1 **Title:** Transcription factor lineages in plant-pathogenic fungi, connecting diversity with fungal
2 virulence

3

4 **Author names and affiliations:**

5 Evan John^{1,3*}, Karam B. Singh^{1,2}, Richard P. Oliver³, Kar-Chun Tan^{1,3*}

6 ¹Centre for Crop and Disease Management, Curtin University, Perth, Australia

7 ²Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation,
8 Perth, Australia

9 ³School of Molecular and Life Sciences, Curtin University, Perth, Australia

10

11 ***Corresponding author:**

12 Evan John - evan.john@curtin.edu.au

13 Kar-Chun Tan – kar-chun.tan@curtin.edu.au

14

15 **Abstract**

16 Plant-pathogenic fungi span diverse taxonomic lineages. Their host-infection
 17 strategies are often specialised and require the coordinated regulation of molecular virulence
 18 factors. Transcription factors (TFs) are fundamental regulators of gene expression,
 19 controlling development and virulence in plant pathogenic fungi. Recent research has
 20 established regulatory roles for several taxonomically conserved fungal TFs, but the
 21 evolution of specific virulence regulators is not well understood. This study sought to explore
 22 the representation of TFs across a taxonomically-diverse range of fungi, with a focus on plant
 23 pathogens. A significant trend was observed among the obligate, host-associated pathogens,
 24 which possess a reduced overall TF repertoire, alluding to a lack of pressure for maintaining
 25 diversity. A novel orthology-based analysis is then presented that refined TF classifications,
 26 traditionally based on the nature of the DNA-binding domains. Using this analysis, cases of
 27 TF over/underrepresentation across fungal pathogen lineages are systematically highlighted.
 28 Specific examples are then explored and discussed that included the TF orthologues of
 29 Ste12, Pf2 and EBR1, plus phytotoxic secondary-metabolite cluster regulators, which all
 30 presented novel and distinct evolutionary insights. Ultimately, as the examples presented
 31 demonstrate, this resource can be interrogated to guide functional studies that seek to
 32 characterise virulence-specific regulators and shed light on the factors underpinning plant
 33 pathogenicity.

34

35 **Keywords**

36 Transcription factor, plant-pathogenic fungi, virulence, regulation, orthologue

37 **Abbreviations**

38 Transcription factor (TF), DNA-binding domain (DBD), complete TF repertoire (TFome),
 39 secondary metabolite (SM)

1. Introduction

Plant-pathogenic fungi have adapted to a wide range of ecological niches. It is now clear that pathogenic lifestyles, broadly defined as biotrophic, necrotrophic or hemibiotrophic, have evolved independently and across distant fungal lineages (Aylward et al., 2017; Möller and Stukenbrock, 2017; Ikeda et al., 2019). As such, the molecular mechanisms underpinning host infections are often specific to a pathogen or its close relatives. In recent years, large-scale comparative genomics and machine learning approaches have demonstrated that the enrichment of particular gene classes can be predictive of a pathogenic lifestyle. This includes carbohydrate-active enzymes, putative effectors, genes related to phytotoxic secondary metabolite (SM) biosynthesis and novel classes that currently lack functional annotations (Collemare and Lebrun, 2011; Pusztahelyi et al., 2016; Plissonneau et al., 2017; Hane et al., 2020; Haridas et al., 2020). For these genes to promote the virulence of a pathogen, their expression requires a coordinated regulation to maximise activity at the appropriate stage of infection (van der Does and Rep, 2017). Hence, the evolution of virulence encoding genes simultaneously requires the pathogen to evolve appropriate regulatory systems.

