New Phytologist

Research

Cross-species analysis between the maize smut fungi *Ustilago maydis* and *Sporisorium reilianum* highlights the role of transcriptional change of effector orthologs for virulence and disease

Weiliang Zuo[1], Jasper R. L. Depotter[1]*, Deepak K. Gupta[2,3,4]*, Marco Thines[2,3,4] and Gunther Doehlemann[1]

[1]Institute for Plant Sciences and Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, Zuelpicher Str. 47a, Cologne 50674, Germany; [2]Department for Biological Sciences, Institute of Ecology, Evolution and Diversity, Goethe University Frankfurt am Main, Frankfurt am Main 60325, Germany; [3]Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main 60325, Germany; [4]Integrative Fungal Research Cluster (IPF), Frankfurt am Main 60325, Germany

Author for correspondence:
Gunther Doehlemann
Email: g.doehlemann@uni-koeln.de

Received: *3 May 2021*
Accepted: *7 July 2021*

New Phytologist (2021) **232:** 719–733
doi: 10.1111/nph.17625

Key words: CRISPR-Cas9 gene conversion, cross-species RNA-sequencing, differentially regulated orthologs, fungal effectors, orthogroup, *Sporisorium reilianum*, *Ustilago maydis*.

Summary

- The constitution and regulation of effector repertoires shape host–microbe interactions. *Ustilago maydis* and *Sporisorium reilianum* are two closely related smut fungi, which both infect maize but cause distinct disease symptoms. Understanding how effector orthologs are regulated in these two pathogens can therefore provide insights into the evolution of different infection strategies.
- We tracked the infection progress of *U. maydis* and *S. reilianum* in maize leaves and used two distinct infection stages for cross-species RNA-sequencing analyses. We identified 207 of 335 one-to-one effector orthologs as differentially regulated during host colonization, which might reflect the distinct disease development strategies.
- Using CRISPR-Cas9-mediated gene conversion, we identified two differentially expressed effector orthologs with conserved function between two pathogens. Thus, differential expression of functionally conserved genes might contribute to species-specific adaptation and symptom development. Interestingly, another differentially expressed orthogroup (*UMAG_05318/Sr10075*) showed divergent protein function, providing a possible case for neofunctionalization.
- Collectively, we demonstrated that the diversification of effector genes in related pathogens can be caused both by alteration on the transcriptional level and through functional diversification of the encoded effector proteins.

Introduction

During symbiosis, plant-colonizing microbes secrete effector proteins to facilitate their interaction with the hosts. Effectors hold various functions, such as the suppression of host immunity and manipulation of host metabolism to promote host infection. The coevolution of microbes with their hosts poses a selection pressure on effector genes, which is an important driver of effector diversification (Jones & Dangl, 2006; De la Concepcion *et al.*, 2018; Tamborski & Krasileva, 2020). Effector polymorphisms are diverse, including presence/absence polymorphisms, amino acid substitutions, epigenetic modification and transcriptional alterations. Every polymorphism has the potential to change the host–microbe interaction and, collectively, polymorphisms in the effector repertoires between strains/species can be responsible for different disease development (Gijzen *et al.*, 2014; Toruño *et al.*, 2016; Franceschetti *et al.*, 2017; Torres *et al.*, 2020). Until now, many effector studies have focused on genome sequencing, effector prediction and variation analysis amongst phylogenetically related pathogens (Raffaele & Kamoun, 2012; Sánchez-Vallet *et al.*, 2018); functional elucidation of single effector proteins (Selin *et al.*, 2016); as well as transcriptional changes during host infection in one-to-one host–pathogen interactions (Lanver *et al.*, 2018). However, the role of transcriptional regulation in effector repertoire evolution has received less attention.

Smut fungi are a group of fungal plant pathogens consisting of more than 1500 species, which infect many grass species including major cereals such as maize, sorghum, wheat and barley (Begerow *et al.*, 2004). A characteristic feature of many smut fungi is their dimorphic growth: saprophytically as yeast, and biotrophically as filamentous hyphae (Kämper *et al.*, 2006). Smuts often infect their host through the root or coleoptile

*These authors contributed equally to this work.

(Martinez *et al.*, 2002; Laurie *et al.*, 2012), from where the fungus proliferates without evident symptoms during the early vegetative growth stage of the host. When plants switch into their reproductive phase, fungal hyphae in floral structures produce massive black/brown teliospores (the notorious 'smut' phenotype). This type of infection style is representative for *Sporisorium reilianum*, which can cause head smut disease on either sorghum (*S. reilianum* f. sp. *reilianum*) or maize (*S. reilianum* f. sp. *zeae*). *Sporisorium reilianum* spreads systemically in the plant while it stays close to the vascular bundles and grows up the main stem axis to reach the cob primordia. By contrast, *Ustilago maydis*, a close relative of *S. reilianum*, deviates in its disease development from other smuts as it has the ability to cause disease symptoms on all aerial tissue, and provoke locally restricted plant tumors within less than 2 wk under controlled conditions.

Ustilago maydis and *S. reilianum* have similar genomes with respect to gene number and synteny (Schirawski *et al.*, 2010). Furthermore, previous studies have suggested that, similar to *U. maydis*, *S. reilianum* can efficiently infect maize leaves to spread systemically and cause head smut disease in floral organs upon seedling inoculation. Several virulence effectors were identified from *S. reilianum* by such an approach (Schirawski *et al.*, 2010; Ghareeb *et al.*, 2015; Schweizer *et al.*, 2018). Gene complementation assays suggested effector orthologs from *U. maydis* and *S. reilianum* to have similar virulence functions (Sharma *et al.*, 2014; Redkar *et al.*, 2015b; Stirnberg & Djamei, 2016), although the molecular mechanism and regulation of these effectors in each species still need further investigation. Together, this renders the two closely related species a well-suited model to study the regulation and function of orthologous genes.

In *U. maydis*, the expression of effector genes is precisely regulated via a network of hierarchical transcription factors (TFs), which modulates transcription of effector genes at distinct infection stages (Skibbe *et al.*, 2010; Lanver *et al.*, 2018). One example is the *b* mating type locus, which encodes bE and bW to form a heterodimer TF. This TF regulates the expression of 38 effectors and also triggers the expression of several other TFs including the C2H2 zinc finger TF Rbf1 (regulator of b-filament) (Heimel *et al.*, 2010a,b), which, in turn, then turns on the TF Hdp2 (homeodomain transcription factor 2) and Biz1 (*b*-dependent zinc finger protein). Hdp2 and Biz1 together activate the expression of effectors important for initialization of the biotrophic interaction (Flor-Parra *et al.*, 2006; Heimel *et al.*, 2010b). At later infection stages, the forkhead TF Fox1 controls effectors necessary for full virulence and inhibition of host defense after establishment of the biotrophic interaction (Zahiri *et al.*, 2010). Another identified TF is Ros1 (Tollot *et al.*, 2016), which regulates effector gene expression related to fungal sporogenesis in mature tumors. A comprehensive set of RNA-sequencing (RNA-seq) data covering the whole biotrophic growth phase of *U. maydis* identified distinct expression patterns of effectors and a novel virulence-related TF Nlt1 (no leaf tumors1) (Lanver *et al.*, 2018), which might function upstream of Ros1 (Lin *et al.*, 2021). Although TFs are often studied in a time-specific manner, their regulation of effectors is probably also spatially distinct within the host, as multiple *U. maydis* effectors are regulated in an organ-specific (Skibbe *et al.*, 2010) or host cell-type-specific manner (Matei *et al.*, 2018).

Reverse genetics studies identified various *U. maydis* effectors holding distinct virulence functions, and some of these effectors have been functionally characterized on a molecular level (Lanver *et al.*, 2017). Examples are the apoplastic effectors Pit2 and Pep1, which inhibit the activity of maize apoplastic Papain-like cysteine proteases (PLCPs) (Mueller *et al.*, 2013; Misas Villamil *et al.*, 2019) and peroxidase (Hemetsberger *et al.*, 2012), respectively. The effectors Cmu1 (Djamei *et al.*, 2011) and Tin2 (Tanaka *et al.*, 2014) are translocated into the host cytoplasm to manipulate host salicylic acid (SA) and lignin metabolism, respectively. The leaf-specific effector See1 reactivates host DNA synthesis to directly promote tumor formation in bundle sheath cells (Redkar *et al.*, 2015a; Matei *et al.*, 2018). In addition, the secreted repetitive effector protein Rsp3 coats the fungal cell wall and protects it against a host antifungal protein (Ma *et al.*, 2018). By contrast, only one effector, SAD1, has been functionally characterized in *S. reilianum*. SAD1 suppresses the apical dominance of infected maize, which increased the amount of branches per plant (Ghareeb *et al.*, 2015). Furthermore, the contribution of several *S. reilianum* effectors and effector clusters to virulence has been verified (Ghareeb *et al.*, 2019).

Comparative genomic analysis from phylogenetically closely related smuts shows that effectors can be classified as core and accessory based on their conservation level (Schuster *et al.*, 2018; Beckerson *et al.*, 2019; Depotter & Doehlemann, 2020; Depotter *et al.*, 2020). Core effectors are conserved in many species and may be involved in conserved biological process vital for infection, leading to stabilization of such effectors during speciation (Hemetsberger *et al.*, 2015; Irieda *et al.*, 2019; Thines, 2019). Conversely, accessory effectors are less conserved and only found in a few or individual pathogen species and these were supposed to have more subtle/specific functions in virulence. However, how these core and accessory effectors are regulated on the transcriptional level between phylogenetically closely related smuts is still unknown.

Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 was first identified in *Streptococcus pyogenes* as part of the bacterial immune system against bacteriophage infection (Barrangou *et al.*, 2007), and has been widely adapted as genome editing tool in various organisms including filamentous fungi and oomycetes (Schuster & Kahmann, 2019). In *U. maydis*, CRISPR-Cas9-mediated mutations are obtained by expressing the codon-optimized Cas9 protein and single guide RNA (sgRNA) in an autonomous replication plasmid (Schuster *et al.*, 2016, 2018), which shows high efficiency. The number of off-target mutations can be reduced by the application of a high-fidelity Cas9 variant (Cas9HF1) (Zuo *et al.*, 2020). Furthermore, CRISPR-Cas9 can also be used for seamless gene conversion but has not been exploited in *U. maydis* before this study. This method allows us to study the functional conservation of orthologs/paralogs while maintaining the *cis*-regulatory of gene of interest.

In this study, we examined the growth of *U. maydis* and *S. reilianum* during leaf infection for recording similar and divergent biotrophic interactions with maize. We investigated the

regulation of orthologous genes by CRISPR-Cas9-mediated gene conversion with an emphasis on effector genes.

