FEMS Microbiology Ecology, 96, 2020, fiaa184

doi: 10.1093/femsec/fiaa184
Advance Access Publication Date: 12 September 2020
Research Article

RESEARCH ARTICLE

Comparative analyses of saprotrophy in *Salisapilia sapeloensis* and diverse plant pathogenic oomycetes reveal lifestyle-specific gene expression

Sophie de Vries[1,*,†], Jan de Vries[1,2,3,4,5,‡], John M. Archibald[1,§] and Claudio H. Slamovits[1,#]

[1]Department of Biochemistry and Molecular Biology, Dalhousie University, 5850 College Street, Halifax, NS B3H 4R2 Canada, [2]Institute of Microbiology, Technische Universität Braunschweig, Spielmannstr. 7, 38106 Braunschweig, Germany, [3]Department of Applied Bioinformatics, Institute for Microbiology and Genetics, University of Goettingen, Goldschmidtstr. 1, 37077 Goettingen, Germany, [4]Goettingen Center for Molecular Biosciences (GZMB), University of Goettingen, Justus-von-Liebig-Weg 11, 37077 Goettingen, Germany and [5]Campus Institute Data Science (CIDAS), University of Goettingen, Goldschmidtstr. 1, 37077 Goettingen, Germany

*Corresponding authors: Department of Biochemistry and Molecular Biology, Dalhousie University, 5850 College Street, Halifax, NS B3H 4R2 Canada. Tel: +1-902-494-7894; E-mail: sophie.devries@dal.ca

One sentence summary: Comparative analyses between the saprotrophic oomycete *Salisapilia sapeloensis* and its pathogenic relatives indicate that distinct gene expression patterns underpin the different lifestyles in oomycetes.

Editor: Angela Sessitsch
†Sophie de Vries, http://orcid.org/0000-0002-5267-8935
‡Jan de Vries, http://orcid.org/0000-0003-3507-5195
§John M. Archibald, http://orcid.org/0000-0001-7255-780X
#Claudio H. Slamovits, http://orcid.org/0000-0003-3050-1474

ABSTRACT

Oomycetes include many devastating plant pathogens. Across oomycete diversity, plant-infecting lineages are interspersed by non-pathogenic ones. Unfortunately, our understanding of the evolution of lifestyle switches is hampered by a scarcity of data on the molecular biology of saprotrophic oomycetes, ecologically important primary colonizers of dead tissue that can serve as informative reference points for understanding the evolution of pathogens. Here, we established *Salisapilia sapeloensis* as a tractable system for the study of saprotrophic oomycetes. We generated multiple transcriptomes from *S. sapeloensis* and compared them with (i) 22 oomycete genomes and (ii) the transcriptomes of eight pathogenic oomycetes grown under 13 conditions. We obtained a global perspective on gene expression signatures of oomycete lifestyles. Our data reveal that oomycete saprotrophs and pathogens use similar molecular mechanisms for colonization but exhibit distinct expression patterns. We identify a *S. sapeloensis*-specific array and expression of carbohydrate-active enzymes and putative regulatory differences, highlighted by distinct expression levels of transcription factors. *Salisapilia sapeloensis* expresses only

Received: 28 April 2020; **Accepted:** 8 September 2020

a small repertoire of candidates for virulence-associated genes. Our analyses suggest lifestyle-specific gene regulatory signatures and that, in addition to variation in gene content, shifts in gene regulatory networks underpin the evolution of oomycete lifestyles.

Keywords: oomycetes; EvoMPMI; plant pathogenicity; lifestyle evolution; saprotrophy; oomycete diversity; CAZymes; gene expression; comparative transcriptomics

INTRODUCTION

Oomycetes encompass a great diversity of ecologically and economically relevant plant pathogens (Thines and Kamoun 2010; Kamoun *et al.* 2015). These include, for example, the (in)famous *Phytophthora infestans*, the cause of the Irish potato famine (Kamoun *et al.* 2015). In addition to the plant pathogenic genera, other oomycetes are pathogens of animals (Phillips *et al.* 2008; Derevnina *et al.* 2016). Oomycete pathogens have different modes of interaction with their hosts, ranging from biotrophs to necrotrophs. To fulfill their life cycle, biotrophs need living hosts, necrotrophs kill and degrade them, and hemibiotrophs have an early biotrophic phase followed by a necrotrophic one. This lifestyle diversity is foremost found in the peronosporalean plant pathogens. Interspersed among diverse pathogenic oomycete lineages are (apparently) non-pathogenic lineages with saprotrophic lifestyles (Diéguez-Uribeondo *et al.* 2009; Beakes, Glockling and Sekimoto 2012; Marano *et al.* 2016) (Fig. 1A). How saprotrophy arises in oomycetes is not yet clear and depends strongly on the character of the last common ancestor of oomycetes. Two hypotheses have been considered: (i) the last common ancestor was a pathogen and the saprotrophic lineages have arisen several times independently and (ii) saprotrophy is the ancestral state in oomycetes (Beakes and Thines 2017).

Saprotrophic oomycetes are thought to play important ecological roles as colonizers and decomposers of organic matter, and by making organic debris more accessible to detritivores (Marano *et al.* 2016). Furthermore, hybridization between saprotrophic and pathogenic *Phytophthora* spp. has been hypothesized to drive distant host jumps of this pathogenic lineage (Marano *et al.* 2016). The lifestyle of a microbe (pathogenic vs non-pathogenic) depends on its environment, and potentially non-pathogenic organisms can be opportunistic pathogens under some environmental conditions (Koide *et al.* 2008; Porras-Alfaro and Bayman 2011; Kuo *et al.* 2014; Charkowski 2016). Hence, any given saprotrophic oomycete lineage might in fact be a latent pathogen; we cannot exclude the possibility that they are able to colonize and exploit living plants. Despite the ecological functions of oomycete saprotrophs and their possible contributions to pathogenicity, how saprotrophs colonize their substrates and whether and how they differ from pathogens remains little understood. Genomic studies have begun to address this question.

Genomic data from *Thraustotheca clavata*, a saprotroph from the Saprolegniales, provided the first insights into the molecular biology of a saprotrophic oomycete (Misner *et al.* 2015). The *in silico* secretome of *T. clavata* is similar in its predicted functions to the secretomes of oomycete pathogens. This agrees with genomic analyses of fungi showing that fungal saprotrophs utilize a similar molecular toolkit as fungal pathogens (Amselem *et al.* 2011; Ohm *et al.* 2012). This toolkit includes, among other things, carbohydrate-active enzymes (CAZymes) for substrate degradation and cell wall remodeling, as well as other degradation-relevant enzymes such as peptidases and lipases, and sterol-binding elicitins (Amselem *et al.* 2011; Ohm *et al.* 2012; Misner *et al.* 2015). Differences between pathogens and saprotrophs are, for example, found in the size of gene families, such as the cutinase family, which is required for successful pathogen infection but is much smaller in the genomes of fungal saprotrophs (Ohm *et al.* 2012). The CAZyme content of fungi also seems to be shaped by lifestyles, hosts and substrates (Zhao *et al.* 2013). In a study of 156 fungal species, degradation profiles were found to be strongly substrate dependent, although on similar substrates the ecological role (pathogen or saprotroph) was relevant (King *et al.* 2011).

One of the principal ideas in studying the genomics of filamentous microbes is that the history of the changes in their lifestyles can be inferred from their genomes (Raffaele and Kamoun 2012; Guttman, McHardy and Schulze-Lefert 2014). Indeed, expansions or reductions of specific gene families can speak to the lifestyle relevance of such genes. Yet, gene presence/absence data do not provide insight into the *usage* of these genes by a given organism. Indeed, a recent publication from fungi highlighted that transcriptomics is a better predictor of ecological strategies of decomposers than their genomic repertoire (Barbi *et al.* 2020). Moreover, as noted above, some saprotrophic oomycetes may be latent pathogens—conclusions based on genomic data alone should be taken with caution. It is here that large-scale transcriptomic and proteomic analyses have the potential to provide deeper insight. Indeed, transcriptomics is a powerful discovery tool for gaining global insight into candidate genes involved in distinct lifestyles, e.g. with virulence-associated and -validated candidates being upregulated during infection (Haas *et al.* 2009; Asai *et al.* 2014; Zuluaga *et al.* 2016; Ah-Fong, Shrivastava and Judelson 2017). Such data underpin the notion that plant pathogenicity is at least partly regulated at the level of gene expression. This is corroborated by the fact that distinct gene expression patterns are associated with host-dependent endophytic vs pathogenic colonization of the same fungus, suggesting that a plastic transcriptomic response may be associated with the capability of fungi to utilize different lifestyles (Baetsen-Young *et al.* 2020; Shahid 2020). Given that saprotrophic interactions of fungi are also at least partially regulated on the transcript level (Barbi *et al.* 2020), we may therefore infer the differences in molecular biology of saprotrophs relative to pathogens via gene expression analyses.

The functional genomic chassis of saprotrophy has been intensively studied in fungi, but not in the 1.5 billion-year-divergent oomycetes. We established an axenic system to study saprotrophy in oomycetes, which entailed using *Salisapilia sapeloensis* grown on sterilized marsh grass litter (MGL). *Salisapilia sapeloensis* was chosen primarily because of its phylogenetic placement, branching next to many well-described oomycete plant pathogens: either as an early-branching peronosporalean or basal to the Peronosporaceae (Hulvey *et al.* 2010; Bennett and Thines 2019). Using this system, we first applied a comparative transcriptomic approach to investigate molecular differences within one saprotrophic interaction across three different conditions and second between this saprotrophic and several pathogenic interactions. For the latter, we started with

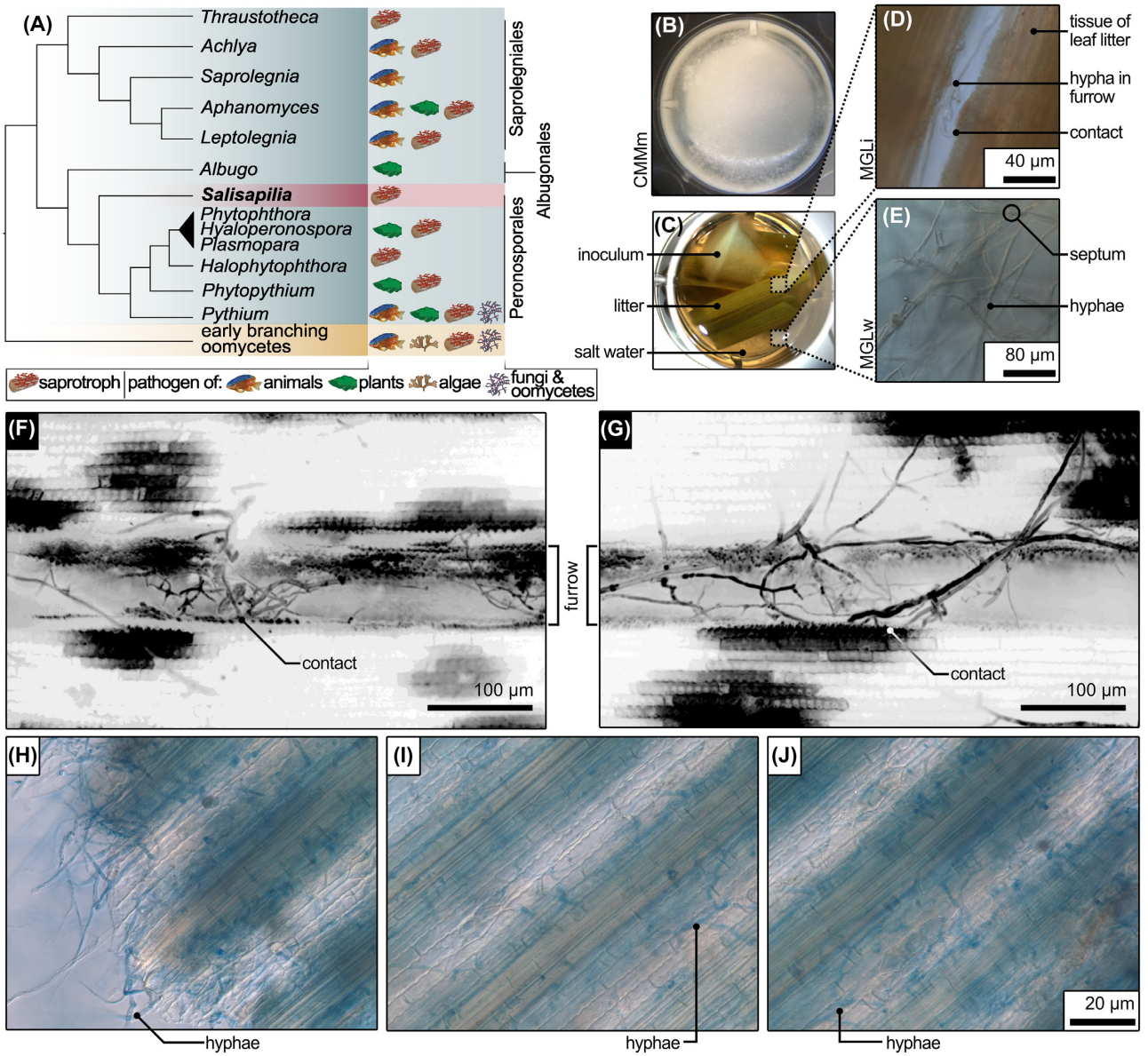


Figure 1. *Salisapilia sapeloensis* colonizes marsh grass litter (MGL). **(A)** Phylogeny of oomycetes. The lifestyles that occur in a specific lineage are indicated on the right of each lineage. Lifestyles include: plant (leaf icon) and animal (fish icon) pathogens, mycoparasitism (mycelium icon) and saprotrophs (log icon) its occurrences are according to Pemberton *et al.* (1990), Ribeiro and Butler (1995), Diéguez-Uribeondo *et al.* (2009), Beakes *et al.* (2012), Marano *et al.* (2016) and Gabrielová *et al.* (2018). The cladogram is based on published phylogenetic studies (Hulvey *et al.* 2010; Beakes *et al.* 2012; McCarthy and Fitzpatrick 2017). We note that based on more recent phylogenetic analyses the Salisapiliaceae, to which the genus *Salisapilia* belongs, could also be sister to the Peronosporaceae, which includes the polyphyletic genus *Phytophthora* (Bennett and Thines 2019), and would hence be embedded between *Phytophthora* and *Pythium*. **(B, C)** An established, axenic system for cultivating *S. sapeloensis*; closeups of *S. sapeloensis* growing in wells with (B) liquid CMM (CMMm) and (C) MGL. **(D)** A micrograph of *S. sapeloensis* growing in the litter's furrows and associated to leaf tissue (MGLi). Hyphae, leaf tissue and contact sites are labeled. **(E)** A micrograph of *S. sapeloensis* growing in water around the MGL (MGLw). Hyphae and septae are labeled. **(F, G)** Calcofluor-white-stained (Fluorescent Brightener 28) *S. sapeloensis* growing into the furrows (labeled) and associating with leaf tissue (labeled as contact; z-stacks reconstructed and projected from confocal micrographs; whole mount staining). It should be noted that calcofluor white stains both the cell walls of plants and oomycetes; more whole mount confocal micrographs can be found in Figure S2 (Supporting Information). **(H–J)** Micrographs of trypan blue-stained leaf litter inoculated with *S. sapeloensis*. Examples for hyphae growing in the leaf tissue are labeled.

a comparison of the transcriptome of *S. sapeloensis* with 22 oomycete genomes and then compared gene expression in our saprotrophic system with that of two well-defined oomycete–host systems: *Pythium ultimum* and *Phytophthora infestans* infecting potato tubers. Because these two pathogens infect a specialized, starch-rich organ, we broadened our analyses to include additional oomycetes from the Albugonales and Peronosporales with diverse ecological backgrounds and hosts to validate our initial data. Our data show *S. sapeloensis* actively degrades MGL in a saprotrophic fashion. Indeed, CAZymes were found to be highly responsive in MGL-associated and MGL-colonizing mycelium. Comparing *S. sapeloensis* with plant pathogenic oomycetes, we identified transcriptomic signatures of saprotrophy and pathogenicity. Our data point to differences in regulatory networks that may shape the evolution of oomycete lifestyles.

MATERIALS AND METHODS

Growth conditions and inoculation of MGL

Salisapilia sapeloensis (CBS 127946) was grown at 21°C in the dark on corn meal agar (1.8% sea salt; 1.5% agar; corn meal extract—produced by cooking corn meal (60 g/l) for 1 h and pressing it through a cheese cloth). *Spartina alterniflora* was collected at 44°40′N, 63°25′W, washed, chopped and autoclaved (MGL). Autoclaved MGL was transferred to 24-well plates containing sterile salt water (1.8% sea salt) and inoculated with agar plugs with 2–6 day old *S. sapeloensis* mycelium. Control mycelium was transferred to 24-well plates containing liquid corn meal medium (CMM; 1.8% sea salt). Inoculations and controls grew in the dark at 21°C for 7 days (three 24-well plates per treatment).

Staining and microscopy

An Axioplan II (Axiocam HRC color camera; Zeiss, Jena, Germany) was used for light microscopy. Trypan blue staining was performed according to (de Vries *et al.* 2017). For confocal microscopy (LSM 710; Zeiss), mycelia were stained with 1% calcofluor white (Fluorescent Brightener 28 powder; Sigma-Aldrich, St. Louis, Missouri, USA) solved in water (Herburger and Holzinger 2016).