Transcription factors (TFs) regulate diverse aspects of fungal development and virulence, targeting either a small number of genes or by exerting broad control as ‘master’ regulators. Conserved TFs originally characterised in model saprophytes are now well-studied in plant-pathogenic fungi (Tan and Oliver, 2017; van der Does and Rep, 2017; John et al., 2021). However, there is extensive variation in the number of TFs belonging to different structural families across the fungal taxa (Park et al., 2008; Todd et al., 2014; Shelest, 2017) and the role of most remains undefined. It is therefore anticipated that a functional exploration of the TFs unique to a fungal lineage or lifestyle will reveal novel regulators that distinguish a pathogen from a non-pathogen. Already, TF gene duplications that are directly implicated in fungal virulence have been described in several lineages of *Fusarium oxysporum*, the causal agent of vascular wilt on a wide range of plants (de Vega-

68 Bartol et al., 2010; Niño-Sánchez et al., 2016; van der Does et al., 2016). On the other
69 hand, broadly conserved regulators in plant-pathogenic fungi, such as the Velvet TFs, are
70 absent in saprophytic yeasts (Bayram and Braus, 2012; Calvo et al., 2016). Hence, in both
71 their presence or absence, TFs are implicated in the distinct regulatory systems that have
72 evolved to control plant-pathogenic lifestyles.

73

74 Random mutagenesis or high-throughput gene knockout studies are traditional
75 approaches that have been used to identify novel virulence-regulating TFs. For example,
76 VdFTF1 in *Verticillium dahliae*, which regulates the expression of secreted virulence factors
77 during infection on cotton, was identified by screening a random mutagenesis library (Zhang
78 et al., 2018). On the other hand, a genome-wide TF knockout screen was undertaken for
79 *Fusarium graminearum*, which identified several virulence regulators otherwise dispensable
80 for saprophytic growth (Son et al., 2011). However, the generation and screening of fungal
81 mutant libraries in this fashion can be labour intensive. An approach that makes use of the
82 annotated fungal genomes now available to disentangle regulators unique to a pathogenic
83 lineage or lifestyle presents a powerful stand-alone approach, or as an initial step to help
84 guide functional TF investigations.

85

86 A novel analysis is presented here that simultaneously assesses the TF repertoires
87 (“TFomes”) from 100 phytopathogenic fungi alongside 20 representative
88 saprophytes/symbionts for comparison. The TFs were refined into orthogroups; the smallest
89 definable set of proteins where all their orthologues are included (Emms and Kelly, 2019). In
90 doing so, the relationships between TFs across plant-pathogenic fungi could be
91 systematically assessed. The distinct aims were to A) explore TF conservation, expansion
92 and loss associated with plant-pathogenic lifestyles and B) provide a novel resource for
93 tracing the evolutionary trajectory of TFs correlated with virulence in distinct fungal lineages.
94 Using the orthogroups defined in the analysis, the evolutionary lineages of several
95 established fungal-virulence regulators are explored in detail. These presented examples of

96 TF conservation, expansion/diversification and loss in key pathogen lineages and also
97 provided novel insights upon which future analyses can build.

98 **2. Methods**

99 **2.1. Compiling fungal TFomes**

100 Non-redundant fungal proteomes were sourced for 100 plant pathogens, 10
 101 saprophytes and 10 symbionts from UniProt (release 2020_05) (Bursteinas et al., 2016) or
 102 MycoCosm (Nordberg et al., 2014) and independent platforms where the proteomes were
 103 available with higher coverage assemblies (**Supplementary item 1**). Pathogens were
 104 selected from diverse fungal taxa of significant research interest (Pedro et al., 2016; Urban et
 105 al., 2020). The saprophytes and symbionts included both close and distant pathogen
 106 relatives across the taxa. Organisms were assigned a five letter organism ID (ORGID) for
 107 brevity (**Supplementary item 1**), adapted from their NCBI Taxonomy database naming
 108 convention (Schoch et al., 2020). General lifestyles (biotroph, hemibiotroph, necrotroph,
 109 saprophyte or symbiont) and plant host-associations/dependency (i.e. obligate, facultative or
 110 not associated) were assigned based on descriptions from the published proteome sources
 111 and/or literature covering the matter (Möller and Stukenbrock, 2017; Hane et al., 2020;
 112 Haridas et al., 2020).

