HAL open science

The Rust Fungus Melampsora larici-populina Expresses a Conserved Genetic Program and Distinct Sets of Secreted Protein Genes During Infection of Its Two Host Plants, Larch and Poplar

Cécile Lorrain, Clémence Marchal, Stéphane Hacquard, Christine Delaruelle, Jeremy Petrowski, Benjamin Petre, Arnaud Hecker, Pascal Frey, Sébastien Duplessis

To cite this version:

Cécile Lorrain, Clémence Marchal, Stéphane Hacquard, Christine Delaruelle, Jeremy Petrowski, et al.. The Rust Fungus Melampsora larici-populina Expresses a Conserved Genetic Program and Distinct Sets of Secreted Protein Genes During Infection of Its Two Host Plants, Larch and Poplar. Molecular Plant-Microbe Interactions, 2018, 31 (7), pp.695-706. 10.1094/MPMI-12-17-0319-R . hal-01893935

HAL Id: hal-01893935

https://hal.science/hal-01893935v1

Submitted on 14 Sep 2023

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

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

Distributed under a Creative Commons Attribution 4.0 International License

MPMI Vol. 31, No. 7, 2018, pp. 695–706. https://doi.org/10.1094/MPMI-12-17-0319-R

e-Xtra*

The Rust Fungus *Melampsora larici-populina* Expresses a Conserved Genetic Program and Distinct Sets of Secreted Protein Genes During Infection of Its Two Host Plants, Larch and Poplar

Cécile Lorrain,[1] Clémence Marchal,[1] Stéphane Hacquard,[1] Christine Delaruelle,[1] Jérémy Pétrowski,[1] Benjamin Petre,[1,2] Arnaud Hecker,[1] Pascal Frey,[1] and Sébastien Duplessis[1,†]

[1]INRA/Université de Lorraine, UMR 1136 Interactions Arbres/Microorganismes, INRA Centre Grand Est-Nancy, F-54280 Champenoux, France; and [2]The Sainsbury Laboratory, Norwich Research Park, NR4 7UH, Norwich, U.K.

Accepted 14 January 2018.

Mechanisms required for broad-spectrum or specific host colonization of plant parasites are poorly understood. As a perfect illustration, heteroecious rust fungi require two alternate host plants to complete their life cycles. *Melampsora larici-populina* infects two taxonomically unrelated plants, larch, on which sexual reproduction is achieved, and poplar, on which clonal multiplication occurs, leading to severe epidemics in plantations. We applied deep RNA sequencing to three key developmental stages of *M. larici-populina* infection on larch: basidia, pycnia, and aecia, and we performed comparative transcriptomics of infection on poplar and larch hosts, using available expression data. Secreted protein was the only significantly overrepresented category among differentially expressed *M. larici-populina* genes between the basidial, the pycnial, and the aecial stages, highlighting their probable involvement in the infection process. Comparison of fungal transcriptomes in larch and poplar revealed a majority of rust genes were commonly expressed on the two hosts and a fraction exhibited host-specific expression. More particularly, gene families encoding small secreted proteins presented striking expression profiles that highlight probable candidate effectors specialized on each host. Our results bring valuable new information about the biological cycle of rust fungi and identify genes that may contribute to host specificity.

Rust fungi (Pucciniales) are obligate biotrophs that establish parasitic associations with their host plants. Pucciniales depict one of the largest orders of plant pathogenic fungi, with more than 8,000 species infecting a wide range of hosts within ferns, gymnosperms, and angiosperms (Aime et al. 2014; Cummins and Hiratsuka 2003). A remarkable feature of some rust fungi is heteroecism—they infect two unrelated host plants to complete their life cycle. To do so, they produce different spore forms, and the life cycle is qualified as macrocyclic. For instance, the poplar rust fungus *Melampsora larici-populina* has a heteroecious and macrocyclic life cycle (Hacquard et al. 2011a; Vialle et al. 2011). It infects two unrelated host plants, Larch (*Larix* spp., conifer, aecial host) and Poplar (*Populus* spp., dicot, telial host). Infection of larch needles is achieved by *M. larici-populina* haploid basidiospores and results in the formation of pycnia and pycniospores (Hacquard et al. 2011a; Pinon and Frey 2005). After fertilization and production of aecia, airbone aecispores infect poplar leaves, leading to the production of urediniospores within a week (Harder 1984; Voegele et al. 2009). Asexual multiplication of urediniospores causes severe epidemics in poplar plantations all summer (Pinon and Frey 2005). In autumn, uredinia differentiate into telia, in which karyogamy occurs and diploid teliospores overwinter on dead poplar leaves (Hacquard et al. 2013). When favorable conditions are met the next spring, basidiospores are produced and the cycle repeats again. The molecular bases of these tightly controlled developmental stages and the genetic programs underlying host alternation in rust fungi remain poorly understood (Bakkeren et al. 2016; Liu et al. 2015; Xu et al. 2011).

During infection, pathogens use effector proteins to modulate the cellular and the immune processes of the host plant (Lo Presti et al. 2015; Win et al. 2012). Effectors are critical to ensure successful establishment of the pathogen in host tissues (Lo Presti et al. 2015). In filamentous plant pathogens (fungi and oomycetes), effector proteins are often small proteins that possess a signal peptide for secretion (Sperschneider et al. 2015). In rust fungi, genome and transcriptome analyses predicted hundreds of secreted proteins (SPs) representing candidate effectors expressed during host infection (Duplessis et al. 2012, 2014). Effectors are specifically adapted to manipulate host plants and they evolve through the pressure of host recognition receptors (Rovenich et al. 2014).

The genetic program used by *M. larici-populina* to colonize poplar leaves has been extensively described (Duplessis et al. 2011b; Hacquard et al. 2010, 2012; Joly et al. 2010; Petre et al. 2012), mostly because of epidemics occurring in poplar plantations (Pinon and Frey 2005). Transcriptome analyses revealed waves of expression for candidate effector genes at different timepoints on poplar, suggesting that the fungus tightly controls

Current address for Clémence Marchal: John Innes Centre, Norwich Research Park, Norwich, U.K.

Current address for Stéphane Hacquard: Max Planck Institute for Plant Breeding Research, Cologne, Germany

†Corresponding author: Sébastien Duplessis;
E-mail: sebastien.duplessis@inra.fr

Funding: This work was supported by the Agence Nationale de la Recherche grant ANR-11-LABX-0002-01, Lab of Excellence ARBRE (Advanced Research on the Biology of FoRest Ecosystem).

*The *e*-Xtra logo stands for "electronic extra" and indicates that seven supplementary figures and four supplementary tables are published online.

© 2018 The American Phytopathological Society

its genetic program during infection (Duplessis et al. 2011b; Hacquard et al. 2012; Petre et al. 2012). These studies helped to prioritize candidate effectors for functional analysis (Germain et al. 2018; Petre et al. 2015, 2016). However, it is not known whether similar or different sets of effectors are acting during the infection of the aecial and the telial host of *M. larici-populina*. This has been a general and recurrent question related to the capacity of heteroecious rust fungi to infect different hosts belonging to unrelated plant taxa (Duplessis et al. 2014; Schulze-Lefert and Panstruga 2011). So far, a limited number of studies addressed gene expression in the different hosts of rust fungi (Cuomo et al. 2017; Liu et al. 2015; Xu et al. 2011) and none clearly determined to which extent a portion of the gene complement may be unique to one host or another.

In this study, we took advantage of the recent ability to complete the life cycle of the poplar rust fungus under laboratory conditions (Pernaci et al. 2014) to investigate the genetic program used by the fungus during larch colonization. To this end, we used an Illumina RNA sequencing (RNA-seq) approach to assess fungal transcript abundance at three different stages of the life cycle, basidia, pycnia, and aecia. We also took advantage of previously published transcriptome data during poplar infection to compare *M. larici-populina* gene expression profiles on the two host plants. We identified genes preferentially expressed on larch or poplar, including SPs that represent potential specific candidate effectors. We further scrutinized the expression features of predefined small SPs (SSPs) gene families (Hacquard et al. 2012) to investigate whether members from SSP families showed a host-specific expression.