Materials and Methods

Strains, growth conditions and plant infections

The mating compatible isolates of *U. maydis* FB1 and FB2 and *S. reilianum* SRZ1 and SRZ2 were used for RNA-seq, and the solopathogenic *U. maydis* strain SG200 and its respective knockout and ortholog conversion mutants were used for virulence tests. All strains were grown in YEPS light liquid medium with 200 rpm shaking or on potato dextrose agar (PD; Difco) plates at 28°C. *Escherichia coli* strain Top10 was used for cloning and grown in dYT liquid medium or YT agar plates with supplementary corresponding antibiotics. Maize variety Early Golden Bantam was used in infection assays. The plants were grown under controlled conditions of 16 h light at 28°C and 8 h dark at 22°C. Seven-day-old seedlings were infected with a mixture of compatible *U. maydis*/*S. reilianum* isolates or SG200. Disease symptoms were scored at 12 d after infection.

Staining and microscopy

For microscopy purposes, 0.1% Tween-20 was added to the inoculum. To visualize the appressorium, leaves were stained 20 h after infection with calcofluor solution (100 μg ml^{-1}) for 1 min and briefly rinsed with water (Lanver *et al.*, 2014). To visualize the fungal growth inside the infected leaves, WGA-AF488 (Wheat Germ Agglutinin, Alexa Fluor 488) and propidium iodide were used for staining the fungal and plant cell wall, respectively, as previously described (Doehlemann *et al.*, 2009). Microscopy was done with a Nikon Eclipse Ti Inverted Microscope using the Nikon NIS-ELEMENTS software (Düsseldorf, Germany). Photos were taken with a HAMAMATSU camera. Confocal microscopy was done with a Leica TCS SP8 confocal laser scanning microscope (Leica, Wetzlar, Germany). The following filters were used for WGA-AF488: excitation 458 nm and emission 470–490 nm, and for propidium iodide: excitation 561 nm and emission 590–603 nm.

RNA preparation and RNA-seq

For RNA sample preparation, a 2-cm-long section of the 3rd leaf from more than 15 individual plants were collected for each sample. To generate biological replicates of infected samples, three plant infections were conducted from three independent fungal cultures. For each infection, more than 50 plants were inoculated for each pathogen. The compatible haploid *U. maydis* and *S. reilianum* cells from cultures with OD_{600} of *c.* 0.8 were spun down, suspended in water to an OD_{600} of 1, mixed in a 1 : 1 ratio and then span down as an axenic culture (AC) control. Three independent cultures from different times were collected in three biological replications. The plant tissues and cell pellets were ground into fine powder with liquid nitrogen. Subsequently, RNA samples were prepared using TRizol (Thermo Fisher, Waltham, MA, USA) according to the manufacturer's protocol followed by DnaseI digestion (Thermo Fisher). RNA libraries were prepared using an Illumina TruSeq Stranded mRNA kit (Illumina, San Diego, CA, USA), and paired-end sequencing was performed on a HiSeq4000 platform to produce 2 × 75 bp reads at the Cologne Center for Genomics (Cologne, Germany).

Data analysis

For individual gene expression analysis of *U. maydis* and *S. reilianum*, reads of three biological replicates were filtered using the TRINITY software (v.2.9.1) option trimmomatic under standard settings (Szegedy *et al.*, 2016). They were mapped to a reference assembly using BOWTIE 2 (v.2.3.5.1) (Langmead & Salzberg, 2012) with the first 15 nucleotides on the 5′-end of the reads being trimmed. The reference genome was either the genome assembly of *U. maydis* (Kämper *et al.*, 2006) or *S. reilianum* (Schirawski *et al.*, 2010) combined with that of *Zea mays* B73 v.3 (Schnable *et al.*, 2009). Reads were counted to the *U. maydis* and *Z. mays* loci using the R package RSUBREAD (v.1.34.7) (Liao *et al.*, 2019). The EDGER package v.3.26.8 was used for statistical analysis of differential gene expression and pairwise comparison was conducted using generalized linear models (GLMs) (Robinson *et al.*, 2009). Genes with $\log_2$ fold change > 1 and $P < 0.05$ were considered as differentially regulated between timepoints, and genes with $\log_2$ fold change < 1, $P < 0.05$ or $\log_2$ fold change > 1 but $P > 0.05$ were considered as not significantly different in the induced pattern analysis. Effector patterns were defined based on the differential regulation relationships between AC, 2 d postinfection (dpi) and 4 dpi samples with the following criteria: pattern 1 are genes induced *in planta* (expression levels at 2 and 4 dpi were significantly higher than AC samples) and the 4 dpi expression level is higher than 2 dpi; genes classified as pattern 2 were induced *in planta*, and the expression levels were similar between 2 and 4 dpi; pattern 3 genes were only induced at 2 dpi and had dropped again at 4 dpi to the AC level; pattern 4 genes were significantly induced *in planta* at both timepoints and the expression level of 2 dpi is higher than 4 dpi; and pattern 5 genes are only induced at 4 dpi.

A phylogenetic tree was constructed including the smut species *Ustilago hordei* (Laurie *et al.*, 2012), *Melanopsychium pennsylvanicum* (Sharma *et al.*, 2014), *Sporisorium scitamineum* (Taniguti *et al.*, 2015), *S. reilianum* (Schirawski *et al.*, 2010) and *U. maydis* (Kämper *et al.*, 2006). *Moesziomyces antarcticus* was used as an outgroup (Morita *et al.*, 2013). A phylogenetic tree was constructed based on 1643 benchmarking universal single-copy orthologs (BUSCOs) from the database 'basidiomycota_odb10' that were present in single copy in all members investigated in the phylogenetic reconstruction (Seppey *et al.*, 2019). For every gene, orthologs were aligned using MAFFT (v.7.464) option '–auto' (Katoh & Standley, 2013). These aligned gene sequences were then concatenated for every species and used for tree construction using RAXML (v.8.2.11) with the substitution model 'GTRGAMMA' and 100 bootstraps (Stamatakis, 2014).

For comparison of ortholog expression, high-quality paired end transcriptomic reads were mapped using BOWTIE2 with

default parameters over all orthologous genes of *U. maydis* and *S. reilianum* as calculated using OrthoMCL (Li *et al.*, 2003). The average coverage of mapping of each gene was calculated from the sorted BAM file by summing the coverage of each base of a gene and then dividing this by the length of that gene. The relative expression value of a gene was calculated by using the following formula:

$$\text{Relative Expression of Gene}(X) = \frac{\frac{\sum_1^n \text{Coverage}(X)\times(N-n)}{n}}{(M-m)/(N-n)},$$

where n = length of the gene (X), N = total number of bases in all genes; m = number of bases mapped to gene (X) and M = number of bases mapped to all genes. Student's *t*-test and/or one sample *t*-test were used for significance testing with Benjamini–Hochberg *P* value correction for multiple comparison. Orthologs with fold change > 3 and adjusted $P < 0.05$ were considered as differentially regulated orthologs between *U. maydis* and *S. reilianum*.

Gene ontology and gene ontology enrichment analysis

The gene ontology (GO) classification and overrepresentation analysis were done by using the *U. maydis* annotation data from PANTHER (pantherdb.org). The overrepresentation analysis was conducted by using Fisher's exact test with Bonferroni correction.

Quantitative real-time PCR

For biomass quantification, DNA of infected leaves from three biological replicates was prepared by using Buffer A (0.1 M Tris-HCl, 0.05 M EDTA, 0.5 M NaCl, 1.5% SDS), then further purified by MasterPure Complete DNA and RNA Purification Kit Bulk Reagents (Epicentre, Madison, WI, USA). In total, 100 ng of DNA was used for quantitative real-time PCR (qPCR), and $2^{-\Delta Ct}$ was calculated to determine the ratio between fungal *peptidylprolyl isomerase* (*ppi*) and maize *GAPDH*.

For gene expression, the total RNA was reverse-transcribed with Oligo(dT) primer by using a RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific), and the expression level of each effector was determined by calculating the $2^{-\Delta Ct}$ ratio between the effector gene and the *ppi* gene. qPCRs were performed using the CFX96 Real-Time PCR Detection System (Bio-Rad) with GoTaq qPCR mix (Promega). The primers designed in this study are listed in Supporting Information Table S1.

Gene conversion in *U. maydis* by CRISPR and disease scoring

CRISPR-Cas9-mediated genome editing was used to generate effector ortholog conversion mutants in *U. maydis* solopathogenic strain SG200 (Kämper *et al.*, 2006). sgRNAs were designed to cause a double strand brake in the coding or the promoter region of effectors, which were subsequently repaired by homology-directed repair using a donor plasmid as template. The design of sgRNA and cloning was done as previously described (Zuo *et al.*, 2020). To construct the donor plasmid, 1 kb fragments flanking the effector coding region or promoter regions and the respective *S. reilianum* ortholog open reading frame or promoters were amplified by Phusion DNA polymerase (New England Biolabs, Ipswich, MA, USA) with primers (Sigma Aldrich) listed in Table S1 and cloned into pAGM1311 vector (Weber *et al.*, 2011) by Gibson assembly (New England Biolabs). The CRISPR plasmid and circular donor plasmid were cotransformed into SG200 protoplasts. The resulted mutants were singled out on a PD plate with 2 μg ml^{-1} carboxin based on the antibiotic selection from CRISPR plasmid first, then transfer to a PD plate without antibiotic to remove CRISPR plasmid. The resulting conversion mutants have no antibiotic resistance. Correct integration of the recombinant DNA was confirmed by southern blot (data not shown).

The conversion mutants and corresponding effector deletion mutants (Schilling *et al.*, 2014) were used for infection. Disease scoring was done as previously described at 12 dpi (Redkar & Doehlemann, 2016). Disease indexes 9, 7, 5, 3, 1 and 0 were assigned to dead, heavy tumor, tumor, small tumor, chlorosis and normal symptom, respectively. The number of diseased plants was multiplied by the corresponding disease index, and the sum was divided by the total number of plants used for infection to give an average disease index. Student's *t*-test was used for significance testing of the disease index from three biological replications.

Results

Maize leaf infection of *S. reilianum* and *U. maydis*

For a comparative transcriptomics approach between *S. reilianum* and *U. maydis*, we first monitored infection and fungal growth inside the maize leaves to record the milestone events during infection and define appropriate timepoints for analysis (Fig. 1). At 18–24 h post-infection, the compatible sporidia cells of both pathogens mated on the leaf surface, and the resulting dikaryotic hyphae formed appressoria to penetrate the leaf surface (Fig. 1a). At 2 dpi, all infected leaves showed chlorosis as a first visible indication of successful infection (Fig. S1a). Microscopic analysis of WGA-AF488-stained hyphae showed similar colonization in the leaf vascular tissue for both smut fungi (Fig. 1b). At 3 dpi, the earliest microscopic difference was observed between *S. reilianum* and *U. maydis*, in agreement with our previous report (Matei *et al.*, 2018). Host bundle sheath cells then underwent *de novo* cell division in *U. maydis*-infected leaves (as more than two cells were observed between leaf vascular tissue), indicating initiation of tumorigenesis. By contrast, *S. reilianum* further accumulated in the leaf veins and showed aggregation as a string of round cells (Fig. 1c). At 4 dpi, the first visible small tumor structures were observed on *U. maydis*-infected leaves, while *S. reilianum* infected leaves showed individual events of necrosis (Fig. S1a) and more frequently fungal cell aggregation (Figs 1d, S1b). In parallel with the microscopic observation, fungal biomass in the infected leaves was quantified for both pathogens. The relative amount of both fungal pathogens increased similarly during infection until 3 dpi.

