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

Weiliang Zuo[1], Deepak K Gupta[2,3,4,*], Jasper RL Depotter[1,*], Marco Thines[2,3,4], Gunther Doehlemann[1]

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

*: D. G and J.D contributed equally to this work

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

Summary

- The constitution and regulation of effector repertoires determines and shapes the outcome of the interaction with the host. *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 to pathogen evolution and host adaption.
- We tracked the infection progress of *U. maydis* and *S. reilianum* in maize leaves, characterized two distinct infection stages for cross species RNA-sequencing analysis and identified 207 out of 335 one-to-one effector orthologs being differentially regulated during host colonization, while transcriptional plasticity of the effector orthologs correlated with the distinct disease development strategies.
- By 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 contributes to species specific adaptation and symptom development. Interestingly, another differentially expressed orthogroup (*UMAG_05318/sr1007)* showed diverged protein function during speciation, providing a possible case for neofunctionalization.
- Together, we showed the diversification of effector genes in related pathogens can be caused both by plasticity on the transcriptional level, as well as through neofunctionalization of the encoded effector proteins.

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

Introduction

During symbiosis, phytopathogenic microbes secret effector proteins to facilitate a compatible interaction with their hosts. Effectors hold various functions, such as the suppression of host immunity and host metabolism manipulation to promote host infection. Co-evolution of pathogens with their hosts is driven by an arms race causing selection pressure on perceived effector genes against their recognition by the host immune system (Jones & Dangl, 2006).

Polymorphisms of effectors are diverse, including presence/absence variation, amino acid substitution, epigenetic modification and transcriptional plasticity, which together determine the outcome of the plant-pathogen interaction (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 and effector prediction and variation amongst phylogenetic related pathogens (Raffaele & Kamoun, 2012; Sánchez-Vallet *et al.*, 2018); functional elucidation of a 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, information is available on the transcriptional regulation of effector orthologs across related pathogen species is scarce.

Smut fungi are a group of fungal plant pathogens, consisting of more than 1,500 species, which infect mainly grasses including major cereals such as maize, sorghum, wheat, barley and sugar cane (Begerow *et al.*, 2004). A characteristic feature of many smut fungi is their dimorphic growth, where distinct growth states are observed: saprophytic growth as yeast, and filamentous hyphae which initiate the biotrophic interaction with the host (Kämper *et al.*, 2006a). Smuts often infect their host through the root or coleoptile (Martinez *et al.*, 2002; Laurie *et al.*, 2012), from where the fungi proliferate without evident symptoms during the early vegetative growth stage of the host. When plants switch into the reproductive phase, fungal hyphae residing in floral may produce massive black/ brown teliospores (the notorious “smut” phenotype). This type of infection style is also found for the smut fungus *Sporisorium reilianum*, which infects either sorghum (*S.□reilianum f. sp. reilianum*) or maize (*S.□reilianum f. sp. zeae*) and cause head smut disease. Closely related to this is *Ustilago maydis*, which causes common maize smut disease (Kämper *et al.*, 2006a; Schirawski *et al.*, 2010). While *U. maydis* has the ability to cause disease symptoms all the aerial maize tissues, *S.□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 initials. *Ustilago maydis* causes locally restricted plant tumors within a short time of less than 2 weeks under laboratory conditions. Comparative genome analysis of *U. maydis* with *S. reilianum* revealed that the two maize smuts share high genome similarity with regard to gene number and synteny (Schirawski *et al.*, 2010). Furthermore, previous studies suggested that, similar to *U. maydis*, *S. reilianum* can efficiently infect maize leaves and also leads to the systemic disease in floral organs (Ghareeb *et al.*, 2015), rendering the two species a well-suited model to study regulation and function of orthologous genes.

In *U. maydis*, the expression of effector genes is precisely regulated via a network of hierarchical transcriptional factors (TF), which modulates activity of the transcription of effectors genes at distinct stages of infection (Skibbe *et al.*, 2010; Lanver *et al.*, 2018). The mating type locus bE/bW encodes a heterodimer TF, which not only regulates the expression of 38 effectors, but also trigger the expression of several TFs including a C2H2 zinc finger TF Rbf1 (regulator of b-filament) (Heimel *et al.*, 2010b,a), which then turn on the TF Hdp2 (homeodomain transcription factor 2) and Biz1 (*b*-dependent zinc finger protein), together activating the expression of effectors important for initialing 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 the host defense after establishment of biotrophic interaction (Zahiri *et al.*, 2010) and the WOPR TF Ros1 (Tollot *et al.*, 2016) regulate effector gene expression related to fungal sporogenesis in maturated tumors. A comprehensive set of 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). The expression of effectors in *U. maydis* was also found to be regulated in an organ-specific manner (Skibbe *et al.*, 2010), and a more recent study using RNA-sequencing from laser capture micro-dissected cells revealed a host cell-type specific expression pattern for several effector genes (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 the molecular level (Lanver *et al.*, 2017). Examples are the apoplastic effectors Pit2 and Pep1, which inhibit the maize apoplastic Papain-like cysteine proteases (Mueller *et al.*, 2013; Misas Villamil *et al.*, 2019) and peroxidase activities (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 SA and lignin metabolism, respectively. The leaf-specific effector See1 re-activates 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 anti-fungal protein (Ma *et al.*, 2018). In contrast, only one *S. relianum* specific effector SAD1, which suppress the apical dominance of infected maize to increase branch number is functional characterized in *S. relianum* (Ghareeb *et al.*, 2015).

Comparative genomic analysis from phylogenetic related plant pathogens suggesting the effector repertories of smut fungi can be classified as core and accessory effector based on the conservation level (Schuster *et al.*, 2018; Beckerson *et al.*, 2019). Core effectors are conserved in all phylogenetic related pathogen and may be involved in conserved biological process vital for infection, leading to the stabilized of such effectors during speciation (Hemetsberger et al., 2015; Irieda et al., 2019; Thines 2019). Complementary, accessory effectors are less conserved and only found in few or individual pathogen specifies were supposed to have more subtle/specific functions in virulence. Previous studies suggested that effector orthologs between *U. maydis* and *S. relianum* are more

functionally conserved compared to *Ustilago hordei* and/or *Melanopsichium pennsylvanicum* (Redkar *et al.*, 2015b; Stirnberg & Djamei, 2016).

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) . Cas9 endonuclease generates DNA double-strand break in a specific target that is recognized by the sgRNA (single guide RNA). This double strand break can be repaired by non-homologous end-joining (NHEJ) which eventually results in a gene knock-out mutant. Alternatively, homology direct repair (HDR) with a donor template can be used for the generation of knock-in or gene conversion mutants. In *U. maydis*, the CRISPR-Cas9 is achieved by expressing the codon-optimized Cas9 protein and sgRNA in an autonomous replication plasmid (Schuster *et al.*, 2016, 2018), which shows high efficiency to generate a selection marker free genome modification mutant. The off-target is further reduced by application of high fidelity variant Cas9HF1 (Zuo *et al.*, 2020).

In this study, we examined the growth of *U. maydis* and *S. reilianum* during leaf infection for recording similar and divergent biotrophic interaction with maize. We investigated the regulation of orthologous genes with an emphasis on effector genes.