DNA extraction and identification of marsh grass

DNA was extracted with Edwards buffer (Edwards, Johnstone and Thompson 1991). *Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit* (*RbcL*) and the *partial18S-ITS1 (internal transcribed spacer 1)-5.8S-ITS2-partial28S* regions were amplified (PrimeSTAR® Max DNA Polymerase; TaKaRa, Mountain View, California, USA) using rbcLa-F (Levin *et al.* 2003)/rbcLa-R (Kress and Erickson 2007) and rbcLa-F/rbcLajf634R (Fazekas *et al.* 2008) and ITS2-S2F (Chen *et al.* 2010)/ITS4 (White *et al.* 1990) primers. PCR products were purified (Monarch® PCR & DNA Cleanup Kit 5 μg; New England BioLabs, Whitby, Ontario, Canada) and sequenced from both sides (Eurofins Genomics, Toronto, Ontario, Canada). After quality assessment, we generated consensus sequences (accessions: MH926040, MH931373). Blastn (Altschul *et al.* 1990) was used against NCBI nt. The *partial18S-ITS1-5.8S-ITS2-partial28S* sequence retrieved two equally good hits (100% coverage, 99.718% identity, e-value 0), both to *Sporobolus alterniflorus* (synonym: *Spartina alterniflora*). The *RbcL* sequence retrieved *Sporobolus maritimus* and *Spartina alterniflora* (99% coverage, 100% identity, e-value 0). Of the two, only *S. alterniflora*'s range includes Nova Scotia.

RNA extraction and sequencing

Biological triplicates per treatment (3 × *S. sapeloensis* grown in CMM, 3 × salt water and 3 × MGL) were harvested for Ribonucleic acid (RNA) extraction. For mycelium grown directly in CMM or salt water close to MGL, four wells per biological replicate were pooled and ground in 1 ml TRIzol using a Tenbroeck homogenizer. For *S. sapeloensis* colonizing MGL, 12 wells of MGL were combined for each replicate and ground in liquid nitrogen using mortar and pestle; 100 mg of powder was extracted using TRIzol. RNA was treated with DNAseI. RNA was assessed via an Epoch spectrophotometer (BioTek, Winooski, Vermont, USA) and a formamide gel, sent to Genome Québec (Montréal, Québec, Canada) and analyzed using a Bioanalyzer. Nine libraries (biological triplicates, three treatments) were prepared with high-quality RNA and 344 505 280 100 bp paired-end (PE) reads were sequenced (Illumina HiSeq4000).

Data processing and taxonomic annotation

Raw data were assessed using FastQC v. 0.11.5 (www.bioinformatics.babraham.ac.uk/projects/fastqc) and trimmed using Trimmomatic v. 0.36 (Bolger, Lohse and Usadel 2014) (settings: ILLUMINACLIP:<custom_adapter_file>:2:30:10:2:TRUE HEADCROP:10 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:36). 312 999 896 PE reads remained after trimming and filtering (Figure S1, Supporting Information). Trimmed reads were re-analyzed with FastQC v. 0.11.5 and are available within Bioproject PRJNA487262. We pooled all read data for a *de novo* assembly using Trinity v. 2.5.0 (Grabherr *et al.* 2011) with Bowtie2 v. 2.3.3.1 (Langmead and Salzberg 2012), producing 44 669 isoforms that were annotated using DIAMOND BLASTx v0.8.34.96 (Buchfink, Xie and Huson 2015): Isoforms were queried against the nr database (June 2017). Statistically significant hits (e-value cutoff 10^{-5}) were kept and sorted according to taxonomy. Isoforms with a hit to an oomycete protein were assigned to the oomycete-affiliated dataset.

Differential expression analyses

Isoform abundance estimates were calculated using Bowtie2 v. 2.3.3.1 and RSEM v. 1.2.18 (Li and Dewey 2011) and transcript per million (TPM) that were trimmed mean of M values (TMM)-normalized ($TPM_{TMM\text{-}normalized}$) values were calculated. Differential gene expression was analyzed based on read counts using edgeR v. 3.20.9 (limma v. 3.34.9; Robinson, McCarthy and Smyth 2010), applying Benjamini–Hochberg false discovery rate correction. Data from all conditions were hierarchically clustered according to their similarity in per-gene expression change (log_2-transformed) from the median-centered expression per transcript. A representative isoform (the isoform with the highest expression) was identified, retrieving 8023 unique genes. After a NCBI contamination screen, 7777 genes were deposited in GenBank (Figure S1, Supporting Information; Bioproject PRJNA487262).

In silico protein prediction

We predicted protein sequences for all major isoforms (i) with a taxonomic affiliation to oomycetes (oomycete-affiliated dataset) and (ii) with no hit in the databases (i.e. potential orphans; unassigned dataset). For the oomycete-affiliated dataset, we predicted all possible reading frames using EMBOSS' getorf (Rice, Longden and Bleasby 2000) and then used BLASTx (Altschul *et al.* 1990) against protein data from (i) all hits retrieved via DIAMOND blast and (ii) 22 oomycete genomes (Table S1, Supporting Information) using the unique genes as queries. Based on the best blast hit we determined the most likely open reading frame and protein sequence. For the unassigned dataset, we used the bioperl script longorf (https://github.com/bioperl/bioperl-live/blob/master/examples/longorf.pl) with the *strict* option.

Functional annotation

Annotations were based on (i) EggNOG-mapper 4.5 (Huerta-Cepas *et al.* 2016), (ii) GhostKOALA (Kanehisa, Sato and Morishima 2016) (database *genus prokaryotes and family eukaryotes*) and (iii) a BLASTp (Altschul *et al.* 1990) against the combined DIAMOND output and oomycete proteome database (e-value

cutoff of 10^{-5}). Differentially expressed genes were manually annotated using these data. If databases produced contradictory results, we used CD search (Marchler-Bauer *et al.* 2017) to identify protein domains and chose the better supported annotation.

Prediction of putative CAZymes

CAZymes of *S. sapeloensis* and 22 oomycetes were predicted using dbCAN (Yin *et al.* 2012) (HMMER3; Mistry *et al.* 2013). We divided sequences assigned to more than one family of CAZymes into two groups: (i) different functional specificity (e.g. CBM and GH domains)—retaining the more specific annotation (i.e. GH, not CBM) and (ii) similar functional specificity (e.g. GT and GH). We used CD search to validate domain presence (Marchler-Bauer *et al.* 2017). If only one of the two domains was found using this approach, the annotation was changed. Otherwise, we kept both annotations (category *mixed*).

CAZyme family enrichment

Distributions of CAZyme families were calculated as percentage of CAZymes in the transcriptome of *S. sapeloensis* and 22 oomycetes genomes (Tyler *et al.* 2006; Haas *et al.* 2009; Baxter *et al.* 2010; Lévesque *et al.* 2010; Kemen *et al.* 2011; Adhikari *et al.* 2013; Jiang *et al.* 2013; Quinn et al. 2013; Misner *et al.* 2015; Sambles *et al.* 2015; Sharma *et al.* 2015; Table S1, Supporting Information). CAZyme profiles of *S. sapeloensis* were compared with those of the other oomycetes. We defined enrichment of expression of CAZyme subfamilies as $\log_2(\text{CAZyme-subfamily}_{Salisapilia}(\%)/(\text{Ø CAZyme-subfamily}_{\text{other oomycetes}}+\text{SD})\ (\%)) \geq 1$.

Prediction of candidate transcription factors

Transcription factor (TF) candidates were predicted using PlantTFDB 4.0 (Jin *et al.* 2017) with the oomycete-affiliated dataset as input. PlantTFDB uses a HMMER-based approach to identify TF domains based on Pfam and published domains and can therefore identify eukaryotic TF candidates outside of plants.

Prediction of putative effector proteins

Putative effectors were predicted using the oomycete-affiliated and unassigned datasets. To predict RxLR-dEER type effectors, we used HMMER 3.1b1 (Mistry *et al.* 2013) with the Hidden Markov Model (HMM) profile from (Win *et al.* 2007) and a heuristic search identifying the RxLR-dEER motif in the first 150 aa (Whisson *et al.* 2007). Putative Crinkler type effectors were identified as in (McGowan and Fitzpatrick 2017) using a heuristic search for LFLA(R/K)X and LYLA(R/K)X between the residues 30 to 70. SSPs were predicted using EffectorP 2.0 (Sperschneider *et al.* 2018). Only considering sequences with a $\text{TPM}_{\text{TMM-normalized}} > 1$, we further required the absence of transmembrane (TM) domains (i.e. no domain predicted by either TMHMM v. 2.0 (Krogh *et al.* 2001), Phobius (Käll, Krogh and Sonnhammer 2007) or TOPCONS2 (Tsirigos *et al.* 2015)) and the presence of a signal peptide (SP; predicted by two algorithms independently).

SPs were predicted using SignalP3-HMM (Bendtsen *et al.* 2004), SignalP4.1-HMM (Petersen *et al.* 2011), Phobius (Käll, Krogh and Sonnhammer 2007) and TOPCONS2 (Tsirigos *et al.* 2015). SignalP3-HMM can outperform SignalP4.1-HMM in SP predictions on oomycete effector proteins (Sperschneider *et al.* 2015). As SP prediction requires intact N-termini, we screened for these using a perl script, resulting in 3242 oomycete-affiliated usable protein sequences. For the unassigned protein sequence dataset, an intact N-terminus was a requirement for translation.

We further analyzed the predicted proteomes of *P. infestans, P. ultimum, Phytophthora parasitica, Phytophthora sojae, Plasmopara halstedii, Hyaloperonospora arabidopsidis, Albugo candida, Albugo laibachii* (Table S1, Supporting Information) and two non-oomycete stramenopiles (*Ectocarpus siliculosus* (Cock *et al.* 2010) and *Phaeodactylum tricornutum* (Bowler *et al.* 2008)) for putative RxLR-dEER type effectors using regular expression and a screen for TM domains and SPs as above.

Comparative transcriptomics

Protein sequences of *S. sapeloensis* and eight oomycetes *A. candida, A. laibachii, H. arabidopsidis, P. infestans, P. ultimum, P. parasitica, P. sojae* and *P. halstedii* (Tyler *et al.* 2006; Haas *et al.* 2009; Baxter *et al.* 2010; Lévesque *et al.* 2010; Kemen *et al.* 2011; Sharma *et al.* 2015) (Table S1, Supporting Information) that found each other reciprocally were designated as RBBHs (BLASTp; e-value cutoff $= 10^{-5}$).

Co-regulation between *S. sapeloensis* and the pathogens *P. infestans* and *P. ultimum* was assessed by comparing differential RNAseq data between RBBHs. Expression data from the colonization of MGL versus growth in CMM were compared with data from early and late infection stages in potato tubers versus early and late growth on Rye medium (RMM) (data from Ah-Fong, Shrivastava and Judelson 2017); $\log_2$(FC) were calculated (Table S1, Supporting Information). We asked whether an RBBH showed co-regulation (i.e. both are upregulated (FC ≥ 1), downregulated (FC ≤ -1) or no change (FC > -1 and <1) in the plant-associated stage compared with the medium growth stage). We further compared the amount of co-regulation between the two pathogens.

Additionally, we compared RNAseq data between *S. sapeloensis* and the eight plant pathogens from the Peronosporales and Albugonales (different lifestyles, different hosts and tissues; Table S1, Supporting Information; Asai *et al.* 2014; Lin *et al.* 2014; Sharma *et al.* 2015; Ah-Fong, Shrivastava and Judelson 2017; Prince *et al.* 2017). Some of these oomycetes are biotrophs and cannot grow outside their hosts. Therefore, we calculated the transcript budget (%TPM or %CPM) to estimate how much transcript an oomycete invests into a gene during infection/colonization and compared it between RBBHs from the different oomycetes.

For *S. sapeloensis* we used $\text{TPM}_{\text{TMM-normalized}}$ values from MGL-colonizing samples. For *P. infestans* and *P. ultimum* we used $\text{CPM}_{\text{TMM-normalized}}$ values (derived from Ah-Fong, Shrivastava and Judelson 2017) for the late infection stage. For the other six oomycetes, we downloaded transcriptome data for progressed infection stages (Table S1, Supporting Information). Here, reads were analyzed as above (trimming settings: HEADCROP:10 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:36) and mapped to the corresponding species' transcript sequences using RSEM v. 1.2.18 (Li and Dewey 2011) with Bowtie v. 2.3.3.1 (Langmead and Salzberg 2012). The data from the different oomycete species are made comparable by using either $\log_2$(FC) or transcript budget (relative transcript investment). Both of these serve to normalize the data within a given dataset before cross-species comparisons.

Phylogenetic analyses

To verify our reciprocal BLAST approach, we queried protein sequences for life cycle markers *Pi*NIFC1,2 and 3, *Pi*CDC14

and PiM90 from *P. infestans* (Judelson and Blanco 2005) against oomycete genomes (*A. candida*, *A. laibachii*, *H. arabidopsidis*, *P. halstedii*, *P. parasitica*, *P. sojae* and *P. ultimum*), *P. infestans*' own genome and the transcriptome of *S. sapeloensis* using BLASTp (Altschul *et al.* 1990). For the TFs, we queried the candidate TFs from *S. sapeolensis* against the genomes of the eight oomycete pathogens using BLASTp (Altschul *et al.* 1990). Sampling in the bZIP phylogeny was increased by adding the additional bZIP identified in *S. sapeloensis* as well as previously characterized bZIPs from *P. infestans* (Blanco and Judelson 2005; Gamboa-Meléndez, Huerta and Judelson 2013). The e-value cutoff for both BLASTp searches was 1×10^{-5}. Sequences were aligned using mafft v.7.3 and 7.4 with either L-INS-I or G-INS-I (depending on the protein family) (Katoh and Standley 2013). Divergent sequences were removed and the data re-aligned. Next, IQTree v.1.5.5 (Nguyen *et al.* 2015) was used to create phylogenies. The best model was selected (Kalyaanamoorthy *et al.* 2017) and the analyses done accordingly. The phylogeny for Myb-related and Nin-like TFs, and the NF-YB TFs and NIFC phylogenies are based on 100 bootstrap replicates; the CDC14, M90 and bZIP TFs phylogenies are based on 500 bootstrap replicates.

Protein sequence and structure analyses

Protein domain structures were predicted using CD search (Marchler-Bauer *et al.* 2017) and InterPro (Mitchell *et al.* 2019).Shrivastava

Co-expression analyses

We used clust (Abu-Jamous and Kelly 2018) to study co-expression in our transcriptome. As input, we used $TPM_{TMM\text{-}normalized}$ values and normalized accordingly from all treatments for all 4592 oomycete-affiliated genes in the oomycete-affiliated dataset. The settings were: cluster tightness 7, minimum cluster size 11, and lowly expressed and flat genes were filtered out prior to clustering, using a 25-percentile cutoff. All retained genes have an expression higher than the cutoff in at least one condition.

Two-step quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

One thousand nanograms of RNA was used for iScript complementary DNA (cDNA) Synthesis (Bio-Rad, Hercules, California, USA). Primers were designed using NCBI primer BLAST (Table S3, Supporting Information), tested on cDNA from *S. sapeloensis* (PrimeSTAR® Max DNA Polymerase; TaKaRa) and amplicons were sequenced by GENEWIZ (South Plainfield, New Jersey, USA).

Expression of four CAZyme-encoding genes (*Salisap1930_c0_g1*, *Salisap3316_c0_g4*, *Salisap3873_c0_g7* and *Salisap4165_c1_g3*) and two TF-encoding genes (*Salisap3868_c0_g1* and *Salisap4629_c0_g1*) across all treatments (3 × biological replicates/treatment; 3 × technical replicates/biological replicate) was assayed on a CFX Connect Real-Time System (BioRad). *H2A* (*Salisap2927_c0_g1*) was selected as a reference gene based on its constant expression across all samples in transcriptome and qRT-PCR. The standard curve was based on pooled cDNA from all conditions. Relative expression was calculated using (Pfaffl 2001). Expression data were tested for normality with a Shapiro–Wilk test (Shapiro and Wilk 1965), equal variance and, depending on the results, significant differences were assessed with a two-sampled *t*-test, a Welch two-sampled *t*-test or a Mann–Whitney U test (Mann and Whitney 1947) using R v. 3.2.1.

RESULTS AND DISCUSSION

Salisapilia sapeloensis: tractable saprotrophy under axenic lab conditions

Oomycetes have evolved pathogenic and saprotrophic lifestyles. While many pathogenic oomycetes are being investigated in depth, little is known about the molecular basis of saprotrophic oomycetes. Comparative functional genomics of plant–microbe interactions can yield vital clues about what distinguishes the lifestyles of pathogens from those of non-pathogenic ones. We are studying *Salisapilia sapeloensis* for three main reasons: (i) it is a peronosporalean oomycete, (ii) it has an interesting phylogenetic position basal to many plant pathogenic oomycetes (Hulvey *et al.* 2010) (Fig. 1A) and (iii) its natural substrate is known and easily accessible.

Salisapilia sapeloensis was originally isolated from MGL of *Spartina alterniflora* (Hulvey *et al.* 2010). To study *S. sapeloensis* in the lab, we created axenic MGL. We then inoculated CMM (control) as well as MGL submersed in salt water with mycelium of *S. sapeloensis* (Fig. 1B and C).