Interestingly, at 4 dpi *S. reilianum* showed a significantly higher abundance compared to *U. maydis*, as the relative amount of *S. reilianum* doubled while the relative amount of *U. maydis* was maintained in comparison to at 3 dpi (Fig. 1e).

Expression profiling of *U. maydis* and *S. reilianum* during maize leaf colonization

Based on our microscopic and biomass analysis, we collected samples at 2 dpi to represent an infection stage where both *U. maydis* and *S. reilianum* successfully establish the biotrophic interaction with the host but do not show detectable differences in microscopical growth and biomass accumulation. Samples were also taken at 4 dpi when *U. maydis* tumorigenesis is initiated, while *S. reilianum* exhibits extensive tissue colonization without inducing morphological changes in the leaf. As control samples, cell pellets from AC of both pathogens were used for RNA extraction. All samples were generated in three independent biological replicates from three independent infections. In total, more than 670 million paired-end reads were generated, of which over 93.7% of reads were uniquely mapped to either *U. maydis*, *S. reilianum* or the maize genome. The changes of fungal transcripts in the RNA-seq samples were consistent with the biomass quantification results by qPCR. At 2 dpi, *c.* 2% of fungal transcripts were mapped from both pathogens, and at 4 dpi, *S. reilianum* reads increased to 12.5% in the infected samples, which is

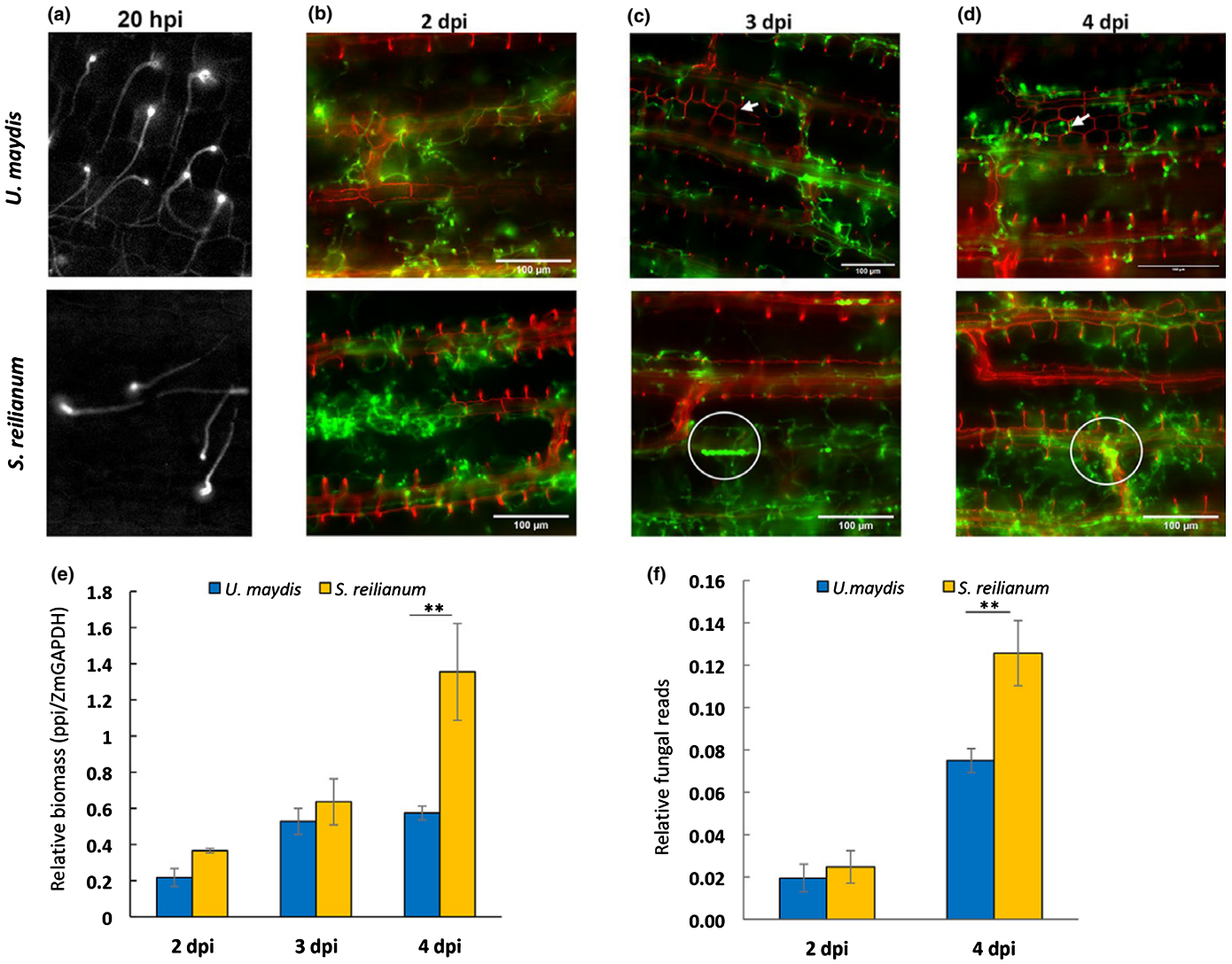

Fig. 1 Comparison of *Ustilago maydis* and *Sporisorium reilianum* growth during leaf infection. (a) Calcofluor white staining of fungal hyphae and appressorium at 20 hpi. (b–d) WGA-AF488-propidium iodide costaining to show the fungal biotrophic growth inside the plant cell. Fungal cells were stained with WGA-AF488 (green) and the cell wall was stained with propidium iodide (red). White arrows indicate the *de novo* divided bundle sheath cells triggered by *U. maydis* infection, as the number of cells between vasculatures increased from two, to three or more. White circles indicate the aggregation of *S. reilianum* cells inside the leaf vein. (e) Relative biomass quantification by qPCR shows the fungal growth during infection, and a significantly higher abundance of *S. reilianum* was detected at 4 dpi. (f) Relative fungal reads from RNA-seq, which was the ratio between reads uniquely mapped to fungi and total reads uniquely mapped (fungi and maize). A significantly higher number of mapped reads was detected from *S. reilianum* at 4 dpi. Data represent the means ± SD from three independent biological replicates. **, $P < 0.05$. Student's *t*-test was used for significance test. hpi, hours post-infection; dpi, days post-infection.

significantly higher compared with *U. maydis*-infected tissues (7.5%) (Fig. 1f). Principal component analysis (PCA) of *U. maydis* and *S. reilianum* samples showed that three biological replications from different conditions form distinct clusters. Moreover, variations between different conditions were higher than the variation of biological replications within the same condition (Fig. 2a). To further analyze the RNA-seq data, we normalized read counts and investigated the differentially expressed genes (DEGs) by EDGER (Robinson *et al.*, 2009). After filtering the low- or not-expressed genes between samples, 6537 of 6765 *U. maydis* genes and 6456 of 6673 *S. reilianum* genes were used to identify DEGs. Compared to AC, dramatic changes in gene expression were observed during seedling infection. Around 19.3–24.2% of genes were upregulated ($\log_2$ fold change > 1, $P < 0.05$) at 2 and 4 dpi in both pathogens, while 12.1–19.1% of genes were downregulated (Fig. 2b). We detected 435 of 467 effector genes from *U.*

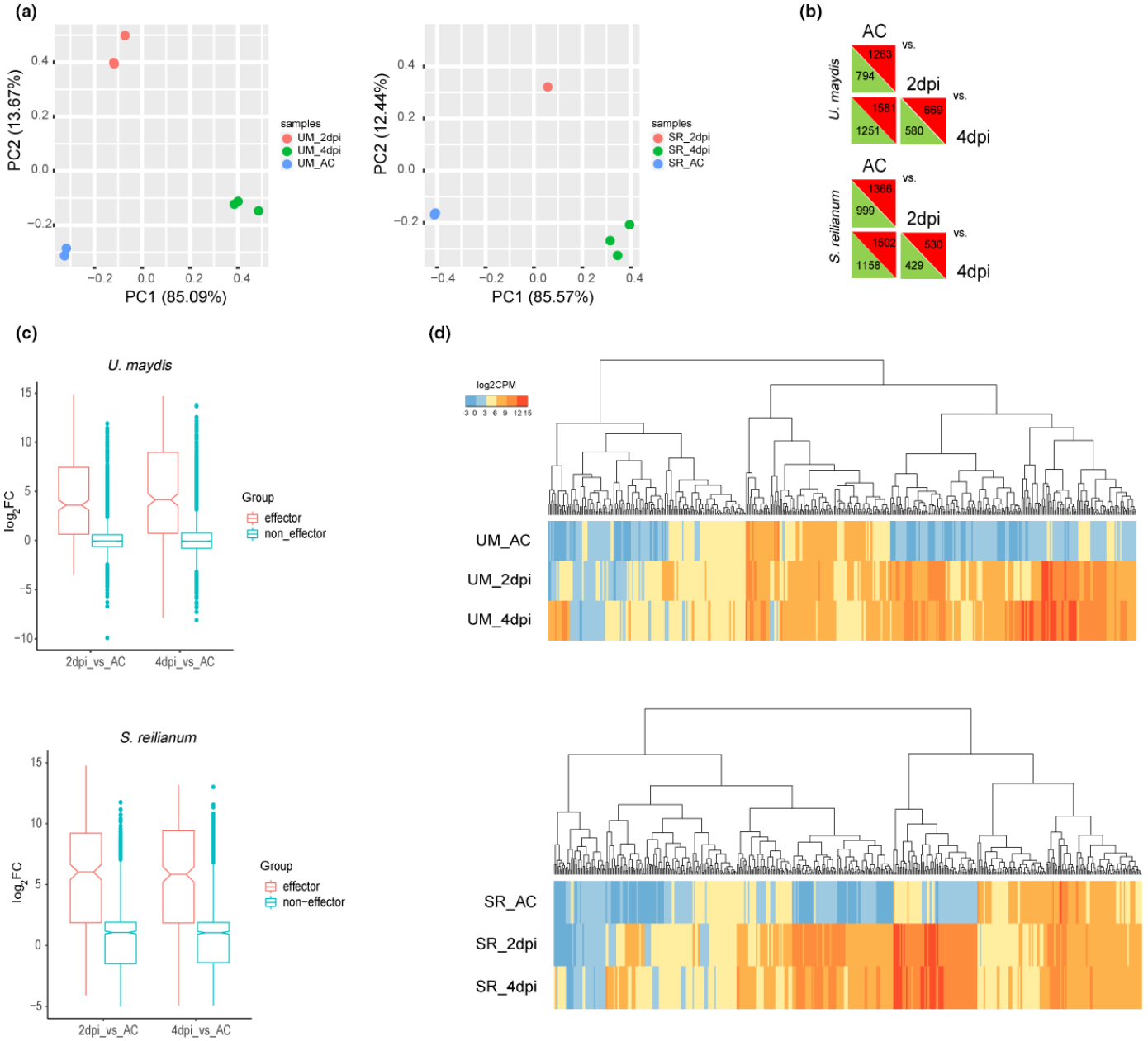

Fig. 2 Overview of *Ustilago maydis* and *Sporisorium reilianum* transcriptome, respectively. (a) PCA of *U. maydis* and *S. reilianum* RNA-seq samples. (b) Expression data from three different growth conditions. The number of differentially expressed genes ($\log_2$ fold change > 1, $P > 0.05$) from pairwise comparison is indicated in the triangles. Red and green triangles represent up- and downregulated genes between 2 dpi vs axenic culture (AC), 4 dpi vs AC and 4 dpi vs 2 dpi, respectively. (c) Boxplot showing transcriptional changes of effector and noneffector genes during biotrophic growth. Effector genes were dramatically upregulated compared to the noneffector genes. The horizontal lines indicate the median of $\log_2$ fold changes from each group, the notched boxes represent the confidence interval and the dots are outlying data. (d) Heatmap showing an overview of the expression changes of effector genes. *n* is the number of effectors detected expressing in RNA-seq samples and the total number of effectors in the genome. dpi, days post-infection.

maydis and 454 of 489 from *S. reilianum*. As expected, most effector genes were specifically expressed in the biotrophic growth stage (311 in *U. maydis* and 307 in *S. reilianum*) and showed higher degrees of transcriptional induction compared to noneffector genes (Fig. 2c,d), which reflects the explicit role of effectors in host infection. Three biological replicates of qRT-PCR were conducted to confirm the expression of several effector genes, including infection markers such as Pit2 and Pep1 from our RNA-seq data (Fig. S2a,b).