Materials and Methods

Strains, growth conditions and plant infection

The mating compatible isolates of *U. maydis* FB1, FB2 and *S. reilianum* SRZ1, SRZ2 were used for RNA-seq, and *U. maydis* solo-pathogenic strain SG200 and its respective knock-out and ortholog conversion mutants were used for virulence tests. All strains were grown in YEPS light liquid medium with 200 rpm shaking or potato dextrose agar (PD, Difco) plate at 28°C. *Escherichia coli* strain Top10 was used for cloning purpose and grown in dYT liquid medium or YT agar plate with supplement of corresponding antibiotics. Maize variety Early Golden Bantam was used for infection. The plants were grown in a controlled condition of 16 hr light at 28 °C and 8 hr dark at 22°C. 7 days-old seedlings were infected with a mixture of compatible *U. maydis*/*S. reilianum* isolates or SG200, and disease symptoms were scored at 12 days after infection.

Staining and microscopy

0.1% tween-20 was added in the inoculum for microscopy purposes. To visualize the appressorium, leaves were stained 20 hours post infection with calcofluor solution (100 µg/ml) 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.*, 2009a). The microscopy was done on a Nikon Eclipse Ti Inverted Microscope with the Nikon NIS-

ELEMENTS software (Düsseldorf, Germany) and the photos were taken by HAMAMATSU camera. Confocal microscopy was done by Leica TCS SP8 confocal laser scanning microscope (Leica, Wetzlar, Germany) and the following filters were used excitation 458 nm and emission 470-490 nM for WGA-AF488 and excitation 561 nm and emission 590-603 nm for propidium Iodide.

RNA preparation and RNA-seq

2-cm length of 3rd leaf sections from more than 15 individual plants were collected for each sample. The compatible haploid *U. maydis* and *S. reilianum* cells from cultures with OD_{600} around 0.8 were spin down and mixed in a 1:1 ratio as axenic culture control. The plant tissues and cell pellets were grinded into fine powder with liquid nitrogen and the RNAs were prepared by TRizol (Thermo Fisher, Waltham, USA) according to the manufacture's protocol and followed by DnaseI digestion (Thermo Fisher, Waltham, USA). The RNA libraries were prepared by Illumina TruSeq Stranded mRNA kit (Illumina, San Diego, USA), and paired-end sequencing was performed on the HiSeq4000 platform to produce 2×75 bp long reads at the Cologne Center for Genomics (Cologne, Germany).

Data analysis

For the gene expression analysis of *U. maydis* and *S. reilianum* individually, reads with filtered using the Trinity software (v2.9.1) option trimmomatic under the standard settings (Grabherr *et al.*, 2013). They were mapped to reference a reference assembly using Bowtie 2 (v2.3.5.1) with the first 15 nucleotides on the 5'-end of the reads being trimmed (Langmead & Salzberg, 2012). The reference genome was either the genome assembly of *U. maydis* (Kämper *et al.*, 2006b) or *S. reilianum* (Schirawski *et al.*, 2010) combined with that of *Z. mays* B73 version 3 (Schnable *et al.*, 2009). Reads were counted to the *U. maydis* and *Z. mays* loci using the R package Rsubread (v1.34.7) (Liao *et al.*, 2019). edgeR package v3.26.8 was used for statistical analysis of differential gene expression and pairwise comparison was conducted using Generalized linear models (glm) (Robinson *et al.*, 2009). Genes with log2 fold change >1 and $p<0.05$ were considered as differentially regulated between treatments. The induced patterns of effectors were defined based on the differentially regulation relationships between AC, 2 dpi and 4 dpi samples with following criterias: pattern1 as 4dpi >2dpi >AC; pattern 2 as 4 dpi > AC, 2dpi >AC and 4 dpi≅ 2 dpi; pattern 3 as 2 dpi > AC, 2 dpi >4 dpi and 4 dpi≅ AC; pattern 4 as 4 dpi >AC, 2 dpi >AC and 2 dpi≅4 dpi; pattern 5 as 2 dpi ≅AC, 4dpi >AC and 4 dpi >2 dpi; where ">" indicated log2 fold change >1 and $p<0.05$ and "≅" indicated log2 fold change <1, $p<0.05$ or log2 fold change >1 but $p>0.05$.

A phylogenetic tree was constructed including the smut species *U. hordei* (Laurie *et al.*, 2012), *M. pennsylvanicum* (Sharma *et al.*, 2014), *S. scitamineum* (Taniguti *et al.*, 2015), *S. reilianum* (Schirawski *et al.*, 2010) and *U. maydis* (Kämper *et al.*, 2006a). *Moesziomyces antarcticus* was used as an outgroup (Morita *et al.*, 2013). The phylogenetic tree was constructed based on 1643 Benchmarking

Universal Single-Copy Orthologs (BUSCOs) from the database "basidiomycota_odb10" that are present in single copy in all members of the tree (Seppey *et al.*, 2019). For every gene, orthologs were aligned using MAFFT (v7.464) option "--auto" (Katoh & Standley, 2013). These aligned gene sequences were then concatenated for every species and used for tree construction using RAxML (v8.2.11) with 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 *Ustilago maydis* and *Sporisorium reilianum* as calculated using OrthoMCL. 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 : $Relative\ Expression\ of\ Gene\ (X) = \frac{\frac{\sum_1^n Coverage\,(X) \times (N-n)}{n}}{(M-m)/(N-n)}$, Where n= Length of the gene (X) , N = Total number of bases in all genes ; m= Number of bases mapped to gene (X) ; M= Number of bases mapped to all genes. *Student t-test* and/or one sample *t-test* were used for significance test with Benjamini-Hochberg *p* value correction for multiple comparison. Orthologs with fold change >3 and adjust $p<0.05$ were considered as differentially regulated ortholog between *U. maydis* and *S. relianum*.

Gene ontology and GO enrichment analysis

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

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

CRISPR-Cas9 mediated genome editing was used to generate the effector ortholog conversion mutants in *U. maydis* solo-pathogenic strain SG200. An sgRNA was designed to target the coding region of the effector to generate the double strand break, which was then 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 donor plasmid, 1 kb fragments flanking the effector coding region and the respective *S. reilianum* ortholog open reading frame were amplified by Phusion DNA polymerase with primers (Sigma Aldrich, St. Louis, US) listed in **Table. S1** and cloned into pAGM1311 vector by Gibson assembly (New England Biolabs, Ipswich, USA). The CRISPR plasmid and donor plasmid were co-transformed into SG200 protoplast and mutants were isolated and confirmed by southern blot (not showed).

The conversion mutants and corresponding effector knockout mutants (Schilling *et al.*, 2014) were used for infection. Disease scoring was done as previous described at 12 dpi. 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 were 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 *t-test* was used for significance test of 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 tracked infection and fungal growth inside the maize leaves to record the milestone events during infection and define appropriate timepoints for analysis (Fig. **1**). 18-24 hours post infection, the sporidia of both pathogens germinated on the leaf surface, compatible cells mated, and dikaryotic hyphae formed appressoria to penetrate the leaf surface (Fig. **1a**). At 2 days post infection (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 alignment to our previous report (Matei *et al.*, 2018), host bundle sheath cells underwent *de novo* cell division in *U. maydis* infected leaves, indicating the initiation of tumorigenesis. In contrast, *S. reilianum* further accumulated in the leaf vein and showed aggregation as a string of round cells (Fig. **1c**). At 4 dpi, the first visible small tumor structures can be observed on *U. maydis* infected leaves, while *S. reilianum* infected leaves showed individual events of necrosis (Fig. S**1a**) and more frequently cell aggregation (Fig. **1d**, Fig. S**1b**). In parallel to 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. relianum* 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 3 dpi (Fig. **1e**).