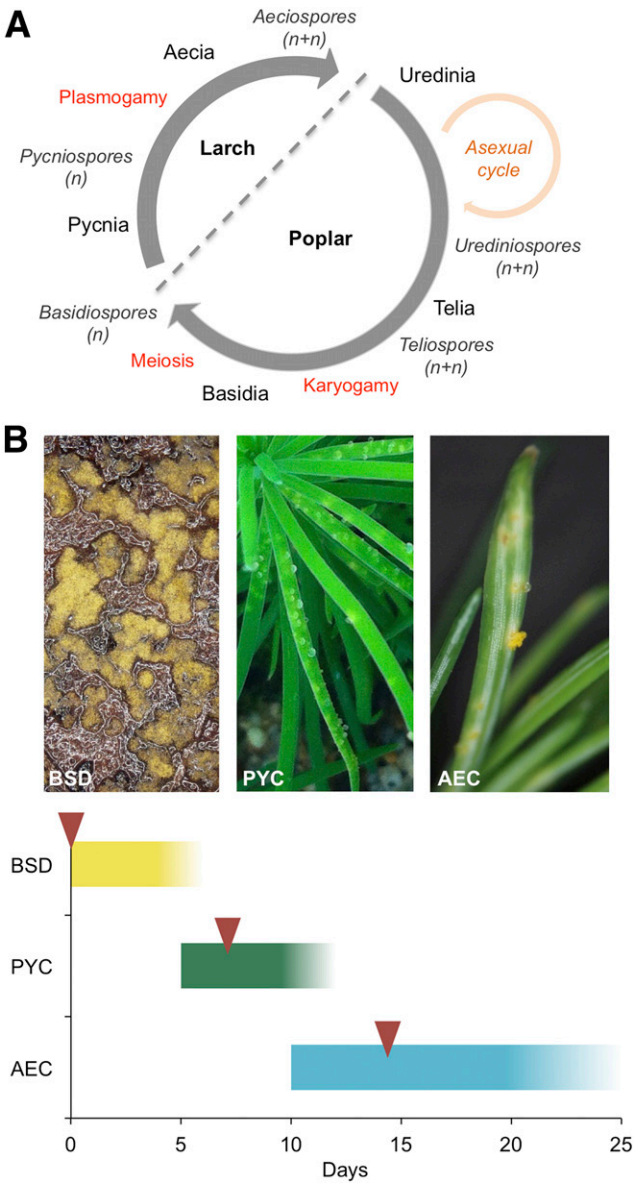


Fig. 1. The *Melampsora larici-populina* life cycle and experimental design for stage selection. **A,** *M. larici-populina* heteroecious and macrocyclic life cycle. Haploid basidiospores (n) infect the aecial host (*Larix* spp.). Pycnia are formed on larch needles, producing nectar droplets containing pycniospores (*n*) with different mating types for fertilization. After plasmogamy, dikaryotic aecia establish. Aeciospores (*n+n*) infect leaves of the telial host (*Populus* spp.) and produce dikaryotic urediniospores (*n+n*). The asexual cycle with continuous production of urediniospores takes place during the summer season. In autumn, telia differentiate and produce teliospores (*n+n*) in which, karyogamy occurs and meiosis begins before dormancy. After winter, meiosis continues and basidia germinate and produce haploid basidiospores (*n*). **B,** Basidia and basidiospores on a dead poplar leaf (left), pycnia nectar droplets on larch needles (middle), and aecia and aeciospores on larch needles (right). Color bars indicate the three larch infection–related stages sampled for RNA sequencing. BSD refers to the basidial stage and the yellow bar represents the time between basidiospore inoculation and pycnia formation. PYC corresponds to the pycnial stage (visible as nectar droplets), and the green bar represents the period of presence of visible PYC on larch needles. AEC refers to the aecial stage, and the blue bar indicates the presence of AEC on larch needles. Red arrows represent the time of collection of the three biological stages.

RESULTS

Transcriptome of *M. larici-populina* larch infection-related stages.

To gain insights into the genetic program of *M. larici populina* during larch infection, we performed deep RNA-seq at three distinct stages, the basidial stage collected on dead poplar leaves, the pycnial stage, and the aecial stage collected on larch needles at 7 and 14 days postinoculation (dpi), respectively (Fig. 1; Supplementary Fig. S1). Large areas covered with apparent fungal structures were targeted in order to maximize the capture of fungal transcripts. An average of 114 million reads per sample and more than 1 billion reads in total were collected (Supplementary Table S1) and were mapped onto 16,399 *M. larici-populina* transcripts. Overall, 16,386 transcripts had at least one read assigned. The proportion of reads mapped ranged from 10.3% in replicate 2 of the pycnial stage to 59.3% in replicate 1 of the basidial stage. This reflected the different proportion of fungal transcripts in plant tissues. Each replicate reached a plateau in rarefaction curves independently of the stage (Supplementary Fig. S2), indicating a sufficient sequencing depth to capture a full transcriptome at each stage. Principal component analysis of DESeq2 rlog-transformed normalized counts showed homogeneous biological replicates (Supplementary Fig. S3). The first axis (94% of the variance) separated the basidial stage from the aecial and the pycnial stages, whereas the second axis (3% of the variance) separated the aecial and the pycnial stage. One aecial replicate was apart from the others, probably because of the higher proportion of null counts observed for this replicate and the overall heterogeneity of the biological material at this stage compared with the basidial and the pycnial stages. Normalized counts distributed in a similar manner across all the samples. Average expression values were then considered for the basidial, pycnial, and aecial stages. Overall, the capture of complete transcriptomes by the RNA-seq approach is suitable for further comparative analysis between stages.

Gene expression profiles correlate in the pycnial and the aecial stages.

To further explore differences between the basidial stage and larch in-planta stages, we analyzed global gene expression profiles and differentially expressed genes (DEGs). Pycnial

and aecial stages occurring in larch needles showed overall closer expression profiles, whereas they are more dispersed when compared with the basidial stage (Fig. 2A to C). In total, 10,204 transcripts (62% of the 16,399 predicted *M. larici-populina* transcripts) were commonly detected in the basidial, pycnial, and aecial stages (Fig. 2D). Only 230, 696, and 48 transcripts were detected exclusively in the basidial, pycnial, and aecial stages, respectively. Pairwise comparisons were performed between the three stages, and transcripts with normalized fold change and *padj* < 0.05 in at least one paired comparison were deemed as significantly differentially expressed (up- or down-regulation). In total, 7,743 transcripts were significantly differentially expressed in at least one stage compared with another (Supplementary Table S2) and 7,031 and 7,079 transcripts were significantly differentially expressed between pycnial and basidial stages or aecial and basidial stages, respectively (Table 1; Supplementary Table S3). Fewer DEGs (2,141) were detected between the aecial and the pycnial stages. We conclude that the closely related profiles recorded for the pycnial and aecial stages support the expression of a similar genetic program in planta in larch.

SPs is the only overrepresented category among DEGs detected on larch.

To discriminate between sets of *M. larci-populina* genes expressed at the basidial stage and during larch infection, we conducted a coexpression network analysis (Langfelder and Horvath 2008). This correlation-based network analysis identifies clusters of genes with expression profiles that are more linked to each other than they are with genes outside the cluster. Among 1,548 significantly DEGs (*padj* < 0.05), we identified six clusters of transcripts sharing coexpression profiles

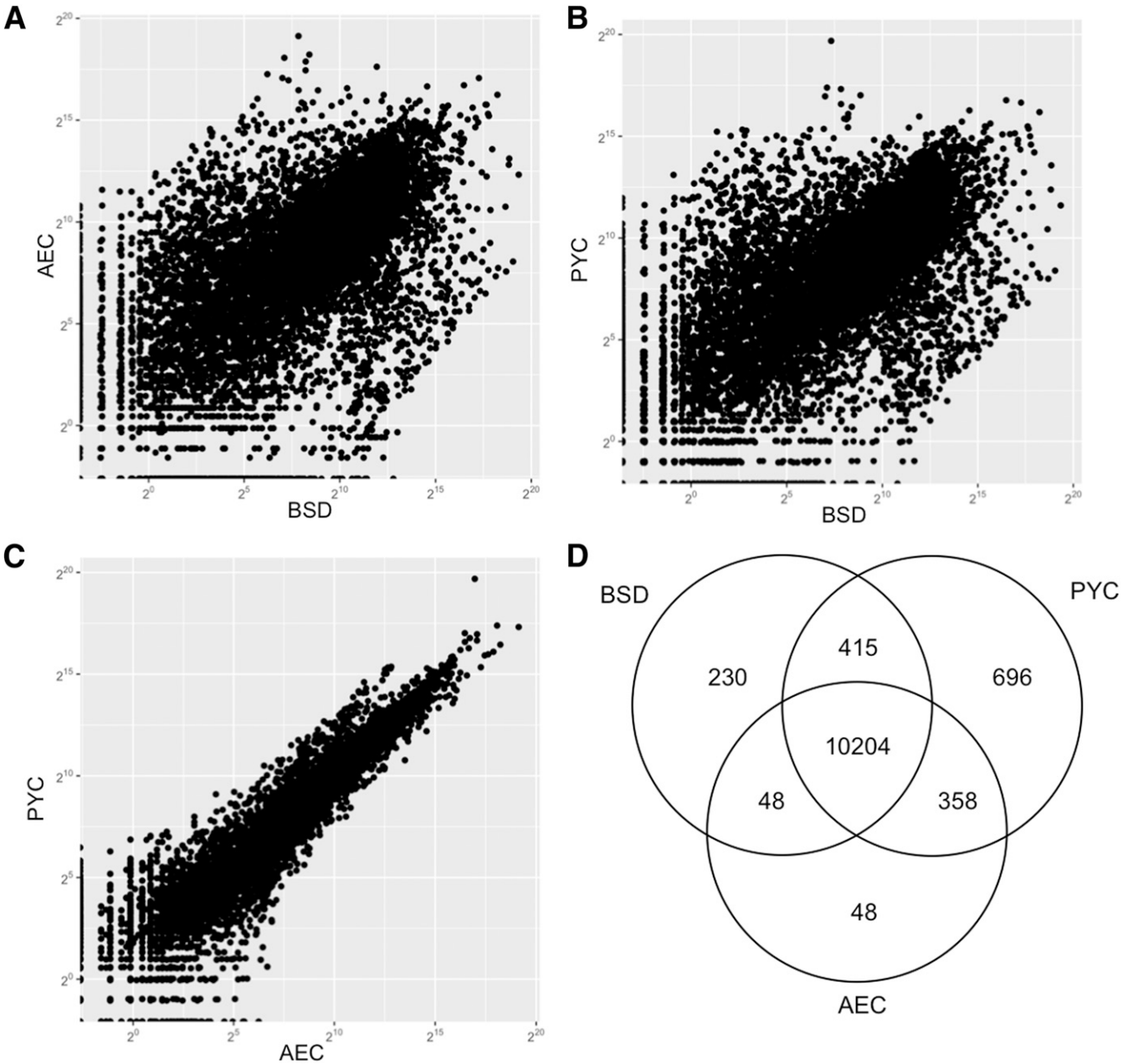


Fig. 2. RNA-seq analysis of the basidial (BSD), pycnial (PYC), and aecial (AEC) stages of the rust fungus *Melampsora larici-populina*. **A,** Scatter plot of rlog-transformed sequencing counts of transcripts expressed in AEC versus BSD, **B,** in PYC versus BSD, and **C,** in PYC versus AEC. **D,** Venn diagram showing the number of genes detected by at least one read per sample in each of the basidial, pycnial, and aecial stages.