Mycelium of *S. sapeloensis* grew well in both medium and salt water next to the litter (Fig. 1B–E). We investigated the inoculated MGL microscopically to study the organism's association with litter. *Salisapilia sapeloensis* was frequently observed around the edges of axenic MGL and grew into its furrows (Fig. 1D and F–J; Figure S2, Supporting Information). The mycelium associated closely with leaf tissue in the furrows (Fig. 1D, F and G) and grew inside the tissue (Fig. 1H–J). Additionally, we noted that the water surrounding inoculated tissue turned yellow (Fig. 2A)—likely a result of litter degradation, supporting the notion that *S. sapeloensis* utilizes MGL as a substrate. Overall, *S. sapeloensis* thrived on litter. Within seven days it produced enough biomass for downstream analyses. Hence, *S. sapeloensis* is capable of colonizing and surviving on MGL, as the MGL contained no additional nutrients but those encased in the tissue of the marsh grass leaves.

Global differential gene expression in *Salisapilia sapeloensis* and a first glimpse into molecular cell biology

To more fully elucidate how *S. sapeloensis* colonizes and lives off litter, we isolated RNA in biological triplicates and performed RNAseq on cultures growing under three conditions: *S. sapeloensis* grown in regular CMM medium (CMMm), in water next to MGL (MGLw) and on MGL (MGLi; Figure S1, Supporting Information). *De novo* assembly resulted in 7777 unique genes (Dataset S1a, Supporting Information). Of those, 4592 were oomycete-affiliated and 2653 were orphans (Dataset S1b and c, Supporting Information); the remaining 532 genes had a taxonomic affiliation to other organisms and were omitted from downstream analyses. Unless otherwise mentioned, the analyses focused on the strict oomycete-affiliated dataset. Overall, 1628 oomycete-affiliated genes were differentially expressed in at least one comparison (MGLi vs CMMm, MGLi vs MGLw and MGLw vs CMMm; Benjamini–Hochberg adjusted *P* (FDR) ≤ 0.05; Dataset S1d, Supporting Information). Of those 1628 genes with differential expression, 90% (1472 genes) are found in the pairwise comparison of MGLi vs CMMm. In agreement with different sets of biochemical cues delivered by the two nutritional sources in our growth conditions, MGLi expression profiles were more similar to MGLw than to CMMm (Fig. 2B). Nonetheless, 145 oomycete-affiliated genes showed differential expression patterns (FDR $\leq$

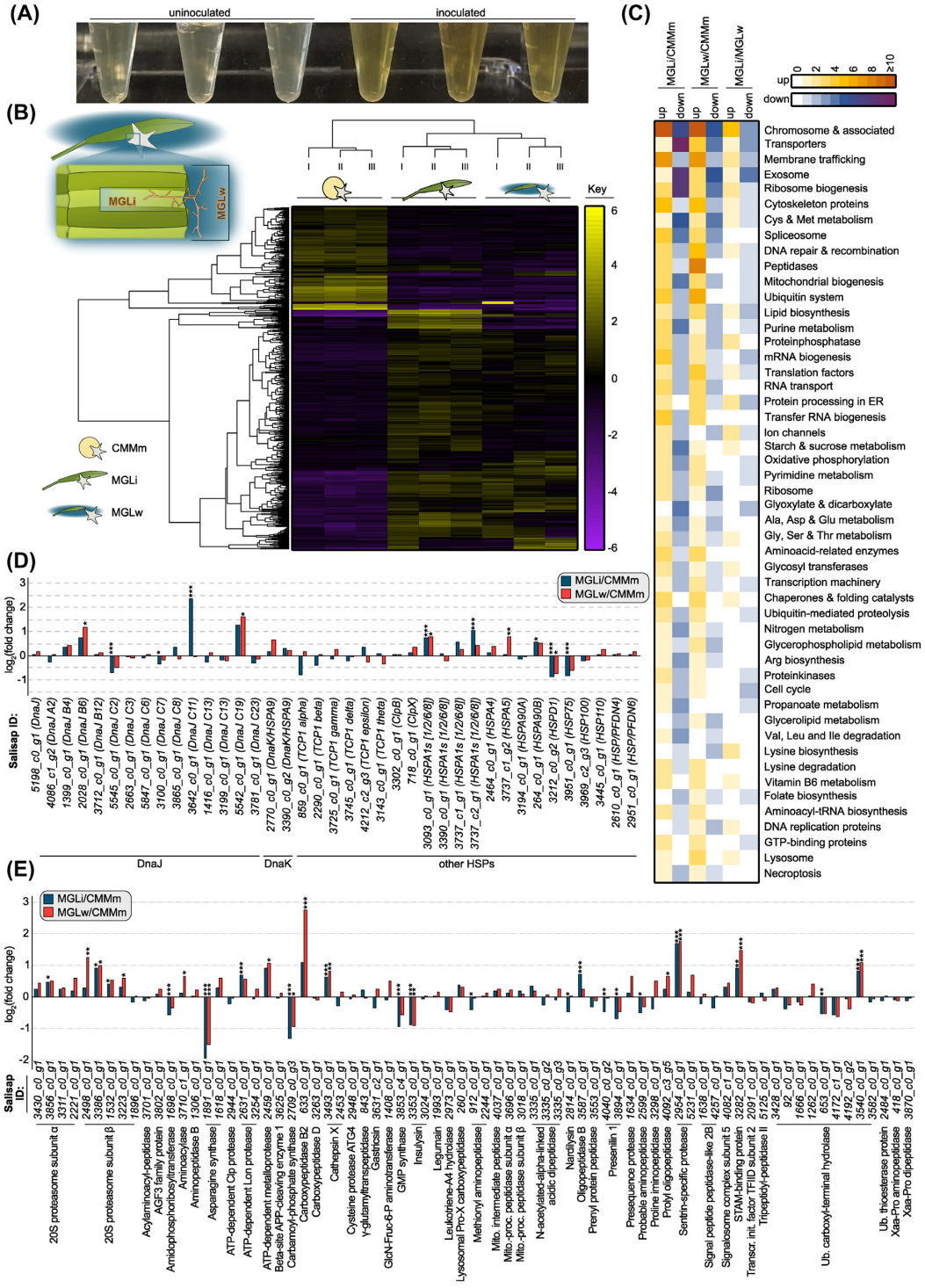


Figure 2. Transcriptomic profiling of the axenic saprotrophic system *Salisapilia sapeloensis* during colonization of litter. **(A)** Visual differences in the water next to uninoculated litter and litter inoculated with *S. sapeloensis*. **(B)** Left: sampling strategy for MGLi and MGLw; right: hierarchically clustered expression values (log_2-transformed; median-centered) for CMMm, MGLw and MGLi. Higher expression than the median-centered expression is indicated in yellow; lower expression is indicated in purple. I, II and III indicate the biological replicates for each condition. **(C)** The 50 most responsive KEGG pathways/BRITE hierarchies, indicating the overall number of KEGG orthologs in a given KEGG pathway/BRITE hierarchy with an induced (white to orange) or reduced (white to dark purple) expression. **(D)** log_2(FC) of the genes corresponding to KEGG orthologs in the umbrella category 'Chaperone and folding catalysts'. Note that one KEGG ortholog can include more than one gene. Only genes with a $TPM_{TMM\text{-}normalized} \geq 1$ in at least one of the three conditions are shown. The y-axis indicates the log_2(FC) when MGLi (blue) or MGLw (red) is compared with CMMm. The gene IDs are given below the bar graph for each gene and in brackets the corresponding KEGG ortholog is indicated. Overall functional categories are given below the gene IDs. Significant differences between MGLi or MGLw and CMMm are indicated by *FDR $\leq$ 0.05, **FDR $\leq$ 0.01 and ***FDR $\leq$ 0.001. **(E)** log_2(FC) of the genes corresponding to KEGG orthologs in the umbrella category 'Peptidases'. The figure shows only genes with a $TPM_{TMM\text{-}normalized} \geq 1$ in at least one treatment. The comparison MGLi vs CMMm is shown in blue and MGLw vs CMMm is shown in red. The gene IDs are indicated below the bars. Below the gene IDs we noted the corresponding KEGG orthologs. Significant differences are given as *FDR $\leq$ 0.05, **FDR $\leq$ 0.01 and ***FDR $\leq$ 0.001.

0.05; Dataset S1d, Supporting Information) between MGLi and MGLw. These are candidate genes specifically associated with the colonization and on-site degradation of litter.

In asexual cultures, peronosporalean oomycetes can form sporangia, which mature by cleaving the cytoplasm to form zoospores (Figure S3A, Supporting Information). Eventually, the zoospores are released and form cysts—triggered by external cues (Tyler 2002; Figure S3A, Supporting Information). *Salisapilia sapeloensis* was observed to produce sporangia after five to 10 days of growth in half-strength seawater (Hulvey *et al.* 2010). After seven days we observed no sporangia independent of treatment/condition.

In *Phytophthora infestans,* induction of cleavage-specific nuclear LIM interactor-interacting factor-encoding genes (*NIFC1,2* and *3*), a dual-specificity protein phosphatase-encoding gene (*CDC14*) and a pumilio-like mating protein-encoding gene (*M90*) is associated with the formation or maturation of sporangia (Cvitanich and Judelson 2003; Judelson and Blanco 2005; Ah-Fong, Xiang and Judelson 2007; Judelson and Tani 2007; Ah-Fong and Judelson 2011). Using data from Ah-Fong, Kim and Judelson 2017, we investigated the expression patterns of those genes during the asexual life cycle of *P. infestans* (Figure S3B–F, Supporting Information). All five genes were expressed at low levels in non-sporulating mycelium (TPM < 2). Expression levels of *PiNIFC1,2 and 3* peak during maturation of sporangia (cleavage); however *PiNIFC* expression already accumulates during sporangia formation. *PiCDC14* and *PiM90* (Figure S3E and F, Supporting Information) are induced in sporangia and *PiM90* showed a second induction in zoospores. Reciprocal BLASTp searches recovered candidates for *SsNIFC2* (*Salisap2115_c0_g1_i2*), *SsCDC14* (*Salisap3692_c0_g1_i19*) and *SsM90* (*Salisap4144_c1_g1_i41*). A phylogenetic approach confirmed these results for CDC14 and M90 (Figure S4A,B, Supporting Information). In the case of RBBH to PiNIFC2, this sequence clustered with PiNIFC1,2,3 and NIFS (Figure S4C, Supporting Information). It is therefore a NIFC/S candidate.

In agreement with our observations, expression levels of *SsNIFC/S* and *SsCDC14* are similar to those of *PiNIFC* and *PiCDC14* in non-sporulating mycelium. *SsM90* expression diverged from this pattern; its expression levels are of similar magnitude to *PiM90* during cleavage. Given that no sporangia formation was observed, we cannot exclude the possibility that *SsM90* may function in other processes than *PiM90*. Comparison between treatments identified an induction of *SsNIFC/S* in MGLi and MGLw (FDR < 0.01) compared with CMMm and a reduction in *SsM90* in litter expression compared with CMM (FDR < 0.01); *PiCDC14* expression was constant between treatments. Despite the differences observed in expression between some treatments, visual assessment and expression levels of *SsNIFC/S* and *SsCDC14* suggest that the cultures of *S. sapeloensis* were not yet sporulating in our treatments.

Molecular signatures of litter colonization

We next determined which pathways the saprotroph requires for litter colonization. To do so, we classified the assembled genes using KEGG pathways and BRITE functional hierarchies (Dataset S2a, Supporting Information), and determined umbrella categories that were most responsive in *S. sapeloensis* (Fig. 2C). We define 'responsiveness' as the number of KEGG orthologs with a cumulative 2-fold change ($\log_2$ (fold change (FC)) ≥ 1 or ≤ -1) in a given umbrella category in pairwise comparisons of MGLi, MGLw and CMMm.

Among the 50 most responsive KEGG pathways/BRITE hierarchy terms, three showed an exclusive induction ($\log_2$ FC $\geq$ 1) upon exposure to MGL. These are 'Chaperone and folding catalysts', 'Vitamin B6 metabolism' and 'Lysosome' (Fig. 2C). 'Chaperone and folding catalysts' may be a stress-associated response. Indeed, of 48 KEGG orthologs present in this pathway, 34 were heat shock proteins (HSPs). The HSP KEGG orthologs were represented by 15 DnaJ-encoding, two DnaK (HSPA09)-encoding and 26 HSP-encoding genes (Dataset S2b, Supporting Information). Of those, 39 genes were expressed at a reliable level ($TPM_{TMM\text{-}normalized} \geq 1$) and seven were significantly induced in litter-associated treatments (i.e. MGLi or MGLw vs CMMm; Fig. 2D). Similarly, Benz *et al.* (2014) observed upregulations of stress-associated proteins, e.g. DnaK, in the saprotrophic fungus *Neurospora crassa* when exposed to different plant cell wall-associated carbon sources. Although MGL is its natural substrate, it is conceivable that toxic degradants and stored plant compounds may trigger a stress response in *S. sapeloensis*. Alternatively, induction of the HSP family could be the result of secretion stress, which has also been observed in saprotrophic fungi (Guillemette *et al.* 2007). The latter is further supported by the responsiveness of other KEGG pathways/BRITE hierarchy terms, such as 'Protein processing in the ER' or 'Membrane trafficking' (Fig. 2C).

Vitamin B6 is a precursor of pyridoxal phosphate, a compound required for diverse biological functions including translation (Jansonius 1998). Several amino acid synthesis-related KEGG pathways/BRITE hierarchy terms were among the 50 most responsive categories (Fig. 2C). Taken together, the elevation of the KEGG pathway 'Vitamin B6 metabolism' in *S. sapeloensis* exposed to MGL may be associated with protein degradation and protein synthesis. Elevation of the KEGG-pathway 'Lysosome' suggests an increase in degradation of biological material, in agreement with saprotrophy. Altogether, the increases in 'Chaperone and folding catalysts', 'Vitamin B6 metabolism' and 'Lysosome' in MGLi or MGLw vs CMMm may point to a change in functional degradation in *S. sapeloensis*. We therefore investigated the direction of responsiveness in KEGG-pathways associated with degradation.

Our experiments on MGLw and MGLi samples show that *S. sapeloensis* can live off litter in the absence of other carbon sources (Fig. 1B–J; Fig. 2A). This observation is underpinned by the responsiveness of the KEGG pathways 'Lysosome' and 'Starch and sucrose metabolism', and BRITE terms 'Transporters', 'Membrane trafficking' and 'Peptidases' (Fig. 2C). Both 'Transporters' and 'Membrane trafficking' are among the three most responsive categories. Both have KEGG orthologs with induced and reduced transcript abundance in pairwise comparisons. Their responsiveness may hence speak to a treatment-specific differential up- and downregulation. Indeed, the putative functions of the KEGG orthologs with the BRITE term 'Transporters' include among others many ABC-transporters, solute carriers involved in, e.g. sugar, fatty acid or amino acid transport, as well as ion transporters (Figure S5, Supporting Information). 'Membrane trafficking' contains 164 KEGG orthologs of which 14 responded to the growth condition: 10 KEGG orthologs (10 genes) were increased in MGLi and MGLw vs CMMm ($\log_2(FC) \geq 1$; Dataset S2, Supporting Information), while four KEGG orthologs (4 genes) were reduced in MGLi and MGLw vs CMMm ($\log_2(FC) \leq -1$; Dataset S2, Supporting Information). Therefore, we observed an overall tendency for an induction of 'Membrane trafficking'. In fungi, the responsiveness of the secretory pathway components is correlated with the degradation of plant material (Benz *et al.* 2014; Coradetti *et al.* 2012). Such connections also likely

exist in oomycetes. Taken together, the visual clues (Fig. 2A) and the responsiveness of degradation-associated KEGG pathways/BRITE hierarchy terms indicate that *S. sapeloensis* actively degrades its litter substrate.

Degradation-associated categories such as the BRITE term 'Peptidase' included 77 KEGG orthologs, 73 of which are expressed at a reliable level ($TPM_{TMM\text{-}normalized} \geq 1$; Fig. 2E). 'Peptidase' is induced in MGLw vs CMMm, with more KEGG orthologs being induced than reduced when comparing MGLi to CMMm (Fig. 2C). 27 peptidase-encoding genes were differentially expressed in one of the three comparisons (FRD $\leq$ 0.05); of those, 56% (15 genes) are upregulated in either MGLi or MGLw vs CMMm (Fig. 2E). A majority of the upregulated peptidase-encoding genes is further attributable to the comparison of MGLw against CMMm, altogether suggesting that protein degradation may be a distinctive feature of litter-associated mycelium.