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

Based on our microscopic and biomass analysis, we decided to collect 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 axenic culture (AC) of both pathogens were used for RNA extraction.

In total, more than 670M 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, respectively. The changes of fungal transcripts in the RNA-seq samples confirmed the biomass quantification results by qPCR. At 2 dpi, around 2% of fungal transcripts were mapped from both pathogen, and at 4 dpi, *S. relianum* reads increased to 12.5% in the infected samples, which is significantly higher compared to *U. maydis*

infected tissues (7.5%) (Fig. **1f**). Principal component analysis (PCA) analysis of *U. maydis* and *S. reilianum* samples showed that three biological replications from different conditions form distinct clusters, and variations between different conditions were larger than the variation of biological replications within the same condition (Fig. **2a**). To further analyze the RNA-seq data, we normalized reads 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 out of 6765 *U. maydis* genes and 6456 out of 6673 *S. reilianum* genes were used to identify DEGs. Compared to axenic culture, dramatic changes in gene expression were observed during seedling infection. Around 19.3-24.2% genes were up-regulated (log2 fold change >1, $p<0.05$) at 2 dpi and 4 dpi in both pathogens, while 12.1-19.1% genes were down-regulated (Fig. **2b**). We detected 435 out of 467 effector genes from *U. maydis* and 454 out of 489 from *S. reilianum* were expressed. As expected, most effector genes were specifically expressed in the biotrophic growth stage (311 in *U. maydis* and 307 in *S. reilianum*, respectively) and showed higher degrees of transcriptional induction compared to non-effector genes (Fig. **2c, d**), which reflects the explicit role of effectors in host infection.

We clustered effector genes into all five possible expression patterns between three conditions (axenic culture, 2 dpi, 4 dpi) (Fig. **3a**). This classified groups of effector genes (i) showing constitutively increasing induction (pattern 1), (ii) induced *in-planta* on a stable level (pattern 2), (iii) being only transiently expressed (pattern3), (iv) dropping at 4 dpi (pattern 4), (v) 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 (pattern 1, 2 and 4), while only 14 *U. maydis* effectors and 8 *S. relianum* effectors were transient induced at 2 dpi (pattern 3) and similar number of effectors were only required in the 4 dpi (23 from *U. maydis* vs. 25 from *S. reilianum* as pattern 5) (Fig. **3b**). The expression patterns are found to reflect previously observed virulence function of known effector genes. For example, the core effector *UmPep1*(*UMAG_01987*) (Doehlemann *et al.*, 2009b; 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 to pattern 1 cluster. Interestingly, *S. reilianum* orthologs of these characterized effectors displayed different induction patterns, keeping a steady level between 2 and 4 dpi. In particular, we identified 100 out 467 of *U. maydis* effectors being induced in pattern 1, which is significantly higher compared to *S. reilianum* (43/488), whereas 184 out of 488 *S. reilianum* effector genes were showing stable expressing levels, which is higher as compared to *U. maydis* (136/467) (Fig. **3b**). This enrichment of different induction patterns of effector genes suggested different requirements in disease for the two species.

Differential regulation of effector orthologues 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. This means that for avoiding 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 analysis based on the relative expression of 6005 ortholog pairs confirmed that plant associated samples form distinct clusters on time and species levels. In contrast, the axenic cultured samples of the different species have minor differences and clustered together (Fig. **4a**). Pearson correlation analysis displayed similar results and suggested the difference between samples increased at 4 dpi compared to 2 dpi, as the correlation decreased (Fig. S**2a**).

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**). The largest number of DEOs (835) was detected from axenic cultured samples (Fig. **4b**), and around 30 and 50% of these DEOs from *U. maydis* and *S. reilianum* were only confined in this stage (Fig. S**2b)**. 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 likely reflects the significant differences in disease development between the two pathogens, i.e. initiation of tumorigenesis by *U. maydis* versus 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 gene ontology terms, except for 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 in the regulation of effector DEOs, particularly at 4 dpi, differential expression was found for virulence related TFs Rbf1, Biz1, Hdp2, Fox1 and Nlt1 (Fig. **S3**). TF Rbf1 was highly induced in *S. relianum*, especially at 4 dpi, the transcript levels were 17 fold higher than that in *U. maydis*, which may lead to the high abundance of Biz1 in *S. relianum* (Fig. **S3**). 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 highly expressed in *U. maydis* during tumorigenesis (Fig. **S3**). Besides, the TF ROS1 showed similar expression level in both species, which may reflect its dedicated function in sporogenesis, a process which needs to be initiated in both pathogens for completing the pathogenic like cycle (Fig. **S3**). Cluster 19A (Fig. **S4a**) is the largest effector cluster, which is a crucial determinant of virulence in both two 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. **S4b**).

In a previous study, Schuster et al. analyzed 12 smut related basidiomycete and suggested that effector repertories of smut fungi comprise sets of core and accessory effector (Schuster *et al.*, 2018). Based on this information, we classified the 335 one-to one effector pairs into plant smut core effectors, which are present in all 4 monocot smut fungi including *Ustilago hordei* and *Sporisorium scitamineum*, and accessory effectors, which present in *U. maydis* and *S. relianum* 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. relianum*, 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 the 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 between two smut fungi by ortholog complement assay, and also these genes show comparable expression levels (Fig. **5a**). Also 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 4 dpi (Fig. **5a**). Another effector DEO which caught our attention is Tin2. SrTin2 neofunctionlized 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

neofunctionalisation, as the transcriptional level divergence could be involved to promote the functional divergence.

From a set of 20 organ-specific effectors, we found 9 effector DEOs and 8 of them were leaf-specific effectors, from which 4 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**), however, the transcripts levels of the orthologs in *U. maydis* were at least 4.8 folds (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. relianum* orthologs (Fig. **S3**). These effectors are specifically required for the induction *U. maydis* induced leaf tumors (Schilling *et al.*, 2014). Thus, their low expression in *S. relianum* might suggest that these genes are of lesser importance for the basal establishment of biotrophic. We assessed the protein function of these DEOs to test if neofunctionalization occurred and associated with differentially expression as we observed 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 HDR with a donor plasmid containing the *S. reilianum* effector ortholog, flanked by 1kb homologs arms (Fig. **5b**). This CRISPR-Cas9 mediated gene conversion eliminated potential ectopic effects and keeps ortholog expressing under the original *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, converted UMAG_05318 into its *S. reilianum* ortholog sr10075 significantly reduced virulence to the similar level of the knockout mutant, which indicates that sr10075 might have a different function, or 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. **S5a, b**). However, also sr10079 converted mutants resembled the virulence defect of the UMAG_05318 knockout mutant, suggesting the functional divergence of UMAG_05318/sr10075.