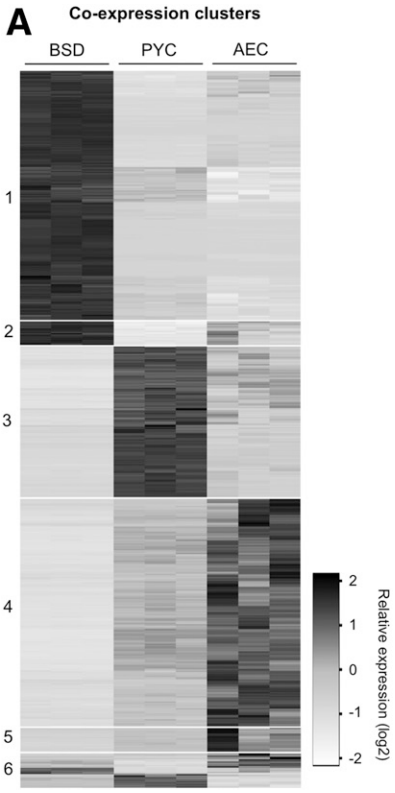
(Fig. 3A; Supplementary Fig. S4). The clusters positively correlated with the basidial stage were negatively correlated with the in-planta pycnial and aecial stages (Fig. 3A). Clusters 1 and 2 were significantly positively correlated with the basidial stage and represented 38.8% of the DEGs (661). Cluster 3 was significantly positively correlated with the pycnial stage, representing 21.3% of the selected DEGs (330). Clusters 4 to 6 were significantly positively coexpressed at the aecial stage with 39.8% of DEGs (617) (Fig. 3A). In total, 61.2% of the DEGs were significantly coexpressed in planta compared with the basidial stage. Clusters 3, 4, and 5 were significantly negatively correlated with the basidial stage while positively correlated with in-planta stages (Fig. 3A).

To further understand the transcriptome reprogramming occurring between basidial and larch in-planta stages, we conducted an enrichment analysis of euKaryotic Orthologous Group (KOGs) in pairwise comparisons of regulated genes. Transcripts regulated between two stages with at least a threefold change and a *padj* < 0.05 were considered (Table 1). We also created a specific KOG-SP category for SPs of unknown function. Among upregulated genes in aecial versus pycnial (1,045 DEGs), aecial versus basidial (3,503 DEGs), and pycnial versus basidial (3,698 DEGs) pairwise comparisons, the only significantly enriched category is KOG-SP (Fig. 3B). SPs represented 28.9, 22.4, and 22.1% of aecial versus pycnial, aecial versus basidial, and pycnial versus basidial upregulated genes, respectively. SP-encoding genes were overrepresented in upregulated genes of aecial and pycnial stages compared with the basidial stage. We also performed reverse transcription-quantitative polymerase chain reaction (RT-qPCR) on 22 randomly selected SP genes to confirm their RNA-seq expression profiles (Supplementary Fig. S5). A good correlation could be observed between RNA-seq and RT-qPCR for pycnial and aecial stages, whereas differences are noticeable for the basidial stage, which may relate to the intrinsic nature of the biological material (dead poplar leaves). Our results show that more SP genes are expressed during larch in-planta stages compared with the basidial stage.

Table 1. Number of *Melampsora larici-populina* genes significantly differentially expressed at the basidial (BSD), pycnial (PYC), and aecial (AEC) stages

Regulated genes[a]	PYC vs. BSD	AEC vs. BSD	AEC vs. PYC
Genes significantly differentially expressed ($P < 0.05$)	7,031	7,079	2,141
Upregulated	3,698	3,503	1,045
>3-fold	2,406	2,280	90
>10-fold	1,526	1,316	5
>100-fold	522	442	0
>1000-fold	105	67	0
Downregulated	3,333	3,576	1,096
<3-fold	1,872	2,128	539
<10-fold	849	1,053	265
<100-fold	319	429	27
<1000-fold	57	126	0

[a] Down- and upregulated genes were classified according to their fold change ratio. Transcripts with a significant *P* value (<0.05) and more than a threefold change in transcript level were considered as significantly differentially expressed.



B Over-represented KOG

KOG Category	FDR DR genes PYC vs BSD	FDR UR genes PYC vs BSD	FDR DR genes AEC vs. BSD	FDR UR genes AEC vs. BSD	FDR DR genes AEC vs. PYC	FDR UR genes AEC vs. PYC
Translation, ribosomal structure and biogenesis	0.9	0.9	0.93	1	1	1
RNA processing and modification	0.81	1	0.73	1	1	1
Transcription	0.8	1	0.51	1	1	1
Replication, recombination and repair	0.9	1	0.4	1	1	1
Chromatin structure and dynamics	0.62	1	0.28	1	1	1
Cell cycle control, cell division, chromosome partitioning	0.81	1	0.55	1	1	1
Nuclear structure	0.9	1	0.92	1	1	1
Defense mechanisms	0.9	1	0.85	1	1	1
Signal transduction mechanisms	0.8	1	0.28	1	1	1
Cell wall/membrane/envelope biogenesis	0.4	1	0.36	1	1	1
Cell motility	0.81	1	0.55	1	1	1
Cytoskeleton	0.4	1	0.36	1	1	1
Secreted protein	0.81	3.2e-47*	0.36	1.85e-47*	0.84	7.8e-05*
Extracellular structures	0.81	1	0.7	1	1	1
Intracellular trafficking, secretion, and vesicular transport	0.8	1	0.85	1	1	1
Posttranslational modification, protein turnover, chaperones	0.9	1	0.9	1	1	1
Energy production and conversion	0.8	1	0.73	1	1	1
Carbohydrate transport and metabolism	0.8	1	0.53	1	1	1
Amino acid transport and metabolism	0.9	1	0.93	1	1	1
Nucleotide transport and metabolism	0.81	1	0.7	1	1	1
Coenzyme transport and metabolism	0.5	1	0.36	1	1	1
Lipid transport and metabolism	0.81	1	0.75	1	1	1
Inorganic ion transport and metabolism	0.62	1	0.36	1	1	1
Secondary metabolites biosynthesis, transport and catabolism	0.4	1	0.36	1	1	1
General function prediction only	0.71	1	0.7	1	1	1
Function unknown	0.81	1	0.73	1	1	1
No hit	0.9	0.98	0.93	1	1	1

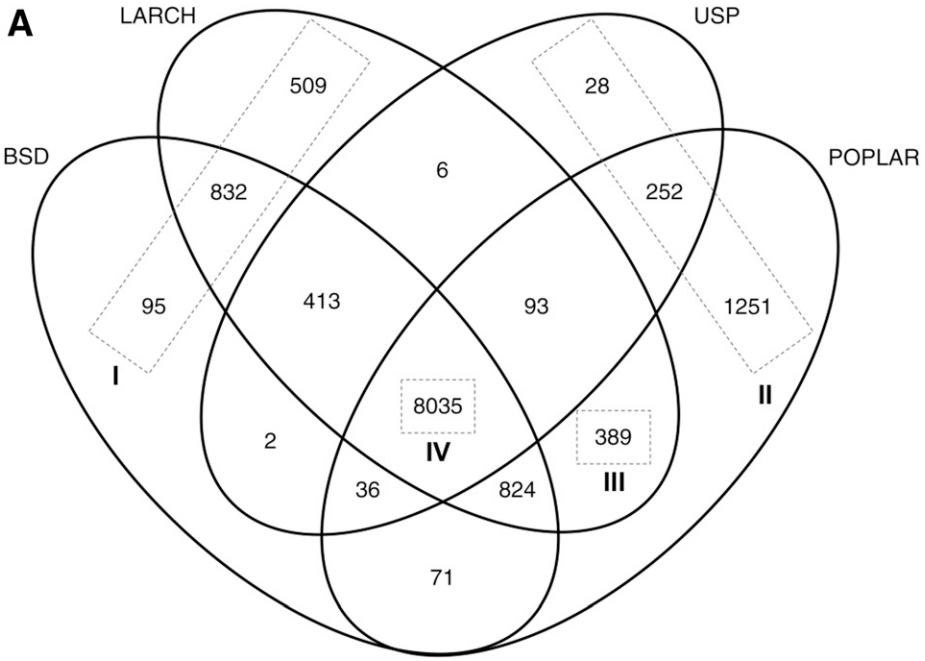
Fig. 3. Coexpression analysis and euKaryotic Orthologous Group (KOG) enrichment analysis of differentially expressed *Melampsora larici-populina* genes in the basidial and the larch-infection stages. **A,** Transcript profiling of 1,548 *M. larici-populina* significantly differentially expressed genes at the basidial (BSD), pycnial (PYC), and aecial (AEC) stages (DESeq2; *padj* < 0.05). Relative expression profiles are represented as $\log_2$ of normalized counts across three replicates of BSD, PYC, and AEC, from overexpression (black) to underexpression (white). Using coexpression network analysis (WGCNA package), genes were grouped into six clusters of coexpression. **B,** False discovery rate (FDR) *P* values matrix from overrepresented KOG categories of threefold change downregulated (DR) genes and threefold change upregulated (UR) genes. Significantly enriched categories (FDR < 0.05) are highlighted with an asterisk (*).

