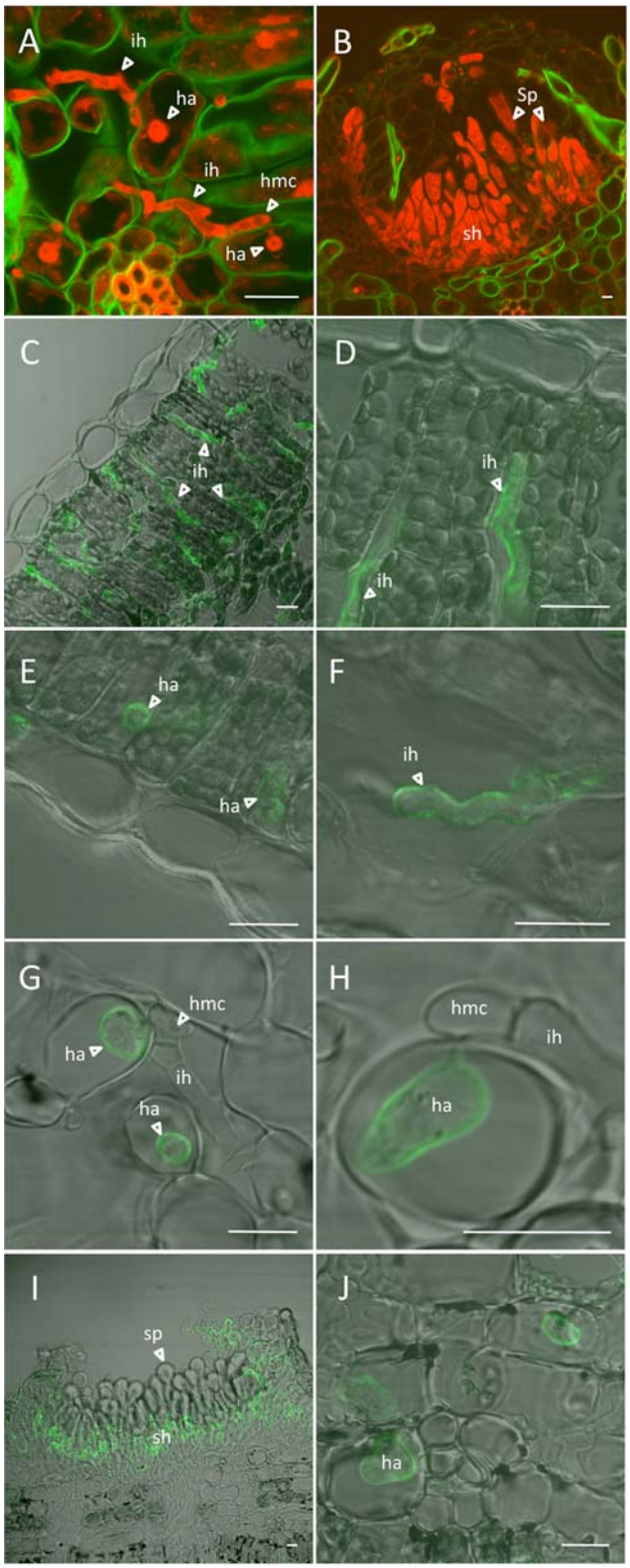


**Fig. 5.** Localization of *Melampsora larici-populina* candidate effectors during poplar leaf infection. **A** and **B,** Confocal laser-scanning microscopy images showing in planta fungal structures using Uvitex 2B (green) and propidium iodide (red) to label fungal–plant cell wall and fungal cells, respectively. **A,** Infection hyphae and haustoria colonizing mesophyll at 96 h postinoculation (hpi) and **B,** uredinium formed after 168 hpi. **C** to **J,** Immunofluorescence localization of rust candidate effectors at late stage of rust infection. Merged-immunofluorescence and phase contrast images are presented. **C** and **D,** Immunolabeling of *Mlp*-RTP1sc31 around infection hyphae at 120 hpi; **E** and **F,** labeling of *Mlp*-HESP-327/RTP1 around haustoria and infection hyphae at 120 hpi; **G** and **H,** *Mlp*-AvrL567 labeling around haustoria at 168 hpi; **I** and **J,** immunolabeling of *Mlp*-SSP15 in sporogenous hyphae and around haustoria at 168 hpi. Abbreviations: ih, infection hyphae; ha, haustorium; hmc, haustorial mother cell; sh, sporogenous hyphae; sp, urediniospores. Pictures were taken using a ×100 objective and scale bars represent 5 µm.

### *M. larici-populina* candidate effectors localize to infection structures.

Visualization of in planta rust structures was achieved in infected poplar leaves using propidium iodide and Uvitex labeling (Fig. 5). Many infection hyphae and haustoria were observed inside both spongy and palisade mesophyll at 96 hpi during the biotrophic growth phase. The fungal terminal cells differentiated to form haustorial mother cells and developed haustorial necks inside host cells, where the haustorial body extended from the neck and invaginated the plant cell plasmalemma (Fig. 5A). At a later stage of rust infection (i.e., 168 hpi), changes in fungal development were observed with aggregates of sporogenous hyphae and newly formed dikaryotic urediniospores (Fig. 5B). A late time frame of poplar leaf infection (120 to 168 hpi) was chosen for immunological detection of SSP to ensure the presence of important fungal biomass and a diversity of fungal cell types (i.e., infection hyphae, haustorial mother cells, haustoria, sporogenous hyphae, and newly formed urediniospores). *M. larici-populina* candidate effectors *Mlp*-AvrL567, *Mlp*-RTP1sc31 (located on scaffold 31), *Mlp*-HESP-327/RTP1, and *Mlp*-SSP15 (protein ID numbers 37347, 123523, 123932, and 123227, respectively) were assayed for localization in infected poplar leaves by immunofluorescence microscopy and revealed distinct labeling patterns (Fig. 5). Strong labeling of *Mlp*-RTP1sc31 was observed at the periphery of infection hyphae colonizing the plant mesophyll (Fig. 5C and 5D; Supplementary Fig. S8) whereas no labeling was observed for haustoria or other fungal structures. In contrast, the *Mlp*-HESP-327/RTP1 fusion protein was immunolocalized at the periphery of haustoria (Fig. 5E) and infection hyphae (Fig. 5F; Supplementary Fig. S9). Specific labeling of *Mlp*-AvrL567 was detected at the periphery of haustoria (Fig. 5G and H). This SSP likely resides in the extrahaustorial matrix and was also detected around the haustorial neck (Fig. 5H) whereas no fluorescence was detected in haustorial mother cells, infection and sporogenous hyphae, or newly formed urediniospores, indicating a localization restricted to haustoria. Moreover, *Mlp*-AvrL567 was also detected at the periphery of haustoria at 48 hpi (Supplementary Fig. S10), which correlates with the peak of transcript expression observed by RT-qPCR (Fig. 4). As observed for the *Mlp*-HESP-327/RTP1 fusion protein, *Mlp*-SSP15 can localize in multiple rust cell types. *Mlp*-SSP15