Alterations in gene expression in the KEGG pathway 'starch and sucrose metabolism' suggest a remodeling of carbon-based degradation in response to a particular substrate. This is expected, because CMMm, MGLi and MGLw represent different carbon sources, with CMMm being enriched in starch compared with MGL or salt water. Likewise, MGL will have long carbon-chains, such as cellulose, which might reach the salt water in an already pre-processed condition. The molecular differences observed when we dissect the components of the 'starch and sucrose metabolism' pathway align with the differences in carbon sources. For example, the KEGG ortholog endo-1,4-β-D-glucanase (EC 3.2.1.4), composed of 10 endo-1,4-β-D-glucanase-encoding genes, is overall induced ($\log_2$ (FC) $>$ 1; Dataset S2, Supporting Information) in MGLi vs MGLw and CMMm. Three of the 10 endo-1,4-β-D-glucanase-encoding genes contribute to both the induction in MGLi vs MGLw and MGLi vs CMMm (*Salisap3873_c0_g6, Salisap3873_c0_g7* and *Salisap3873_c0_g8*; Dataset S1d, Supporting Information). Endo-1,4-β-D-glucanase is involved in the degradation of glucans found in leaves (Kubicek, Starr and Glass 2014), pointing to an ability of *S. sapeloensis* to degrade leaf litter. This is further illustrated by the extent to which the three endo-1,4-β-D-glucanase-encoding genes that caused the overall induction were upregulated: All three genes were on average 87-fold upregulated (FDR $\leq$ 0.05) in MGLi vs CMMm and 49-fold upregulated (FDR $\leq$ 0.05) in MGLi vs MGLw. One of the three genes (*Salisap3873_c0_g8*) even increased by 248-fold in MGLi vs CMMm (FDR $= 5.8 \times 10^{-46}$) and 109-fold in MGLi vs MGLw (FDR $= 5.3 \times 10^{-33}$).

In contrast to the overall induction of endo-1,4-β-D-glucanase, the KEGG ortholog phosphoglucomutase (represented by one gene, *Salisap4157_c0_g2*) has a reduced transcript abundance in MGLi vs CMMm ($\log_2$(FC) $\leq$ -1; Dataset S2a, Supporting Information). This is substantiated by the significant downregulation by 2-fold (FDR $= 5.7 \times 10^{-10}$) of the single phosphoglucomutase-encoding gene in MGLi vs CMMm. Phosphoglucomutase is associated with starch catabolism (Colowick and Sutherland 1942; Preiss 1982), and downregulation of phosphoglucomutase is thus a potential readout of the reduced availability of starch in litter compared with corn meal-based medium. Taken together, the visual and molecular data suggest that *S. sapeloensis* is a saprotroph in our axenic setup, making it an apt system for studying oomycete saprotrophy *in vitro*.

Notably, all four degradation-associated categories had KEGG orthologs specifically responsive (log2 FC $>$ 1/ $<$ -1) in MGLi vs MGLw (Fig. 2C). This suggests that *S. sapeloensis* in MGLw has a slightly different metabolism than in MGLi. This makes sense given that the water next to the litter presumably contains pre-processed organic compounds. Environmental differences may result in a different regulation of biochemical and metabolic pathways, in which mycelium growing on litter (i.e. proximal) mainly degrades large carbon-components into smaller pieces that are further processed by mycelium grown in the surrounding water (distal). Similar patterns have been observed across different mycelial parts in *Sclerotinia sclerotiorum*, an ascomycete growing as a septate mycelium (Peyraud *et al.* 2019). Peyraud *et al.* (2019) hypothesized that a cooperation of distinct parts in a multicellular mycelium may be of an advantage to the pathogen. They coined this a division of labor. Indeed, their study indicated that the division in metabolic tasks would improve plant colonization. Given that the *S. sapeloensis* mycelium is also septate (Hulvey *et al.* 2010) (Fig. 1E), such cooperation between distinct mycelial parts with different environments is conceivable. Similar to *S. sclerotiorum,* distinct mycelia regions (i.e. 'in litter' and 'next to litter' mycelium) of *S. sapeloensis* show significant differences in gene expression associated with the degradation of plant material, a capability that may give it an advantage in colonizing the dead plant tissue.

Salisapilia sapeloensis expresses a distinct profile of degradation-relevant enzymes

The transcriptomic 'fingerprints' of *S. sapeloensis* indicate that it actively degrades MGL. Degradation of plant material is, however, not restricted to saprotrophy, but is also essential for plant pathogenicity (Kubicek, Starr and Glass 2014). To analyze whether and how degradation processes differ between *S. sapeloensis* and oomycete plant pathogens, we explored putative orthologous genes between *S. sapeloensis* and 16 plant pathogens from Albugonales and Peronosporales and then analyzed those genes unique to *S. sapeloensis*. We identified on average 1376 genes in *S. sapeloensis* with no detectable RBBH in at least one pathogen; 705 of these genes (51.2%) have no RBBH with any of the plant pathogens from the Peronosporales (biotrophs, hemibiotrophs and necrotrophs) and Albugonales (biotrophs), suggesting that, while possibly belonging to larger gene families, they have no putative 1:1-ortholog in *S. sapeloensis* (Dataset S3a–c, Supporting Information). Among these unique genes, 191 (27% of 705) show significantly different expression (FDR $\leq$ 0.05) in MGLi vs CMMm (Dataset S3d, Supporting Information). Similarly, 22% (153 genes) are differentially expressed in MGLw vs CMMm (FDR $\leq$ 0.05), but only 31 (4.3%) are differentially expressed in MGLi vs MGLw. To identify the gross molecular functions of these differentially expressed, unique genes we performed a Gene Ontology (GO)-enrichment analysis for the categories MGLi vs CMMm and MGLw vs CMMm. This analysis pointed to a specific enrichment in hydrolase activity (Fig. 3A), implying that *S. sapeloensis* has a novel profile of degradation enzymes compared with oomycete plant pathogens.

To further investigate the mechanisms underlying the degradation of litter and how they may differ from those involved in the degradation of living plant material, we predicted the functional composition of the CAZyme repertoire of *S. sapeloensis*. *Salisapilia sapeloensis* expressed in total 149 genes predicted to encode CAZymes (134 with $TPM_{TMM\text{-}normalized} \geq 1$ in at least one treatment; Dataset S4, Supporting Information). Given that our dataset is a transcriptome, *S. sapeloensis* likely harbors more CAZyme-encoding genes in its genome than it expresses. However, the relative representation of CAZyme-encoding genes in the transcriptome of *S. sapeloensis* can still provide insight into

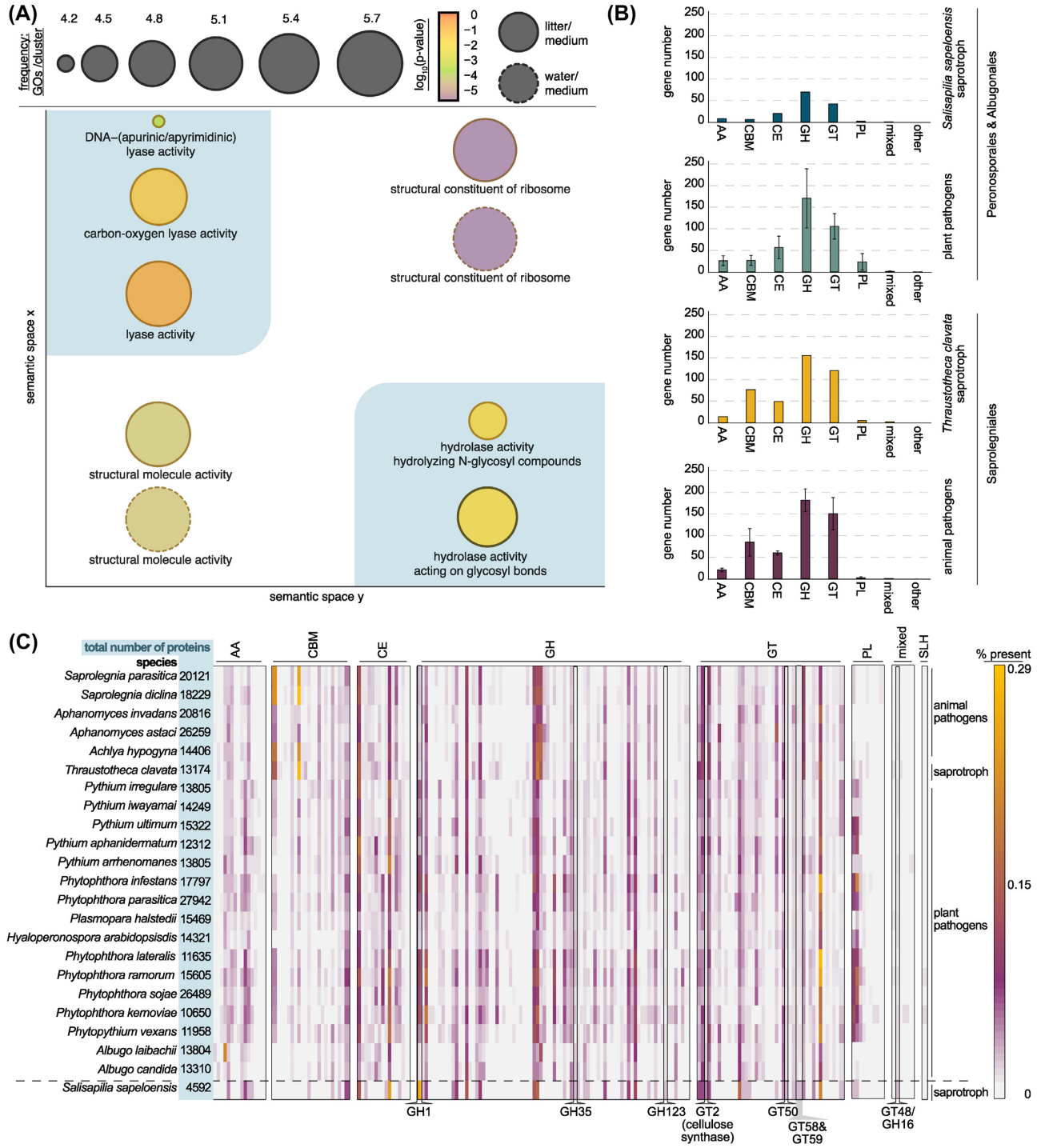


Figure 3. Comparative analyses of the CAZyme repertoire of *Salisapilia sapeloensis*. **(A)** GO-term enrichment of genes unique to *S. sapeloensis*, which are upregulated in MGLi or MGLw vs CMMm (GO-term frequency = circle size, $\log_2$(*P*-value) is indicated by the colors in the circle). The solid lines around the circles indicate an enrichment in MGLi/CMMm and the dotted lines around a circle indicate an enrichment in MGLw/CMMm. The blue background highlights lyase and hydrolase activity. **(B)** CAZyme distribution across oomycetes. **(C)** Heatmap showing the relative number of CAZymes in relation to the overall number of all analyzed *in silico* translated proteins (%) in an oomycete species (the total number of predicted proteins in a genome is given to the right of each species name). The 23 oomycete species are noted on the left; their lifestyles are noted on the right. Every vertical line represents a CAZyme family. Highlighted are CAZyme-encoding gene families with an increased relative presence (i.e. $\log_2(FC) \geq 1$) in *S. sapeloensis*' transcriptome vs the oomycete genomes. These are hence gene families that are expressed more than expected. The relative presence is indicated from white = 0% to yellow = 0.29%.

their relevance for its saprotrophic lifestyle in our *in vitro* system. To get a measure on this relative representation (in absence of genomic data from *S. sapeloensis*), we analyzed 22 genomes (see Table S1, Supporting Information) from diverse oomycetes to estimate the average amount of CAZyme-encoding genes.

We detected between 242 (*A. candida*, a biotroph) and 731 (*P. parasitica*, a hemibiotroph) CAZyme-encoding genes in the oomycete genomes (representing 1.8 and 2.6% of their protein-coding repertoires, respectively; Dataset S4a, Supporting Information). On average, the oomycete genomes harbor 431 ± 139 CAZyme-encoding genes. Assuming a similar amount for the saprotroph, *S. sapeloensis* expresses 38.2 ± 12.2% of the average CAZyme repertoire. The numbers of CAZyme-encoding genes detected in the *S. sapeloensis* transcriptomes were similar to that encoded in the genomes of plant pathogenic oomycetes (from the Peronosporales and Albugonales), with the exception of polysaccharide lyases (PLs) (Fig. 3B). The most abundant PLs in oomycetes (PL1 and PL3; Dataset S4a, Supporting Information) are involved in pectin and pectate degradation (Kubicek, Starr and Glass 2014) and most plant pathogens in our dataset infect dicots. The reduced representation of expressed PLs in the transcriptome of saprotrophic *S. sapeloensis* (relative to the plant pathogenic oomycetes) may reflect a generally lower pectin content in its monocotyledonous substrate compared with their dicot counterparts (Cosgrove 1997; Voragen *et al.* 2009). It is also possible that some PLs simply are not required to degrade leave litter or autoclaved leave tissue. However, apart from the reduced PLs, *S. sapeloensis* expresses an overall representative number of CAZyme-encoding genes.

In contrast to the Peronosporales and Albugonales, all Saprolegniales (regardless of trophism) showed an increase in relative abundance of carbon binding module (CBM)-containing proteins (Fig. 3B). Overall, the CAZyme repertoire of *S. sapeloensis* is more similar to that of the plant pathogens than to the saprolegnian saprotroph. All plant pathogens analyzed here are more closely related to *S. sapeloensis* than *T. clavata* (Fig. 1A), a saprolegnian saprotroph. The representation of the six major CAZyme categories in a genome may, therefore, rather be determined by phylogenetic position than lifestyle (Fig. 1A). Nevertheless, our GO-term analyses showed that *S. sapeloensis'* unique genes were associated with CAZyme-encoding genes. While we currently cannot exclude the possibility that these unique genes are lineage-specific, we observe that this picture mirrors what is known from fungi, where lifestyle does shape the CAZyme repertoire (Zhao *et al.* 2013). We hence predicted that differences in the *S. sapeloensis* CAZyme repertoire might be CAZyme family-specific.

Pathogenicity-associated CAZyme families are not expressed during the saprotrophic interaction

In peronosporalean plant pathogens, CAZymes in general and glycosyl hydrolases (GHs) in particular, constitute a large fraction of genes suggested to have been horizontally acquired from fungi (Richards *et al.* 2011). *Salisapilia sapeloensis,* whose phylogenetic placement, despite not being fully clear (Hulvey *et al.* 2010; Bennett and Thines 2019), appears basal to many peronosporalean plant pathogens (Fig. 1A), does not express these particular GH families at detectable levels during any of the saprotrophic interactions tested herein (Dataset S4, Supporting Information; although this does not exclude their presence in the un-sequenced *S. sapeloensis* genome). To further explore the possibility that *S. sapeloensis* does not express pathogenicity-associated CAZYmes during saprotrophy, we mined the transcriptome of *S. sapeloensis* for CAZyme gene families induced during infection in *Phytophthora infestans,* a peronosporalean plant pathogen. Of the wide variety of *P. infestans* CAZyme family-encoding genes that are responsive during infection of potato tubers, many, such as Carbohydrate esterase 8 (CE8), GH28, GH53, GH78 and PL2 (Ah-Fong, Shrivastava and Judelson 2017), are not expressed in *S. sapeloensis* (Dataset S4a, Supporting Information). Hence, in agreement with our GO-term enrichment analyses (Fig. 3A), these data suggest that the saprotroph may use different CAZyme families compared to plant pathogens.

To identify CAZyme families with specific relevance to saprotrophy, we determined which genes were upregulated during litter colonization. *Salisapilia sapeloensis* expressed mainly GHs, glycosyltransferases (GTs) and CEs (Fig. 3B and C). In agreement with litter degradation, 19.0% of GH-encoding and 27.8% of CE-encoding genes showed a $TPM_{TMM\text{-}normalized} > 100$ in MGLi (Fig. 4A; Dataset S4c, Supporting Information). Overall, 37 CAZyme-encoding genes were upregulated in litter-associated growth vs growth on CMMm (FDR ≤ 0.05). We analyzed the relative expression of four of those genes via qRT-PCR (*Salisap4165_c1_g3 (CE10), Salisap3316_c0_g4 (GH1), Salisap3873_c0_g7 (GH6)* and *Salisap1930_c0_g1 (GT57)*) using qRT-PCR (Fig. 4B). The upregulation in MGLi vs CMMm of all four genes was confirmed for three of them (*Salisap4165_c1_g3 (CE10), Salisap3316_c0_g4 (GH1)* and *Salisap3873_c0_g7 (GH6)*). *Salisap1930_c0_g1 (GT57)* was the only gene upregulated in MGLw vs CMMm in the transcriptomic data; this was likewise confirmed by qRT-PCR. One of the most highly expressed genes in the entire CAZyme dataset was *Salisap3873_c0_g7,* one of three *GH6* genes identified in our KEGG analyses. *Salisap3873_c0_g7* was not only upregulated in MGLi vs CMMm (FDR = 4.8×10^{-15}) but also MGLi vs MGLw (FDR = 1×10^{-5}; Fig. 4A and B; Dataset S1d, Supporting Information). Across all CAZyme categories, 7.6% (10 of 131 genes with a calculable $\log_2$(FC)) were specifically upregulated in MGLi vs MGLw; only 2.3% (3 of 131 genes) were downregulated (Fig. 4C). Together this amounts to 8.9% of the genes that are differentially regulated in MGLi vs MGLw. These data underpin the importance of differential regulation of CAZyme-encoding genes in oomycete saprotrophy, at least the system analyzed herein. Among these litter-specific upregulated genes in *S. sapeloensis* are two CEs, six GHs and all PLs. The latter may appear to be in contrast to our previous observation that PLs are underrepresented in the transcriptome of *S. sapeloensis*. However, expressing fewer PLs does not disagree with the possibility that those few PLs that are expressed may have a role in the degradation of litter. Taken together, these results suggest that particular PLs, GHs and CEs may play important roles in litter degradation.