We clustered effector genes into all five possible expression patterns between three conditions (AC, 2 dpi, 4 dpi) (Fig. 3a): showing constitutively increasing induction (pattern 1), induced *in planta* on a stable level (pattern 2), being only transiently expressed (pattern3), dropping at 4 dpi (pattern 4) or only being induced at 4 dpi (pattern 5). In *U. maydis* and *S. reilianum*, 227 effectors from each species were highly expressed across the whole biotrophic growth phases (patterns 1, 2 and 4), while only 14 *U. maydis* effectors and eight *S. reilianum* effectors were transient induced at 2 dpi (pattern 3) and similar number of effectors were only required at 4 dpi (23 from *U. maydis* vs 25 from *S. reilianum* as pattern 5) (Fig. 3b). The expression patterns were found to reflect previously observed virulence function of known effector genes. For example, the core effector *UmPep1* (*UMAG_01987*) (Doehlemann *et al.*, 2009; Hemetsberger *et al.*, 2015) was stably expressed and clustered as pattern 2, while genes coding for the known virulence factors *UmPit2* (*UMAG_01375*) (Misas Villamil *et al.*, 2019), *UmCmu1* (*UMAG_05731*) (Djamei *et al.*, 2011), *UmTin2* (*UMAG_05302*) (Tanaka *et al.*, 2014) and *UmRsp3* (*UMAG_03274*) (Ma *et al.*, 2018) group as pattern 1. Interestingly, *S. reilianum* orthologs of these characterized effectors displayed different induction patterns, maintaining a steady level between 2 and 4 dpi. In particular, we identified 100 of 467 *U. maydis* effectors being *in planta* induced according to pattern 1, which is significantly higher compared to *S. reilianum* (43/488), whereas 184 out of 488 *S. reilianum* effector genes showed stable expression levels (pattern 2), which is a higher fraction than in *U. maydis* (136/467) (Fig. 3b). These differences in representation of effector genes across these expression patterns suggest different effector expression requirements for the two species.

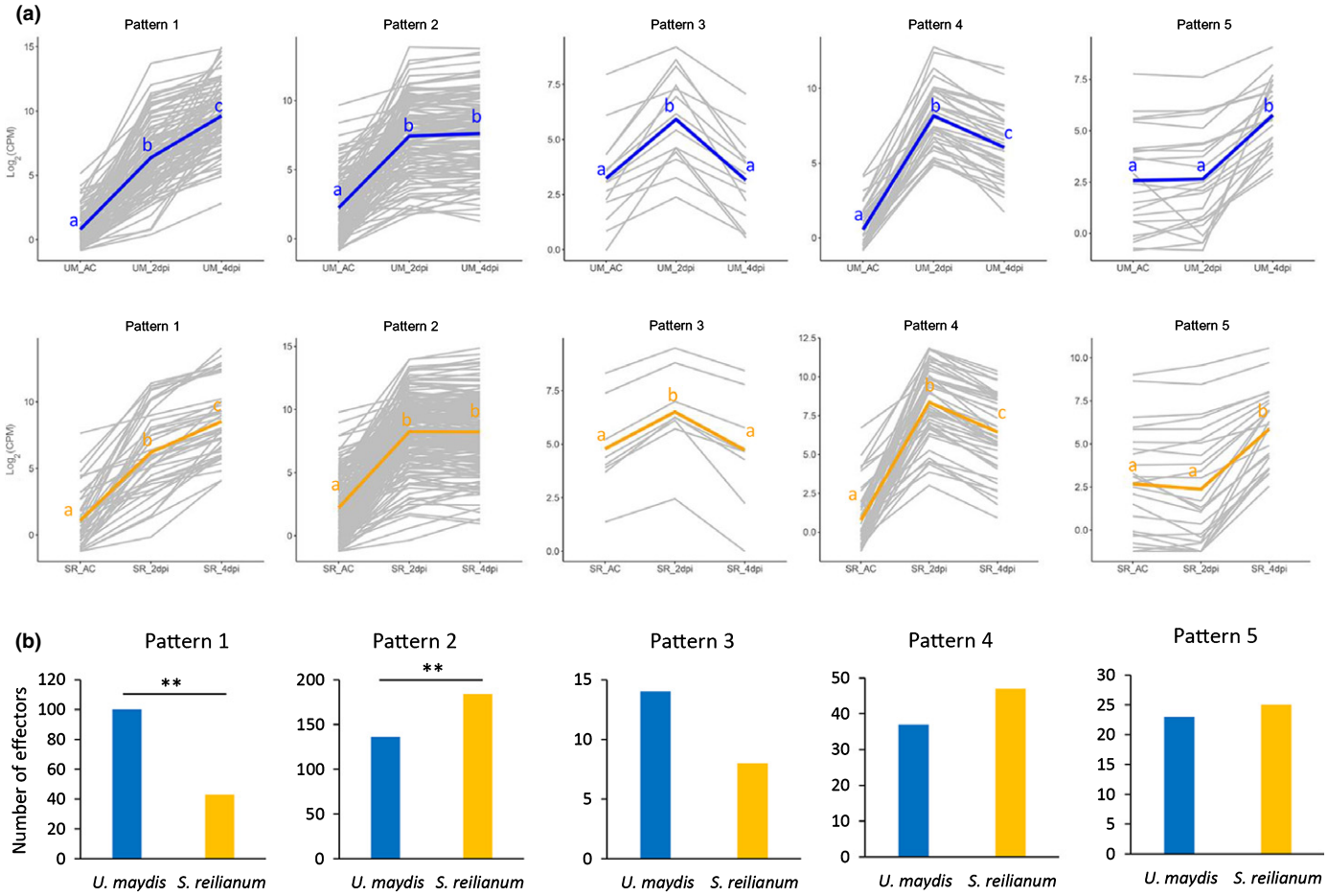

Fig. 3 The induced pattern shows the expression changes of effectors. (a) The induced effectors are clustered into five patterns based on the expression changes between axenic culture (AC), and 2 and 4 dpi conditions. The gray lines represent individual genes, and blue and orange lines are the mean of all genes in each cluster. a, b and c indicate the significance levels between each condition. Genes with $\log_2$ fold change > 1 and $P < 0.05$ were considered significantly induced in pairwise comparisons. (b) The number of induced genes of each pattern. The chi-square test was used as the statistical test of significance. **, $P < 0.01$; dpi, days post-infection.

Differential regulation of effector orthologs in *U. maydis* and *S. reilianum*

To better understand how *U. maydis* and *S. reilianum* orchestrate their respective effector repertoires with respect to their differential disease development, we directly compared the relative expression levels of effector orthologs. We used ORTHOMCL to identify one-to-one ortholog pairs between *U. maydis* and *S. reilianum* and obtained 6005 one-to-one ortholog pairs including 335 effectors (Table S3). This means that to avoid ambiguous results due to functional redundancy or paralog compensation, one-to-many or many-to-many ortholog pairs were not included in this analysis. Relative expression was corrected for putative differences in length between the orthologs. PCA based on the relative expression of 6005 ortholog pairs confirmed that plant-associated samples are distinct on time and species levels. By contrast, the AC samples of the different species had minor differences and clustered together (Fig. 4a). Pearson correlation analysis displayed similar results to the PCA and suggested the difference between samples increased at 4 dpi compared to 2 dpi, as the correlation decreased (Fig. S3a).

Next, we compared the relative ortholog expression levels to identify differentially expressed orthologs (DEOs) in the two smut fungi. In total, we found 769 DEOs with comparatively higher expression in *U. maydis* and 737 DEOs being more highly expressed in *S. reilianum*, respectively (Fig. 4b; Table S3). The largest number of DEOs (835) was detected from AC samples (Fig. 4b), and *c.* 30 and 50% of these DEOs from *U. maydis* and *S. reilianum* were only confined in this stage (Fig. S3b). The DEO number dropped to 323 at 2 dpi when the pathogens had switched from *in vitro* yeast growth to the initial filamentous, biotrophic growth in the host tissue. At 4 dpi the number increased to 747, which probably reflects the significant differences in disease development between the two pathogens (i.e. initiation of tumorigenesis by *U. maydis* vs extended proliferation of *S. reilianum* for systemic spreading). We also detected more *U. maydis*-induced DEOs during infection, particularly at 2 dpi, when 227 DEOs were *U. maydis*-induced, but only 97 DEOs for *S. reilianum* (Fig. 4b). However, we could not detect overrepresented GO terms, except at 4 dpi, where an enrichment of cellular lipid catabolic process (GO:0044242) was detected in the higher expressed *U. maydis* DEOs. The DEOs were involved in different biological processes related to pathogen growth (Fig. 4c), suggesting that during the yeast and biotrophic growth no particular cellular or metabolic pathway was selected to promote cell growth in either *U. maydis* or *S. reilianum*.