showed a very strong labeling at the periphery of the basal sporogenous hyphae in uredinia (Fig. 5I; Supplementary Fig. S11) whereas the newly formed urediniospores remained exempt of any signal. Interestingly, a labeling was also observed at the periphery of haustoria but not in infection hyphae (Fig. 5J) in the mesophyll, suggesting an accumulation of the protein at the interface between fungal haustoria and infected host cells. For each immunological detection, controls with preimmune serum did not show labeling of rust infected tissue.

## DISCUSSION

Sequencing genomes of pathogens provides an invaluable resource for biologists, and successful mining of sequenced genomes is a key to new research avenues. Here, we have used the data from the *M. larici-populina* genome to analyze, in detail, the 1,184 genes encoding SSP (Duplessis et al. 2011a). In order to identify candidate effectors from these secreted proteins, we designed a multifaceted approach based on gene family annotation, screening for positive selection, expression profiling, and protein immunolocalization in the pathosystem.

Duplessis and colleagues (2011a) described the largest SSP-encoding gene family in *M. larici-populina* as Cys-rich proteins with a conserved gene structure and short amino acid motifs. Annotation of other Cys-rich SSP classes revealed that a total of 228 genes shared this common gene structure as well as YXC and $CX_{2\text{-}3}Y$ motifs for many of them. We show that these motifs occur at higher frequencies at specific positions in the eight largest Cys-rich classes. Even though blastp 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 harboring YXC motifs were found to be expressed in haustoria and two classes were homologous to HESP from the flax rust fungus, *M. lini*. Recently, a new class of candidate effectors with a similar motif (Y/F/WXC) was identified in the barley powdery mildew, *B. graminis* f. sp. *hordei*, as well as in two cereal rusts, *P. graminis* f. sp. *tritici* and *Puccinia triticina* (Godfrey et al. 2010). The presence of this motif was confirmed in the secretome of *P. graminis* f. sp. *tritici* by genome-wide analysis (Duplessis et al. 2011a). Consistent with our observation in *M. larici-populina*, overall sequence identity was low and these proteins had a highly similar exon-intron structure. The in-depth analysis of *M. larici-populina* Cys-rich SSP classes strengthens previous observations (Duplessis et al. 2011a), indicating that YXC motifs are not restricted to the N-terminal region of these proteins, contrary to *B. graminis* (Godfrey et al. 2010), and that several YXC motifs can be detected. Moreover, other motifs can be present in Cys-rich SSP, such as the $CX_{2\text{-}3}Y$ motif found in many of the YXC-containing SSP. In the barley powdery mildew, putative effectors containing Y/F/WXC motifs were highly expressed in haustoria, suggesting that such secreted proteins could serve effector functions (Godfrey et al. 2010; Spanu et al. 2010).

In many plant pathogens, secreted proteins show accelerated evolutionary rates, consistent with the model of an "arms race" between the plant immune system and the secreted effectors of these pathogens (Dodds and Rathjen 2010; Terauchi and Yoshida 2010). Using a comparative genomic approach on EST data, Joly and associates (2010) reported accelerated evolution of genes encoding secreted proteins from four *Melampsora* spp. Although they identified positively selected genes encoding Cys-rich secreted proteins with no close homologue outside of the genus, their study focused mostly on orthologous genes expressed at the surface of poplar leaves (mostly in urediniospores, germ tubes, appressoria, and uredinia). Most positively selected *M. larici-populina* SSP-encoding paralogous families identified in the present study were also lineage specific and Cys rich but the majority were expressed in planta, suggesting possible interactions with host factors. Cys-rich proteins are common in effectors of fungal pathogens in which disulfide bridges could enhance stability in the plant apoplast (Stergiopoulos and de Wit 2009). Thus, uncovering such a large content of Cys-rich proteins in the genome of the poplar rust fungus was not surprising and strengthens this general feature (Duplessis et al. 2011a). Rather than providing resistance to apoplastic proteases, the presence of disulfide bonds in secreted proteins could be important for their structure and function (Kamoun 2006; Stergiopoulos and de Wit 2009). In several *M. larici-populina* SSP families under positive selection, Cys patterns were conserved despite low sequence identity. Although 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 SSP ideal for performing roles in recognition and specificity (Povolotskaya and Kondrashov 2010; Templeton et al. 1994). Thus, a high content in Cys residues could 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 of 95 CPG showed robust evidence of diversifying selection, 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 R proteins continues to expand (Jiang et al. 2008). In such cases, rapid sequence divergence in newly formed effector paralogs allows the number of effector genes in the pathogen to increase while minimizing the likelihood of host recognition (Jiang et al. 2008). 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, suggesting that this region interacts with host components, 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 (Allen et al. 2004; Barrett et al. 2009; Rehmany et al. 2005; Van der Merwe et al. 2009; Win et al. 2007). 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 (Schornack et al. 2009; Win et al. 2007).