Together, the results of the gene conversion approach showed that functional diversification of effector orthologs can be caused on two levels: (i) the differential expression of orthologous genes which encode functionally conserved proteins (as found for orthogroups UMAG_05306/sr10059 and UMAG_11060/sr14941), and (ii) the neofunctionlization of effectors on the protein level, which further fine-tuned on transcription levels.

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 right timing for sampling to exclude that the observed differences result rather from the different disease progression states than from a differential regulation as such. Sucher *et al* (2020) collected samples from the edge of disease lesions, which were smaller than 25 mm in diameter, as 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 plasticity on different hosts (Baetsen-Young *et al.*, 2020). In our experiment, we conducted a series of microscopy combined with qPCR-based biomass quantification, and selected two biotrophic interaction phases representing a comparable (2 dpi) and a distinct colonization stage (4 dpi) for RNA-seq.

When comparing the induction profiles of effectors in *U. maydis* and *S. reilianum* respectively, we found a significant 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. 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 induction patterns and differential regulation of effector DEOs are consistent with the revelation of different symptoms. At 4 dpi, *U. maydis* triggers local host cell proliferation to produce tumors, which will require the high transcript abundance of effector genes related to tumorigenesis; in contrast, *S. reilianum* sustaining stable expression levels of effectors likely reflects maintenance of the biotrophic interaction in already colonized tissue and continuation of systemic proliferation through the leaf veins with inducing structural changes in the host tissue

Not only effector genes, but also pathogenesis related transcription factors were found to be differentially regulated, which in turn is probably playing 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 significantly higher expressed in *U. maydis*, indicating the complicated regulation of TF cascade. Until now, only few direct target genes of TFs were identified by CHIP-seq in smut fungi (Tollot *et al.*, 2016). To elucidate the binding motif of these pathogenesis related TFs and 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 which 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 both 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, which knockout resulted in reduced virulence in *U. maydis* but are functional conserved of the respective *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 changed protein function. However, UMAG_05318 provides another example for combination of transcriptional plasticity and functional diversification of the protein level. While the orthogroups UMAG_05318/sr10075 and Umtin2/Srtin2 were expressed during leaf infection both in *U. maydis* and *S. reilianum*, they also show 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 mechanism of these differentially regulated organ specific effectors are not elucidated in *U. maydis*, and whether they also contribute to virulence in *S. reilianum* is largely unknown. Deletion of sr10075 alone did not seem affect virulence, while a knock-out 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* and *M. pennsylvanicum*. Elucidate the function of this ortholog group and its role in smut fungi evolution and speciation will be our next aim. Interestingly, both effectors showing neofunctionalization are located in effector cluster 19A, which is the largest virulence cluster in both pathogens 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. Thus, it could be possible that these effectors were highly expressed in a different time/spatial pattern to contribute to the systemic spreading of *S. relianum*. On the other hand, it would be tempting to study if highly 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 of how evolution of effector orthologs promote the emergence of new pathogens.

With only a limited set of virulence DEOs can be tested in gene conversion experiments, this study identified three effector DEOs as promising candidates for effectors with virulence function in 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 compared to UmPit2. However, in a previous study we found that SrPit2 only partially complemented a *U. maydis* pit2 knock out mutant, which was linked with a reduced ability of SrPit2 to inhibit SA-associated PLCPs (Misas Villamil *et al.*, 2019). Furthermore, a recent apoplastic proteome analysis suggested that 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 plasticity and neofunctionalization 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.

Acknowledgments

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. Weiliang Zuo and Jasper Depotter are supported by the Research Fellowship Programme for Postdoctoral Researchers of the Alexander von Humboldt Foundation. MT and DKG received support by 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: W.Z. and G.D. designed the research; W.Z. performed the data analysis, molecular experiment and virulence assay; J.D. and D.K.G. mapped the RNA-seq data; J.D. conducted phylogenetic analysis; D.K.G. and M.T. provided ortholog expression normalization; W.Z. and G.D. wrote the paper, with contributions from the other authors.

Data availability

RNA sequencing data has been submitted to NCBI Genbank and are available under the following links: https XXXXXXXXX

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*.

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*.

Begerow D, Göker M, Lutz M, Stoll M. **2004**. On the evolution of smut fungi on their hosts. *Frontiers in basidiomycote mycology*.

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*.

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*.

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

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

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*.

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*.

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

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

Grabherr MG., Brian J. Haas, Moran Yassour Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, Lin Fan, Raktima Raychowdhury, Qiandong Zeng, Zehua Chen, Evan Mauceli, Nir Hacohen, Andreas Gnirke, Nicholas Rhind, Federica di Palma, Bruce W. N, Friedman and AR. **2013**. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology* **29**: 644–652.

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*.

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*.

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**.

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*.

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 of the United States of America*.

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.* 2006a. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*.

Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O, *et al.* 2006b. 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 version 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*.

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*.

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

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*.

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**.

Ma LS, 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*.

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*.

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**.

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*.

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**.

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

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*.

Redkar A, Villajuana-Bonequi M, Doehlemann G. **2015b**. Conservation of the *Ustilago maydis* effector see1 in related smuts. *Plant Signaling and Behavior*.

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

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*.

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*.

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**.

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*.

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

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

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

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. In: Methods in Molecular Biology.

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*.

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

Stamatakis A. **2014**. RAxML version 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*.

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.
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*.
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*.
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*.
Tollot M, Assmann D, Becker C, Altmüller J, Dutheil JY, Wegner CE, 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**.
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*.
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.
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*.
Zuo W, Depotter JR, Doehlemann G. **2020**. Cas9HF1 enhanced specificity in *Ustilago maydis*. *Fungal Biology* **124**: 228–234.

Figure Legends

Fig.1. Comparison of *U. maydis* and *S. reilianum* growth during leaf infection. a, calcofluor white staining of fungal hyphae and appressorium at 20hpi. **b-d**, WGA-AF488-propidium Iodide co-staining to show the fungal biotrophic growth inside the plant cell. Fungal cells were stained with WGA-AF488 (green), the vascular cell wall was stained with Propidium Iodide (red). White arrows indicate the *de-novo* divided bundle sheath cells triggered by *U. maydis* infection. White cycle showed the aggregation of *S. reilianum* cells inside the leaf vein. **e**, the relative biomass quantification by qPCR shows the fungal growth during infection, and a significant higher abundance of *S. reilianum* was detected at 4 dpi. **f**, the 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 amount of mapped reads was detected from *S. reilianum* at 4 dpi. "**", $p<0.05$. *Student t-test* was used for significance test.

Fig.2. Overview of *U. maydis* and *S. reilianum* transcriptome, respectively. a, PCA analysis of *U. maydis* and *S. reilianum* RNA-seq samples. **b,** expression data from three different growth conditions. The number of differentially expressed genes (log2 fold change>1, p>0.05) from pairwise comparison were indicated in the triangles. Red and green triangles represent up-regulated and down-regulated genes between 2 dpi vs. axenic culture (AC), 4 dpi vs. AC and 4 dpi vs. 2 dpi, respectively. **c,** boxplot shows transcriptional changes of effector and non-effector genes during biotrophic growth. Effector genes were dramatically up-regulated compared to the non-effectors genes. **d,** heatmap shows the overview of the expression changes of effector genes. "n" shows the number of effectors detected expressing in RNA-seq sample and the total number of effectors in the genome.