From the 335 one-to-one effector orthologs, more than 60% were differentially regulated, including 100 *U. maydis* and 102 *S. reilianum* highly expressed effectors, from which 32 and 35 showed consistently enhanced expression during the two biotrophic time points (Fig. 4d). The number of effector DEOs also increased from 99 at 2 dpi to 170 at 4 dpi as disease developed. Previously identified pathogenesis-associated TFs might be responsible for the regulation of effector DEOs, particularly at 4 dpi, since differential expression was found for virulence-related TFs *Rbf1*, *Biz1*, *Hdp2*, *Fox1* and *Nlt1* (Fig. S4). TF *Rbf1* was highly induced in *S. reilianum*, especially at 4 dpi, and transcript levels were 17-fold higher than those in *U. maydis*, which may lead to the high abundance of *Biz1* in *S. reilianum* (Fig. S4). On the other hand, the expression level of TFs *Hdp2*, *Fox1* and *Nlt1* were significantly higher in *U. maydis* at 4 dpi, which could indicate that these TFs play a role in the regulation of effectors that are highly expressed in *U. maydis* during tumorigenesis (Fig. S4). Besides these, the TF *Ros1* showed a similar expression level in both species, which may reflect its dedicated function in sporogenesis, a process which needs to be initiated in both pathogens to complete the pathogenic life cycle (Fig. S4). The expression of *Ros1* is dramatically increased after 6 dpi in *U. maydis*, which corresponds to the initiation of sporogenesis (Tollot *et al.*, 2016), and may explain why we did not observe its differential expression between the two species. Cluster 19A (Fig. S5a) is the largest effector cluster, which is a crucial determinant of virulence in both pathogens (Brefort *et al.*, 2014; Ghareeb *et al.*, 2019). We observed that 10 out of 14 effectors residing in cluster 19A were *U. maydis*-enhanced, whereas only two were highly expressed in *S. reilianum,* including the neofunctionalized Tin2 (Tanaka *et al.*, 2019) (Fig. S5b).

In a previous study, Schuster *et al.* (2018) analyzed 12 smut-related basidiomycetes and suggested that effector repertoires of smut fungi comprise sets of core and accessory effectors. Based on this information, we classified the 335 one-to one effector pairs into plant smut core effectors, which were present in all four monocot smut fungi, including *Ustilago hordei* and *Sporisorium scitamineum*, and accessory effectors, which present in *U. maydis* and *S. reilianum* but not in all other plant smut fungi (Fig. 4e). *Melanopsychium pennsylvanicum* was included in this analysis due to its close phylogenetic relationship to monocot smut pathogens, which suggests a recent host jump event (Sharma *et al.*, 2014). In total, we identified 191 core effectors and 144 accessory effectors. At 2 dpi, 40 core effectors and 31 accessory effectors were differentially expressed between *U. maydis* and *S. reilianum*, and these numbers were significantly increased to 67 and 61 at 4 dpi, respectively (Fig. 4f). We did not find accessory effectors subjected to more intensive transcription regulation compared to core effectors (Fig. 4f), suggesting that differential regulation of both core and accessory effectors is needed for the distinct pathogenic development in the two pathogens.

Functional conservation and neofunctionalization of differentially expressed effector orthologs

Several *U. maydis* effectors have been functionally characterized, which provides a reference to link ortholog expression with their biological functions. The conserved core effector Pep1 (Hemetsberger *et al.*, 2012, 2015; Sharma *et al.*, 2019), as well as the virulence factors Pit2 (Misas Villamil *et al.*, 2019) and Cmu1 (Djamei *et al.*, 2011) were highly induced during biotrophic growth (Table S2) and showed similar expression level between *U. maydis* and *S. reilianum*, which is consistent with their crucial roles in maize immunity inhibition (Fig. 5a). The effectors Rsp3 (Ma *et al.*, 2018) and ApB73 (Stirnberg & Djamei, 2016) had been validated in previous studies for their conserved function

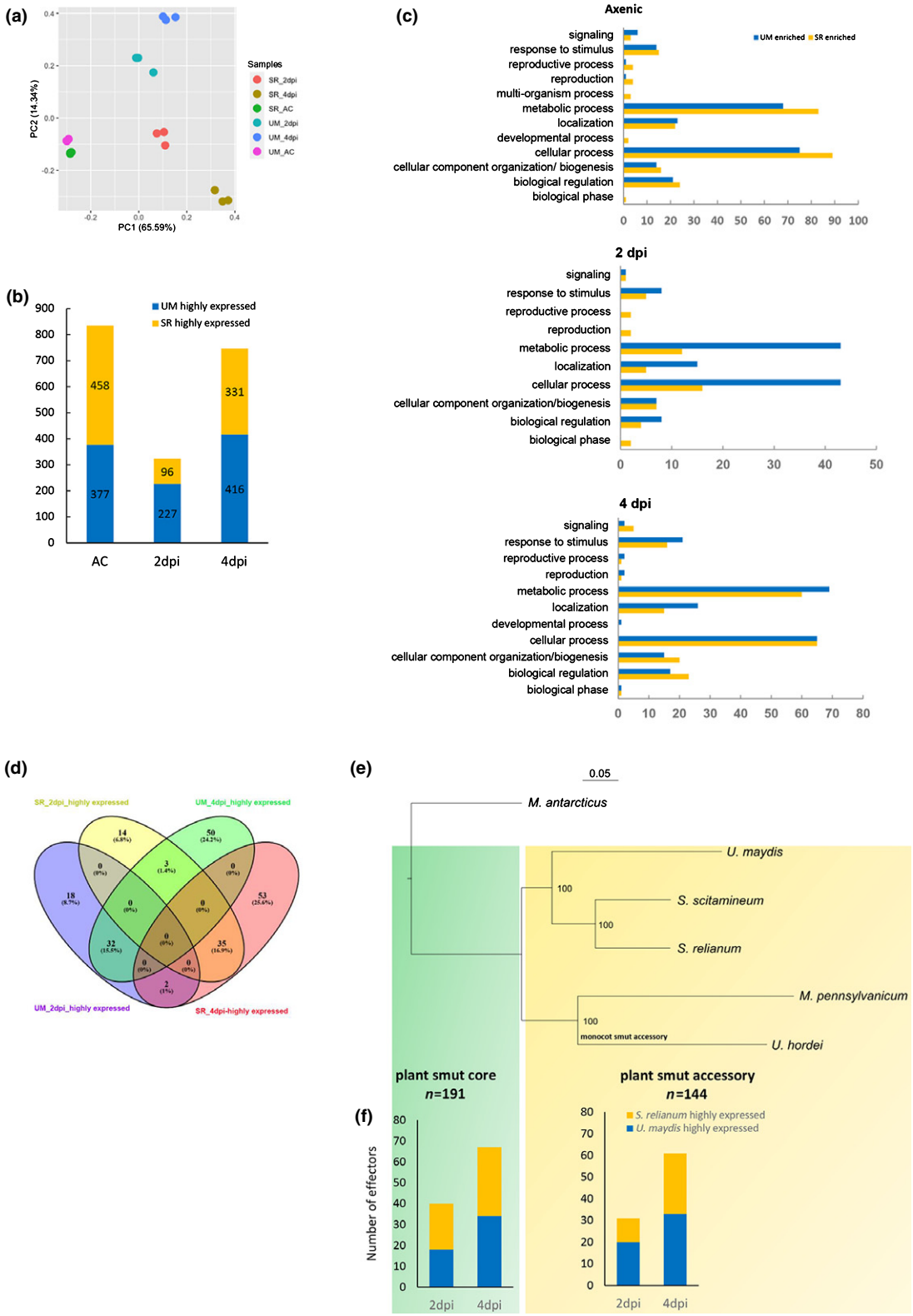

Fig. 4 Orthologs were differentially regulated between *Ustilago maydis* and *Sporisorium reilianum*. (a) PCA based on the relative ortholog expression of 6005 one-to-one ortholog pairs from *U. maydis* and *S. reilianum*. (b) The number of DEOs detected in different conditions. Fold change > 3 and $P < 0.05$ were used to identify DEOs. (c) The GO classification of DEOs in three different conditions. (d) Venn diagram showing the number of differentially expressed effector orthologs during biotrophic growth. (e–f) Differential regulation of effectors clustered by the conservation between five smut fungi. (e) Phylogenetic tree of five smut fungi, including *Ustilago maydis*, *Sporisorium reilianum*, *Ustilago hordei*, *Sporisorium scitamineum* and *Melanopsychium pennsylvanicum*. (f) The number of effector DEOs in each category. The chi-square test was used for significance testing. **, $P < 0.01$; ***, $P < 0.001$. AC, axenic culture; dpi, days post-infection; UM, *U. maydis*; SR, *S. reilianum*.

between two smut fungi by ortholog complement assays, and also these genes show comparable expression levels (Fig. 5a). In addition, the See1 effector, which in *U. maydis* has a virulence function specifically for leaf tumor formation (Redkar *et al.*, 2015a), is expressed in *S. reilianum* both at 2 and at 4 dpi (Fig. 5a). Another effector DEO of note is Tin2. SrTin2 neofunctionalized during speciation between *U. maydis* and *S. reilianum* (Tanaka *et al.*, 2019). We found a significantly enhanced expression of *SrTin2* compared to *UmTin2* (Fig. 5a), in line with the neofunctionalization, as the transcriptional level divergence could be involved to promote the functional divergence.

From a set of 20 organ-specific effectors, we found nine effector DEOs and eight of them were leaf-specific effectors, from which four were validated as virulence factors from a deletion mutant screen (Schilling *et al.*, 2014) (Fig. 5a). Orthogroups *UMAG_11060/Sr14941*, *UMAG_05306/Sr10059* and *UMAG_05318/Sr10075* were particularly interesting as they were induced in both pathogens (Table S2), but the transcripts levels of the orthologs in *U. maydis* were at least 4.8 (for orthogroup *UMAG_11060/Sr14941* at 2 dpi) to more than 50 times (orthogroup *UMAG_05318/Sr10075* at 4 dpi) higher compared to their *S. reilianum* orthologs (Table S2). These effectors are specifically required for *U. maydis*-induced leaf tumors (Schilling *et al.*, 2014). Thus, their low expression in *S. reilianum* might suggest that these genes are of lesser importance for the basal establishment of biotrophic growth. We assessed the protein function of these DEOs to test if functional diversification occurred and is associated with differentially expression, as in the case of Tin2. We converted the open reading frame of the candidate effectors in *U. maydis* to their respective *S. reilianum* orthologs. For the generation of *in situ* seamless mutants we used a CRISPR-Cas9 approach that allowed marker-free selection of gene-replacement mutants in *U. maydis* (Fig. 5b). In each case, an sgRNA was designed to target the effector gene to induce a double strand break, which was then repaired by homology-directed repair with a donor plasmid containing the *S. reilianum* effector ortholog, flanked by 1 kb homolog arms (Fig. 5b). This CRISPR-Cas9-mediated gene conversion eliminated potential