Comparison of *M. larici-populina* transcriptomes on larch and poplar highlights subsets of commonly and specifically expressed genes.

To identify *M. larici-populina* genes specific to larch or poplar infection, we compared transcriptome profiling of larch stages to previous expression studies performed on the telial host (Duplessis et al. 2011b). We simplified our dataset to four explicit stages: the basidial stage as a larch-infecting spore stage, LARCH as the in-planta stage on the aecial host, Urediniospores (USP) as a poplar-infecting spore stage, and POPLAR as the in-planta stage on the telial host. Expression data recorded for poplar-related stages were obtained with custom oligoarrays. Thus, direct calculation of fold change was not possible between larch- and poplar-related stages due to the different technologies used. However, the transcriptomes were complete on both hosts (Duplessis et al. 2011b), allowing direct comparison to determine the presence of specifically expressed genes (at least a count of 1 by RNA-seq or above background value on oligoarrays). Overall, 14,883 transcripts were detected in poplar- or larch-related stages, and 8,035 transcripts are expressed in all four fungal stages (Fig. 4A). A total of 1,436 transcripts were only detected at the basidial stage and LARCH stage, and 1,531 transcripts were detected in USP and POPLAR stages. Interestingly, 389 genes were specifically detected in LARCH and POPLAR stages, revealing a set of in planta–specific genes. In total, 2,967 (20%) genes are specifically expressed in larch or poplar in in-planta stages.

Detailed analysis of gene categories expressed on larch and poplar reveals SP genes specifically expressed on each host.

To determine whether *M. larici-populina* uses specific pathways or functions during colonization of larch and poplar, we surveyed expression profiles of gene categories. First, no



B

	I	II	III	IV	Expressed
CAZymes	35	11	8	310	454
Lipases	10	3	-	61	88
Proteases	20	8	8	304	380
Secreted Proteins	245	103	98	688	1645
Transporters	8	7	1	282	327
Other	1118	1399	274	6390	11989
Total	1436	1531	389	8035	14883

Fig. 4. Comparative transcriptomic analysis between four stages of the *Melampsora larici-populina* life cycle. **A,** A Venn diagram showing the number of genes expressed in BSD (basidial stage), LARCH (pycnial and aecial stages, i.e., during larch needle infection), USP (dormant and germinating urediniospores), and POPLAR (during poplar leaf infection). Group I refers to genes specifically expressed in larch-related stages, group II refers to genes specifically expressed in poplar-related stages, group III refers to the genes specifically expressed in planta, and group IV refers to genes detected in all four stages. **B,** Number of genes expressed in selected annotated gene categories related to host infection, i.e., carbohydrate-active enzymes (CAZymes), lipases, proteases, secreted proteins, and transporters in selected groups I, II, III and IV. The total numbers of expressed genes for each selected annotated category is shown on the right. Total numbers of genes belonging to selected groups I, II, III and IV are shown below the table.

major difference could be determined on one host or another regarding the expression of KOG cellular categories (Supplementary Fig. S6). Manual exploration of Kyoto Eukaryotic Genes and Genomes (KEGG) annotations revealed a similar result for metabolic pathways. Second, of 6,660 genes of unknown function (excluding the SP genes addressed below) with expression information, 58% were expressed at all stages, whereas 13% (863) and 14.5% (967) presented a specific expression pattern on larch and poplar, respectively (Supplementary Fig. S7). We scrutinized cellular categories previously annotated in the *M. larici-populina* genome (Duplessis et al. 2011a). Among those, 282 transporters, 310 carbohydrate-active enzymes (CAZymes), and 304 proteases were expressed at all four stages. Few genes in these surveyed categories were specifically detected in larch- and poplar-related stages. Likewise, only eight transporters were detected only at the basidial stage and the LARCH stage, including three auxin efflux carriers and one oligopeptide transporter, which belong to expanded rust fungal gene families (Duplessis et al. 2011a) (Fig. 4B). A total of 35 CAZyme genes were detected at the basidial and LARCH stages, including 22 glycoside hydrolases, and 20 proteases were detected at these stages, including eight peptidase inhibitors. Even fewer genes were specifically detected in USP and POPLAR stages, with seven, 11, and eight transporters, CAZymes, and proteases, respectively. Overall, less than 10% of CAZymes, proteases, lipases, or transporters are specifically expressed in polar or larch stages. Among mating-related genes, pheromone receptor genes were expressed in both hosts and one gene showed a very high expression in LARCH. Four pheromone precursor genes showed a very high expression in LARCH compared with other stages and were in the top 1% most highly expressed genes at the aecial stage, whereas mating compatibility is established between pycniospores and receptive hyphae. In total, 60% of the specifically expressed genes on poplar and on larch are of unknown function.

Since SP genes were significantly overrepresented in *M. larici-populina* genes expressed on larch (this study) and on poplar (Duplessis et al. 2011b), we particularly considered their expression profiles during larch and poplar infection. Among 1,645 expressed SP genes, 688 (42%) were detected at all developmental stages (Fig. 4B). Transcripts specifically detected either on larch or poplar contained an important proportion of SP-encoding genes, accounting for 17 and 25% of the total number of genes specifically expressed at these stages, respectively. A lower proportion of SP genes were specifically detected on the telial host poplar compared with the aecial host larch. Even if the proportion of SP genes expressed at all stages was important (688; 42% of 1,645 SP genes), it represented only 8.6% of the 8,035 transcripts expressed at all four stages. Moreover, of the top 10% genes preferentially highly expressed on larch or on poplar in planta, 18% are SPs and, among them, 14% are SSPs (Fig. 5). In comparison, SPs and SSPs represent only 4 and 5%, respectively, of other expressed genes. Altogether, these results indicate an enrichment of SSPs among genes highly expressed during poplar or larch infection.

Gene members of *M. larici-populina* SSP families present specific patterns of expression on larch and on poplar.

We took the opportunity of the detailed analysis of SSP gene families in *M. larici-populina* (Duplessis et al. 2011a; Hacquard et al. 2012) to further explore their specific expression patterns on the two host plants. RNA-seq and oligoarray expression data were quantile-normalized and were compared to determine whether entire families or members of families are host-specific (Fig. 6). Overall, 205 and 351 SSP family genes were preferentially expressed in planta on larch and on poplar, respectively, when the general expression profile of all SSP genes was considered (Fig. 6A). When focusing on singletons or families of SSPs grouping two, three, or more members, three-fourths of the genes were detected on the two hosts (Fig. 6B). The remaining fourth corresponded to genes expressed specifically in one or the other host (basidial and LARCH or USP and POPLAR). We further considered large SSP families to identify family-specific expression patterns. Among 64 families of at least four members, 17 (26%) showed a preferential expression on one given host (nine and eight SSP families preferentially expressed in larch or poplar in-planta stages, respectively). For instance, the SSP family 1 (111 members) presented 24 members expressed specifically at the basidial and LARCH stages and four only at USP and POPLAR stages, the remaining members showed a

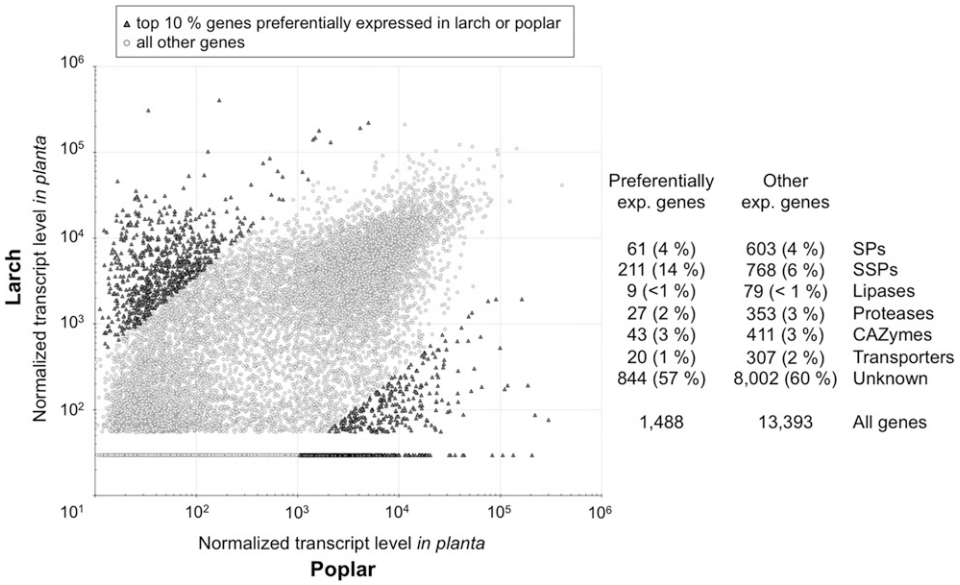


Fig. 5. *Melampsora larici-populina* genes preferentially expressed in planta on larch and on poplar. Scatter plot representing the normalized gene expression level of 14,883 *M. larici-populina* genes expressed in planta on poplar (*x* axis) and on larch (*y* axis). The dark gray triangles represent the top 10% genes preferentially expressed on poplar (5%) and on larch (5%). Numbers and percentage of genes in secreted proteins (SPs), small SPs (SSPs), lipases, proteases, carbohydrate-active enzymes (CAZymes), transporters, and genes of unknown function of the top 10% genes and of other expressed genes are detailed. The original expression data for rust-infected poplar leaves and rust-infected larch needles were quantile normalized to allow comparison.