Filamentous pathogens of plants with a hemibiotrophic and necrotrophic lifecycle likely grow saprotrophically in the plant tissue that they have just killed. For their saprophytic stages CAZymes will be important, too. Yet, their CAZyme-profile may differ from their early infection phases. Such analyses have been done for the fungus *Zymoseptoria* (Brunner *et al.* 2013); distinct CAZyme expression profiles were seen during its biotrophic, necrotrophic and saprotrophic infection phases. Likewise, oomycete plant pathogens, such as the peronosporalean pathogens *P. infestans* and *P. ultimum*, as well as the saprolegnian pathogen *A. euteiches*, have dynamically responding repertoires of CAZyme-encoding genes throughout their infection progress (Ah-Fong, Shrivastava and Judelson 2017; Gaulin *et al.* 2018). Whether the induction of specific CAZyme-encoding

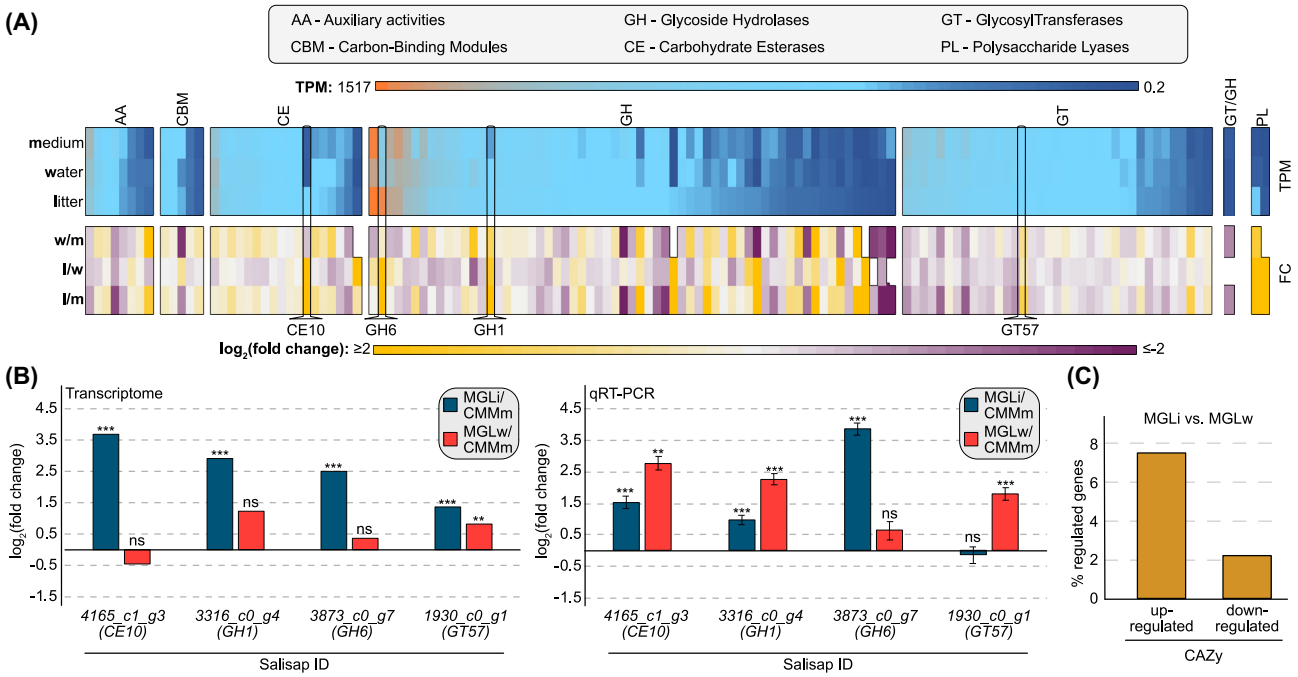


Figure 4. Expression of CAZyme-encoding genes in *Salisapilia sapeloensis*. **(A)** Expression of CAZyme-encoding genes in *S. sapeloensis*. Top: heatmap of $TPM_{TMM\text{-}normalized}$ values. The heatmap is sorted according to the expression in litter from highest (orange) to lowest (dark blue) in each CAZyme category. Medium indicates the expression in CMMm, water in MGLw and litter in MGLi. Each vertical line corresponds to one gene. Bottom: heatmap of the $\log_2$(FC) in w/m (MGLw/CMMm), l/w (MGLi/MGLw) and l/m (MGLi/CMMm). The order of genes corresponds to the order of genes in the heatmap for $TPM_{TMM\text{-}normalized}$ values on top. Four genes, which were further tested using qRT-PCR and mentioned in the main text, are highlighted. **(B)** Average relative expression ($\log_2$[FC]) of four genes upregulated in MGLi and/or MGLw vs CMMm in the transcriptome (left) and tested using qRT-PCR (right). The reference gene for the qRT-PCR was *SsH2A*. The error bars indicate the standard error of the mean (SEM). Gene IDs of the tested genes are given below the bar graphs. Significant differences are given above the bars as asterisk: *FDR/P-value ≤ 0.05, **FDR/P-value ≤ 0.01, ***FDR/P-value ≤ 0.001. **(C)** Overall up- and downregulation of CAZyme-encoding genes in litter colonization (MGLi vs MGLw).

gene families is similar between (a) full saprotrophy and (b) a saprotrophic phase that occurs at the end of an infection cycle remains unknown.

To find additional CAZyme families relevant to the here studied saprotrophic interaction, we analyzed if and which CAZyme-encoding gene families show enriched expression in *S. sapeloensis*. Given the similar distribution pattern of CAZyme categories in the transcriptome and oomycete genomes (Fig. 3B), we reasoned that the relative amount of any CAZyme family specifically required for saprotrophy would be enhanced compared with the average relative representation in oomycete genomes. To that end we calculated the percentage of CAZyme-encoding genes in the transcriptome of *S. sapeloensis* and in the genomes of the other oomycetes (Fig. 3C; Dataset S4b, Supporting Information). We defined an overrepresentation of a CAZyme-encoding family in the transcriptome as an at least 2-fold increase in relative abundance of this family compared with its average abundance in the oomycete genomes (averaged over 22 genomes (from oomycetes with different ecological and phylogenetic histories)). Of the 191 CAZyme families detected, eight were more abundantly expressed than expected (Fig. 3C). These eight families correspond to 14 GH-, five GT- and one mixed GH/GT-encoding gene(s) (Fig. 3C). Of these, seven were upregulated in association with litter (MGLi and MGLw vs CMMm; FDR ≤ 0.05; Dataset S1d, Supporting Information). Six induced genes corresponded to the GH1 family (*Salisap3316_c0_g4*, *Salisap3316_c0_g3*, *Salisap3316_c0_g1*, *Salisap3477_c0_g1*, *Salisap3477_c0_g3* and *Salisap3920_c3_g1*) and one to GT58 (*Salisap 2235_c0_g1*). Only one gene, a *GH1* (*Salisap3316_c0_g4*), is expressed specific to the colonization of litter, showing a 3.2-fold upregulation (FDR = 0.026; Dataset S1b, Supporting Information). The GH1 family is generally involved in cellulose degradation (Coradetti *et al.* 2012), and as such the family was not only relatively abundant in the transcriptome of *S. sapeloensis* but also in many genomes of hemibiotrophic and necrotrophic plant pathogens (Fig. 3C). However, its overrepresentation in the transcriptome of *S. sapeloensis*, together with many GH1-encoding genes being induced in MGLi vs CMMm, indicates that this family—and especially *Salisap3316_c0_g4*—is an interesting candidate among the CAZymes for being important for the saprotrophic lifestyle of *S. sapeloensis*.

In summary, a large fraction of *S. sapeloensis*' differentially regulated genes (i.e. FDR ≤ 0.05) in MGLi vs MGLw encode CAZymes. Many of these were highly responsive in litter-associated growth relative to growth on CMM. This can be attributed to the changes in carbon sources used in MGLi, MGLw and CMMm. Likewise, CAZyme-encoding genes of fungi contribute a major fraction to the differential transcriptome after changes in carbon sources (Eastwood *et al.* 2011; Benz *et al.* 2014). This indicates that fungi and oomycetes may have co-opted their molecular biology in similar ways to similar lifestyles. The evolution of saprotrophy may hence entail the convergent recruitment of similar degradation mechanisms in these two groups of organisms.

Salisapilia sapeloensis expresses few candidates for virulence-associated factors

To be able to infect living plants, pathogens counteract the plant immune system by secreting proteins, called effectors, into the host cells or their apoplast (Whisson *et al.* 2007; Wang *et al.* 2017). *Salisapilia sapeloensis* is classified as a saprotroph and our

above-described analyses indicate that with regard to CAZyme-encoding genes, it does not express pathogen-associated families. However, an organism may act as a saprotroph in some environments and as a pathogen in others. We, thus, screened our *S. sapeloensis* data for candidates of virulence-associated genes. Using a BLASTp approach with sequences from the peronosporalean plant pathogen *P. infestans* as queries, we investigated whether the transcriptome of *S. sapeloensis* encodes candidates for (i) protease inhibitor-like-encoding genes (Epi and EpiC), (ii) elicitin-like sequences, (iii) members of the Nep1-like (NLP) family and (iv) cellulose binding elicitor lectin (CBEL)-encoding genes. All of these families have been shown to induce disease symptoms and/or are required for successful plant infection (Gaulin *et al.* 2002, 2006; Khatib *et al.* 2004; Kanneganti *et al.* 2006; Tian *et al.* 2007; Kaschani *et al.* 2010; Jiang and Tyler 2012). We found one EpiC-like sequence, eight elicitin-like, 10 CBEL-like, but no Epi- or NPP-like sequences in *S. sapeloensis*.

Protein structure analyses confirmed the presence of all family-specific domains in the EpiC-like, four of the eight elicitin-like and three of the 10 CBEL-like candidates (Fig. 5A; Dataset S5a, Supporting Information). Four of the 10 CBEL-like sequences contained only a PAN/Apple domain and lacked a cellulose-binding domain, suggesting that they are either involved in protein-protein or protein–carbohydrate interactions (Dataset S5a, Supporting Information). Many sequences also encoded an SP, supporting secretion (Dataset S5a, Supporting Information).

Expression analyses of these putative virulence-associated factors showed that *SsEpiC-like*, three *Elicitin-like* (*Salisap2907_c0_g2_i1* (*SsElicitin-like1*), *Salisap4220_c11_g2_i3* (*SsElicitin-like2*) and *Salisap4220_c12_g1_i1* (*SsElicitin-like3*)) and two *CBEL* (*Salisap3205_c0_g1_i1* and *Salisap3882_c0_g5_i1*) transcripts were expressed at a reliable level ($TPM_{TMM\text{-}normalized} \geq 1$ in at least one treatment; Dataset S5b, Supporting Information). All transcripts expressed at a reliable level were either not differentially expressed (*SsEpiC-like*, *SsCBEL2*, *SsElicitin-like1* and 3; Fig. 5B) or downregulated (*SsElicitin-like2*: $\log_2(FC) = -1.3$, FDR $= 3.5 \times 10^{-7}$; *SsCBEL1*: $\log_2(FC) = -1.7$, FDR $= 1.1 \times 10^{-17}$) in MGLi vs CMMm. All three *Elicitin-like* transcripts (*SsElicitin-like1*: $\log_2(FC) = 1.2$, FDR $= 1 \times 10^{-5}$; *SsElicitin-like2*: $\log_2(FC) = 0.9$, FDR $= 0.003$; *SsElicitin-like3*: $\log_2(FC) = 1.3$, FDR $= 0.028$) were induced in MGLw compared with CMMm (Fig. 5B). None were significantly induced in MGLi vs MGLw (Fig. 5B).

We also investigated the presence of candidates for cytoplasmic effectors. Crinklers (CRNs), RxLRs and small secreted peptides (SSPs) are present in several oomycetes (Schornack *et al.* 2010; McGowan and Fitzpatrick 2017; Gaulin *et al.* 2018). Because only few orthologous effectors exist between the different oomycetes (Win *et al.* 2007; Haas *et al.* 2009), we used a heuristic and HMM search to identify initial RxLR candidates and a heuristic search to identify CRN candidates. For SSPs, we used EffectorP2.0 (Sperschneider *et al.* 2018). Effectors are secreted from the pathogen cell, hence we required the candidate effector genes to encode an SP and no TM domain.

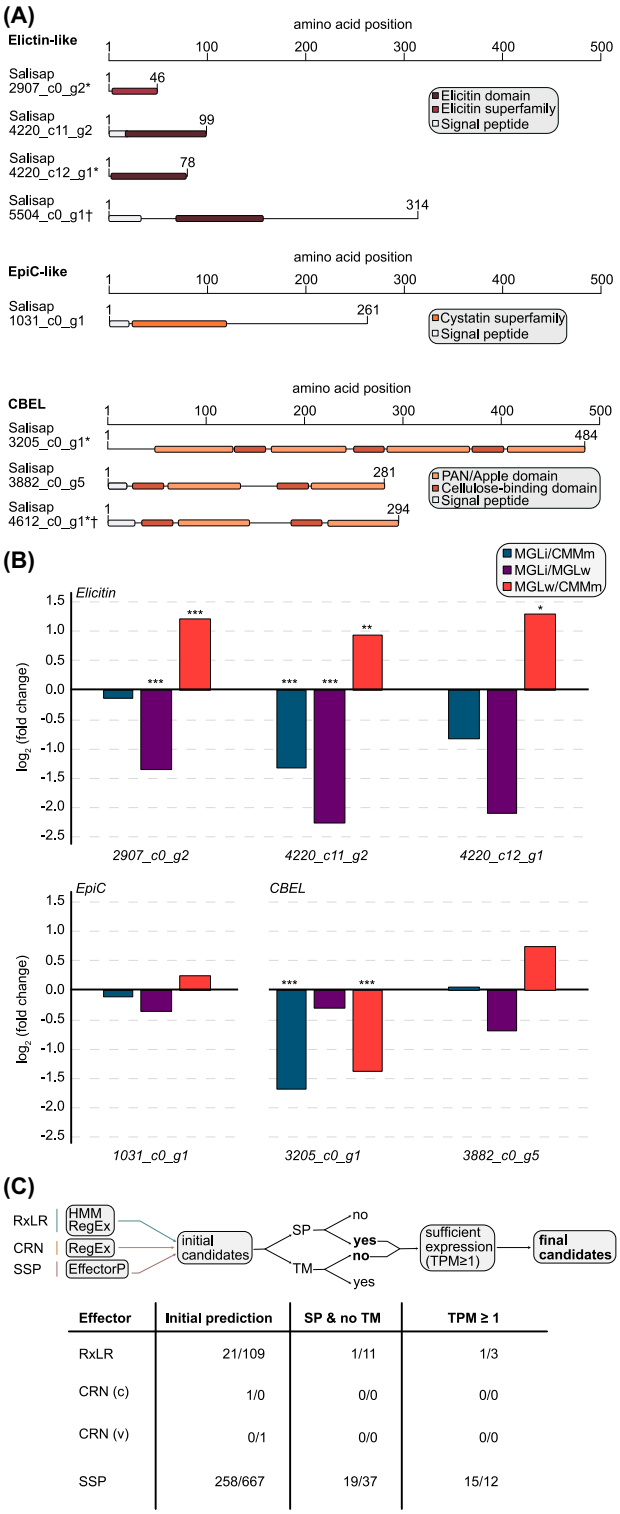
Using this approach, we identified candidates for four RxLR and 27 SSP but no CRN across the oomycete-affiliated and orphan dataset in *S. sapeloensis* (Fig. 5C). One putative RxLR was also annotated as an SSP candidate, resulting in 30 effector candidates whose expression was then analyzed. In total, 10/30 effector candidates showed significant upregulation (FDR $\leq$ 0.05) in association with litter (MGLi vs MGLw and/or vs CMMm) (Fig. 5D; Dataset S5c–e, Supporting Information). *Salisapilia sapeloensis* thus expresses several candidates for cytoplasmic effector-encoding genes.

While SSPs have recently been reported as a novel class of cytoplasmic effectors in oomycetes (Gaulin *et al.* 2018), the abundance of putative RxLRs and CRNs has been explored in 37 oomycetes (McGowan and Fitzpatrick 2017). The latter two types of effectors show substantial variation across oomycetes (McGowan and Fitzpatrick 2017): predicted by a heuristic search, their numbers range from one candidate in the genus *Saprolegnia* to 167 in *P. infestans* for CRNs and one candidate in *Saprolegnia parasitica* to 359 candidates in *P. infestans* for RxLRs. Functional analyses are restricted to a few genera. To put our data on putative RxLRs of *S. sapeloensis* into perspective, we predicted RxLR candidates in eight oomycete genomes from the Peronosporales and Albugonales; additionally, we included the diatom *Phaeodactylum tricornutum* and the brown alga *Ectocarpus siliculosus* as representatives of non-pathogenic stramenopiles. The numbers predicted by our heuristic search are in agreement with those predicted by McGowan and Fitzpatrick (2017). We find that the highest numbers of putative RxLRs are in species of the polyphyletic genus *Phytophthora* followed by the closely related biotrophs *Hyaloperonospora arabidopsidis* and *Plasmopara halstedii* (Table S2, Supporting Information). *P. ultimum* and the two species of the genus *Albugo* had only few candidates (ranging from two to six) and the two non-pathogenic stramenopiles showed 10 predicted candidates each (Table S2, Supporting Information). This means that whether stemming from a pathogen or not, all genomic data investigated here harbored at least some candidate RxLRs.