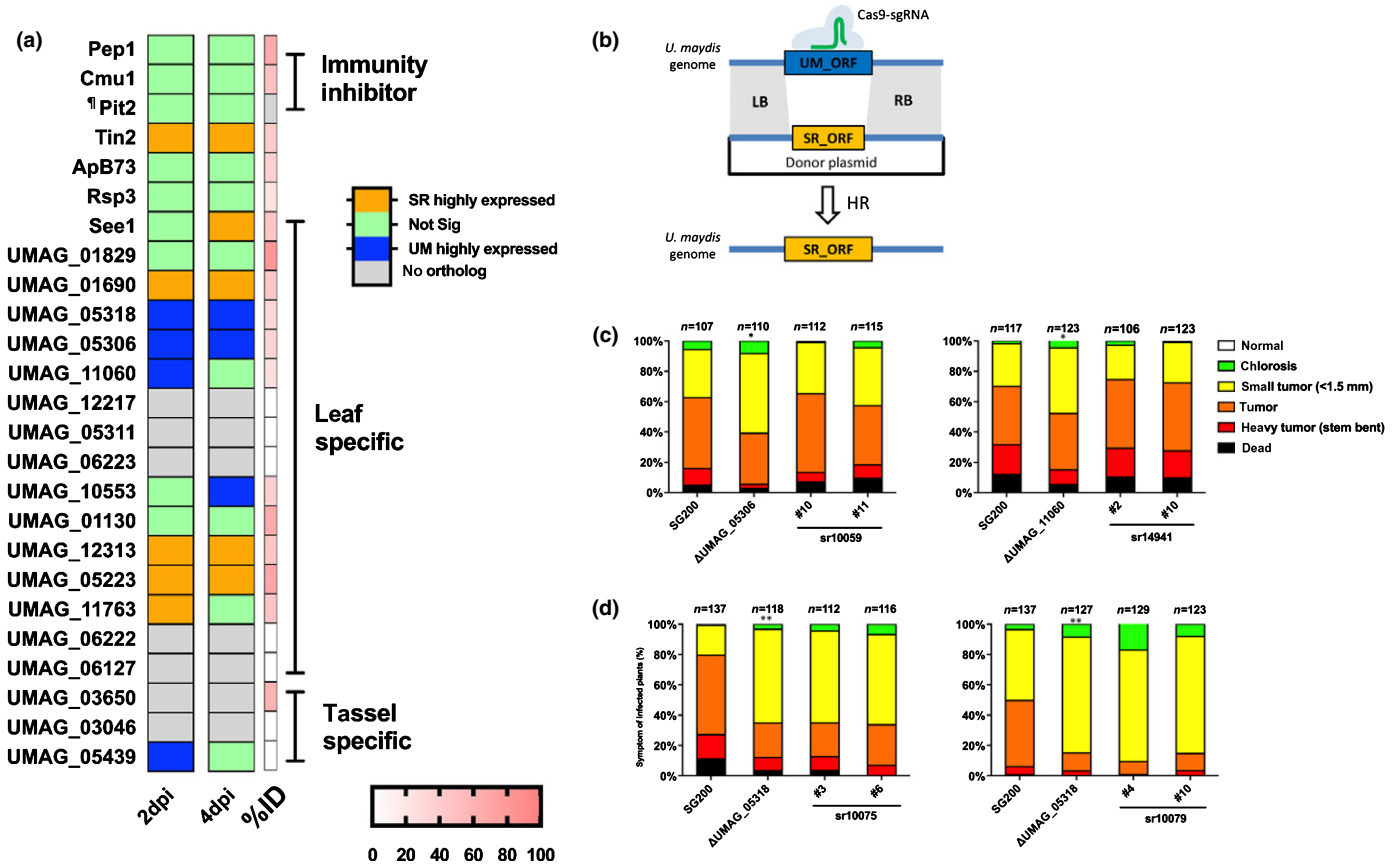

Fig. 5 Effector orthologs between *Ustilago maydis* and *Sporisorium reilianum* are functional conserved. (a) Heatmap showing the expression difference and protein sequence identity of effector ortholog pairs which have been knocked out and subjected to virulence tests in *U. maydis*. ¶, Pit2 only conserved in the functional PID motif which failed to be detected during ortholog identification; %ID, not shown. (b) Scheme demonstrating the process of CRISPR-Cas9-mediated gene conversion. sgRNA was designed to target the coding sequence of the *U. maydis* effector, and the donor plasmid contains *S. reilianum* ortholog coding sequence flanked by a 1 kb homologous region. (c, d) Disease scoring results of gene conversion mutants and corresponding *U. maydis* knockout mutant. (c) Conversion of UMAG_05306 and UMAG_11060 to the *S. reilianum* orthologs did not affect the virulence in SG200. (d) Conversion of UMAG_05318 to its *S. reilianum* ortholog Sr10075 or Sr10079, the paralog of Sr10075, compromised SG200 virulence, and show reduced virulence as the deletion mutant. Experiments were performed in three independent biological replicates. *n* indicates the total number of plants tested in three replicates. Significance as determined by Student's *t*-test with disease index: *, $P < 0.05$; **, $P < 0.01$. UM, *U. Maydis*; SR, *S. Reilianum*; LB, left border; RB, right border; HR, homologous recombination.

ectopic effects and allowed expression of the orthologs under the native *in situ cis*- or *trans*-regulatory elements.

In maize infection assays, conversion of UMAG_05306 and UMAG_11060 to their *S. reilianum* orthologs did not affect the virulence compared to the *U. maydis* progenitor strain SG200, while the deletion of these two effectors resulted in significantly reduced virulence (Fig. 5d). However, conversion of UMAG_05318 into its *S. reilianum* ortholog Sr10075 significantly reduced virulence to a similar level of the deletion mutant, which indicates that Sr10075 might have a different function, or has lost its activity during speciation (Fig. 5d). In *S. reilianum*, Sr10075 has a paralog Sr10079, which has only low sequence identity to UMAG_05318 and therefore was not detected in our one-to-one ortholog analysis (Fig. S6a,b). Deletion of Sr10079 was previously shown to reduce virulence in *S. reilianum* (Ghareeb *et al.*, 2019). Conversion of UMAG_05318 into Sr10079 resembled the virulence defect of the UMAG_05318 deletion mutant. Thus, neither Sr10075 nor Sr10079 can functionally replace UMAG_05138, suggesting the potential functional divergence of the orthogroup *UMAG_05318/Sr10075 – Sr10079* during speciation, which identifies them as an interesting objective to test for possible neofunctionalization. However, it needs to be stated at this point that the production and stability of these proteins in *U. maydis* was not directly confirmed in this study.

To further confirm and investigate how promoters regulate the expression of DEOs, we used the CRIPSR-Cas9-mediated conversion approach to replace the promoter of the *U. maydis*-induced DEOs *UMAG_11060* and *UMAG_05306* with their respective *S. reilianum* counterparts in SG200 (Fig. S7a), and qRT-PCR was used to detect the chimeric gene expression. Infected samples at the time points when the differential expression levels were detected in RNA-seq (2 dpi for *UMAG_11060* and 4 dpi for *UMAG_05306*, Fig. 5a) were tested. Expressing *UMAG_11060* under $Pro^{sr14941}$ showed lower expression at 2 dpi (Fig. S7b), which was similar to our RNA-seq analysis, indicating the *cis*-regulatory elements in the promoter region determined their differential expression. However, $Pro^{Sr10059}$ surprisingly drove a similar expression level of *UMAG_05306* at 4 dpi in the *U. maydis* genetic background, suggesting a *trans* regulator or distal enhancer, which is variable between the two species, controls the differential expression (Fig. S7b).

Together, the results of the gene conversion approach showed that functional diversification of effector orthologs can be caused on two levels: the differential expression of orthologous genes that encode functionally conserved proteins (as found for orthogroups UMAG_05306/Sr10059 and UMAG_11060/Sr14941); the functional divergency of effectors on the protein level, which are further fine-tuned on transcription levels for functional adaption; and the regulation of DEO is complicated, involving both *cis*- and *trans*-regulatory elements.

Discussion

In this study we applied a cross-species transcriptome comparison between *U. maydis* and *S. reilianum* to investigate transcriptomic changes of effector orthologs during maize leaf infection.

A crucial challenge in cross-species RNA-seq is to determine the comparable fungal states for sampling to exclude that the observed differences result from the obvious different disease progression rather than from differential regulation. Sucher *et al.* (2020) collected samples from the edge of disease lesions, which were < 25 mm in diameter, as an indication of comparable infection stage to study the different host responses to *Sclerotinia sclerotiorum* for quantitative disease resistance (Sucher *et al.*, 2020). Time-resolved microscopes were applied in *Fusarium virguliforme* to facilitate RNA-seq to study its transcription alteration on different hosts (Baetsen-Young *et al.*, 2020). In our experiment, we conducted a series of microscopy combined with qPCR-based biomass quantification, and based on these results two biotrophic interaction phases representing a comparable (2 dpi) and a distinct colonization stage (4 dpi) were selected for RNA-seq.

When comparing the induction profiles of effectors in *U. maydis* and *S. reilianum* respectively, we found a significantly higher number of *S. reilianum* effectors showing stable expression between 2 and 4 dpi, while in *U. maydis* more effectors displayed an increasing expression over time. We further focused on the expression of one-to-one effector orthogroups. Overall, 207 out of 335 one-to-one effector orthologs were differentially expressed between *U. maydis* and *S. reilianum* during biotrophic interaction, and the number of effector DEOs increased at 4 dpi. The difference in effector *in planta* induced expression patterns and differential regulation of effector DEOs are consistent with the appearance of different symptoms. At 4 dpi, *U. maydis* triggers local host cell proliferation to produce tumors, which results in the high expression of effectors related to tumorigenesis; by contrast, *S. reilianum* sustaining stable expression levels of effectors probably reflects maintenance of the biotrophic interaction in already colonized tissue and continuation of systemic proliferation through the leaf veins without inducing structural changes in the host tissue. It is of note that the DEOs, especially the effector DEOs we identified from the 2 dpi samples, may reveal molecular differences between *U. maydis* and *S. reilianum*, which at this early stage of infection are not yet visible on the microscopic level, or in biomass quantification experiments. However, to what extent such early regulated genes can serve as markers for pathogenic development, or to predict further fungal invasive growth, still needs to be investigated.

Not only effector genes, but also pathogenesis-related TFs were found to be differentially regulated, which in turn probably plays a crucial role in the control of effector DEO expression. TFs Rbf1 and Biz1 were highly expressed in *S. reilianum*, which may be responsible for the expression of *S. reilianum*-enriched effector DEOs. However, the abundance of TF Hdp2 transcripts, which is activated by Rbf1 (Heimel *et al.*, 2010b), was more highly expressed in *U. maydis*, indicating the complicated regulation of TF cascade. The promoter conversion of $Pro^{Sr10059}$ showed comparable expression with Pro^{UMAG_05306} in *U. maydis* (Fig. S7), indicating potential *trans* effects such as TFs or distal enhancers might be driving differential regulation while for orthogroup UMAG_11060/sr14941, *cis*-regulatory elements in the promoter region probably determine the expression pattern. Until now, only

few direct target genes of TFs were identified by CHIP-seq in smut fungi (Tollot *et al.*, 2016). Elucidation of the binding motif of these pathogenesis-related TFs combined with the promoter analysis of effector DEOs, which were obtained in this study, could be a starting point for a more detailed study on the specific roles of TFs in the orchestration of effector expression with respect to the adaptation to different infection styles in pathogen speciation.