higher expression at the basidial and LARCH stages (Fig. 6C). On the contrary, SSP family 9 (11 members) showed a preferential expression on poplar, and only two members presented noticeable expression level on the other host. SSP family 7 presented a striking profile, with almost half the gene members expressed in the aecial host larch whereas the other half was expressed in the telial host poplar. Some of these members presented a sequence identity of 75% and showed opposite expression profiles on the two host plants (e.g., SSP genes 86274 and 91075 expressed in larch and in poplar, respectively [data not shown]). In smaller SSP gene families, the different situations exemplified in Figure 6C were also observed. Expression profiles of *M. larici-populina* SSP gene families proved to be diverse, with 60% of the genes expressed on the two host plants, 31% showing preferential expression on one host or the other, and 9% divided families (six SSP families of more than three members).

DISCUSSION

In the present study, we report on gene expression profiling at three biological stages that are related to infection of larch, the aecial host of the poplar rust fungus *M. larici-populina*, taking

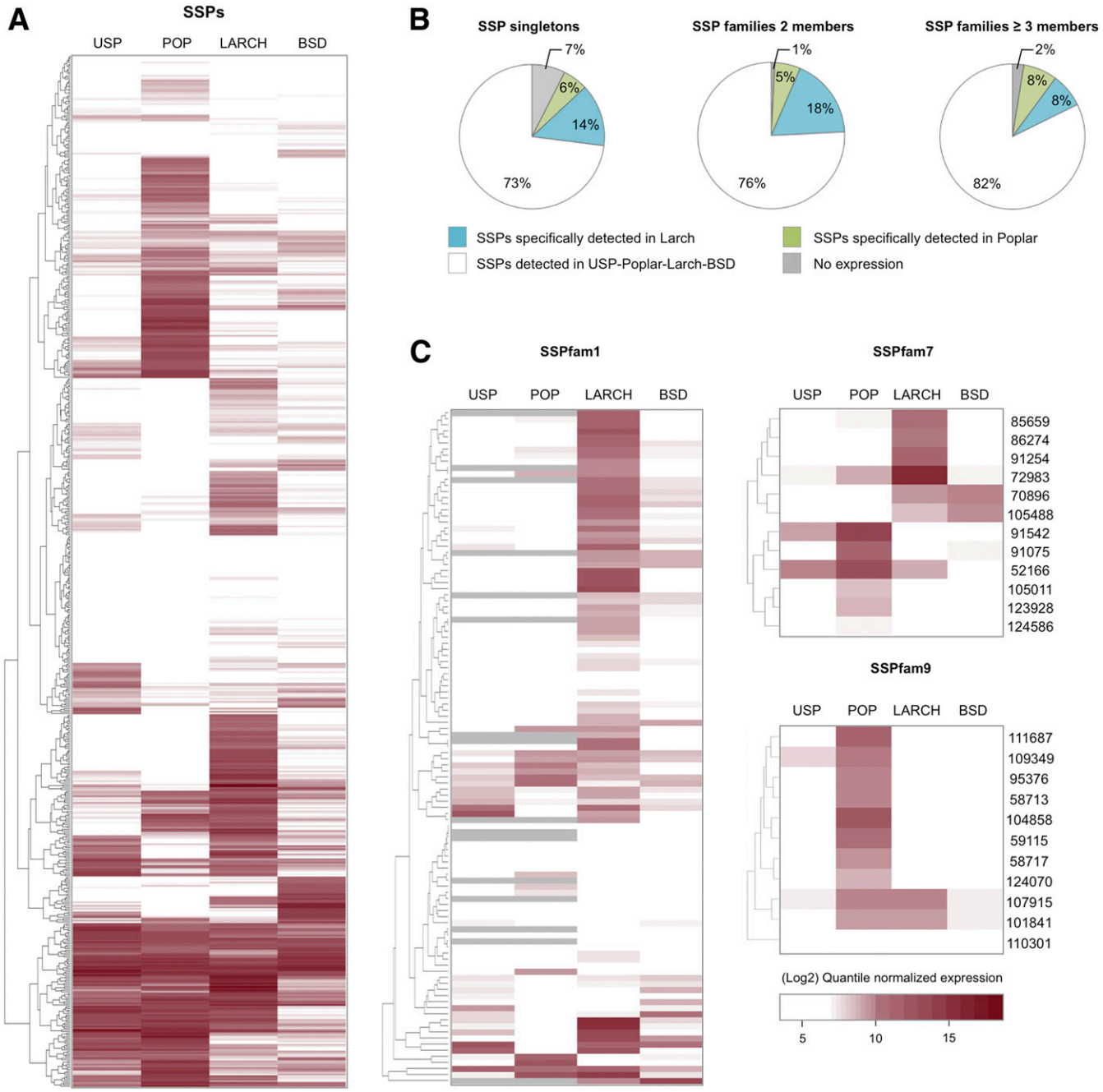


Fig. 6. Transcript profiling of *Melampsora larici-populina* small secreted protein (SSP) genes on the aecial and the telial hosts. **A,** Heatmap of expression profiles of *M. larici-populina* SSP transcripts. Expression levels (white to red) are shown as quantile-normalized expression values (white corresponding to the lowest expression and dark red corresponding to the greatest expression). **B,** Diagrams presenting the expression patterns for SSP singletons, SSP gene families with two members, and SSP gene families with three or more members detected in larch- and poplar-related stages (i.e., basidial stage, BSD; pycnial or aecial stages, LARCH; Urediniospores, USP; and poplar infection, POP). **C,** Examples of expression profiles for selected SSP families: SSP family 1, preferentially expressed in larch; SSP family 7, expressed both in poplar and larch; SSP family 9, preferentially expressed in poplar. Joint Genome Institute protein ID numbers are indicated on the right of profiles for SSPfam7 and SSPfam9. Gray on the SSPfam1 expression profile indicates nonavailable information (no specific oligomers represented on the oligoarray).

advantage of an experimental setup in semi-controlled laboratory conditions (Pernaci et al. 2014). A full transcriptome was recovered at each of the stages through deep RNA-seq, allowing comparisons between stages. In total, 73% of the *M. larici-populina* genes were expressed at the basidial, pycnial, and aecial stages. This is similar to previous findings (76%) in which stages related to infection of poplar, the telial host of the fungus, were considered (Duplessis et al. 2011b; Hacquard et al. 2010, 2012). Altogether, 62% of the genes were commonly expressed in the basidial, pycnial and aecial stages; however, the general profile was closer between the two larch-infection stages than with the basidial stage. This suggests an important reprogramming of the fungal transcriptome after the transition to host infection. Although the pycnial and aecial stages occur before and after fertilization, major in-planta structures, such as protoaecia, are formed before the fertilization between pycniospores and receptive hyphae of compatible mating types (Harder 1984; Voegele et al. 2009), which may explain the similarity of the general expression profile at these stages. It is worth noticing that mating-related genes were among the most highly expressed genes detected at the pycnial stage. The analysis of the wheat leaf rust genome recently resolved the mating system and loci for a rust fungus (Cuomo et al. 2017). Exploring expression of mating genes on the host on which sexual reproduction occurs would allow a better understanding of the underlying genetic processes.

Heteroecious rust fungi exhibit the most complex life cycle known in fungi. They involve infection of two taxonomically unrelated host plants and production of five different types of spores (Aime et al. 2017). Transcriptomics of rust has been widely applied to asexual stages and spores responsible for epidemics and damages in plantations (Duplessis et al. 2014). There is still scarce information about expression along the whole life cycle of rust fungi, and only a few studies looked at transcript profiles on alternate host plants. In *Puccinia triticina*, expressed sequence tag sequencing and RNA-seq identified sets of genes expressed on the two host plants, meadow rue (*Thalictrum* spp.) and wheat, however, at a limited scale (Cuomo et al. 2017; Xu et al. 2011). RNA-seq was applied at a higher depth to *Cronartium ribicola* spores collected from the two host plants *Pinus monticola* (aeciospores) and *Ribes nigrum* (urediniospores) and compared with two infection stages in western white pine (Liu et al. 2015). Although no reference genome is available for *C. ribicola*, the RNA-seq approach could identify a large proportion of *C. ribicola* genes expressed in all situations (61%) with only a few (4 to 5%) specific to each separate stage (Liu et al. 2015).