Is the number of putative RxLR-encoding genes in a genome a good predictor for the function of the respective gene products to act as virulence factors? Likely not. A recent functional analysis found that RxLRs in different *Pythium* species induce necrosis or defense responses (Ai *et al.* 2020). The analyzed *Pythium* species (Ai *et al.* 2020) returned only a few predicted RxLRs (Table S2, Supporting Information; McGowan and Fitzpatrick 2017). In turn, this does also not mean that the few candidates for putative RxLRs found in the datasets analyzed here play any role in pathogenicity. With regard to the induced effector candidates of *S. sapeloensis* this may mean that their induction is due to the fact that *S. sapeloensis* senses leaf tissue or these proteins may have other saprotrophy-related functions. Without functional analyses one simply cannot tell.

Gene expression profile of *Salisapilia sapeloensis* during litter decomposition differs from that of two plant pathogenic relatives during infection

In addition to what gene families are used in a saprotroph, it is worth asking whether gene families present in saprotrophs and pathogens are used differently in the distinct lifestyles during plant colonization. While this can ultimately only be answered on the protein level, studying expression profiles of gene families is an important first step in exploring this question. We therefore compared the global gene expression data from *S. sapeloensis* in MGLi vs CMMm with published transcriptome-derived expression data from two peronosporalean plant pathogens: (i) *Phytophthora infestans* and (ii) *Pythium ultimum* infecting potato tubers vs growth on RMM (RMMm) (Ah-Fong, Shrivastava and Judelson 2017) (Fig. 6). *Salisapilia sapeloensis* is ideal for this comparison because of its phylogenetic position either basal to *Phytophthora* and *Pythium* (Hulvey *et al.* 2010) or between *P. infestans* and *P. ultimum* (Bennett and Thines 2019) (Fig. 1A). Moreover, the two pathogens have different modes of pathogenic lifestyles: *P. infestans* is a hemibiotrophic pathogen.



Effector	Initial prediction	SP & no TM	TPM ≥ 1
RxLR	21/109	1/11	1/3
CRN (c)	1/0	0/0	0/0
CRN (v)	0/1	0/0	0/0
SSP	258/667	19/37	15/12

Figure 5. Candidate repertoire of virulence factors and pathogenicity-associated genes and their expression profile in *Salisapilia sapeloensis*. **(A)** Domain structure of *in silico* translated pathogenicity-associated genes from the protein families Elicitin, EpiC and CBEL. **(B)** Relative expression profile of candidate genes from the Elicitin, EpiC and CBEL gene families that have an average $TPM_{TMM\text{-}normalized} \geq 1$ in at least one treatment. Expression is compared between MGLi and CMMm, MGLi and MGLw, and MGLw and CMMm. Significant up-/downregulation is indicated by *FDR ≤ 0.05, **FDR ≤ 0.01 and ***FDR ≤ 0.001. **(C)** On top: the process used to identify effector candidates using HMMR (HMM) and a heuristic search with regular expression (RegEx) or EffectorP2.0 (EffectorP). At the bottom: the number of identified effector candidates in the oomycete-affiliated (left of '/') and orphan gene (right of '/') datasets. On the left the type of effector candidate is given: RxLR, CRN with the canonical CRN motif (c), CRN with an alternative CRN motif (v) and SPP candidates. The table shows the number of effector candidates identified in the 'initial prediction' (i.e. compiled results of the HMM and heuristic searches), after filtering for the signal peptides (SP) as well as requiring the absence of transmembrane domains (TM), and after applying an expression cutoff of $TPM_{TMM\text{-}normalized} \geq 1$ in at least one treatment. **(D)** Relative expression [log_2(FC)] of effector candidates detected in *S. sapeloensis*. Significant up-/downregulation is indicated by *FDR ≤ 0.05, **FDR ≤ 0.01 and ***FDR ≤ 0.001.

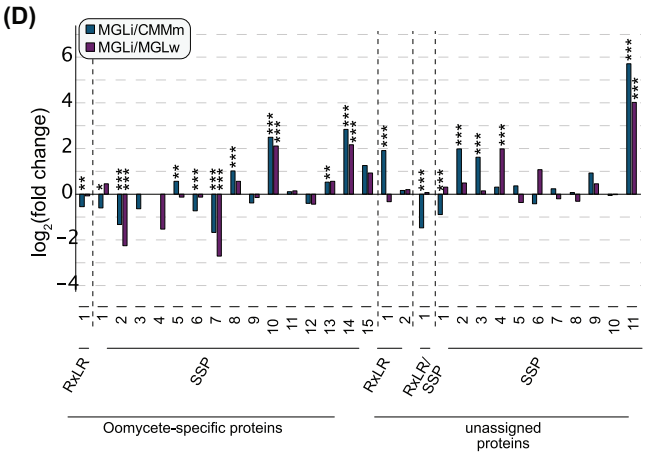


Figure 5. – *continued*

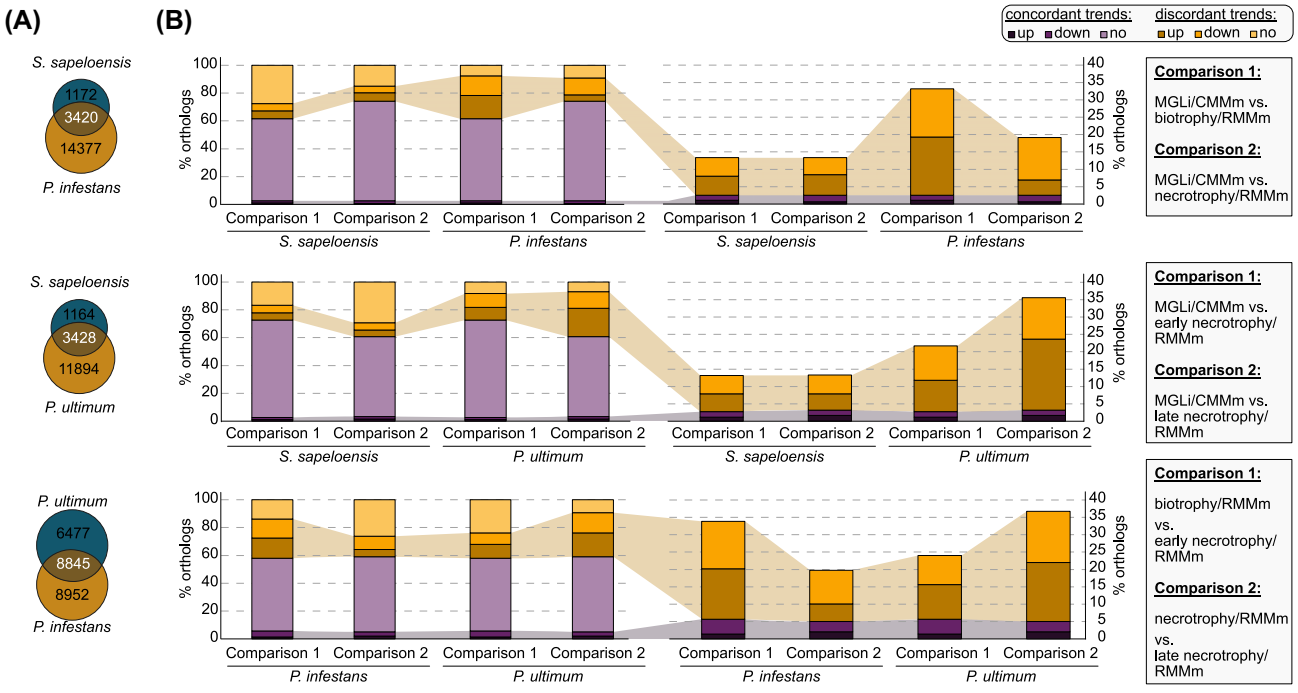


Figure 6. Comparative transcriptomics of *Salisapilia sapeloensis* with two oomycete plant pathogens. **(A)** The Venn diagrams show the shared RBBHs between *P. infestans*, *P. ultimum* and *S. sapeloensis*. **(B)** The bar graphs give the percentage concordantly (purple) and discordantly (orange) regulatory trends of RBBHs between *S. sapeloensis* and *P. infestans* (top), *S. sapeloensis* and *P. ultimum* (middle) and *P. infestans* and *P. ultimum* (bottom). We compared regulatory trends of RBBHs in *S. sapeloensis* MGLi vs CMMm with early infection vs RMMm (Comparison 1) or with late infection vs RMMm (Comparison 2). In the comparisons between the two pathogens, Comparison 1 corresponds to early infection vs RMMm of *P. infestans* with early infection vs RMMm of *P. ultimum*. Comparison 2 corresponds to the comparisons between the late infection phases and RMMm. 'Up' indicates that a gene is induced ($\log_2$[FC] $\geq$ 1); 'down' indicates that a gene is reduced ($\log_2$[FC] ≤ -1) and 'no' indicates that a gene shows no changes ($-1 >$ $\log_2$[FC] < 1). Right: overview of all RBBHs. Left: overview of RBBHs that are induced or reduced.

It requires a living host in its early infection stage (biotrophic phase) and later induces host cell death (necrotrophic phase). In contrast, *P. ultimum* is a necrotrophic pathogen. It immediately kills its host after infection to utilize the degradants.

To gain a first glimpse of how similar gene expression is between *S. sapeloensis* and the two peronosporalean pathogens when comparing relative expression data from infection vs litter colonization, we first used a RBBH approach between *S. sapeloensis*, *P. infestans* and *P. ultimum* (Fig. 6A). Second, we calculated the $\log_2$(FC) between MGLi and CMMm for *S. sapeloensis* and between infection and RMMm for the two peronosporalean plant pathogens. Using these data, we asked whether the expression of a gene in the three species was higher (up; $\log_2$(FC) $\geq$ 1), lower (down; $\log_2$(FC) ≤ -1) or similar (no change; $-1 <$ $\log_2$(FC) < 1) upon encountering plant tissue compared with CMMm or RMMm. Next, we calculated how many of the RBBHs showed concordant trends in expression: An RBBH was counted as showing a concordant trend in two species if expression was higher (up) in both species, lower (down) in both species or similar (no change) in both species (Fig. 6B) upon treatment. We analyzed concordant trends in expression between litter colonization vs CMMm and an early infection phase vs RMMm and between litter colonization vs CMMm and a late infection phase vs RMMm. This was done to account for the fact that *P.*

infestans is a hemibiotrophic pathogen that changes from biotrophy to necrotrophy during infection, which is also accompanied by a switch in gene expression (Ah-Fong, Shrivastava and Judelson 2017).

The majority of RBBHs in *S. sapeloensis* and the two peronosporalean plant pathogens had similar patterns of gene expression (62.0 ± 5.9%; Fig. 6B; Dataset S6a, Supporting Information). Hereby, the early necrotrophic phase of the two pathogens (i.e. in the hemibiotroph *P. infestans* the late infection stage; in the necrotroph *P. ultimum* the early infection stage) showed more concordant trends to saprotrophy in *S. sapeloensis* than the biotrophic phase (early infection stage of *P. infestans*) and late necrotrophic phase (late infection stage of *P. ultimum*) (Fig. 6B). These similarities are not surprising, but align with the fact that in saprotrophy and necrotrophy the oomycete is eventually confronted with dead tissue. It is thus only logical that the transcriptional responses to these cues are partially overlapping. In contrast, the transition from biotrophy to necrotrophy requires a switch from a transcriptional program aimed at keeping a host alive to one designed to kill a host. Indeed, hemibiotrophic pathogens show strong transcriptional shifts between their biotrophic and necrotrophic phases (Jupe *et al.* 2013; Ah-Fong, Shrivastava and Judelson 2017; Evangelisti *et al.* 2017). Hence, a strong transcriptional difference between the here-studied saprotrophic interaction, in which *S. sapeloensis* feeds on dead plant tissue, compared with the biotrophic phase in the hemibiotroph *P. infestans*, where its physiology is optimized to keep the host alive, is only expected.

The main constituent of the RBBHs with concordant regulatory trends between the two plant pathogens and the saprotroph *S. sapeloensis* were those showing no change (Fig. 6B). Only 0.69–1.85% of the concordantly regulated RBBHs were responsive (up- or downregulated) upon colonization or infection. Hence a large fraction of the litter-responsive transcriptome of *S. sapeloensis* has distinct regulatory trends compared with the two pathogens. This means that homologous genes that are induced during litter degradation in this saprotroph are not induced during infection in these two plant pathogens.

Conducting the comparison between the hemibiotroph *P. infestans* and the necrotroph *P. ultimum* we found that the relative amount of concordant regulatory trends between the two pathogens was two times higher than what was observed for the comparison between *S. sapeloensis* and either pathogen (Fig. 6B). This indicates that the molecular toolkit required by the two different pathogens for infecting a living plant is more similar to each other than to that required for the colonization of dead plant material by *S. sapeloensis*.

To gain insight into the responsiveness of genes with discordant trends, i.e. genes that are uniquely-regulated in the different organisms, we first identified such genes and second asked whether they are induced, reduced or showed no change in expression between infection or colonization vs medium (Dataset S6a–c, Supporting Information). Of all RBBHs identified in *P. infestans* and *S. sapeloensis*, as well as *P. ultimum* and *S. sapeloensis*, 32.7 ± 6.2% were uniquely regulated across all four comparisons. This corresponds to 938 ± 178 RBBHs. The two pathogens had more uniquely regulated RBBHs that were induced or reduced compared with the saprotroph *S. sapeloensis*: 16.8 ± 3.6% (152 ± 11 RBBHs) of all uniquely-regulated RBBHs were induced or reduced in the saprotroph colonizing litter (MGLi) vs CMMm, while 37.1 ± 10.2% (354 ± 131 genes) were induced or reduced in the pathogens infecting their hosts vs RMMm (Dataset S6a, Supporting Information). On balance, our data suggest that *S. sapeloensis*' gene expression profile differs from that of the two oomycete pathogens.

Gene expression of transcription factor candidates is associated with saprotrophy in *Salisapilia sapeloensis*

Differences in gene expression patterns between the saprotroph *S. sapeloensis* and the two peronosporalean plant pathogens suggest the presence of different gene regulatory networks (GRNs), or that the *S. sapeloensis* GRNs respond differently to the colonization of dead plant material than do homologous GRNs of *P. infestans* and *P. ultimum* when infecting living tubers. Indeed, we found a C3H TF that was specifically induced in both pathogens during an early infection phase vs RMMm, while its RBHH (*Salisap4203_c0_g1*) was not differentially expressed in *S. sapeloensis* in MGLi vs CMMm (Dataset S1d, Supporting Information). TFs are upstream in a hierarchical GRN and thus slight changes in their expression can result in (i) strong changes of the expression of their target genes or (ii) activation of entirely different GRNs; the latter can be mediated by TF–TF interactions where binding specificity or affinity is defined by the combination of TFs (Bemer *et al.* 2017). If the expression of a TF is no longer co-opted, other TF–TF interactions may occur, which in turn will cause different downstream expression patterns. Hence, we next focused on a comparison of TF expression in oomycetes with different lifestyles.

Phytophthora infestans and *Pythium ultimum* are of course only two of many pathogens and are possibly more closely related to each other than to *S. sapeloensis*. Additionally, the transcriptomic data we used to estimate the concordant regulatory trends were generated from infections of starch-rich tubers, not leaf. In addition to offering a different substrate, the process of tuber infection presents a special case of oomycete plant pathogenicity. For a more unbiased view, we extended our analyses to eight oomycete pathogens, chosen because of their phylogenetic position (Albugonales and Peronosporales) and the fact that both genomic and transcriptomic data from late infection stages are available. These eight pathogens cover three different pathogenic lifestyles (four biotrophs, three hemibiotrophs, one necrotroph), five different hosts (*Arabidopsis*, potato, tomato, soybean, sunflower) and four different tissue types (roots, leaves, tubers, hypocotyl; Fig. 7; Dataset S6d and e, Supporting Information). Using such diverse ecological conditions will introduce 'noise', but by asking how all pathogens together differ from *S. sapeloensis*, we made use of these data to strengthen the robustness of our analyses—that is, if all pathogens respond similarly despite their different ecologies and if this response is different to *S. sapeloensis,* we have identified a candidate related to the saprotrophic biology of *S. sapeolensis*.