A major conclusion that can be drawn from the comparative transcriptome analysis is that the expression levels of effector orthologs between *U. maydis* and *S. reilianum* are related to their function: effectors with general virulence functions showed similar expression levels between the two smut fungi, whereas effectors adapted to execute specific processes such as the leaf-tumor specific effectors of *U. maydis* showed significant differences in their transcriptional regulation. UMAG_05306 and UMAG_11060, whose disruption resulted in reduced virulence in *U. maydis*, were functionally conserved with respect to the corresponding *S. reilianum* orthologs. This implies that the functional diversification of effectors can be driven on the transcriptional level and is not necessarily linked to a change in protein function. Recently, Navarrete *et al.* (2021) showed that a cluster of 10 effectors from *U. maydis* interferes with the generation of plant ROS (reactive oxygen species) by different mechanisms, which suggests that phenotypic complementation of effectors does not necessarily imply an identical molecular function, as a host process could be modulated at different levels with similar outcome. By contrast, the orthologs of the Tin2 effector are functional diverse by targeting different paralogs in the same host protein family (Tanaka *et al.*, 2019). Thus, the underlying molecular mechanisms of effector orthologs need further investigation in every individual case to conclude on their functional diversification, or conservation. On the other hand, UMAG_05318 provides another example for the combination of transcriptional alteration and functional diversification at the protein level. While the orthogroups UMAG_05318/Sr10075 or Sr10079 and Umtin2/Srtin2 were expressed during leaf infection in both *U. maydis* and *S. reilianum*, they also show a similar induction pattern during biotrophic interaction. However, when comparing the expression levels of the orthologs, UMAG_05318 is more than 50 times higher than its *S. reilianum* ortholog, while Srtin2 has a 10 times higher expression compared to its respective ortholog in *U. maydis*. Until now, the molecular mechanisms of these differentially regulated organ-specific effectors have not been elucidated in *U. maydis*, and whether they also contribute to virulence in *S. reilianum* is largely unknown. Deletion of Sr10075 alone did not seem to affect virulence, while a knockout Sr10079 reduced *S. reilianum* virulence (Ghareeb *et al.*, 2019). The orthologs of UMAG_05318 were found in *S. scitamineum* and *S. reilianum* f. sp. *reilianum*, but not in *U. hordei* or *M. pennsylvanicum*. Elucidation of the function of this ortholog group and its role in smut fungus evolution and speciation will be our next aim. Interestingly, both effectors showing functional divergence are located in effector cluster 19A (Brefort *et al.*, 2014; Ghareeb *et al.*, 2019), and several *U. maydis* highly expressed effector DEOs were also identified in this cluster from our study. However, the impact of 19A deletion on *S. reilianum* virulence was monitored by the head smut phenotype in the maize inflorescence (Ghareeb *et al.*, 2019). Thus, it is possible that these effectors were highly expressed in a different time/spatial pattern to contribute to the systemic spread of *S. reilianum*. On the other hand, it would be tempting to study if higher expression of cluster 19A effectors could promote the adaptation of *S. reilianum* to leaf infection and eventually even trigger the formation of leaf tumor by this pathogen. Therefore, to elucidate the molecular functions of these DEOs and the identification of their host targets will be instrumental to gain new insights into how the evolution of effector orthologs promotes the emergence of new pathogens. Only a few effectors have been shown to exhibit phenotypic reduction *in U. maydis* (Lanver *et al.*, 2017; Zuo *et al.*, 2019), which limits the virulence DEOs that can be tested in gene conversion experiments. In this study, we identified three effector DEOs as promising candidates for effectors with adapted virulence function during speciation. Despite these findings, one should not neglect the importance of effectors that evolved new functions without fundamental changes in expression patterns. For this, the Pit2 effector is a potential example: SrPit2 showed a similar expression level to UmPit2. However, in a previous study we found that SrPit2 only partially complemented a *U. maydis* pit2 knockout mutant, which was linked to a reduced ability of SrPit2 to inhibit SA-associated PLCPs (Misas Villamil *et al.*, 2019). Furthermore, a recent apoplastic proteome analysis suggested that the composition of PLCPs is different between maize leaf and root (Schulze Hüynck *et al.*, 2019). Given that, in natural conditions, *S. reilianum* infects maize through the root while *U. maydis* infects aerial tissues, one could speculate that Srpit2 adapted to be a more efficient inhibit of root PLCPs, while UmPit2 adapted to inhibit PLCPs in the maize leaf. This reflects that effector function can specialize to different hosts of tissues. In addition to this mechanism of gradual effector specialization, our approach to compare the transcriptional levels of orthologs across two different pathogen species documented that both the transcriptional alteration and functional divergence of effector proteins contribute to pathogen speciation. Including more phylogenetically related smut fungi infecting different hosts, such as *S. reilianum* f. sp. *reilianum* for sorghum and *U. hordei* for barley in future research will provide a more comprehensive map of the emergence and adaptation of pathogenicity effectors during host adaptation and specialization. Besides the plastic regulation of effector orthogroups, one should not ignore the role of species-specific effectors in virulence, which is not discussed in our study. Studies on how pathogens organize their effector warehouse in a temporal and spatial manner to shape pathogenesis, and how these effectors were adapted during speciation will shed a light on the evolution and mechanistic basis of plant–pathogen interactions.

Acknowledgements

This work was funded by the European Research Council under the European Union's Horizon 2020 research and innovation program (consolidator grant conVIRgens, ID 771035), as well as funding by the Deutsche Forschungsgemeinschaft (DFG,

German Research Foundation) under Germany's Excellence Strategy-EXC-2048/1-Project ID: 390686111. WZ and JD are supported by the Research Fellowship Programme for Postdoctoral Researchers of the Alexander von Humboldt Foundation. MT and DKG received support from the LOEWE initiative of the government of Hesse, in the framework of the LOEWE Centre for Translational Biodiversity Genomics (TBG) and the LOEWE Cluster for Integrative Fungal Research (IPF).

Author contributions

WZ and GD designed the research; WZ performed the data analysis, molecular experiment and virulence assay; JRLD and DKG mapped the RNA-seq data; JRLD conducted phylogenetic analysis; DKG and MT provided ortholog expression normalization; WZ and GD wrote the paper with contributions from the other authors. JRLD and DKG contributed equally to this work.

ORCID

Jasper R. L. Depotter https://orcid.org/0000-0002-6950-7438
Gunther Doehlemann https://orcid.org/0000-0002-7353-8456
Deepak K. Gupta https://orcid.org/0000-0001-9732-7533
Marco Thines https://orcid.org/0000-0001-7740-6875
Weiliang Zuo https://orcid.org/0000-0001-5241-1440

Data availability

RNA sequencing data have been submitted to NCBI GenBank and are available under the following accession: BioProject ID PRJNA674747.

References

Baetsen-Young A, Wai CM, VanBuren R, Day B. 2020. *Fusarium virguliforme* transcriptional plasticity is revealed by host colonization of maize versus soybean. *Plant Cell* **32**: 336–351.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.

Beckerson WC, Rodríguez de la Vega RC, Hartmann FE, Duhamel M, Giraud T, Perlin MH. 2019. Cause and effectors: whole-genome comparisons reveal shared but rapidly evolving effector sets among host-specific plant-castrating fungi. *mBio* **10**: e02391-19.

Begerow D, Göker M, Lutz M, Stoll M. 2004. On the evolution of smut fungi on their hosts. In: Agerer R, Piepenbring M, Blanz P, eds. *Frontiers in basidiomycote mycology*. Eching, Germany: IHW-Verlag & Verlagsbuchhandlung, 81–98.

Brefort T, Tanaka S, Neidig N, Doehlemann G, Vincon V, Kahmann R. 2014. Characterization of the largest effector gene cluster of *Ustilago maydis*. *PLoS Pathogens* **10**: e1003866.

De la Concepcion JC, Franceschetti M, Maqbool A, Saitoh H, Terauchi R, Kamoun S, Banfield MJ. 2018. Polymorphic residues in rice NLRs expand binding and response to effectors of the blast pathogen. *Nature Plants* **4**: 576–585.

Depotter JRL, Doehlemann G. 2020. Target the core: durable plant resistance against filamentous plant pathogens through effector recognition. *Pest Management Science* **76**: 426–431.

Depotter JRL, Zuo W, Hansen M, Zhang B, Xu M, Doehlemann G. 2020. Effectors with different gears: divergence of *Ustilago maydis* effector genes is associated with their temporal expression pattern during plant infection. *Journal of Fungi* **7**: 16.

Djamei A, Schipper K, Rabe F, Ghosh A, Vincon V, Kahnt J, Osorio S, Tohge T, Fernie AR, Feussner I *et al.* 2011. Metabolic priming by a secreted fungal effector. *Nature* **478**: 395–398.

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

Flor-Parra I, Vranes M, Kämper J, Pérez-Martín J. 2006. Biz1, a zinc finger protein required for plant invasion by *Ustilago maydis*, regulates the levels of a mitotic cyclin. *Plant Cell* **18**: 2369–2387.

Franceschetti M, Maqbool A, Jiménez-Dalmaroni MJ, Pennington HG, Kamoun S, Banfield MJ. 2017. Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiology and Molecular Biology Reviews* **81**: e00066–e00116.

Ghareeb H, Drechsler F, Löfke C, Teichmann T, Schirawski J. 2015. SUPPRESSOR OF APICAL DOMINANCE1 of *Sporisorium reilianum* modulates inflorescence branching architecture in maize and Arabidopsis. *Plant Physiology* **169**: 2789–2804.

Ghareeb H, Zhao Y, Schirawski J. 2019. *Sporisorium reilianum* possesses a pool of effector proteins that modulate virulence on maize. *Molecular Plant Pathology* **20**: 124–136.

Gijzen M, Ishmael C, Shrestha SD. 2014. Epigenetic control of effectors in plant pathogens. *Frontiers in Plant Science* **5**: 1–4.

Heimel K, Scherer M, Schuler D, Kämper J. 2010a. The *Ustilago maydis* Clp1 protein orchestrates pheromone and b-dependent signaling pathways to coordinate the cell cycle and pathogenic development. *Plant Cell* **22**: 2908–2922.

Heimel K, Scherer M, Vranes M, Wahl R, Pothiratana C, Schuler D, Vincon V, Finkernagel F, Flor-Parra I, Kämper J. 2010b. The transcription factor rbf1 is the master regulator for b-mating type controlled pathogenic development in *Ustilago maydis*. *PLoS Pathogens* **6**: 17–18.

Hemetsberger C, Herrberger C, Zechmann B, Hillmer M, Doehlemann G. 2012. The *Ustilago maydis* effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. *PLoS Pathogens* **8**: e1002684.

Hemetsberger C, Mueller AN, Matei A, Herrberger C, Hensel G, Kumlehn J, Mishra B, Sharma R, Thines M, Hückelhoven R *et al.* 2015. The fungal core effector Pep1 is conserved across smuts of dicots and monocots. *New Phytologist* **206**: 1116–1126.

Irieda H, Inoue Y, Mori M, Yamada K, Oshikawa Y, Saitoh H, Uemura A, Terauchi R, Kitakura S, Kosaka A *et al.* 2019. Conserved fungal effector suppresses PAMP-triggered immunity by targeting plant immune kinases. *Proceedings of the National Academy of Sciences, USA* **116**: 496–505.