Many oomycete effectors harbor the bipartite RXLR-EER amino acid motif that is dispensable for exocytosis but required for plant access (Dou et al. 2008; Whisson et al 2007). The RXLR motif appears to mediate entry of effectors into host cells independently of any pathogen-derived machinery (Dou et al. 2008). Some fungal effectors can also enter plant cells in the absence of the pathogen and, although they lack an RXLR domain and do not share obvious conserved peptide motifs, sequences loosely related to the RXLR motif are often found (Catanzariti et al. 2006; Kale et al. 2010; Manning and Ciuffeti 2005; Rafiqi et al. 2010). Some positively selected CPG encoding different families of small Cys-rich SSP were shown to possess various N-terminal motifs reminiscent of the RXLR-EER translocation signal, and several of these SSP-encoding genes were specifically expressed in planta. The RXLR motif is overrepresented and positionally constrained in the

secretome of *Phytophthora* and *Hyaloperonospora* spp. 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 (Jiang et al. 2008; Win et al. 2007). A systematic search to uncover such overrepresented motifs in other eukaryotes has revealed the overrepresentation of LXRR in *Puccinia* spp.-secreted proteins (E. Kemen and J. D. G. Jones, *personal communication*), and our data suggest that this motif is also represented in *M. larici-populina* candidate effectors.

In several plant–microbe interactions such as *Magnaporthe oryzae*–rice (Mosquera et al. 2009), *Ustilago maydis*–maize (Kämper et al. 2006), *B. graminis*–barley (Spanu et al. 2010), and *Laccaria bicolor*–poplar (Martin et al. 2008; Plett et al. 2011), secretome prediction along with transcript profiling led to the identification of candidate fungal effectors specifically expressed during colonization of plant tissues. Transcript profiling of *M. larici-populina* during time-course infection of poplar leaves revealed temporal and tissue-specific expression of fungal SSP (Duplessis et al. 2011b). This finding was corroborated in the present study by RT-qPCR for 34 selected SSP. This result highlights the intricate processes of cellular control developed by plant pathogens during host colonization through delivery of signal molecules. Most of the SSP transcripts assessed in the present study showed specific accumulation during or after haustoria formation and several encode homologues of rust effectors such as *U. fabae* RTP1 (Kemen et al. 2005) and *M. lini* avirulence factors (Dodds et al. 2004; Ellis et al. 2009). Interestingly, several SSP transcripts induced at 96 hpi in planta showed lower expression levels at 168 hpi. These results are consistent with the transcriptional switch observed at the uredinial stage of *M. larici-populina* between biotrophy- and sporulation-related areas isolated using laser-capture microdissection (Hacquard et al. 2010). The panel of tested genes included seven members of CPGH1 and CPG5464 (*M. lini AvrP4* homologues) that were not predicted in the automatic annotation of the poplar rust genome and that were not represented on the *M. larici-populina* custom oligoarray. RT-qPCR profiles obtained with specific primers validated their expression during host infection, and showed distinct expression profiles within the same gene family, as previously shown for homologues of *M. lini* HESP-417 in *M. larici-populina* (Duplessis et al. 2011b). Thus, the gene diversity among CPG assessed by positive selection analyses seems to be mirrored in the diversity of expression profiles observed by RT-qPCR during infection. Because our study specifically focused on the infection of the telial host (i.e., poplar leaves), it remains to be determined to what extent these SSP-encoding genes might also be expressed during infection of larch needles, the aecial host.

To manipulate host cell functions, many effector proteins target host cell components during plant tissue invasion (Dodds et al. 2009). In *M. larici-populina*, several putative rust effectors immunolocalize at the periphery of haustorial structures during colonization of poplar leaves. For many fungal effectors, translocation inside host cell has been hypothesized because most Avr proteins are recognized by cytoplasmic R proteins (Dodds et al. 2006; Houterman et al. 2009). Only a few examples reported direct evidence of effector translocation inside the plant cell (Kemen et al. 2005; Khang et al. 2010; Plett et al. 2011; Rafiqi et al. 2010) or into the plant apoplast (Doehlemann et al. 2009). In rust fungi, immunolocalization of RTP1 from *U. fabae* and AvrM from *M. lini* showed effector accumulation in the extrahaustorial matrix but also inside the plant cell, suggesting that haustorial structure might be involved in effector secretion (Kemen et al. 2005; Rafiqi et al. 2010). Interestingly, accumulation of SSP (*Mlp*-HESP-327/RTP1, *Mlp*-AvrL567, and *Mlp*-SSP15) was observed at the periphery of *M. larici-populina* haustorial structure. The absence of labeling inside the host cell does not seem to support protein internalization. Nevertheless, we cannot exclude the possibility that a low quantity of fungal effector could be released inside host cells at undetectable levels by immunodetection (Rafiqi et al. 2010) or that their delivery into host cells is temporally regulated. The candidate effectors *Mlp*-SSP15 and *Mlp*-HESP-327/RTP1 displayed localization in different fungal cell types. Such a localization at distinct sites was previously shown for *M. lini* AvrM, which was detected both at hyphal tips and in haustoria during flax infection (Rafiqi et al. 2010). In spite of the cytoplasmic and nuclear localization reported for *U. fabae* RTP1 (Kemen et al. 2005), the localization of *Mlp*-HESP-327/RTP1 was restricted to the periphery of haustoria and infection hyphae and was not detected inside plant cells. During flax–flax rust interaction, intracellular recognition between AvrL567 and the cognate resistance genes suggests effector translocation from haustoria toward infected host cell (Dodds et al. 2004, 2006). *M. lini* AvrL567 uptake in the cytoplasm of plant cells is dependent on an N-terminal region (amino acid position between 26 and 50) that contains an RFYR motif reminiscent of the RXLR motif found in oomycete effectors (Kale et al. 2010; Rafiqi et al. 2010). *M. lini* and *M. larici-populina* AvrL567 homologues show secondary structure similarities; however, the RFYR motif is not conserved and the 26- to 50-amino acid region shows poor similarity. Beside, *Mlp*-AvrL567 displayed a specific localization at the periphery of haustoria and was not detected in the host cytoplasm, consistent with an accumulation in the extrahaustorial matrix. By sequencing *Mlp-AvrL567* alleles in 32 poplar rust isolates of known pathotypes, we were not able to detect either extensive polymorphism or any signatures of diversifying selection, contrasting with the positive selection identified in *M. lini AvrL567* (Dodds et al. 2006). The striking differences between *M. lini* and *M. larici-populina* AvrL567 homologues might illustrate the specific evolution of the effector arsenal within the *Melampsora* genus.