113
 114 Interpro protein-domain annotations were also downloaded with the respective fungal
 115 proteome sources. Interproscan (release 82.0) was used to annotate proteomes from the
 116 independent servers where these were not available (Blum et al., 2020). TFs were selected
 117 from the annotated proteomes by cross-referencing a list of Interpro domain accessions that
 118 represent sequence-specific DNA-binding domains (DBDs) (**Supplementary item 2**). Unique
 119 protein identifiers were assigned to each TF following the format 'ProteinID | ORGID' where
 120 the 'ProteinID' was derived from the downloaded proteome source.

121

122 **2.2. TFome regression analysis**

123 The total occurrences of each TF domain across the annotated fungal proteomes
 124 (which included cases where multiple domains exist within a single protein) were counted
 125 using a custom bash script. Linear regressions were then produced to model the association

between proteome size and TFome size for the entire set of 120 fungi, as well as for the lifestyle and host-association sub-groups. The data analysis, formatting and plotting was undertaken in R (version 4) using the ggplot2, ggpubr, rstats, rstatix and emmeans packages (R Core Team, 2020). Pearson's correlation coefficients (Mukaka, 2012) were calculated for each linear regression and used as the test statistic to assess whether the association was significant ($P < 0.05$). A covariance analysis (ANCOVA) was also undertaken to test whether the regression-slopes (i.e. the rate of TFome size increase relative to the proteome size) were different between the lifestyle or host-association groups (Bonferroni Padj < 0.05).

2.3. TF orthology analysis and species-tree inference

An orthology analysis for the entire TF set compiled from the TFomes was undertaken using OrthoFinder (release 2.4.0) (Emms and Kelly, 2019), with MMseqs2 (Steinegger and Söding, 2017) invoked for the sequence search stage. The TF-distance based species-tree (which included here several sub-species taxonomic rankings), produced as part of the OrthoFinder analysis by STAG (Emms and Kelly, 2018), was manually rooted at the most distantly related pathogen *Synchytrium endobioticum* (SYNEN), inferred from the phylogeny documented in a previous study (Choi and Kim, 2017). OrthoFinder was subsequently invoked using the '-s' option and supplied the rooted species-tree to resolve the orthogroup 'protein-trees'. The protein-trees were built using the DendroBLAST distance measure (Kelly and Maini, 2013) as part of the OrthoFinder analysis. Tree visualisations and formatting were subsequently undertaken using iTOL (Letunic and Bork, 2019). NCBI taxonomic rankings for each respective phylum, class, order, family and genus were then sourced using Taxize (version 0.9.99) to demonstrate the clades in the TF-distance based species-tree corresponded with fungal taxonomies (Chamberlain et al., 2020).

Fungal TF-orthogroup datasets were ordered as per the leaf-node order derived from the species-tree (beginning with the root at *S. endobioticum*) for subsequent visualisation and analyses. Data formatting and the production of heatmaps was undertaken in R (version

4) using the phylogram, complex heatmap, ggplot2, ggpubr and the rstats packages (R Core Team, 2020). A hierarchical clustering analysis was conducted to group the TF orthogroups based on their relative counts across the 120 fungi analysed. The Z-scores were first calculated similarly to a previous TF analysis (Charoensawan et al., 2010a), which represented a clustering measure independent of the differences in orthogroup sizes. Clustering distances were calculated based on Pearson's correlation and the orthogroups were clustered using the Average linkage function.

2.4. Independent analysis of established fungal-virulence regulators

Orthogroups harbouring established virulence regulators were determined by cross-referencing the respective ProteinIDs with experimentally validated TFs from the scientific literature (John et al., 2021). TFs of evolutionary interest were compiled by inspection of the respective orthogroup protein-trees. A multiple-sequence alignment was then undertaken on the selected protein sequences using Clustal Omega (Sievers and Higgins, 2018) and conserved regions corresponding to Interpro-domain annotations (Blum et al., 2020) were assessed across the protein set. These alignments were then used to produce Neighbour-Joining trees with the Jukes-Cantor distance measure in Geneious Prime (version 2021) as a complementary method to assess the orthogroup clade-topologies.