In our study, the significant overrepresentation of SPs among regulated genes between stages was striking and mirrors previous findings on the poplar host (Duplessis et al. 2011b; Hacquard et al. 2010, 2012). This shows that the expression of SP genes is important for larch infection and central to biotrophy on both hosts. Many effectors of fungal pathogens are SPs, including avirulence effectors identified in the flax rust fungus *Melampsora lini* (Duplessis et al. 2012; Sharpee and Dean 2016; Tyler and Rouxel 2012). Rust fungi exhibit large secretomes composed of expanded gene families and SP and SSP genes specifically expressed in planta represent priority candidate effectors (Aime et al. 2017; Petre et al. 2014). Previous expression profiling reports of poplar infection by *M. larici-populina* have identified long lists of putative effectors that are key candidates for functional analysis (Duplessis et al. 2011b; Hacquard et al. 2010, 2012, 2013; Joly et al. 2010; Petre et al. 2012). No expression data were available before for infection of larch needles by *M. larici-populina*. Whether specific sets of effectors may be required for rust fungi to infect their different hosts was so far unaddressed (Schulze-Lefert and Panstruga 2011). Only a fraction of SSP genes was specifically expressed on poplar or on larch, in a similar fashion for singletons or multigene families, and a larger portion is expressed on the two hosts. Our results confirm and largely extend the observations made in *Puccinia triticina* or *C. ribicola* regarding specific gene expression (Cuomo et al. 2017; Liu et al. 2015; Xu et al. 2011). Considering that rust effectors are likely to reside among these genes, our results could indicate that a considerable portion of *M. larici-populina* effectors may share similar functions on the two hosts.

Some *M. larici-populina* SSP families showed a preferential expression on a single host and more striking was the profile of families exhibiting members whose expression was host-specific. Families showing host-specific expression could reflect a diversification under the pressure of resistance genes or the evolution of a new effector function targeting a particular process not represented in the alternate host. Regarding gene families combining specificity of expression toward the two hosts, such as the *M. larici-populina* SSP7 family, we may observe adaptation of subsets of effectors with i) different functions in the cells of the two hosts, or with ii) a conserved function but different targets in the two plants, or iii) having the same target that evolved differently in the two plant lineages from a unique ancestral target. These results open many avenues for speculation about the evolution of SSP and candidate effector families in rust fungi, but it remains to be determined whether these particular families correspond to *bona fide* effectors. With the knowledge of preferential or specific expression on the two hosts, it is now possible to narrow the list of candidate effectors that are only expressed in the main host, where most of the damages are caused, and prioritize future functional investigations. In addition, the subset of candidate effectors expressed on the two host plants (i.e., core effectors) may be important in the rust general infection process and can represent useful targets for resistance breeding. These findings also open perspectives for the development of resistance in the alternate host in which the rust fungus accomplishes its sexual reproduction.

Beyond candidate effectors, other SP genes, such as the CAZymes, showed specific expression in poplar or in larch. Similar observations were made for *C. ribicola* infecting *Ribes* spp. or pine (Liu et al. 2015). It is known that plant cell wall varies in terms of composition, architecture, and integrity across plant species (Kubicek et al. 2014). CAZymes are essential enzymes for decomposing polysaccharides from plant cell walls and leaf or needle cuticles (Henrissat et al. 2015). Rust fungal hyphae need to penetrate through the plant cuticle in the aecial host and to go through the plant cell wall to form haustoria within the host cell cavity in both hosts (Ragazzi et al. 2005). Different small sets of CAZymes were expressed in poplar and larch, which could reflect the adaptation to the host polysaccharide composition in a conifer and a dicot. On the contrary, the survey of central metabolic pathways or cellular categories did not reveal major differences. *M. larici-populina* seems to express a conserved genetic program during the biotrophic growth in larch and poplar tissues, suggesting a common strategy for the acquisition and utilization of nutrients derived from the two hosts. A surprising feature of rust fungal genomes is the proportion of genes of unknown function with no homology in the Pucciniales (Aime et al. 2017; Duplessis et al. 2011a). Among 6,660 *M. larici-populina* genes of unknown function with expression records, more than 10% are specifically expressed on poplar or on larch, whereas 58% are expressed on both. A better dissection of the infection process on the two hosts could help to draw hypotheses about processes not yet unraveled and in which these genes could be involved, e.g., during early colonization, biotrophic growth, or sporulation. Moreover, systematic comparative transcriptomics in

different heteroecious rust fungi as well as with autoecious rust fungi may provide new leads to unravel new functions related to biotrophy.

When the transcriptome is considered across the whole *M. larici-populina* life cycle, we could detect expression for 86% of the genes and a total of 2,047 genes were never detected. Although this relatively small number of genes may contain annotation errors, it is likely that some may be specific for sections of the rust biological life cycle not explored in this or in previous studies. Telia are overwintering structures adapted to drastic environmental conditions in which important biological processes such as karyogamy and meiosis occur (Mendgen 1984). A small proportion of genes specifically expressed in telia were previously reported in *M. larici-populina* (Hacquard et al. 2013), which could be explained by the fact that transcriptome profiling was carried out in early differentiating telia. This telia stage was not included for comparison in the present study because it was obtained from environmental samples that were different from the reference strain 98AG31 (Hacquard et al. 2013). When compared with the present study, only 117 genes could be detected as telia-specific (data not shown). More specific profiles can be expected at later stages of telia differentiation, when karyogamy and meiosis occur, or during entry and exit of winter dormancy. Moreover, the infection of the aecial host is far more complex than in the telial host, with the establishment of distinct fungal cell types in a coordinated manner (Harder 1984; Littlefield and Heath 1979). Only three specific stages were selected for this work and specific expression of other genes may be expected during the course of the interaction with the aecial host, as previously illustrated in poplar (Duplessis et al. 2011b; Hacquard et al. 2010, 2012).

Other fungal biotrophs possess the capacity to interact with multiple host species with various levels of specificity. Mycorrhizal fungi are biotrophs that engage in mutualistic associations with their host plants by forming specific infection structures (van der Heijden et al. 2015). During the colonization process, they deliver symbiosis effectors to establish long-lasting beneficial interactions (Martin et al. 2016). Large-scale transcriptome studies have shown that ectomycorrhizal fungi also present specific expression of large sets of SSP genes of unknown function (Kohler et al. 2015). Similar to our observations with *M. larici-populina*, mycorrhizal fungi express common and specific sets of SSP genes during their interaction with different host plants. This has been shown for ectomycorrhizal fungi in different interactions between *Suillus* spp. and *Pinus* spp. (Liao et al. 2016) and *Laccaria bicolor* with poplar and Douglas fir (Plett et al. 2015). The same was observed for the arbuscular mycorrhizal fungi *Rhizophagus irregularis* and *Gigaspora rosea* colonizing different host plants (Kamel et al. 2017). The fungal root endophyte *Piriformospora indica* establishes different colonization strategies in a host-dependant manner, i.e., biotrophy in *Arabidopsis* and saprotrophy in barley (Lahrmann et al. 2013). Transcriptome analysis of root colonization in these plants shows the expression of common and specific sets of SSP genes. Although the identified SSPs are different in these different fungal species, there is a common trend for their expression during infection, particularly during the establishment of biotrophic interactions, suggesting that they may participate in defining the level of specificity with different host plants.

MATERIALS AND METHODS

Experimental setup and inoculation procedures.

Specific *M. larici-populina* isolate 98AG31 stages were collected on the aecial host larch (*Larix decidua*), following the procedure reported by Pernaci et al. (2014). Figure 1 illustrates the experimental set up. The basidial stage was collected from dead poplar leaves and contained remains of telia and freshly produced basidia and basidiospores. Large areas of dead leaves fully covered with yellow fungal material were targeted. The pycnial stage was collected from larch needles 7 dpi with basidiospores and are made up of different cell types, such as in-planta infection hyphae, flexuous hyphae, and pycniospores. The aecial stage was collected from larch needles 2 weeks after inoculation and corresponded to freshly produced aecia and remaining pycnia. For both the pycnial and the aecial stages, sections of leaves fully covered with apparent fungal material were targeted and portions of leaves with no apparent fungus were removed. Three biological replicates were harvested at each stage; the basidial replicates came from different poplar leaves, while the pycnial and aecial replicates were produced from different boxes of inoculated larch seedlings. For each replicate (i.e., R1, R2, R3), the aecial and pycnial stages were harvested from different larch seedlings present in the same inoculated box, ensuring synchronization of these stages after a single initial basidiospore inoculation.

RNA isolation and Illumina RNA-seq.

Total RNA from the basidial stage samples were isolated with the RNeasy plant mini kit (Qiagen, Courtaboeuf, France), using RLC buffer with β-mercaptoethanol, following manufacturer's recommendations. For each basidial replicate, several rounds of RNA isolation were performed, using low inputs of material (10 to 20 mg) to avoid RNA degradation from decayed leaf material, and RNA samples representing a total of 100 mg of material were pooled together. Total RNA from the pycnial and the aecial stages were isolated from 100 mg of infected larch needles, using a protocol adapted to RNA isolation from conifers (Chang et al. 1993). A DNase I (Qiagen) treatment was applied to all samples according to the manufacturer's recommendations, to eliminate traces of genomic DNA. Electrophoretic RNA profiles and RNA quality were assessed with the Experion analyzer, using the Experion Standard-sens analysis kit (Bio-Rad, Marnes la Coquette, France) before sample shipment and in Beckman's facilities (Beckman Coulter Genomics, Grenoble, France), prior to RNA-seq. The following RIN numbers were recorded before Illumina libraries preparation: 8.6, 8.3, 8.8, 6.9, 7, 6.8, 7.6, 8.6, and 8.3, corresponding to replicates R1, R2, R3 of basidial, pycnial, and aecial samples, respectively. cDNA synthesis and RNA-seq library construction were performed by Beckman Coulter Genomics, using the Illumina TruSEQ RNA seq kit following the manufacturer's recommendations. Each library was quantified by qPCR and was sequenced, with the Illumina HiSeq2500 platform, as paired-end 100-nt reads. Data acquisition was performed in Beckman Coulter Genomics facilities according to standard Illumina's procedures. On average, 114 million paired-end reads were produced per sample (ranging from 95 to 129 million reads per sample). The complete expression dataset is available at the National Center for Biotechnology Information Gene Expression Omnibus repository under the series number GSE106863.