The recent analysis of the genome sequence of the poplar leaf rust fungus *M. larici-populina* paved the way for detailed investigations of gene families related to biotrophic lifestyle and pathogenicity (Duplessis et al. 2011a). In this study, we presented a detailed analysis of the 1,184 SSP in *M. larici-populina.* Genome-wide patterns of SSP expression and variation helped to describe a large reservoir of candidate effectors in the poplar rust fungus. Hopefully, the genetic transformation system developed for *U. fabae* (Djulic et al. 2011) should help to validate the function of these candidates in the poplar–poplar rust pathosystem. A better understanding of effectors' evolution will allow the rational design of experiments to assess functions of positively evolving gene families. Globally, effector identification in rust pathogens and their functional characterization should facilitate the comprehension of obligate biotrophy. Information regarding the function of *M. larici-populina* effectors, combined with accurate data on the *Populus* immune system, will be crucial for elucidating mechanisms underlying processes such as virulence and ETI and is likely to lead to new breeding strategies. More rust genomes are currently being sequenced (Broad Institute and Joint Genome Institute), which will clearly enhance our capacity to discern patterns in related effector repertoires.

## MATERIALS AND METHODS

**Genomic resources.**

Genome sequence data of the *M. larici-populina* 98AG31 (i.e., sequences of assembled scaffolds, sets of predicted genes and deduced protein sequences, and EST) were downloaded

from the Joint Genome Institute and additional sequence data (i.e., in planta *M. larici-populina* EST, basidiomycete genomic data for comparative analyses) were available to the *Melampsora* Genome Consortium at the MycorWeb website. Comparisons to the *P. graminis* f. sp. *tritici* genome sequence and nucleic or proteic datasets were performed through the Puccinia Group database webportal at the Broad Institute.

**Sequence analysis.**

In silico predictions of secreted proteins were carried out using combinations of SignalP 3.0, TargetP 1.1, and TMHMM 2.0 (Emanuelsson et al. 2007) as previously reported by Duplessis and associates (2011a), based on different criteria for SignalP-HMM Sprob score, SignalP-NN Smax and D scores, and TargetP signal peptide prediction (Klee and Ellis 2005), and SSP were selected based on an arbitrary cut-off of 300 amino acids (Duplessis et al. 2011a). 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 (Altschul et al. 1997). Multiple alignments were conducted using the programs ClustalW (Thompson et al. 1994) and MUSCLE (Edgar 2004) and adjusted manually. Sequence alignments were submitted to the WebLogo server to generate graphical consensus displays. Phylogenetic analyses of the AvrP4 family was performed using Bayesian inference and Markov chain Monte Carlo simulations (B/MCMC) implemented in MrBayes version 3.0 (Ronquist and Huelsenbeck 2003). The models for nucleotide substitutions were selected prior to the MCMC using the likelihood ratio test implemented in the Modeltest 3.7 program (Posada and Crandall 1998). One cold and seven incrementally heated chains were run for one 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.

**Sequence randomization analysis.**

The conservation of the YXC and $CX_{2\text{-}3}Y$ motifs in each Cys-rich SSP class was assessed by permutating amino acids in SSP using two type of sequence randomization. In a first test, after removal of predicted secretory leader, SSP sequences were randomly shuffled 1,000 times. Positions of YXC and $CX_{2\text{-}3}Y$ motifs in each class (true SSP class and permutated SSP set) were detected by the FUZZPRO program in the EMBOSS package. A permutation test was applied to determine whether the observed occurrences of YXC and $CX_{2\text{-}3}Y$ motifs in the different classes happened by chance. Using the reshuffled sequences with the same amino acid composition as the real proteins, the observed motif frequencies per class were compared with the expected values based on 1,000 interactions. The expected motif occurrences reflecting the null model of motif conservation were used to determine a *P* value for motif constraint. Based on the conservation profiles, the third and eighth Cys residues with the surrounding peptides always compose the YXC motif (except class II) and the sixth Cys residue composes the $CX_{2\text{-}3}Y$ motif. To further confirm the importance of the conserved Cys residues, protein sequences of five amino acids upstream and downstream of the motifs were defined as core regions. Therefore, classes I, III, and XXIV have three core regions and class II has two core regions. For example, the conserved Cys position in the YXC motif was always kept at position eight of the core region ($X_7CX_5$) and the sixth position in the $CX_{2\text{-}3}Y$ motif core region was always Cys ($X_5CX_{7\text{-}8}$). Peptides in the core region of each SSP classes were extracted to build the HMM (Eddy 1998). The significance of the Cys and surrounding residues were measured by the aligned HMM score. Peptides in the core region were randomly shuffled 1,000 times whereas the Cys position was fixed, and the frequencies of the HMM scores between the original core peptides and the randomized core peptides were compared.

**Positive selection analyses.**

*M. larici-populina* multigene families were generated by Tribe-MCL (Duplessis et al. 2011a) and enlarged using recursive tblastn searches (*E* value < 1e-6). Using this approach, 536 *M. larici-populina* genes encoding secreted proteins were grouped into CPG that contained a minimum of three sequences. CPG were manually edited to remove either poorly aligned regions (e.g., large indels) or sequences for which similarity was not found throughout the majority of the coding sequence (i.e., less than 50%). The resulting 95 CPG were then submitted to positive selection analyses using a suite of programs regrouped in a single Python script, as previously described (Joly et al. 2010).