Fig.3. The induced pattern shows the expression changes of effectors. a, the induced effectors are clustered into 5 patterns based on the expression changes between axenic culture (AC), 2 dpi and 4 dpi conditions. The gray lines represent individual gene, and blue and orange lines are the mean of all genes in each cluster. "a", "b" and "c" were used to indicate the significance levels between each condition. Genes with log2 fold change>1 and p<0.05 were considered significantly induced in pairwise comparison. **b,** the number of induced

genes of each pattern. The *chi-square test* was used for statistic test for the significance. "**", $p<0.01$.

Fig.4. Orthologs were differentially regulated between *U. maydis* and *S. reilianum*. a, PCA analysis 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. The fold change>3 and p<0.05 were used to identify DEO. **c,** the GO classification of DEOs in three different conditions. **d,** the Venn diagram shows the number of differentially expressed effector orthologs during biotrophic growth. **e-f,** the differentially regulation of effectors clustered by the conservation between 5 smut fungi. **e,** the phylogenetic tree of 5 smut fungi, including *U. maydis*, *S. relianum*, *Ustilago hordei*, *S. scitamineum* and *M. pennsylvanicum*. **f,** the number of effector DEOs in each category. *Chi-square test* was used for significance test. "**", $p<0.01$, "***", $p<0.001$.

Fig.5. Effector orthologs between *U. maydis* and *S. reilianum* are functional conserved. a, a heatmap shows the expression difference and protein sequence identity of effectors ortholog pairs which have been knocked out and subjected to virulence tests in *U. maydis*. "¶", Pit2 only conserved in functional PID motif which failed to be detected during ortholog identification and "%ID" was not showed. **b,** the scheme demonstrated the process of CRISPR-Cas9 mediated gene conversion. The sgRNA was designed to target the coding sequence of *U. maydis* effector, and the donor plasmid contains *S. reilianum* ortholog coding sequence flanked by 1kb homolog region. **c-d,** Disease scoring results of gene conversion mutants and corresponding *U. maydis* knock out mutant. **c,** Converted UMAG_05306 and UMAG_11060 to the *S. reilianum* orthologs did not affect the virulence in SG200. **d,** Converted UMAG_05318 to its *S. reilianum* ortholog sr10075 or sr10079, the paralog of sr10075, comprise SG200 virulence, and show reduced virulence as knockout mutant. The significance was conducted by *student t-test* with disease index "*", $p<0.05$, "**", $p<0.01$.

Supplementary Captions

Supplementary Fig.1 The symptom of *U. maydis* and *S. relianum* infected leaves. a, The photos shows the symptom development from 2 dpi to 4 dpi. The dashed circle shows the first visible tumor from *U. maydis* infected leaf while the dashed rectangle shows the necrosis spot caused by *S. relianum* infection. The back *of U. maydis* infected leaf from 4 dpi is shown to better visualize the tumor. **b,** Confocal microscopy shows the details of the *S. reilianum* aggregation inside the leaf vascular at 4 dpi after WGA-AF488- Propidium Iodide staining. The fungal cells aggregate inside the vascular tissue. Scale bar=100 µm.

Supplementary Fig.2 Differentially regulation of orthologs between *U. maydis* and *S. relianum* from different samples. a, Pearson correlation analysis between *U. maydis* and *S. reilianum* samples based on relative ortholog expression from 6005 one-to-one orthologs. The *r* values indicate a high to low correlation between *U. maydis* and *S. reilianum* samples: AC>2 dpi>4 dpi. **b,** Venn diagram shows the number of differentially expressed ortholog in *U. maydis* and *S. relianum* from different conditions. About 30% and 50% of detected DEOs were only detected from axenic culture.

Supplementary Fig.3 Expression of pathogenic development related transcription factors between *U. maydis* and *S. relianum* during biotrophic growth from RNA-seq data. The transcription factors rbf1 and biz1 were highly expressed in *S. reilianum* while transcription factors hdp2, fox1 and nlt1 were enhanced expressed in *U. maydis*. The

transcription factor ros1 shows similar expression between these two smut fungi. *Student t-test* was used for significance test. "**", $p < 0.01$.

Supplementary Fig.4 The distribution of effector DEOs on chromosomes. **a,** the distribution of DEOs in 22 chromosomes based on the *U. maydis* genome organization. The blue bar shows *U. maydis* highly expressed DEOs, and the yellow bar shows the *S. reilianum* highly expressed DEOs. 2 dpi DEOs and 4 dpi DEOs were plotted above and below the chromosome, respectively. The red rectangle shows effector cluster 19A. **b,** the diagram shows the overview of synteny of effectors in cluster 19A. The effector genes are showed in green arrows; non-effector genes are shown in dark gray. The synteny of effector genes are in grey shade, the outline color of the shade indicates the DEO. The *U. maydis* highly expressed effector ortholog pairs are outlined in blue, the *S. relianum* highly expressed genes are outlined in yellow.

Supplementary Fig.5 Sequence comparison of UMAG_05318, sr10075 and sr10079. a, the maximum likelihood phylogenetic tree of UMAG_05318, sr10075 and sr10079 based on the amino acid sequences. sr10079 is distant from UMAG_05318 and sr10075. **b,** the amino acid sequence alignment of UMAG_05318, sr10075 and sr10079. the identical amino acids were shaded in grey.

Supplement Table.1 Oligo sequences used.

Supplement Table.2 Log2 CPM (count per million) of fungal reads.

Supplement Table.3 Relative ortholog expression data of 6005 one-to-one orthologs between *U. maydis* and S. *reilianum*.