The *S. sapeloensis* transcriptome featured 47 candidates for TF-encoding genes, 40 of which had a $TPM_{TMM\text{-}normalized} \geq 1$. These 40 candidates correspond to 14 TF families. The genomes of the eight pathogen species encoded on average 35 ± 3 TF candidates that had RBBHs to the 40 TF candidates of *S. sapeloensis*. We detected the most RBBHs in the necrotroph *P. ultimum* (38) and the least in the biotroph *Albugo laibachii* (30). Two TF candidate genes of *S. sapeloensis,* both belonging to the Myb-related family, had no RBBHs in any of the eight plant pathogens (Fig. 7A). Five TF candidate genes were, when their RBBHs were present in the pathogens, more highly expressed in *S. sapeloensis* than in the pathogens ($log_2(FC) \geq 1$; Fig. 7A; Dataset S6d, Supporting Information). These included a bZIP-, two Myb-related-, a NF-YB- and a Nin-like candidate. In contrast, no

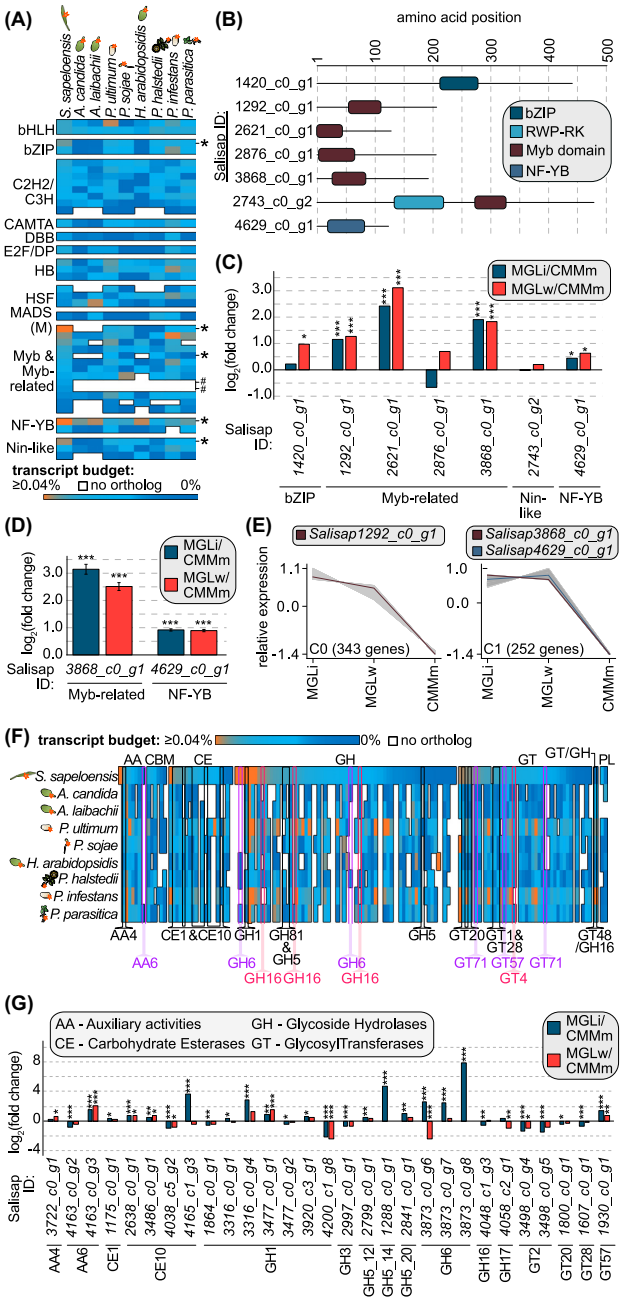


Figure 7. Transcript budget analyses of candidates for transcription factor- and CAZyme-encoding genes across plant colonizing oomycetes. **(A)** Transcript budget (given in %TPM of the total TPM) invested into candidates for TF-encoding genes from dark blue (0% transcript investment) to orange (≥0.04% transcript investment). Every line indicates a candidate for a TF-encoding gene in *S. sapeloensis* and its RBBHs in eight plant pathogenic oomycetes. White spaces indicate that no RBBH was identified. The heatmap is sorted according to the transcript budget in *S. sapeloensis* from high to low and per TF category. Asterisks show those candidates for TF-encoding candidate gene with an increased transcript budget ($\log_2$[FC] ≥ 1) in *S. sapeloensis* vs all oomycetes with an RBBH, and the hashtag highlights those TF candidates with no RBBH in the other eight oomycetes. The tissue that was colonized/infected by the oomycete is indicated with the symbol to the right of their names. **(B)** Domain structure of the *in silico* translated seven TF-encoding gene candidates with increased transcript budget or no RBBH in the plant pathogens. **(C)** $\log_2$(FC) of the seven candidates for TF-encoding genes in MGLi (blue) or MGLw (red) vs CMMm in the transcriptome data. Significant differences are indicated as *FDR ≤ 0.05, **FDR ≤ 0.01, ***FDR ≤ 0.001. **(D)** qRT-PCR of a Myb-related- and the NF-YB-encoding gene from the seven candidates for TF-encoding genes. The average $\log_2$(FC) for MGLi (blue) or MGLw (red) vs CMMm is shown as a bar diagram. Error bars indicate the SEM. The reference gene was *SsH2A*. Significant differences are given as ***P-value ≤ 0.001. **(E)** The clusters C0 and C1 include three of the five TF-encoding genes with an upregulation in MGLi or MGLw vs CMMm. Candidates for Myb-related-encoding genes are indicated with a dark red line and the NF-YB-encoding gene is indicated with a dark blue line. **(F)** Comparison of the transcript budget invested in CAZyme-encoding genes. The heatmap shows the genes from *S. sapeloensis* and their respective RBBHs in eight plant pathogens that infect five different hosts and four different tissues (symbolized by the icons to the left of the species name). Heatmaps are sorted within each CAZyme category and from high transcript investment (orange) to low (dark blue) in *S. sapeloensis*. Each line represents one gene in the saprotroph and its RBBH in the pathogens. White indicates that no RBBH was identified in a given pathogen. Highlighted CAZyme-encoding genes have a $\log_2$(FC) ≥ 1 difference in %TPM between either *S. sapeloensis* compared with all pathogens or vice versa. **(G)** $\log_2$(FC) of gene expression of all CAZyme-encoding genes with increased transcript budget or no RBBH in other oomycetes. Only genes with an expression on a reliable level in all conditions ($TPM_{TMM\text{-}normalized}$ ≥ 1) are shown. Significant differences are indicated as *FDR ≤ 0.05, **FDR ≤ 0.01, ***FDR ≤ 0.001.

TF-candidate had a consistently higher expression across all pathogens compared with *S. sapeloensis*. Phylogenetic analyses confirmed the orthologous relationships of the TF candidates, which had either RBBHs in the pathogens with reduced expression or no orthologs at all (Figure S6, Supporting Information). Only a few exceptions occurred (Figure S6A, Supporting Information). Overall, our RBBH approach had an accuracy of on average 91% for the TF candidates identified herein. The few additional co-orthologs identified with our phylogenetic approach also had a much lower transcript investment compared with the corresponding candidates for TF-encoding gene in *S. sapeloensis* ($\log_2$(FC) $\geq$ 1 *S. sapeloensis* vs pathogens), the only exception being *PYU1_T006437*.

Considering the large number of genes that are responsive to either colonization or infection and exhibit discordant regulatory trends, our TF candidate data warrant attention. They suggest that GRNs are differently wired for saprotrophy of *S. sapeloensis* and plant pathogenicity in oomycetes. Either similar components of the networks have to interact differently (if orthologs exist) because of the distinct expression levels or new components can be added or can replace existing factors in the GRNs. Both situations will alter downstream expression patterns and are in fact not mutually exclusive. Our data on transcript investment in candidates for TF-encoding genes across several pathogenic oomycetes with different lifestyles, hosts and tissue-specialization and *S. sapeloensis* hence represent additional support for the usage of slightly altered GRNs by the saprotroph and the plant pathogens. The data further yield seven TF candidates that we analyzed in more depth. Domain predictions using InterPro (Fig. 7B) identified one Myb domain in the four Myb-related TF candidates, which is typical for Myb-related TFs (Xiang and Judelson 2010; Wilhelmsson *et al.* 2017). Likewise, we confirmed the presence of the bZIP domain in the bZIP candidate, the NF-YB domain in the NF-YB candidate and one RWP-RK and Myb domain in the Nin-like candidate, typical for these TFs (Schausser *et al.* 1999; Wilhelmsson *et al.* 2017). InterPro further predicted DNA binding sites in all seven candidates.

Higher transcript investment does not necessarily mean that those genes are also upregulated during litter colonization. We thus investigated their expression in *S. sapeloensis*. Of the seven TF candidates, five were upregulated in either MGLi or MGLw vs CMMm (FDR $\leq$ 0.05). The other two showed no differential expression. Of the five candidate TF-encoding genes, the bZIP-encoding gene was only upregulated in MGLw vs CMMm, while three candidates for Myb-related-encoding genes and the candidate for a NF-YB-encoding gene were upregulated in both MGLi and MGLw vs CMMm (Fig. 7C). This included the two candidates for Myb-related-encoding genes that had no orthologs in the plant pathogens (*Salisap1292_c0_g1* and *Salisap2621_c0_g1*). We tested two TF-encoding genes (*Salisap3868_c0_g1* and *Salisap4629_c0_g1*) using qRT-PCR (Fig. 7D); both were upregulated in MGLi and MGLw compared with CMMm (*P*-value $<$ 0.001), confirming the transcriptomic data. Overall, these data suggest that the majority of the newly identified TF candidates are also important in litter colonization and/or degradation.

While only functional analyses can ultimately determine whether these candidates are true TFs and in which pathways they may influence, we sought to gain insight into the biological processes they might associate with. To do so, we analyzed co-expression patterns of the five TF candidates upregulated in MGLi and MGLw vs CMMm with other genes in our transcriptome. Co-expression analyses can help identify genes involved in the same biological processes, including potential regulatory relationships between TFs and downstream target genes (Persson *et al.* 2005; Aoki, Ogata and Shibata 2007; Hirai *et al.* 2007; Usadel *et al.* 2009; Fu and Xue 2010). In most cases, co-expression analyses consider positively co-expressed genes. The caveat of such analyses with regard to TF-target relationships is that it can only identify TFs that are positive regulators with immediate impact on the expression of the target genes. The idea is that in these cases the target gene has a similar gene expression pattern as its regulating TF. Despite this caveat, co-expression analyses have successfully identified regulatory relationships between TFs and target genes (Hirai *et al.* 2007; Fu and Xue 2010). Here, we used this approach to further investigate the gene expression neighborhood in which these TF candidates are expressed to further explore their putative role in litter degradation.

In total, the transcriptomic data were broken into 19 clusters of co-expressed genes (C0–C18). These 19 clusters contained 2165 of the 4592 oomycete-affiliated genes (Figure S7, Supporting Information); another 1434 genes were not co-expressed with a sufficient number of genes and were disregarded. Three candidates of the TF-encoding genes were co-expressed with other genes in the transcriptome, and two had unique expression profiles. The three genes included two Myb-related (*Salisap1292_c0_g1* and *Salisap3868_c0_g1*) and the NF-YB (*Salisap4629_c0_g1*) candidate genes (Fig 7E).

The candidate for the Myb-related-encoding gene *Salisap1292_c0_g1* clustered with 342 other genes in cluster C0, while the candidates for the Myb-related-encoding gene *Salisap3868_c0_g1* and the NF-YB-encoding gene *Salisap4629_c0_g1* are co-expressed with 250 other genes in cluster C1 (Fig. 7E). In agreement with our previous data, both clusters encompass genes that exhibit an elevated expression in MGLi and MGLw compared with CMMm (Fig. 7C–E). We visualized the annotation of all differentially expressed genes that are upregulated in MGLi vs CMMm (FDR $\leq$ 0.05) for C0 and C1 (excluding all hypothetical and uncharacterized proteins; Figure S8, Supporting Information). Several distinct functional groups were represented in C0 and C1, including putative CAZyme- and transporter-encoding genes. This corroborates our previous analyses, further indicating an involvement of these TF candidates during litter colonization.

Comparative transcriptomics suggests a putative link between expression of CAZyme- and transcription factor-encoding genes

Our co-expression clustering results primed us to explore the link between CAZyme and TF candidates in more detail. In total, 72 of the 149 CAZyme-encoding genes are present in one of the 19 clusters. Of those, 20 (27.8%) are found in either C0 or C1 and are hence associated with both litter and litter-associated growth of *S. sapeloensis*. In both clusters are CAZymes from the families AA, CBM, CE, GH and GT, but not PL. In agreement with their clustering, most of the genes in C0 (6 of 7) and C1 (7 of 13) were differentially upregulated in MGLi and/or MGLw vs CMMm (FDR $\leq$ 0.05).

We next asked whether CAZyme-encoding genes show a lifestyle-associated transcript investment (Fig. 7F). We first searched for RBBHs of the 149 CAZyme-encoding genes from *S. sapeloensis* in the eight genomes of the plant pathogens from the Peronosporales and Albugonales. For 24 (16.1%) we found no RBBH across these plant pathogens, suggesting that they are either undetectable by this approach due to an extremely high-sequence divergence or are lineage-specific acquisitions

in *S. sapeloensis*. These putatively lineage-specific genes were distributed across most CAZyme-categories (except GH/GT and PL). In agreement with the GO-term analyses for genes unique to *S. sapeloensis* (Fig. 3A), the majority of these genes (16 of 24) encoded GHs (Fig. 7F; Dataset S6e, Supporting Information), underscoring the unique CAZyme-repertoire of *S. sapeloensis*. In five of the seven categories (AA, CE, GH, GT and GH/GT) we found CAZyme-encoding genes in which either *S. sapeloensis* or all pathogens (independent of their ecology) invest a much higher fraction of the transcript budget (log2 (FC) $\geq$ 1; Fig. 7F; Dataset S6e, Supporting Information). This included 26 genes (belonging to AA, CE, GH, GT) with a higher transcript investment in *S. sapeloensis* and 1 (belonging to GH/GT) with a higher transcript investment in all 8 plant pathogens. Of the 50 genes that had higher expression levels in *S. sapeloensis* or no RBBH in any of the pathogens, 18 were upregulated in MGLi and/or MGLw vs CMMm (FDR $<$ 0.05; Fig. 7G). Taken together, these 18 genes may be relevant for saprotrophy in *S. sapeloensis* and, if present, may be used differently in oomycete pathogens.

Finally, we related the transcript investment data from CAZyme- and TF-encoding gene candidates with the clustering analyses. This identified five CAZyme-encoding genes (i) into which *S. sapeloensis* invests much higher transcript levels compared with the eight plant pathogens from the Peronosporales and Albugonales and (ii) that are co-expressed with one of the Myb-related or NF-YB-encoding candidate genes. These five genes include one AA6-, two CE1- and two CE10-encoding genes (*AA6*: *Salisap4163_c0_g3*; *CE1*: *Salisap3051_c0_g2*, *Salisap3972_c0_g1*; *CE10*: *Salisap2638_c0_g1*, *Salisap3486_c0_g1*). Each cluster (C0 and C1) contains one CE1 and one CE10 gene. The AA6-encoding gene is clustered in C1. Three genes were expressed at a reliable level in all three conditions (*Salisap4163_c0_g3 (AA6)*, *Salisap2638_c0_g1 (CE10)* and *Salisap3486_c0_g1 (CE10)*). Of those, all were significantly upregulated in MGLi and MGLw vs CMMm (FDR $\leq$ 0.05; Fig. 7G). Taken together, this indicates that the AA6- and the four CE-encoding gene candidates may be relevant for tissue degradation in this case of saprotrophy. It also supports the involvement of the three TF candidates in litter colonization.

Comparisons between saprotrophic and pathogenic interactions can reveal vital clues about what, at the molecular and biochemical level, distinguishes a pathogen from a non-pathogen. We have established a first reference point for the differential regulation of gene expression that occurs during oomycete saprotrophy. Investigation of *S. sapeloensis* allowed us to contrast pathogenic and non-pathogenic gene expression patterns. Due to the distinct expression responses to infection and litter colonization, we hypothesize that *S. sapeloensis* has evolved slight alterations in its GRNs.

ACKNOWLEDGMENTS

We thank Emma Blanche for technical assistance.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

FUNDING

SdV thanks the Killam Trusts for an Izaak Walton Killam Postdoctoral Fellowship. JdV thanks the German Research Foundation (DFG) for a Research Fellowship (VR132/1-1) and a Return Grant (VR132/3-1), and the European Research Council (ERC)for funding under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 852725; ERC-StG 'TerreStriAL'). CHS and JMA acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (CHS: Discovery grant RGPIN/05754-2015; JMA: RGPIN-2014-05871).

Conflicts of interest. None declared.

REFERENCES

Abu-Jamous B, Kelly S. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. *Genome Biol* 2018;**19**:172.

Adhikari BN, Hamilton JP, Zerillo MM *et al.* Comparative genomics reveals insight into virulence strategies of plant pathogenic oomycetes. *PLoS One* 2013;**8**:e75072.

Ah-Fong AM, Judelson HS. New role for Cdc14 phosphatase: localization to basal bodies in the oomycete *Phytophthora* and its evolutionary coinheritance with eukaryotic flagella. *PLoS One* 2011;**6**:e16725.

Ah-Fong AMV, Kim KS, Judelson HS. RNA-seq of life stages of the oomycete *Phytophthora infestans* reveals dynamic changes in metabolic, signal transduction and pathogenesis genes and a major role for calcium signaling in development. *BMC Genomics* 2017;**18**:198.

Ah-Fong AMV, Shrivastava J, Judelson HS. Lifestyle, gene gain and loss, and transcriptional remodeling cause divergence in the transcriptomes of *Phytophthora infestans* and *Pythium ultimum* during potato tuber colonization. *BMC Genomics* 2017;**18**:764.

Ah-Fong AMV, Xiang Q, Judelson HS. Architecture of the sporulation-specific *Cdc14* promoter from the oomycete *Phytophthora infestans*. *Eukaryot Cell* 2007;**6**:2222–30.

Ai G, Yang K, Ye W *et al.* Prediction and characterization of RXLR effectors in *Pythium* species. *Mol Plant Microb Int* 2020;**33**:1046–58.

Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.