**Plant material and fungal inoculation procedure.**

Plant infection experiments were performed using strain 98AG31 (pathotype 3-4-7) of *M. larici-populina* on *Populus trichocarpa* × *P. deltoides* 'Beaupré' leaves (compatible interaction). Urediniospores of *M. larici-populina* were propagated on detached leaves of susceptible *P. deltoides* × *P. 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 h 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 were immediately fixed in 4% (wt/vol) paraformaldehyde for microscopy analyses or snap frozen in liquid nitrogen and kept at –80°C for further nucleic acid isolation.

**RNA isolation.**

Isolation of total RNA was performed with the RNeasy Plant Mini kit (Qiagen, Courtaboeuf, France) from 1 mg of resting and germinated spores, and from 100 mg of infected leaf tissues during the compatible interaction (2 to 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).

**Expression profiling during time-course infection.**

Expression profiling of 34 SSP-encoding genes was monitored by RT-qPCR and using *M. larici-populina* custom oligoarrays as previously described (Duplessis et al. 2011b; Hacquard et al. 2010, 2011b). For RT-qPCR, specific primers amplifying fragments ranging from 84 to 307 nucleotides were designed with the Primer 3 and Amplify 3X programs. Specific primers for genes encoding protein ID numbers 58459, 70587, 70656, 71404, and 123524 were previously described (Hacquard et al. 2010). 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 synthesis and PCR amplifications were performed as previously described (Duplessis et al. 2011b; Hacquard et al. 2010, 2011b). Expression of the *M. larici-populina* candidate genes was determined using the $2^{-\Delta Ct}$ calculation and calibrated to the highest level of expression ob-

served (Livak and Schmittgen 2001). *M. larici-populina* transcript levels were normalized with α-tubulin (*Mlp-aTUB*) and elongation factor (*Mlp-ELF1α*) (Hacquard et al. 2011b). Time-course expression data obtained with *M. larici-populina* custom oligoarrays (Duplessis et al. 2011b) are available under the National Center for Biotechnology Information Gene Expression Omnibus accession GSE21624.

**Laser-scanning confocal microscopy.**

Fungal structures were visualized in infected leaves at 96 and 168 hpi. Leaves were cut and fixed for 3 h at 4°C in 4% (wt/vol) paraformaldehyde prepared in phosphate-buffered saline (PBS), pH 7, embedded in 6% agarose (wt/vol), and cut into 15- and 20-μm sections using a vibratome VT1000S (Leica, Nanterre, France). Propidium iodide and Uvitex staining were performed on 20-μm sections as previously described (Hacquard et al. 2010). For indirect immunofluorescent localization, peptides were synthesized for *Mlp*-AvrL567, *Mlp*-SSP15, *Mlp*-RTP1sc31, and *Mlp*-HESP-327/RTP1 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 immunoglobulin G (IgG) fractions as well as preimmune sera were purified and desalted using the MabTrap kit (GE Healthcare, Orsay, France) according to the manufacturer's recommendations. Transversal sections of 15 μm from infected leaves of Beaupré poplar were fixed and embedded in agarose as described above. Immunolocalization was performed essentially as previously described (Martin et al. 2008), with the following modifications. Sections were digested for 10 min in PBS buffer supplemented with 0.1% (wt/vol) cellulase, 0.01% (wt/vol) pectolyase, and 0.1% (wt/vol) bovine serum albumen (BSA). After digestion, sections were washed five times for 5 min each with PBS buffer and then incubated in PBS containing 1% (wt/vol) 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, 0.023, 0.022, and 0.022 mg/ml, respectively, in PBS buffer with 1% BSA (wt/vol). Sections were incubated with IgG purified from preimmune sera at the same concentrations for control observation. After five PBS washes, sections were incubated in PBS containing the goat anti-rabbit IgG Alexa fluor 488 conjugate (1:400) used as secondary antibody (Invitrogen, Cergy Pontoise, France). 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.

***Mlp*-AvrL567 polymorphism analysis.**

In all, 32 dikaryotic *M. larici-populina* isolates displaying various combinations of virulences were selected in a collection available at INRA Nancy (Barrès et al. 2006). DNA was isolated using the DNeasy plant mini kit (Qiagen). *AvrL567* alleles were amplified with specific primers (5′-ATGAAGATTC ATCTATCATTGAAAGCC and 3′-TCACCACTTCTTGGGTT TTGGTA) defined after JGI protein ID number 37347. Amplified fragments were purified with the QIAquick PCR purification kit (Qiagen) and directly sequenced according to GenomeLab Dye terminator cycle sequencing with Quick Start kit (Beckman Coulter, Villepinte, France). For isolates showing allele polymorphism at the *AvrL567* locus, amplicons were purified and cloned in pCR 4-TOPO using the TOPO TA cloning kit (Invitrogen). Plasmidic DNA was isolated with the GeneJet plasmid miniprep kit (Fermentas Life Science, St. Rémy Les Chevreuse, France) and *AvrL567* alleles were sequenced as mentioned above using pCR 4-TOPO M13 5′ and 3′ primers.

## ACKNOWLEDGMENTS

This research was supported by grants from the INRA and the Région Lorraine council to S. Duplessis and F. Martin, from Natural Resources Canada to R. Hamelin, and from IUAP P6/25 (BioMaGNet) to P. Rouzé and Y. Van de Peer. We thank the Joint Genome Institute (JGI) for allowing early access to the *M. larici-populina* genome sequence. JGI sequencing is supported by the Office of Science of the United States Department of Energy under contract number DE-AC02-05CH11231. We thank E. Morin at INRA Nancy (France) for bioinformatic support and K. Vandepoele at Ghent University (Belgium) for biostatistic support. S. Hacquard, D. L. Joly, F. Martin, R. C. Hamelin, and S. Duplessis conceived and designed the experiments; S. Hacquard, D. L. Joly, E. Tisserant, Y.-C. Lin, N. Feau, B. Petre, C. Delaruelle, and P. Tanguay performed the experiments; S. Hacquard, D. L. Joly, Y.-C. Lin, and S. Duplessis analyzed the data; V. Legué, A. Kohler, P. Frey, Y. Van de Peer, and P. Rouzé contributed reagents, materials, or analysis tools; and S. Hacquard, D. L. Joly, Y.-C. Lin, B. Petre, N. Feau, and S. Duplessis wrote the article.