3. Results & Discussion

3.1. Correlations between TFome and proteome sizes in plant pathogenic fungi

Fungal TFomes were defined as the set of proteins harbouring at least one sequence-specific DBD in line with previous analyses (Park et al., 2008; Shelest, 2008; Todd et al., 2014; Shelest, 2017). The list of corresponding Interpro domains was updated (**Supplementary item 2**) from the most recent analysis (Shelest, 2017) by incorporating 16 additional sequence-specific DBDs such as the Velvet domain (IPR037525), the high-mobility-group box domain (IPR009071), the APSES/Swi6 domain (IPR001606) and the Sant/Myb domain (IPR001005). Conversely, two domains were excluded that represented single-stranded nucleic-acid binding proteins (IPR012340 and IPR001878). Hence, the 120 fungal TFomes were assembled from refined TF-annotation criteria. Their relative size across each of the 120 proteomes was first assessed to gauge any broader trends in the TF-regulatory capacity (**Fig. 1**). These ranged from 189 TFs identified (for *Blumeria graminis* f. sp. *tritici*; BLUGT) to 1,493 (for the *F. oxysporum* biocontrol strain Fo47; FUSO4). The average TFome size was 473 from an average proteome size of 12,355. This corresponded to a median of 437.5 and 12,296.5 respectively. The TF fraction of the proteomes ranged between 1.2% (for *Puccinia triticina*; PUCTR) to 6.9% (for FUSO4) with an average of 3.9%. The highest proportion of TFs was generally observed among the *Fusarium* lineages, while rust pathogens such as *Melampsora* and *Puccinia* spp. shared a relatively low TF content. Overall, a moderate-linear correlation ($R = 0.58$) was observed between TFome size and proteome size. A previous study that assessed a broad range of eukaryotes reported a stronger correlation ($R = 0.78$) (Charoensawan et al., 2010b). This suggested TF-content variation was higher among the fungi assessed in this study.

Fig. 1 - Fungal transcription factor content (TFome) relative to proteome size

TFome vs proteome sizes for the 120 fungi used in this analysis. The five-letter fungal ORGIDs are used to label the respective points which provide an indication of the respective lifestyle and host-association (**Supplemental item A**). The correlation coefficient (R) and P -value (p) used to test for any association ($P < 0.05$) and linear regression (y) for the relationship are presented at the upper-left of the plot.

The TF content of fungal sub-groups, defined by their lifestyle or association with plant hosts, were then compared to explore any distinct trends relevant to phytopathogenicity. Linear correlations were observed for TFome vs proteome counts from each pathogenic (i.e. biotrophic, hemibiotrophic or necrotrophic) and non-pathogenic (i.e. saprohytic or symbiotic) fungal lifestyle (**Fig. 2**). No significant differences were observed across the linear models derived from these groups. This was not the case when fungi were grouped based on their host-associations (i.e. facultative, obligate or not host-associated). The TF content of obligate-host fungi relative to the proteome size was significantly lower than facultative pathogens and non-host saprophytes (**Fig. 2**). These were largely represented by the rust and mildew pathogens from the Pucciniales (1.7% average TF fraction) and Erysiphales (3.2% average TF fraction) fungal orders. This observation provides evidence that a restricted regulatory capacity exists where fungal pathogens are intimately connected with their host for survival. Conversely, fungi which occupy broader ecological niches are likely to encounter diverse selection pressures that maintain TF diversity. Some outliers to these trends were observed in distinct fungal lineages. For example, the average TF fraction in the Agaricomycetes class, which included several facultative necrotrophs and non-pathogenic saprophytes, was just 2.5%. TF-content comparisons within lineages are therefore necessary to resolve the specific host-lifestyle differences that have evolved.