Amselem J, Cuomo CA, van Kan JAL *et al.* Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLos Genet* 2011;**7**:e1002230.

Aoki K, Ogata Y, Shibata D. Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant Cell Physiol* 2007;**48**:381–90.

Asai S, Rallapalli G, Piquerez SJM *et al.* Expression profiling during *Arabidopsis*/Downy mildew interaction reveals a highly-expressed effector that attenuates response to salicylic acid. *PLoS Pathog* 2014;**10**:e1004443.

Baetsen-Young A, Wai CM, VanBurenb R *et al.* *Fusarium virguliforme* transcriptional plasticity is revealed by host colonization of corn vs. soybean. *Plant Cell* 2020;**32**:336–51.

Barbi F, Kohler A, Barry K *et al.* Fungal ecological strategies reflected in gene transcription – a case study of two litter decomposers. *Environ Microbiol* 2020;**22**:1089–103.

Baxter L, Tripathy S, Ishaque N *et al.* Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 2010;**330**:1549–51.

Beakes GW, Glockling SL, Sekimoto S. The evolutionary phylogeny of the oomycete 'fungi'. *Protoplasma* 2012;**249**:3–19.

Beakes GW, Thines M. Hyphochytriomycota and oomycetes. In: Archibald JM, Simpson AGB, Slamovits CH (eds). *Handbook of the Protists*. Cham: Springer, 2017, 435–505.

Bemer M, van Dijk ADJ, Immink RGH *et al.* Cross-family transcription factor interactions: an additional layer of gene regulation. *Trends Plant Sci* 2017;**22**:66–80.

Bendtsen JD, Nielsen H, von Heijne G *et al.* Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;**340**:783–95.

Bennett RM, Thines M. Revisiting *Salisapiliaceae*. *Fungal Syst Evol* 2019;**3**:171–84.

Benz JP, Chau BH, Zheng D *et al.* A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptions. *Mol Microbiol* 2014;**91**:275–99.

Blanco FA, Judelson HS. A bZIP transcription factor from Phytophthora interacts with a protein kinase and is required for zoospore motility and plant infection. *Mol Microbiol* 2005;**56**:638–48.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.

Bowler C, Allen AE, Badger JH *et al.* The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 2008;**456**:239–44.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;**12**:59–60.

Charkowski AO. Opportunistic pathogens of terrestrial plants. In: Hurst CJ (ed). *The Rasputin Effect: When Commensals and Symbionts Become Parasitic*. Switzerland: Springer International Publishing, 2016, 147–68.

Chen S, Yao H, Han J *et al.* Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* 2010;**5**:e8613.

Cock M, Sterck L, Rouzé P *et al.* The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* 2010;**465**:617–21.

Colowick SP, Sutherland EW. Polysaccharide synthesis from glucose by means of purified enzymes. *J Biol Chem* 1942;**144**: 423–37.

Coradetti ST, Craig JP, Xiong Y *et al.* Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. *Proc Natl Acad Sci USA* 2012;**109**: 7397–402.

Cosgrove DJ. Assembly and enlargement of the primary cell wall in plants. *Annu Rev Cell Dev Biol* 1997;**13**:171–201.

Cvitanich C, Judelson HS. A gene expressed during sexual and asexual sporulation in *Phytophthora infestans* is a member of the Puf family of translational regulators. *Eukaryot Cell* 2003;**2**:465–73.

Derevnina L, Petre B, Kellner R *et al.* Emerging oomycete threats to plants and animals. *Phil Trans R Soc B* 2016;**371**:20150459.

de Vries S, von Dahlen JK, Uhlmann C *et al.* Signatures of selection and host-adapted gene expression of the *Phytophthora infestans* RNA silencing suppressor PSR2. *Mol Plant Pathol* 2017;**18**:110–24.

Diéguez-Uribeondo J, García MA, Cerenius L *et al.* Phylogenetic relationships among plant and animal parasites, and saprotrophs in *Aphanomyces* (Oomycetes). *Fungal Genet Biol* 2009;**46**:365–76.

Eastwood DC, Floudas D, Binder M *et al.* The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* 2011;**333**:762–5.

Edwards K, Johnstone C, Thompson C. A simple and rapid method for the preparation of plant genomic DNA for PCR. *Nucleic Acids Res* 1991;**19**:1349.

Evangelisti E, Gogleva A, Hainaux T *et al.* Time-resolved dual transcriptomics reveal early induced *Nicotiana benthamiana* root genes and conserved infection-promoting *Phytophthora palmivora* effectors. *BMC Biol* 2017;**15**:39.

Fazekas AJ, Burgess KS, Kesanakurti PR *et al.* Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* 2008;**3**:e2802.

Fu F-F, Xue H-W. Coexpression analysis identifies Rice Starch Regulator1, a rice AP2 /EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol* 2010;**154**:927–38.

Gabrielová A, Mencl K, Suchánek M *et al.* The oomycete *Pythium oligandrum* can suppress and kill the causative agents of Dermatophytes. *Mycopathologia* 2018;**183**:751–64.

Gamboa-Meléndez H, Huerta AI, Judelson HS. bZIP transcription factors in the oomycete *Phytophthora infestans* with novel DNA-binding domains are involved in defense against oxidative stress. *Eukaryot Cell* 2013;**12**:1403–12.

Gaulin E, Dramé N, Lafitte C *et al.* Cellulose binding domains of a *Phytophthora* cell wall protein are novel pathogen-associated molecular patterns. *Plant Cell* 2006;**18**:1766–77.

Gaulin E, Jauneau A, Villalba F *et al.* The CBEL glycoprotein of *Phytophthora parasitica* var. *nicotianae* is involved in cell wall deposition and adhesion to cellulosic substrates. *J Cell Sci* 2002;**115**:4565–75.

Gaulin E, Pel MJC, Camborde L *et al.* Genomics analysis of *Aphanomyces* spp. identifies a new class of oomycete effector associated with host adaptation. *BMC Biol* 2018; **16**:43.

Grabherr MG, Haas BJ, Yassour M *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011;**29**:644–52.

Guillemette T, van Peij N, Goosen T *et al.* Genomic analysis of secretion stress response in the enzyme-producing cell factory *Aspergillus niger*. *BMC Genomics* 2007;**8**:158.

Guttman DS, McHardy AC, Schulze-Lefert P. Microbial genome-enabled insights into plant–microorganism interactions. *Nat Rev Genet* 2014;**15**:797–813.

Haas BJ, Kamoun S, Zody MC *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 2009;**461**:393–8.

Herburger K, Holzinger A. Aniline blue and Calcofluor white staining of callose and cellulose in the streptophyte green algae *Zygnema* and *Klebsormidium*. *Bio Protoc* 2016;**6**:e1969.

Hirai MY, Sugiyama K, Sawada Y *et al.* Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc Natl Acad Sci USA* 2007;**104**:6478–83.

Huerta-Cepas J, Szklarczyk D, Forslund K *et al.* eggnog 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 2016;**44**:D286–93.

Hulvey J, Telle S, Nigrelli L *et al.* Salisapiliaceae – a new family of oomycetes from marsh grass litter of southeastern North America. *Persoonia* 2010;**25**:109–16.

Jansonius JN. Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struc Biol* 1998;**8**:759–69.

Jiang RHY, de Bruijn I, Haas BJ *et al.* Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. *PLoS Genet* 2013;**9**:e1003272.

Jiang RHY, Tyler BM. Mechanisms and evolution of virulence in oomycetes. *Annu Rev Phytolpathol* 2012;**50**:295–318.

Jin JP, Tian F, Yang DC *et al.* PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res* 2017;**45**:D1040–5.

Judelson H, Tani S. Transgene-induced silencing of the zoosporogenesis-specific *NIFC* gene cluster of *Phytophthora infestans* involves chromatin alterations. *Eukaryot Cell* 2007;**6**:1200–9.

Judelson HS, Blanco FA. The spores of *Phytophthora*: weapons of the plant destroyer. *Nat Rev Microbiol* 2005;**3**:47–58.

Jupe J, Stam R, Howden AJM *et al*. *Phytophthora capsici*–tomato interaction features dramatic shifts in gene expression associated with a hemi-biotrophic lifestyle. *Genome Biol* 2013;**14**:R63.

Kalyaanamoorthy S, Minh BQ, Wong TKF *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9.

Kamoun S, Furzer O, Jones JDG *et al*. The top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol* 2015;**16**:413–34.

Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 2016;**428**:726–31.

Kanneganti T-D, Huitema E, Cakir C *et al*. Synergistic interactions of the plant cell death pathways induced by *Phytophthora infestans* Nep1-like protein PiNPP1.1 and INF1 Elicitin. *Mol Plant Microbe* 2006;**19**:854–63.

Kaschani F, Shabab M, Bozkurt T *et al*. An effector-targeted protease contributes to defense against *Phytophthora infestans* and is under diversifying selection in natural hosts. *Plant Physiol* 2010;**154**:1794–804.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80.

Kemen E, Gardiner A, Schultz-Larsen T *et al*. Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. *PLoS Biol* 2011;**9**: e1001094.

Khatib M, Lafitte C, Esquerré-Tugayé M-T *et al*. The CBEL elicitor of *Phytophthora parasitica* var. *nicotianae* activates defence in *Arabidopsis thaliana* via three different signalling pathways. *New Phytol* 2004;**162**:501–10.

King BC, Waxman KD, Nenni NV *et al*. Arsenal of plant cell wall degrading enzymes reflects host preference among plant pathogenic fungi. *Biotechnol Biofuels* 2011;**4**:4.

Koide RT, Sharda JN, Herr JR *et al*. Ectomycorrhizal fungi and the biotrophy–saprotrophy continuum. *New Phytol* 2008;**178**: 230–3.

Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* 2007;**2**:e508.

Krogh A, Larsson B, von Heijne G *et al*. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;**305**:567–80.

Kubicek CP, Starr CP, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Ann Rev Phytopathol* 2014;**52**:427–51.

Kuo H-C, Hui S, Choi J *et al*. Secret lifestyles of *Neurospora crassa*. *Sci Rep* 2014;**4**:5135.

Käll L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. *Nucleic Acids Res* 2007;**35**:W429–32.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9.

Levin RA, Wagner WL, Hoch PC *et al*. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *Am J Bot* 2003;**90**:107–15.

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011;**12**:323.

Lin F, Zhao M, Baumann DD *et al*. Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. *BMC Genomics* 2014;**15**:18.

Lévesque CA, Brouwer H, Cano L *et al*. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol* 2010;**11**:R73.

Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 1947;**18**:50–60.

Marano AV, Jesus AL, de Souza JI *et al*. Ecological roles of saprotrophic Peronosporales (Oomycetes, Straminipila) in natural environments. *Fungal Ecol* 2016;**19**:77–88.

Marchler-Bauer A, Bo Y, Han L *et al*. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 2017;**45**:D200–3.

McCarthy CGP, Fitzpatrick DA. Phylogenomic reconstruction of the oomycete phylogeny derived from 37 genomes. *mSphere* 2017;**2**:e00095–17.

McGowan J, Fitzpatrick DA. Genomic, network, and phylogenetic analysis of the oomycete effector arsenal. *mSphere* 2017;**2**:e00408–17.

Misner I, Blouin N, Leonard G *et al*. The secreted proteins of *Achlya hypogyna* and *Thraustotheca clavate* identify the ancestral oomycete secretome and reveal gene acquisitions by horizontal gene transfer. *Genome Biol Evol* 2015;**7**:120–35.

Mistry J, Finn RD, Eddy SR *et al*. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 2013;**41**:e121.

Mitchell AL, Attwood TK, Babbitt PC *et al*. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* 2019;**47**:D351–60.

Nguyen L-T, Schmidt HA, von Haeseler A *et al*. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Ohm R, Feau N, Henrissat B *et al*. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen *Dothideomycetes* fungi. *PLoS Pathog* 2012;**8**:e1003037.

Pemberton CM, Davey RA, Webster J. et al. Infection of *Pythium* and *Phytophthora* species by *Olpidiopsis gracilis* (Oomycetes). *Mycol Res* 1990;**94**:1081–5.

Persson S, Wei H, Milne J *et al*. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci USA* 2005;**102**:8633–8.

Petersen TN, Brunak S, von Heijne G *et al*. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011;**8**:785–6.

Peyraud R, Mbengue M, Barbacci A *et al*. Intercellular cooperation in a fungal plant pathogen facilitates host colonization. *Proc Natl Acad Sci USA* 2019;**116**:3193–201.

Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;**29**:e45.

Phillips AJ, Anderson VL, Robertson EJ *et al*. New insights into animal pathogenic oomycetes. *Trends Microbiol* 2008;**16**:13–9.

Porras-Alfaro A, Bayman P. Hidden fungi, emergent properties: endophytes and microbiomes. *Ann Rev Phytopathol* 2011;**49**:291–315.

Preiss J. Regulation of the biosynthesis and degradation of starch. *Ann Rev Plant Physiol* 1982;**33**:431–54.

Prince DC, Rallapalli G, Xu D *et al*. *Albugo*-imposed changes to tryptophan-derived antimicrobial metabolite biosynthesis may contribute to suppression of non-host resistance to *Phytophthora infestans* in *Arabidopsis thaliana*. *BMC Biol* 2017; **15**:20.

Quinn L, O'Neill PA, Harrison J *et al*. Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. *FEMS Microbiol Lett* 2013;**344**:179–85.

Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat Rev Microbiol* 2012;**10**:417–30.

Ribeiro WRC, Butler EE. Comparison of the mycoparasites *Pythium periplocum, P. acanthicum* and *P. oligandrum*. *Mycol Res* 1995;**99**:963–8.

Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 2000;**16**:276–7.

Richards TA, Soanes DM, Jones MDM *et al*. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in oomycetes. *Proc Natl Acad Sci USA* 2011;**108**:15258–63.

Robinson MD, McCarthy DJ, Smyth GK. edeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–40.

Sambles C, Schlenzig A, O'Neill P *et al*. Draft genome sequences of *Phytophthora kernoviae* and *Phytophthora ramorum* lineage EU2 from Scotland. *Genom Data* 2015;**6**:193–4.

Schausser L, Roussis A, Stiller J *et al*. A plant regulator controlling development symbiotic root nodules. *Nature* 1999;**402**: 191–5.

Schornack S, van Damme M, Bozkurt TO *et al*. Ancient class of translocated oomycete effectors targets the host nucleus. *Proc Natl Acad Sci USA* 2010;**107**:17421–6.

Shahid S. To be or not to be pathogenic: transcriptional reprogramming dictates a fungal pathogen's response to different hosts. *Plant Cell* 2020;**32**:289–90.

Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika* 1965;**52**:591–611.

Sharma R, Xia X, Cano LM *et al*. Genome analyses of the sunflower pathogen *Plasmopara halstedii* provide insights into effector evolution in downy mildews and *Phytophthora*. *BMC Genomics* 2015;**16**:741.

Sperschneider J, Dodds PN, Gardiner DM *et al*. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol Plant Pathol* 2018;**19**:2094–110.

Sperschneider J, Williams AH, Hane JK *et al*. Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors. *Front Plant Sci* 2015;**6**:1168.

Thines M, Kamoun S. Oomycete–plant coevolution: recent advances and future prospects. *Curr Opin Plant Biol* 2010;**13**:427–33.

Tian M, Win J, Song J *et al*. A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. *Plant Physiol* 2007;**143**:364–77.

Tsirigos KD, Peters C, Shu N *et al*. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* 2015;**43**:W401–7.

Tyler BM, Tripathy S, Zhang X *et al*. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 2006;**313**:1261–6.

Tyler BM. Molecular basis of recognition between *Phytophthora* pathogens and their hosts. *Annu Rev Phytopathol* 2002;**40**: 137–67.

Usadel B, Obayashi T, Mutwil M *et al*. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant Cell Environ* 2009;**32**:1633–51.

Voragen AGJ, Coenen G-J, Verhoef RP *et al*. Pectin, a versatile polysaccharide present in plant cell walls. *Struct Chem* 2009;**20**:263–75.

Wang S, Boevink PC, Welsh L *et al*. Delivery of cytoplasmic and apoplastic effectors from *Phytophthora infestans* haustoria by distinct secretion pathways. *New Phytol* 2017;**216**:205–15.

Whisson SC, Boevink PC, Moleleki L *et al*. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 2007;**450**:115–8.

White TJ, Bruns T, Lee S *et al*. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds). *PCR Protocols: A Guide to Methods and Applications*. New York: Academic Press Inc, 1990, 315–22.

Wilhelmsson PKI, Mühlich C, Ulrich KK *et al*. Comprehensive genome-wide classification reveals that many plant-specific transcription factors evolved in streptophyte algae. *Genome Biol Evol* 2017;**9**:3384–97.

Win J, Morgan W, Bos J *et al*. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 2007;**19**:2349–69.

Xiang Q, Judelson HS. Myb transcription factors in the oomycete *Phytophthora* with novel diversified DNA-binding domains and developmental stage-specific expression. *Gene* 2010;**453**:1–8.

Yin Y, Mao X, Yang J *et al*. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2012;**40**:W445–51.

Zhao Z, Liu H, Wang C *et al*. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 2013;**14**:274.

Zuluaga AP, Vega-Arreguín JC, Fei Z *et al*. Transcriptional dynamics of *Phytophthora infestans* during sequential stages of hemibiotrophic infection of tomato. *Mol Plant Pathol* 2016;**17**:29–41.