Reference genome and annotation.

The *M. larici-populina* isolate 98AG31 genome (version 1.0, United States Department of Energy Joint Genome Institute) (Duplessis et al. 2011a) was used as a reference. The RNA-seq analysis was performed using the catalog of 16,399 *M. larici-populina* isolate 98AG31 reference transcripts. Automatic annotations such as Gene Ontology, KEGG, and KOG annotations were recovered from the reference genome website. The expert manual annotation of specific gene categories (e.g., CAZymes, lipases, proteases, transporters) performed by the Poplar Rust Genome Consortium was used (Duplessis et al. 2011a). The

repertoire of SSPs previously described (Hacquard et al. 2012) was extended to include all SPs of unknown function independently of their protein length, as described by Lorrain et al. (2015).

Gene expression analysis.

All reads per sample were mapped onto the 16,399 *M. larici-populina* transcripts using the RNA-seq analysis tool of CLC Genomics Workbench 6.5.1 (Qiagen) with default parameters, except for the similarity fraction, which was set at 90% across 90% of the read length. Reads showed an average per base Phred quality score higher than 30 over their 100-nt length in all samples, and no trimming was applied before mapping. The number of unique reads mapped per transcript (counts) was recorded for each. A saturation curve analysis was performed with the rarefy function of the VEGAN package (Oksanen et al. 2017) to determine whether the sequencing depth of our data were sufficient. R packages DESeq2 (v1.10.1) and edgeR (3.12.0) were used to conduct differential expression analysis of *M. larici-populina* transcripts (Love et al. 2014; Robinson et al. 2010). Counts were normalized using the size factor method proposed by Anders and Huber (2010). DEGs were then identified using both packages (*padj* < 0.05), and the intercept was considered for further analyses with DESeq2 normalized count values as the reference dataset. The R package WGCNA was used to generate clusters of coexpression among significantly DEGs (Langfelder and Horvath 2008). For this purpose, only DEGs found between all three stages were considered.

KOG enrichment analysis.

KOG enrichment analysis was performed as previously described by Hacquard et al. (2013). SPs of unknown function were extracted from the KOG categories "No Hit" and "Function unknown" to define a new KOG category that we termed "Secreted proteins". Overrepresented KOG categories in the basidial, pycnial, and aecial stage up- or downregulated genes (threefold change was considered significantly differentially expressed) were determined in respect to global gene distribution. Significance of overrepresented KOG categories in each condition was estimated using the Fisher's Exact test and false discovery rate (FDR) corrected by Benjamini-Hochberg test (FDR < 0.05).

Comparison between *M. larici-populina* larch and poplar infection transcriptomes.

Expression data for *M. larici-populina* dormant and germinating urediniospores and for time-course infection of poplar leaves were recovered from a previous transcriptome analysis performed with whole-genome custom oligoarrays (Duplessis et al. 2011b). Expression data were recovered for 14,883 *M. larici-populina* transcripts showing expression values greater than the background level. The expression sets were reduced for comparison purposes as follows: expression values in dormant and germinating urediniospores were grouped in a unique "Urediniospores" stage (highest expression value of the two stages retained), and expression values during the poplar leaf time-course infection at 24, 48, 96, and 168 h postinoculation were pooled into a unique "Poplar" stage (highest expression value during the time-course retained). Larch infection expression datasets were similarly reduced into two sets: expression value in the basidial stage was considered as a unique larch-infecting spores set and expression values during larch needles infection (i.e., pycnial and aecial stages) were grouped into a unique "LARCH" stage (highest expression value retained). We used a quantile normalization to estimate and compare transcript distributions within the oligoarrays and RNA-seq datasets with the R package Limma (Ritchie et al. 2015). Regarding the comparison of specific gene families, some gene members or gene families had nonavailable expression records from oligoarrays. This was due to the oligonucleotide inability to discriminate between transcript species. These genes were, thus, not included in the comparative expression analysis.

RTqPCR.

A random set of 22 SSP genes was selected to assess RNA-seq expression profiles by RT-qPCR. Primer sequences and efficiency are presented in Supplementary Table S4. Primer design, first-strand cDNA synthesis, and PCR amplification were performed as previously described (Duplessis et al. 2011b; Hacquard et al. 2010). Transcript expression levels were normalized with *M. larici-populina* α-tubulin and elongation factor reference genes (Hacquard et al. 2011b), considering the specific efficiency of each primer (Pfaffl 2001). For each gene, to compare profiles measured by oligoarrays and RTqPCR, expression ratios were calculated between relative levels in each experimental condition and the average expression level across conditions.

ACKNOWLEDGMENTS

The authors acknowledge the Joint Genome Institute (Office of Science of the United States Department of Energy under contract number DE-Ac02-05cH11231) for the sequencing and annotation of the genome of the poplar rust fungus *Melampsora larici-populina* and access made through the MycoCosm portal. E. Morin and A. Kohler at the Institut National de la Recherche Agronomique (INRA) Nancy are acknowledged for their bioinformatic support. Constructive remarks and suggestions from the editor and two anonymous reviewers, as well as from B. Schwessinger are also acknowledged. C. Lorrain is supported by a young scientist grant (CJS) from INRA. S. Duplessis, A. Hecker, and P. Frey planned and designed the research; C. Delaruelle, S. Hacquard, and J. Pétrowski performed experiments; C. Lorrain, C. Marchal, B. Petre, and S. Duplessis analyzed data; C. Lorrain, C. Marchal, B. Petre, and S. Duplessis wrote the manuscript; all authors read and approved the final manuscript.

LITERATURE CITED

Aime, M. C., Toome-Heller, M., and McLaughlin, D. J. 2014. Pucciniomycotina. Pages 271-294 in: The Mycota VII, Part A. Systematics and Evolution, 2nd ed. D. McLaughlin and J.W. Spatafora, eds. Springer-Verlag, Heidelberg.

Aime, M. C., McTaggart, A. R., Mondo, S. J., and Duplessis, S. 2017. Phylogenetics and phylogenomics of rust fungi. Adv. Genet. 100:267-307.

Anders, S., and Huber, W. 2010. Differential expression analysis for sequence count data. Genome Biol. 11:R106-R106.

Bakkeren, G., Joly, D. L., and Duplessis, S. 2016. Editorial: Genomics research on non-model plant pathogens: Delivering novel insights into rust fungus biology. Front. Plant Sci. 7:216.

Chang, S., Puryear, J., and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine tree. Plant Mol. Biol. Report. 11:113-116.

Cummins, G., and Hiratsuka, Y. 2003. Illustrated genera of rust fungi. Pages 1-225 in: Illustrated Genera of Rust Fungi. APS Press, St Paul, MN, U.S.A.

Cuomo, C. A., Bakkeren, G., Khalil, H. B., Panwar, V., Joly, D., Linning, R., Sakthikumar, S., Song, X., Adiconis, X., Fan, L., Goldberg, J. M., Levin, J. Z., Young, S., Zeng, Q., Anikster, Y., Bruce, M., Wang, M., Yin, C., McCallum, B., Szabo, L. J., Hulbert, S., Chen, X., and Fellers, J. P. 2017. Comparative analysis highlights variable genome content of wheat rusts and divergence of the mating loci. G3 (Bethesda) 7:361-376.

Duplessis, S., Bakkeren, G., and Hamelin, R. 2014. Advancing knowledge on biology of rust fungi through genomics. Adv. Bot. Res. 70:173-209.

Duplessis, S., Cuomo, C. A., Lin, Y. C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D. L., Hacquard, S., Amselem, J., Cantarel, B. L., Chiu, R., Coutinho, P. M., Feau, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M. G., Kodira, C. D., Kohler, A., Kües, U., Lindquist, E. A., Lucas, S. M., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J. L., Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Salamov, A. A., 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 unraveled 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.

Duplessis, S., Joly, D. L., and Dodds, P. N. 2012. Rust Effectors. Pages 155-193 in: Effectors in Plant-Microbe Interactions. John Wiley & Sons Inc., New York.

Germain, H., Joly, D. L., Mireault, C., Plourde, M. B., Letanneur, C., Stewart, D., Morency, M.-J., Petre, B., Duplessis, S., and Séguin, A. 2018. Infection assays in *Arabidopsis* reveal candidate effectors from the poplar rust fungus that promote susceptibility to bacteria and oomycete pathogens. Mol. Plant Pathol. 19:191-200.

Hacquard, S., Delaruelle, C., Frey, P., Tisserant, E., Kohler, A., and Duplessis, S. 2013. Transcriptome analysis of poplar rust telia reveals overwintering adaptation and tightly coordinated karyogamy and meiosis processes. Front. Plant Sci. 4:456.

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., Joly, D. L., Lin, Y.-C., Tisserant, E., Feau, N., Delaruelle, C., Legué, V., Kohler, A., Tanguay, P., Petre, B., Frey, P., Van de Peer, Y., Rouzé, P., Martin, F., Hamelin, R. C., and Duplessis, S. 2012. A comprehensive analysis of genes encoding small secreted proteins identifies candidate effectors in *Melampsora larici-populina* (poplar leaf rust). Mol. Plant-Microbe Interact. 25:279-293.