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

Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O *et al.* 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **444**: 97–101.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software v. 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lanver D, Berndt P, Tollot M, Naik V, Vranes M, Warmann T, Münch K, Rössel N, Kahmann R. 2014. Plant surface cues prime *Ustilago maydis* for biotrophic development. *PLoS Pathogens* **10**: e1004272.

Lanver D, Müller AN, Happel P, Schweizer G, Haas FB, Franitza M, Pellegrin C, Reissmann S, Altmüller J, Rensing SA *et al.* 2018. The biotrophic development of *Ustilago maydis* studied by RNA-seq analysis. *Plant Cell* **30**: 300–323.

Lanver D, Tollot M, Schweizer G, Lo Presti L, Reissmann S, Ma L-S, Schuster M, Tanaka S, Liang L, Ludwig N *et al.* 2017. *Ustilago maydis* effectors and their impact on virulence. *Nature Reviews Microbiology* **15**: 409–421.

Laurie JD, Ali S, Linning R, Mannhaupt G, Wong P, Güldener U, Münsterkötter M, Moore R, Kahmann R, Bakkeren G *et al.* 2012. Genome

comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. *Plant Cell* **24**: 1733–1745.

Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* **13**: 2178–2189.

Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* **47**: e47.

Lin JS, Happel P, Kahmann R. 2021. Nuclear status and leaf tumor formation in the *Ustilago maydis*–maize pathosystem. *New Phytologist* **231**: 399–415.

Ma L-S, Wang L, Trippel C, Mendoza-Mendoza A, Ullmann S, Moretti M, Carsten A, Kahnt J, Reissmann S, Zechmann B *et al.* 2018. The *Ustilago maydis* repetitive effector Rsp3 blocks the antifungal activity of mannose-binding maize proteins. *Nature Communications* **9**: 1711.

Martinez C, Roux C, Jauneau A, Dargent R. 2002. The biological cycle of *Sporisorium reilianum f.sp. zeae*: an overview using microscopy. *Mycologia* **94**: 505–514.

Matei A, Ernst C, Günl M, Thiele B, Altmüller J, Walbot V, Usadel B, Doehlemann G. 2018. How to make a tumour: cell type specific dissection of *Ustilago maydis*-induced tumour development in maize leaves. *New Phytologist* **217**: 1681–1695.

Misas Villamil JC, Mueller AN, Demir F, Meyer U, Ökmen B, Schulze Hüynck J, Breuer M, Dauben H, Win J, Huesgen PF *et al.* 2019. A fungal substrate mimicking molecule suppresses plant immunity via an inter-kingdom conserved motif. *Nature Communications* **10**: 1576.

Morita T, Koike H, Koyama Y, Hagiwara H, Ito E, Fukuoka T, Imura T, Machida M, Kitamoto D. 2013. Genome sequence of the basidiomycetous yeast *Pseudozyma antarctica* T-34, a producer of the glycolipid biosurfactants mannosylerythritol lipids. *Genome Announcements* **1**: e00064–e00113.

Mueller AN, Ziemann S, Treitschke S, Aßmann D, Doehlemann G. 2013. Compatibility in the *Ustilago maydis*–maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2. *PLoS Pathogens* **9**: e1003177.

Navarrete F, Grujic N, Stirnberg A, Saado I, Aleksza D, Gallei M, Adi H, Alcântara A, Khan M, Bindics J, *et al.* 2021. The Pleiades are a cluster of fungal effectors that inhibit host defenses. *PLoS Pathogens* **17**: e1009641.

Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* **10**: 417–430.

Redkar A, Doehlemann G. 2016. *Ustilago maydis* virulence assays in maize. *Bio-Protocol* **6**: e1760.

Redkar A, Hoser R, Schilling L, Zechmann B, Krzymowska M, Walbot V, Doehlemann G. 2015a. A secreted effector protein of *Ustilago maydis* guides maize leaf cells to form tumors. *Plant Cell* **27**: 1332–1351.

Redkar A, Villajuana- Bonequi M, Doehlemann G. 2015b. Conservation of the *Ustilago maydis* effector See1 in related smuts. *Plant Signaling & Behavior* **10**: e1086855.

Robinson MD, McCarthy DJ, Smyth GK. 2009. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The genome biology of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology* **56**: 21–40.

Schilling L, Matei A, Redkar A, Walbot V, Doehlemann G. 2014. Virulence of the maize smut *Ustilago maydis* is shaped by organ-specific effectors. *Molecular Plant Pathology* **15**: 780–789.

Schirawski J, Mannhaupt G, Münch K, Brefort T, Schipper K, Doehlemann G, Di Stasio M, Rössel N, Mendoza-Mendoza A, Pester D *et al.* 2010. Pathogenicity determinants in smut fungi revealed by genome comparison. *Science* **330**: 1546–1548.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA *et al.* 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**: 1112–1115.

Schulze Hüynck J, Kaschani F, van der Linde K, Ziemann S, Müller AN, Colby T, Kaiser M, Misas Villamil JC, Doehlemann G. 2019. Proteases underground: analysis of the maize root apoplast identifies organ specific papain-like cysteine protease activity. *Frontiers in Plant Science* **10**: 473.

Schuster M, Kahmann R. 2019. CRISPR-Cas9 genome editing approaches in filamentous fungi and oomycetes. *Fungal Genetics and Biology* **130**: 43–53.

Schuster M, Schweizer G, Reissmann S, Kahmann R. 2016. Genome editing in *Ustilago maydis* using the CRISPR-Cas system. *Fungal Genetics and Biology* **89**: 3–9.

Schuster M, Schweizer G, Kahmann R. 2018. Comparative analyses of secreted proteins in plant pathogenic smut fungi and related basidiomycetes. *Fungal Genetics and Biology* **112**: 21–30.

Schweizer G, Münch K, Mannhaupt G, Schirawski J, Kahmann R, Dutheil JY. 2018. Positively selected effector genes and their contribution to virulence in the smut fungus *Sporisorium reilianum*. *Genome Biology and Evolution* **10**: 629–645.

Selin C, de Kievit TR, Belmonte MF, Fernando WGD. 2016. Elucidating the role of effectors in plant–fungal interactions: progress and challenges. *Frontiers in Microbiology* **7**: 1–21.

Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods in Molecular Biology* **1962**: 227–245.

Sharma R, Mishra B, Runge F, Thines M. 2014. Gene loss rather than gene gain is associated with a host jump from monocots to dicots in the smut fungus *Melanopsichium pennsylvanicum*. *Genome Biology and Evolution* **6**: 2034–2049.

Sharma R, Ökmen B, Doehlemann G, Thines M. 2019. Saprotrophic yeasts formerly classified as *Pseudozyma* have retained a large effector arsenal, including functional Pep1 orthologs. *Mycological Progress* **18**: 763–768.

Skibbe DS, Doehlemann G, Fernandes J, Walbot V. 2010. Maize tumors caused by *Ustilago maydis* require organ-specific genes in host and pathogen. *Science* **328**: 89–92.

Stamatakis A. 2014. RAxML v. 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stirnberg A, Djamei A. 2016. Characterization of ApB73, a virulence factor important for colonization of *Zea mays* by the smut *Ustilago maydis*. *Molecular Plant Pathology* **17**: 1467–1479.

Sucher J, Mbengue M, Dresen A, Barascud M, Didelon M, Barbacci A, Raffaele S. 2020. Phylotranscriptomics of the pentapetalae reveals frequent regulatory variation in plant local responses to the fungal pathogen *Sclerotinia sclerotiorum*. *Plant Cell* **32**: 1820–1844.

Szegedy C, Vanhoucke V, Ioffe S, Shlens J & Wojna Z 2016. Rethinking the Inception Architecture for Computer Vision. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December: 2818–2826.

Tamborski J, Krasileva KV. 2020. Evolution of plant NLRs: from natural history to precise modifications. *Annual Review of Plant Biology* **71**: 355–378.

Tanaka S, Brefort T, Neidig N, Djamei A, Kahnt J, Vermerris W, Koenig S, Feussner K, Feussner I, Kahmann R. 2014. A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. *eLife* **3**: e01355.

Tanaka S, Schweizer G, Rössel N, Fukada F, Thines M, Kahmann R. 2019. Neofunctionalization of the secreted Tin2 effector in the fungal pathogen *Ustilago maydis*. *Nature Microbiology* **4**: 251–257.

Taniguti LM, Schaker PDC, Benevenuto J, Peters LP, Carvalho G, Palhares A, Quecine MC, Nunes FRS, Kmit MCP, Wai A *et al.* 2015. Complete genome sequence of *Sporisorium scitamineum* and biotrophic interaction transcriptome with sugarcane. *PLoS ONE* **10**: e0129318.

Thines M. 2019. An evolutionary framework for host shifts – jumping ships for survival. *New Phytologist* **224**: 605–617.

Tollot M, Assmann D, Becker C, Altmüller J, Dutheil JY, Wegner C-E, Kahmann R. 2016. The WOPR protein Ros1 is a master regulator of sporogenesis and late effector gene expression in the maize pathogen *Ustilago maydis*. *PLoS Pathogens* **12**: e1005697.

Torres DE, Oggenfuss U, Croll D, Seidl MF. 2020. Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fungal Biology Reviews* **34**: 136–143.

Toruño TY, Stergiopoulos I, Coaker G. 2016. Plant–pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. *Annual Review of Phytopathology* **54**: 419–441.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. 2011. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **6**: 16765.

Zahiri A, Heimel K, Wahl R, Rath M, Kämper J. 2010. The *Ustilago maydis* forkhead transcription factor Fox1 is involved in the regulation of genes required for the attenuation of plant defenses during pathogenic development. *Molecular Plant–Microbe Interactions* **23**: 1118–1129.

Zuo W, Depotter JR, Doehlemann G. 2020. Cas9HF1 enhanced specificity in *Ustilago maydis*. *Fungal Biology* **124**: 228–234.

Zuo W, Ökmen B, Depotter JRL, Ebert MK, Redkar A, Misas Villamil J, Doehlemann G. 2019. Molecular interactions between smut fungi and their host plants. *Annual Review of Phytopathology* **57**: 411–430.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Symptoms of *U. maydis*- and *S. reilianum*-infected leaves.

Fig. S2 qPCR confirms the expression of effector genes from *U. maydis* and *S. reilianum*.

Fig. S3 Differential regulation of orthologs between *U. maydis* and *S. reilianum* from different samples.

Fig. S4 Expression of pathogenic development-related transcription factors between *U. maydis* and *S. reilianum* during biotrophic growth from RNA-seq data.

Fig. S5 The distribution of effector DEOs on chromosomes.

Fig. S6 Sequence comparison of UMAG_05318, Sr10075 and Sr10079.

Fig. S7 The regulation of *S. reilianum* promoters in *U. maydis* by CRIPSR-Cas9-mediated promoter conversion.

Table S1 Oligo sequences.

Table S2 Log_2 CPM (count per million) of fungal reads.

Table S3 Relative ortholog expression data of 6005 one-to-one orthologs between *U. maydis* and *S. reilianum*.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.