Figure 1

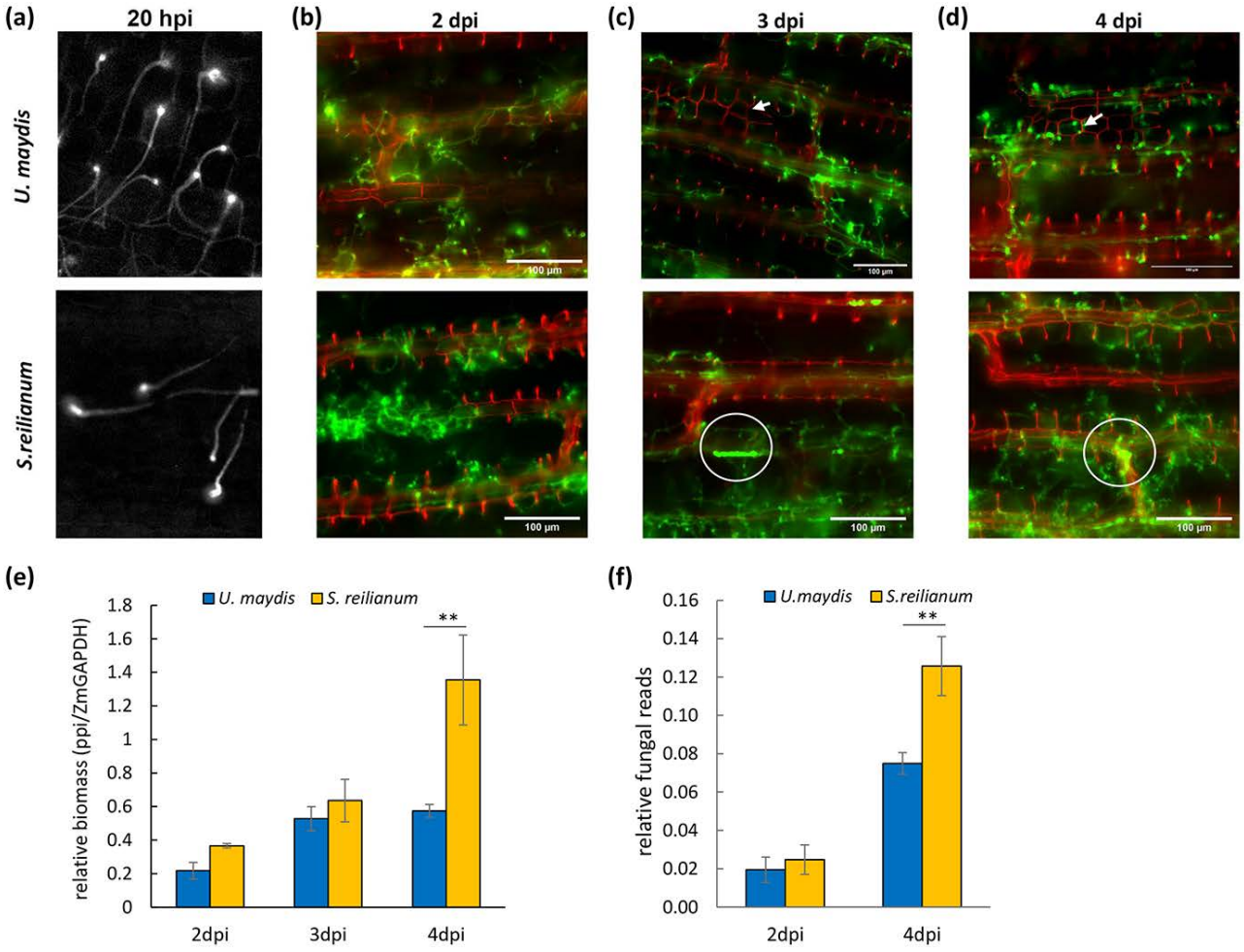

(a)
20 hpi
(b)
2 dpi
(c)
3 dpi
(d)
4 dpi
U. maydis
S.reilianum
100 µm
(e)
U. maydis
S. reilianum
relative biomass (ppi/ZmGAPDH)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
**
2dpi
3dpi
4dpi
(f)
U.maydis
S.reilianum
relative fungal reads
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
**
2dpi
4dpi

Figure 2

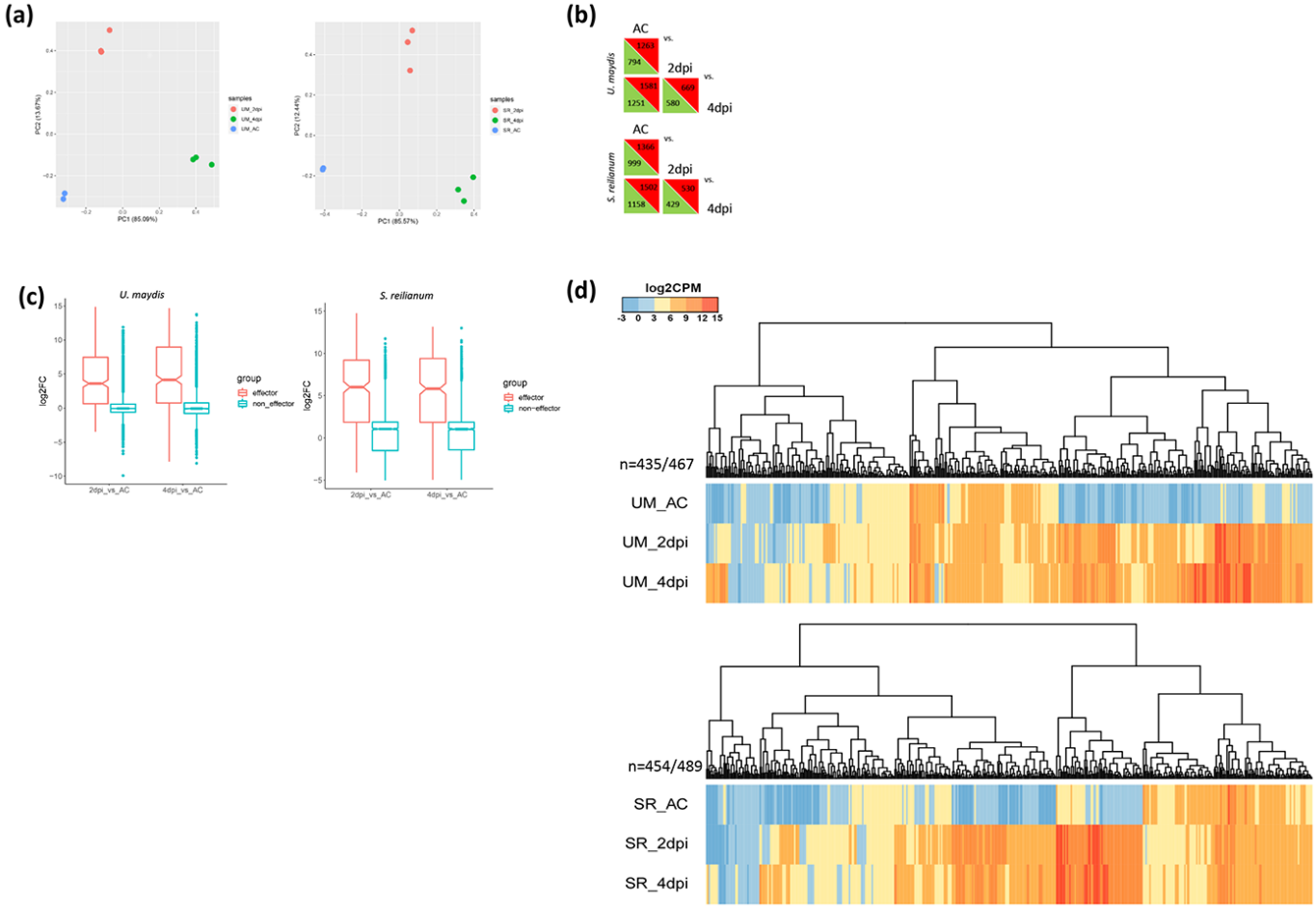

(a)
PC1 (85.06%)
PC2 (13.87%)
samples
UM_2dpi
UM_4dpi
UM_AC
PC1 (85.57%)
PC2 (12.44%)
samples
SR_2dpi
SR_4dpi
SR_AC
(b)
U. maydis
AC
vs.
1763
794
2dpi
vs.
1581
669
1251
580
4dpi
S. reilianum
AC
vs.
1366
999
2dpi
vs.
1502
530
1158
429
4dpi
(c)
U. maydis
S. reilianum
log2FC
group
effector
non_effector
2dpi_vs_AC
4dpi_vs_AC
(d)
log2CPM
-3 0 3 6 9 12 15
n=435/467
UM_AC
UM_2dpi
UM_4dpi
n=454/489
SR_AC
SR_2dpi
SR_4dpi