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 2011:716041.

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 of poplar leaves. Physiol. Mol. Plant Pathol. 75: 106-112.

Harder, D. 1984. Developmental ultrastructure of hyphae and spores. Pages 333-361 in: The Cereal Rust. W. R. Bushnell and A. P. Roelfs, eds., Academic Press, New York.

Henrissat, B., Surolia, A., and Stanley, P. 2015. A Genomic View of Glycobiology. Pages 2015-2017 in: Essentials of Glycobiology. A. Varki, R. D. Cummings, J. D. Esko, eds. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.

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.

Kamel, L., Tang, N., Malbreil, M., San Clemente, H., Le Marquer, M., Roux, C., and Frei Dit Frey, N. 2017. The comparison of expressed candidate secreted proteins from two arbuscular mycorrhizal fungi unravels common and specific molecular tools to invade different host plants. Front. Plant Sci. 8:124.

Kohler, A., Kuo, A., Nagy, L. G., Morin, E., Barry, K. W., Buscot, F., Canbäck, B., Choi, C., Cichocki, N., Clum, A., Colpaert, J., Copeland, A., Costa, M. D., Doré, J., Floudas, D., Gay, G., Girlanda, M., Henrissat, B., Herrmann, S., Hess, J., Högberg, N., Johansson, T., Khouja, H.-R., LaButti, K., Lahrmann, U., Levasseur, A., Lindquist, E. A., Lipzen, A., Marmeisse, R., Martino, E., Murat, C., Ngan, C. Y., Nehls, U., Plett, J. M., Pringle, A., Ohm, R. A., Perotto, S., Peter, M., Riley, R., Rineau, F., Ruytinx, J., Salamov, A., Shah, F., Sun, H., Tarkka, M., Tritt, A., Veneault-Fourrey, C., Zuccaro, A., Mycorrhizal Genomics Initiative Consortium, Tunlid, A., Grigoriev, I. V., Hibbett, D. S., and Martin, F. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat. Genet. 47:410-415.

Kubicek, C. P., Starr, T. L., Glass N. L. 2014. Plant cell wall–degrading enzymes and their secretion in plant-pathogenic fungi. Annu. Rev. Phytopathol. 52:427-451.

Lahrmann, U., Ding, Y., Banhara, A., Rath, M., Hajirezaei, M. R., Döhlemann, S., von Wirén, N., Parniske, M., and Zuccaro, A. 2013. Host-related metabolic cues affect colonization strategies of a root endophyte. Proc. Natl. Acad. Sci. U.S.A. 110:13965-13970.

Langfelder, P., and Horvath, S. 2008. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

Liao, H.-L., Chen, Y., and Vilgalys, R. 2016. Metatranscriptomic study of common and host-specific patterns of gene expression between pines and their symbiotic ectomycorrhizal fungi in the genus *Suillus*. PLoS Genet. 12:e1006348.

Littlefield, L.J., and Heath, M.C. 1979. Ultrastructure of Rust Fungi. Academic Press, New York.

Liu, J.-J., Sturrock, R. N., Sniezko, R. A., Williams, H., Benton, R., and Zamany, A. 2015. Transcriptome analysis of the white pine blister rust pathogen *Cronartium ribicola*: *de novo* assembly, expression profiling, and identification of candidate effectors. BMC Genomics 16:678.

Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., and Kahmann, R. 2015. Fungal effectors and plant susceptibility. Annu. Rev. Plant Biol. 66:513-545.

Lorrain, C., Hecker, A., and Duplessis, S. 2015. Effector-mining in the poplar rust fungus *Melampsora larici-populina* secretome. Front. Plant Sci. 6:1051.

Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550.

Martin, F., Kohler, A., Murat, C., Veneault-Fourrey, C., and Hibbett, D. S. 2016. Unearthing the roots of ectomycorrhizal symbioses. Nat. Rev. Microbiol. 14:760-773.

Mendgen, K. 1984. Development and physiology of teliospores. Pages 375-384 in: The Cereal Rust. W. R. Bushnell and A. P. Roelfs, eds., Academic Press, New York.

Oksanen, J., Blanchet, G. F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, L., Solymos, P., Stevens, M. H. H. S., Szoecs, E., and Wagner, H. 2017. The vegan package. Community ecol. package. The CRAN R Project.

Pernaci, M., De Mita, S., Andrieux, A., Pétrowski, J., Halkett, F., Duplessis, S., and Frey, P. 2014. Genome-wide patterns of segregation and linkage disequilibrium: The construction of a linkage genetic map of the poplar rust fungus *Melampsora larici-populina*. Front. Plant Sci. 5:454.

Petre, B., Joly, D. L., and Duplessis, S. 2014. Effector proteins of rust fungi. Front. Plant Sci. 5:416.

Petre, B., Lorrain, C., Saunders, D. G. O., Win, J., Sklenar, J., Duplessis, S., and Kamoun, S. 2016. Rust fungal effectors mimic host transit peptides to translocate into chloroplasts. Cell. Microbiol. 18:453-465.

Petre, B., Morin, E., Tisserant, E., Hacquard, S., Da Silva, C., Poulain, J., Delaruelle, C., Martin, F., Rouhier, N., Kohler, A., and Duplessis, S. 2012. RNA-seq of early-infected poplar leaves by the rust pathogen *Melampsora larici-populina* uncovers PtSultr3;5, a fungal-induced host sulfate transporter. PLoS One 7:e44408.

Petre, B., Saunders, D. G. O., Sklenar, J., Lorrain, C., Win, J., Duplessis, S., and Kamoun, S. 2015. Candidate effector proteins of the rust pathogen *Melampsora larici-populina* target diverse plant cell compartments. Mol. Plant-Microbe Interact. 28:689-700.

Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:e45.

Pinon, J., and Frey, P. 2005. Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. Pages 139-154 in: Rust Diseases of Willow and Poplar. CAB International, Wallingford, U.K.

Plett, J. M., Tisserant, E., Brun, A., Morin, E., Grigoriev, I. V., Kuo, A., Martin, F., and Kohler, A. 2015. The mutualist *Laccaria bicolor* expresses a core gene regulon during the colonization of diverse host plants and a variable regulon to counteract host-specific defenses. Mol. Plant-Microbe Interact. 28:261-273.

Ragazzi, A., Longo, N., Naldini, B., Moricca, S., and Dellavalle, I. 2005. Basidiospore-derived penetration by species of Cronartium and Melampsora an outline. Pages 161-174 in: Rust Diseases of Willow and Poplar. M. H. Pei and A.R McCracken, eds. CABI, London.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43:e47.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Rovenich, H., Boshoven, J. C., and Thomma, B. P. H. J. 2014. Filamentous pathogen effector functions: Of pathogens, hosts and microbiomes. Curr. Opin. Plant Biol. 20:96-103.

Schulze-Lefert, P., and Panstruga, R. 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci. 16:117-125.

Sharpee, W. C., and Dean, R. A. 2016. Form and function of fungal and oomycete effectors. Fungal Biol. Rev. 30:62-73.

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Manners, J. M., Singh, K. B., and Taylor, J. M. 2015. Advances and challenges in computational prediction of effectors from plant pathogenic fungi. PLoS Pathog. 11:e1004806.

Tyler, B. M., and Rouxel, T. 2012. Effectors of fungi and oomycetes: their virulence and avirulence functions and translocation from pathogen to host cells. Pages 123-167 in: Molecular Plant Immunity. Wiley-Blackwell, Hoboken, NJ, U.S.A.

van der Heijden, M. G. A., Martin, F. M., Selosse, M.-A., and Sanders, I. R. 2015. Mycorrhizal ecology and evolution: The past, the present, and the future. New Phytol. 205:1406-1423.

Vialle, A., Frey, P., Hambleton, S., Bernier, L., and Hamelin, R. 2011. Poplar rust systematics and refinement of *Melampsora* species delineation. Fungal Divers. 50:227-248.

Voegele, R. T., Hahn, M., and Mendgen, K. 2009. The Uredinales: Cytology, biochemistry, and molecular biology. Pages 69-94 in: The Mycota V Plant Relationships. Springer, Berlin.

Win, J., Chaparro-Garcia, A., Belhaj, K., Saunders, D. G., Yoshida, K., Dong, S., Schornack, S., Zipfel, C., Robatzek, S., Hogenhout, S. A., and Kamoun, S. 2012. Effector biology of plant-associated organisms: Concepts and perspectives. Cold Spring Harb. Symp. Quant. Biol. 77:235-247.

Xu, J., Linning, R., Fellers, J., Dickinson, M., Zhu, W., Antonov, I., Joly, D. L., Donaldson, M. E., Eilam, T., Anikster, Y., Banks, T., Munro, S., Mayo, M., Wynhoven, B., Ali, J., Moore, R., McCallum, B., Borodovsky, M., Saville, B., and Bakkeren, G. 2011. Gene discovery in EST sequences from the wheat leaf rust fungus *Puccinia triticina* sexual spores, asexual spores and haustoria, compared to other rust and corn smut fungi. BMC Genomics 12:161.

AUTHOR-RECOMMENDED INTERNET RESOURCE

Joint Genome Institute MycoCosm Melampsora larici-populina v1.0 webpage: http://genome.jgi.doe.gov/Mellp1/Mellp1.home.html