Fungal division	ANCOVA	Regression line groupings
Lifestyle	Pval = 0.561	Biotroph ^A , Hemibiotroph ^A , Necrotroph ^A , Saprophyte ^A , Symbiont ^A
Host dependence	Pval = 9.13E-5*	Facultative ^A , Not host-associated ^A , Obligate ^B

Fig. 2 - A comparison of the transcription factor proportion of fungal proteomes grouped by lifestyle or host-association

TFome vs proteome sizes are depicted along with the correlation coefficients (R) and P -values (p) for the respective linear regressions (y) in each plot. The panel on the left depicts the relationships by lifestyle and the right panel by host-association for the 120 fungal

234 proteomes assessed in this study. The bottom table summarises where a statistically
 235 significant difference ($P_{\text{val}} < 0.05$)* exists across the regression rates determined by
 236 ANCOVA. Superscript letters A or B indicate where individual differences occur within the
 237 groups.

3.2. Lineage-specific TF orthogroup expansions among plant-pathogenic fungi

A cursory look at the total set of fungal TFs classified by their DBD-based families (**Supplementary item 2**) revealed that the Zn2Cys6 and C2H2 zinc finger domains were the most prevalent among the 120 fungal species/strains, which paralleled previous analyses (Todd et al., 2014; Shelest, 2017). A total of 32,023 Zn2Cys6, 11,436 C2H2, 4,777 homeodomain, 3,168 bZIP and 2,932 CCCH-type domains represented the top five among the 70 TF families where at least 10 instances were identified overall. The large size of such families motivated an orthology analysis with the aim to better classify the TFs represented in distinct fungal lineages. This resolved the 56,754 fungal TFs into 855 orthogroups (**Supplementary item 3**), which presented a large refinement of the DBD-based classification system. The TF-distance based species-tree derived from the fungal orthogroups was then compared with the fungal taxonomic clade-architecture to assess the overall accuracy of the analysis. Consistent topologies were observed (**Fig. 3**; **Supplementary items 4 & 5**) which suggested that sequence divergence within the TF orthogroups correlated well with species divergence across fungal lineages.

The TF orthogroups were then clustered based on their normalised counts across the fungal species/strains (**Supplementary item 6**). This sought to assess the extent to which the respective TFs are expanded or specific to distinct pathogen lineages, a potential indicator for virulence-associated selection. Inspection of the corresponding heatmap (**Fig. 4**) highlighted widespread cases of overrepresentation down to the genus, species or sub-species level. While the relevant consequences in the capacity for individual fungi to regulate virulence will need to be functionally assessed, the analysis demonstrated evolutionarily recent shifts are traceable to specific TF lineages.

Fig. 3 - Fungal taxonomy relative to the transcription factor (TF)-distance based species-tree

Clade architectures derived from NCBI taxonomic divisions (node dots) are presented in the left tree, demonstrating shared topologies with the TF-distance based species-tree. Fungal ORGIDs are used at branch tips due to space constraints (**Supplementary item 1**). The corresponding trees using the full species names are available for higher resolution viewing as **Supplementary item 4 & 5**.

Fig. 4 - Relative proportion of transcription factors (TFs) within orthogroups across the 120 fungi assessed in this study

A heatmap depicting TF expansion based on relative counts (Z scores) for each of the 855 orthogroups defined in this analysis (arrayed on the X axis and clustered using the Euclidean distance measure). Fungal ORGIDs are listed on the Y-axis based on the TF-species-tree order depicted in **Fig. 3**. Coloured markers adjacent to ORGIDs indicate the respective fungal lifestyle and plant host-association. The TF counts used to produce this figure are provided (**Supplementary item 6**) along with a high-resolution version of the heatmap that details the respective IDs for the clustered orthogroups (**Supplementary item 7**).