Figure 3

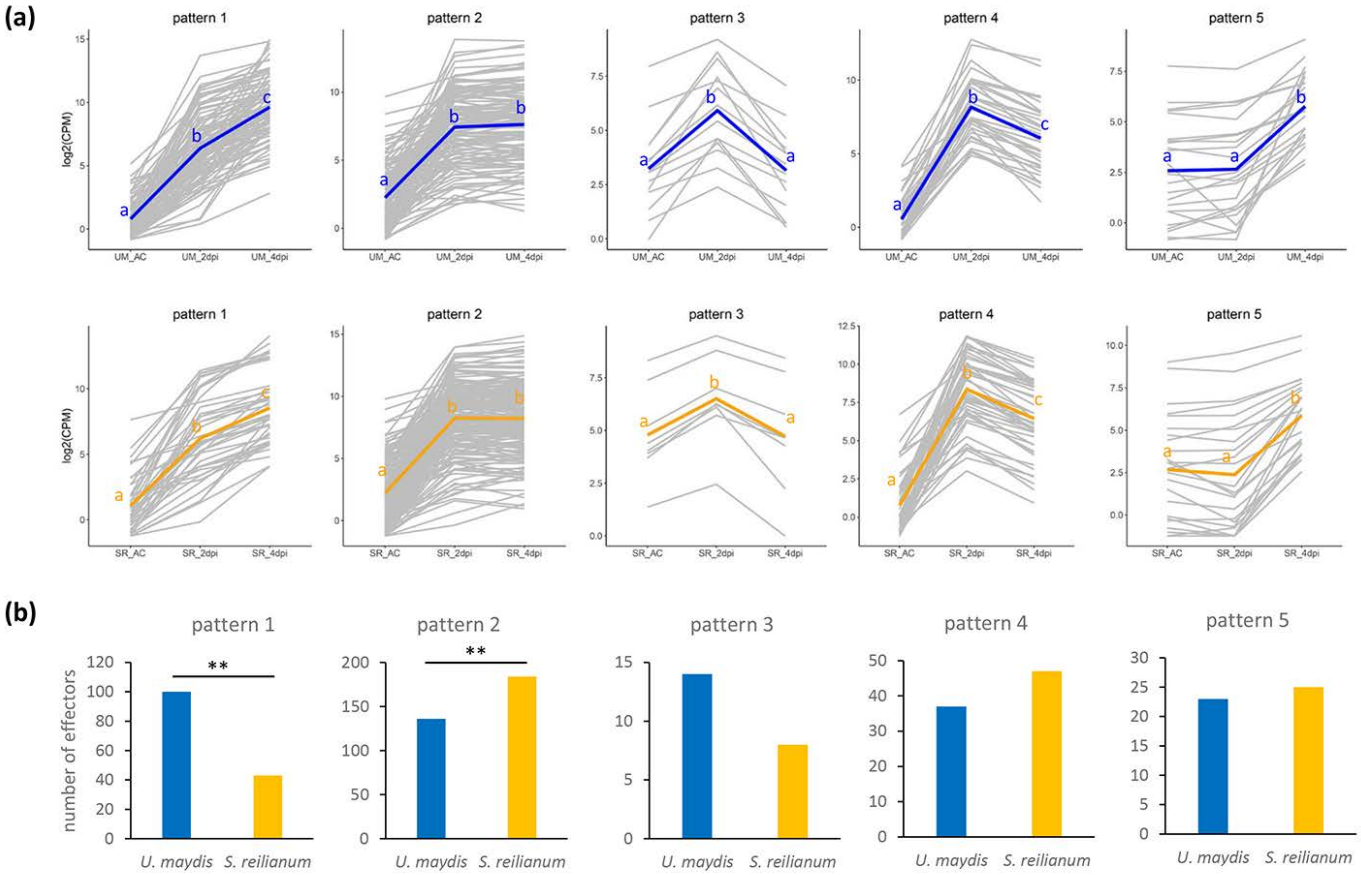

(a)
pattern 1
pattern 2
pattern 3
pattern 4
pattern 5
log2(CPM)
UM_AC
UM_2dpi
UM_4dpi
SR_AC
SR_2dpi
SR_4dpi
(b)
number of effectors
**
U. maydis
S. reilianum