## LITERATURE CITED

Allen, R. L., Bittner-Eddy, P. D., Grenville-Briggs, L. J., Meitz, J. C., Rehmany, A. P., Rose, L. E., and Beynon, J. L. 2004. Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. Science 306:1957-1960.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Barrès, B., Dutech, C., Andrieux, A., Caron, H., Pinon, J., and Frey, P. 2006. Isolation and characterization of 15 microsatellite loci in the poplar rust fungus, *Melampsora larici-populina*, and cross-amplification in related species. Mol. Ecol. Notes 6:60-64.

Barrett, L. G., Thrall, P. H., Dodds, P. N., van der Merwe, M., Linde, C. C., Lawrence, G. J., and Burdon, J. J. 2009. Diversity and evolution of effector loci in natural populations of the plant pathogen *Melampsora lini*. Mol. Biol. Evol. 26:2499-2513.

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, J. L., Chibucos, M. C., Coates, M., Dehal, P., Delehaunty, K., Dong, S., Downtown, P., Dumas, B., Fabro, G., Fronick, C., Fuerstenberg, S. I., Fulton, L., Gaulin, E., Govers, F., Hughes, L., Humphray, S., Jiang, R. H. Y., 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, S. W., Rogers, J., Van den Ackerveken, G., Jones, J. D. G., McDowell, J. M., Beynon, J., and Tyler, B. M. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science 330:1549-1551.

Bolton, M. D., Van Esse, H. P., Vossen, J. H., De Jonge, R., Stergiopoulos, I., Stulemeijer, I. J. E., Van Den Berg, G. C. M., Borrás-Hidalgo, O., Dekker, H. L., De Koster, C. G., De Wit, P. J. G. M., Joosten, M. H. A. J., and Thomma, B. P. H. J. 2008. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. Mol. Microbiol. 69:119-136.

Brefort, T., Doehlemann, G., Mendoza-Mendoza, A., Reissmann, S., Djamei, A., and Kahmann, R. 2009. *Ustilago maydis* as a pathogen. Annu. Rev. Phytopathol. 47:423-445.

Brunner, P. C., Keller, N., and McDonald, B. A. 2009. Wheat domestication accelerated evolution and triggered positive selection in the β-xylosidase enzyme of *Mycosphaerella graminicola*. PLoS ONE 4:e7884. Published online.

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.

de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, T., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H. A. J., and Thomma, B. P. H. J. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science 329:953-955.

Djulic, A., Schmid, A., Lenz, H., Sharma, P., Koch, C., Wirsel, S. G. R., and Voegele, R. T. 2011. Transient transformation of the obligate biotrophic rust fungus *Uromyces fabae* using biolistics. Fungal Biol. 115:633-642.

Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: Towards an integrated view of plant-pathogen interactions. Nat. Rev. Genet. 11:539-548.

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., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A., Kobe, B., and Ellis, J. G. 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.

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.

Doehlemann, G., van der Linde, K., Assmann, D., Schwammbach, D., Hof, A., Mohanty, A., Jackson, D., and Kahman, R. 2009. Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. PLoS Pathog. 5:e1000290. Published online.

Dou, D., Kale, S. D., Wang, X., Jiang, R. H. Y., Bruce, N. A., Arredondo, F. D., Zhang, X., and Tyler, B. M. 2008. RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. Plant Cell 20:1930-1947.

Duplessis, S., Major, I., 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.

Duplessis, S., Cuomo, C. A., Lin, Y, C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D. L., Hacquard, S., Amselem, J., Cantarel, B., Chiu, R., Coutinho, P., Feau, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M., Kodira, C., Kohler, A., Hües, U., Lindquist, E. A., Lucas, S., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J. L., Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Schmutz, J., Selles, B., Shapiro, H., Tanguay, P., Tuskan, G.A., Henrissat, B., Van de Peer, Y., Rouzé, P., Ellis, J. G., Dodds, P. N., Schein, J. E., Zhong, S., Hamelin, R. C., Grigoriev, I. V., Szabo, L. J., and Martin, F. 2011a. Obligate biotrophy features unravelled by the genomic analysis of rust fungi. Proc. Natl. Acad. Sci. U.S.A. 108:9166-9171.

Duplessis, S., Hacquard, S., Delaruelle, C., Tisserant, E., Frey, P., Martin, F., and Kohler, A. 2011b. *Melampsora larici-populina* transcript profiling during germination and timecourse infection of poplar leaves reveals dynamic expression patterns associated with virulence and biotrophy. Mol. Plant-Microbe Interact. 24:808-818.

Eddy, S.R. 1998. Profile hidden Markov models. Bioinformatics 14:755-763.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Ellis, J. G., and Dodds, P. N. 2011. Showdown at the RXLR motif: Serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. Proc. Natl. Acad. Sci. U.S.A. 108:14381-14382.

Ellis, J. G., Rafiqi, M., Gan, P., Chakrabarti, A., and Dodds, P. N. 2009. Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. Curr. Opin. Plant Biol. 12:399-405.

Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. Nat. Prot. 2:953-971.

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.

Fernandez, D., Tisserant, E., Talhinhas, P., Azinheira, H., Vieira, A., Petitot, A. S., Loureiro, A., Poulain, J., Da Silva, C., Silva, M. d. C., and Duplessis, S. 2012. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. Mol. Plant. Pathol. 13:17-37.