Figure 4

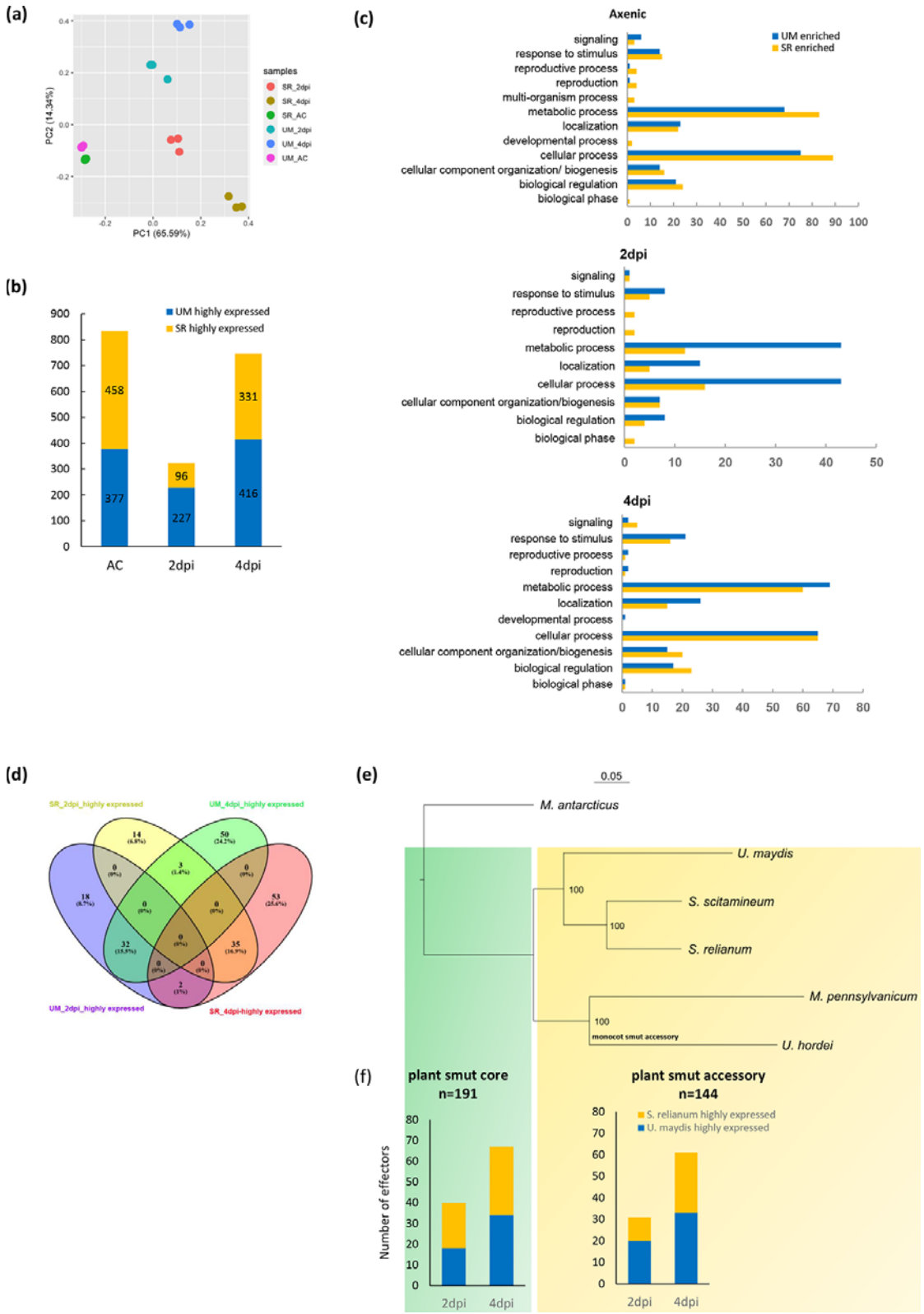
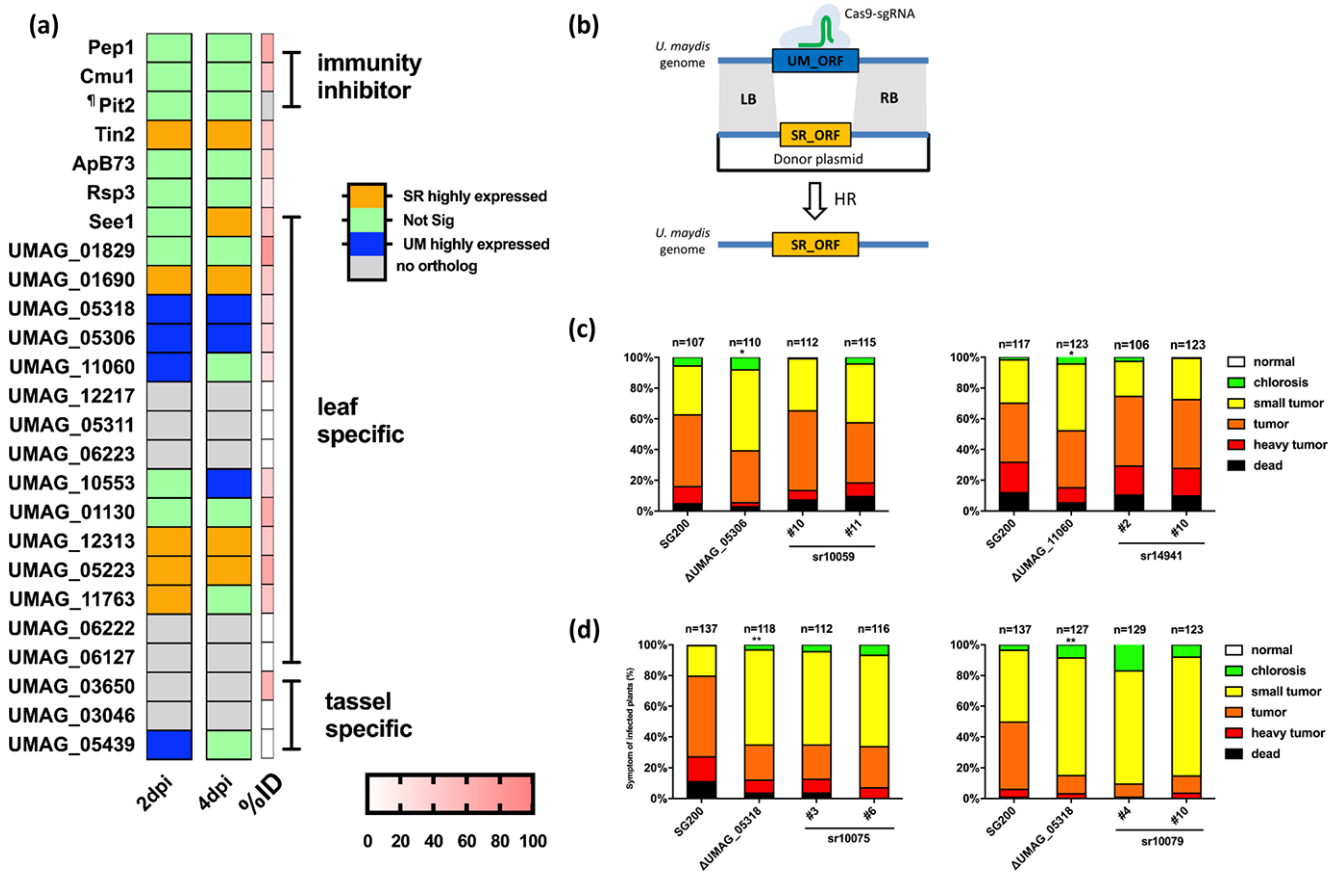

(a)
PC2 (14.34%)
PC1 (65.59%)
samples
SR_2dpi
SR_4dpi
SR_AC
UM_2dpi
UM_4dpi
UM_AC
(b)
UM highly expressed
SR highly expressed
458
377
96
227
331
416
AC
2dpi
4dpi
(c)
Axenic
UM enriched
SR enriched
signaling
response to stimulus
reproductive process
reproduction
multi-organism process
metabolic process
localization
developmental process
cellular process
cellular component organization/ biogenesis
biological regulation
biological phase
2dpi
4dpi
(d)
SR_2dpi_highly expressed
UM_4dpi_highly expressed
UM_2dpi_highly expressed
SR_4dpi-highly expressed
(e)
0.05
M. antarcticus
U. maydis
S. scitamineum
S. relianum
M. pennsylvanicum
U. hordei
monocot smut accessory
(f)
plant smut core
n=191
plant smut accessory
n=144
S. relianum highly expressed
U. maydis highly expressed
Number of effectors
2dpi
4dpi

Figure 5


(a)
Pep1
Cmu1
¶Pit2
Tin2
ApB73
Rsp3
See1
UMAG_01829
UMAG_01690
UMAG_05318
UMAG_05306
UMAG_11060
UMAG_12217
UMAG_05311
UMAG_06223
UMAG_10553
UMAG_01130
UMAG_12313
UMAG_05223
UMAG_11763
UMAG_06222
UMAG_06127
UMAG_03650
UMAG_03046
UMAG_05439
2dpi
4dpi
%ID
immunity inhibitor
leaf specific
tassel specific
SR highly expressed
Not Sig
UM highly expressed
no ortholog
0 20 40 60 80 100
(b)
Cas9-sgRNA
U. maydis genome
UM_ORF
LB
RB
SR_ORF
Donor plasmid
HR
U. maydis genome
SR_ORF
(c)
n=107
n=110
n=112
n=115
SG200
ΔUMAG_05306
#10
#11
sr10059
n=117
n=123
n=106
n=123
SG200
ΔUMAG_11060
#2
#10
sr14941
normal
chlorosis
small tumor
tumor
heavy tumor
dead
(d)
Symptom of infected plants (%)
n=137
n=118
n=112
n=116
SG200
ΔUMAG_05318
#3
#6
sr10075
n=137
n=127
n=129
n=123
SG200
ΔUMAG_05318
#4
#10
sr10079
normal
chlorosis
small tumor
tumor
heavy tumor
dead