Flor, H. H. 1955. Host-parasite interaction in flax rust—its genetics and other implications. Phytopathology 45:680-685.

Godfrey, D., Böhlenius, H., Pedersen, C., Zhang, Z., Emmersen, J., and Thordal-Christensen, H. 2010. Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genomics 11:317.

Göhre, V., and Robatzek, S. 2008. Breaking the barriers: Microbial effector molecules subvert plant immunity. Annu. Rev. Phytopathol. 46:189-215.

Guttman, D. S., Gropp, S. J., Morgan, R. L., and Wang, P. W. 2006. Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae*. Mol. Biol. Evol. 23:2342-2354.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y, Handsaker, R. E., Cano, L. M, Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M. V., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I. B., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grünwald, N. J., Horn, K., Horner, N. R., Hu, C. H., Huitema, E., Jeong, D. H., Jones, A. M. E., Jones, J. D. G., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L. J., MacLean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J. G., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzòn, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J. I., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R. J., Whisson, S. C., Judelson, H. S., and 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., 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., Petre, B., Frey, P., Hecker, A., Rouhier, N., and Duplessis, S. 2011a. The poplar-poplar rust interaction: Insights from genomics and transcriptomics. J. Pathogens. doi:10.4061/2011/716041. Published online.

Hacquard, S., Veneault-Fourrey, C., Delaruelle, C., Frey, P., Martin, F., and Duplessis, S. 2011b. Validation of *Melampsora larici-populina* reference genes for in planta RT-quantitative PCR expression profiling during time-course infection. Physiol. Mol. Plant Pathol. 75:106-112.

Houterman, P. M., Ma, L., van Ooijen, G., de Vroomen, M. J., Cornelissen, B. J., Takken, F. L., and Rep, M. 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.

Jiang, R. H. Y., Tripathy, S., Govers, F., and Tyler, B. M. 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.

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.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Kale, S. D., Gu, B., Capelluto, D. G., Dou, D., Feldman, E., Rumore, A., Arredondo, F. D., Hanlon, R., Fudal, I., Rouxel, T., Lawrence, C. B., Shan, W., and Tyler, B. M. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell 142:284-295.

Kamoun, S. 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. Annu. Rev. Phytopathol. 44:41-60.

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-Martin, J., Feldbrügge, M., Basse, C. W., Steinberg, G., Ibeas, J. I., Holloman, W., Guzman, P., Farman, M., Stajich, J. E., Sentandreu, R., Gonzaléz-Prieto, J. M., Kennell, J. C., Molina, L., Shirawski, J., Mendoza-Mendoza, A., Greilinger, D., Münch, K., Rössel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E. C. H., Cahill, M.J., Boyce, K. J., Klose, J., Klosterman, S. J., Deelstra, H. J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P. H., Häuser-Hahn, I., Vaupel, M., Koopman, 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., 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., and Voegele, R. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. Mol. Plant-Microbe Interact. 18:1130-1139.

Khang, C. H., Berruyer, R., Giraldo, M. C., Kankanala, P., Park, S. Y., Czymmek, K., Kang, S., and Valent, B. 2010. Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell 22:1388-1403.

Klee, E. W., and Ellis, L. B. M. 2005. Evaluating eukaryotic secreted protein prediction. BMC Bioinform. 6:256.

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

Livak, K. J., and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(–Delta Delta C(T)) method. Methods 25:402-408.

Manning, V. A., and Ciuffetti, L. M. 2005. Localization of Ptr ToxA produced by *Pyrenophora tritici-repentis* reveals protein import into wheat mesophyll cells. Plant Cell 17:3203-3212.

Martin, F., Aerts, A., Ahrén, D., Brun, A., Danchin, E. G. J., Duchaussoy, F., Gibon, J., Kohler, A., Lindquist, E., Pereda, V., Salamov, A., Shapiro, H. J., Wuyts, J., Blaudez, D., Buée, M., Brokstein, P., Canbäck, B., Cohen, D., Courty, P. E., Coutinho, P. M., Delaruelle, C., Detter, J. C., Deveau, A., DiFazio, S., Duplessis, S., Fraissinet-Tachet, L., Lucic, E., Frey-Klett, P., Fourrey, C., Feussner, I., Gay, G., Grimwood, J., Hoegger, P. J., Jain, P., Kilaru, S., Labbé, J., Lin, Y. C., Legué, V., Le Tacon, F., Marmeisse, R., Melayah, D., Montanini, B., Muratet, M., Nehls, U., Niculita-Hirzel, H., Oudot-Le Secq, M. P., 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, G. K., Polle, A., Pukkila, P. J., Richardson, P. M., Rouzé, P., Sanders, I. R., Stajich, J. E., Tunlid, A., Tuskan, G. A., and Grigoriev, I. V. 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452:88-92.

Morgan, W., and Kamoun, S. 2007. RXLR effectors of plant pathogenic oomycetes. Curr. Opin. Microbiol. 10:332-338.

Mosquera, G., Giraldo, M. C., Hyun Khang, C., Coughlan, S., and 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.

Pallaghy, P. K., Norton, R. S., Nielsen, K. J., and Craik, D. J. 1994. A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides. Protein Sci. 3:1833-1839.

Panstruga, R., and Dodds, P. N. 2009. Terrific protein traffic: The mystery of effector protein delivery by filamentous plant pathogens. Science 324:748-750.

Pinon, J., and Frey, P. 2005. Interaction between poplar clones and *Melampsora* populations and their implication for breeding for durable resistance. Pages 139-154 in: Rust Diseases of Willow and Poplar. M. H. Pei and A. R. McCracken, eds. CABI Publishing, Cambridge.

Plett, J. M., Kemppainen, M., Kale, S. D., Kohler, A., Legué, V., Brun, A., Tyler, B. M., Pardo, A. G., and Martin, F. 2011. A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. Curr. Biol. 21:1197-1203.

Posada, D., and Crandall, K. A. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817-818.

Povolotskaya, I. S., and Kondrashov, F. A. 2010. Sequence space and the ongoing expansion of the protein universe. Nature 465:922-926.

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.

Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., Jiang, R. H. Y., Zody, M. C., Kunjeti, S. G., Donofrio, N. M., Meyers, B. C., Nusbaum, C., and Kamoun, S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330:1540-1543.

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.

Ravensdale, M., Nemri, A., Thrall, P. H., Ellis, J. G., and Dodds, P. N. 2011. Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. Mol. Plant Pathol. 12:93-102.

Rehmany, A. P., Gordon, A, Rose, L. E., Allen, R. L., Armstrong, M. R., Whisson, S. C., Kamoun, S., Tyler, B. M., Birch, P. R. J., and Beynon, J. L. 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. 2005. Small proteins of plant-pathogenic fungi secreted during host colonization. FEMS (Fed. Eur. Microbiol. Soc.) 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., 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.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

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.

Spanu, P. D., Abbott, J. C., Amselem, J., Burgis, T. A., Soanes, D. M., Stüber, K., Ver Loren van Themaat, E., Brown, J. K. M., Butcher, S. A., Gurr, S. J., Lebrun, M. H., Ridout, C. J., Schulze-Lefert, P., Talbot, N. J., Ahmadinejad, N., Ametz, C., Barton, G. R., Benjdia, M., Bidzinki, P., Bindschedler, L. V., Both, M., Brewer, M. T., Cadle-Davidson, L., Cadle-Davidson, M. M., Collemare, J., Cramer, R., Frenkel, O., Godfrey, D., Harriman, J., Hoede, C., King, B. C., Klages, S., Kleeman, J., Knoll, D., Kotti, P. S., Kreplak, J., Lopez-Ruiz, F. J., Lu, X., Maekawa, T., Mahanil, S., Micali, C., Milgroom, M. G., Montana, G., Noir, S., O'Connell, R. J., Oberhaensli, S., Parlange, F., Pedersen, C., Quesneville, H., Reinhardt, R., Rott, M., Sacristan, S., Schmidt, S. M., Schön, M., Skamnioti, P., Sommer, H., Stephens, A., Takahara, H., Thordal-Christensen, H., Vigouroux, M., Weßling, R., Wicker, T., and Panstruga, R. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543-1546.

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

Stergiopoulos, I., and de Wit, P. J. G. M. 2009. Fungal effector proteins. Annu. Rev. Phytopathol. 47:233-263.

Stergiopoulos, I., van den Burg, H. A., Ökmen, B., Beenen, H. G., van Liere, S., Kema, G. H. J., and de Wit, P. J. G. M. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. Proc. Natl. Acad. Sci. U.S.A. 107:7610-7615.

Templeton, M. D., Rikkerink, E. H. A., and Beever, R. E. 1994. Small, cysteine-rich proteins and recognition in fungal-plant interactions. Mol. Plant-Microbe Interact. 7:320-325.

Terauchi, R., and Yoshida, K. 2010. Towards population genomics of effector-effector target interactions. New Phytol. 187:929-939.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. 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.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M. B., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K.H., Lee, M. K., McDonald, W. H., Medina, M., Meijer, H. J. G., Nordberg, E. K., MacLean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K. C., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W. S., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261-1266.

van den Hooven, H. W., van den Burg, H. A., Vossen, P., Boeren, S., de Wit, P. J., and Vervoort, J. 2001. Disulfide bond structure of the AVR9 elicitor of the fungal tomato pathogen *Cladosporium fulvum*: Evidence for a cystine knot. Biochemistry 40:3458-3466.

Van der Merwe, M. M., Kinnear, M. W., Barrett, L. G., Dodds, P. N., Ericson, L., Thrall, P. H., and Burdon, J. J. 2009. Positive selection in *AvrP4* avirulence gene homologues across the genus *Melampsora*. Proc. R. Soc. B 276:2913-2922.

Voegele, R. T., and Mendgen, K. 2003. Rust haustoria: Nutrient uptake and beyond. New Phytol. 159:93-100.

Voegele, R. T., Hahn, M., and Mendgen, K. 2009. The Uredinales: Cytology, biochemistry, and molecular biology. Pages 69-98 in: The Mycota, Volume 5: Plant Relationships. H. B. Deising, ed. Springer, Berlin.

Wang, C. I., Guncar, G., Forwood, J. K., The, T., Catanzariti, A. M., Lawrence, G. J., Loughlin, F. E., Mackay, J. P., Schirra, H. J., Anderson, P. A., Ellis, J. G., Dodds, P. N., and 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.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., Armstrong, M. R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I. K., Leighton, P., and Birch, P. R. J. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature 450:115-118.

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A., Ammar, R., Staskawicz, B. J., and Kamoun, S. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. Plant Cell 19:349-2369.

Yaeno, T., Li, H., Chaparro-Garcia, A., Schornack, S., Koshiba, S., Watanabe, S., Kigawa, T., Kamoun, S., and Shirasu, K. 2011. Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. Proc. Natl. Acad. Sci. U.S.A. 108:14682-14687.

Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586-1591.

Zhou, J. M., and Chai, J. 2008. Plant pathogenic bacterial type III effectors subdue host responses. Curr. Opin. Microbiol. 11:179-185.

## AUTHOR-RECOMMENDED INTERNET RESOURCES

Broad Institute Puccinia Group database: www.broadinstitute.org/annotation/genome/puccinia_group/MultiHome.html

Department of Energy Joint Genome Institute fungi portal: genome.jgi-psf.org/programs/fungi/index.jsf

Department of Energy Joint Genome Institute *Melampsora* genome portal: genome.jgi-psf.org/Mellp1/Mellp1.home.html

INRA MycorWeb Genome Resources: mycor.nancy.inra.fr/genomeResources.php

National Center for Biotechnology Information Gene Expression Omnibus database: www.ncbi.nlm.nih.gov/geo

University of California Berkeley WebLogo server: weblogo.berkeley.edu