3.3. Regulators of virulence in conserved and recently evolved TFs

The breadth of the 855 orthogroups defined in the analysis constrained the extent to which individual fungal lineages could be assessed in detail. Instead, the evolutionary trajectories for several established virulence regulators are explored here in the context of their respective orthogroups to provide examples where novel insights can be gained into fungal plant pathogenicity. This makes use of the orthogroup ‘protein-trees’ produced from the analysis which allow the predicted origins of TF duplication and/or loss to be traced across fungal lineages. The respective protein-trees for each orthogroup (OG0000000 > OG0000854) are provided as supplementary material (**Supplementary item 8**). The virulence regulators explored here include TFs which regulate either a broad or restricted set of genetic pathways, as well as those that are conserved across fungal taxa and those where orthologues are difficult to trace.

3.3.1. Pf2-orthologues underpin virulence within a lineage of carbohydrate metabolic regulators

Orthogroup OG0000017 comprises 370 annotated TFs that belong to the Zn2Cys6 DBD family. These are represented across the Ascomycota, in most cases with multiple TFs identified for each fungus. This includes MAL13, ZNF1, MAL33 and YGR288W characterised in *Saccharomyces cerevisiae* that target either the aerobic and/or anaerobic pathways regulating carbohydrate metabolism (Akache et al., 2001; Tangsombatvichit et al., 2015). Analysis of the OG0000017 protein-tree revealed these four TFs were adjacent to each other in a single clade, suggesting they are paralogues in the context of this orthogroup (**Supplementary item 9**). Further inspection of the protein-tree revealed another clade harboured the AmyR regulator (ANIA_02016) from the filamentous saprophyte *Aspergillus nidulans*. AmyR is known to play a fundamental role in starch/polysaccharide hydrolysis (Tani et al., 2001; Nakamura et al., 2006). The common link to carbohydrate metabolism for these regulators is also conserved in a separate clade that is of significant interest from the perspective of phytopathogenicity. This encompasses Pf2 (Pleosporales spp.) and several

other virulence-regulating TFs that include ART1 (*Fusarium* spp.), MoCod1 (*Magnaporthe oryzae*), PdeR (*Botrytis cinerea*) and Zt107320 (*Zymoseptoria tritici*) (Cho et al., 2013; Chung et al., 2013; Oh et al., 2016; Jones et al., 2019; Habig et al., 2020; Han et al., 2020). The analysis suggested each of these virulence-regulators are traceable as direct taxonomic orthologues to *Neurospora crassa* Col-26 (**Supplementary item 9**). Recent studies have established the important role of Col-26 in polysaccharide metabolism and plant cell-wall degradation (Xiong et al., 2017; Li et al., 2021).

It was intriguing that the Pf2/Col-26 monophyletic clade within OG0000017 did not include TFs from the Eurotiomycetes or Saccharomycetes, such as AmyR or MAL13, but is conserved among the other Ascomycete lineages. A protein sequence alignment/phylogenetic tree was therefore built from the functionally-characterised MAL13, AmyR and Pf2-orthologues to highlight the common and divergent features (**Fig. 5**). A relatively high degree of sequence conservation was observed at the Zn2Cys6 DBD (IPR001138) and the 'Fungal transcription factor' domain (IPR007219) that are present across the proteins. However, a region unique to the Pf2-orthologues was identified at a sequence centred on the C-terminus of the Zn2Cys6 DBD (**Fig. 5**), suggesting differentiation relevant to DNA-binding. Previous studies that screened Zn2Cys6 gene-deletion-mutant libraries in both *M. oryzae* and *F. graminearum* (Son et al., 2011; Lu et al., 2014) support a distinct role for the Pf2-orthologues. In these analyses no obvious virulence-defects were associated with any of the 12 other TFs identified in OG0000017, which suggested the Pf2-orthologues (MoCod1 and FgART1 in these species) are the prominent regulators in this orthogroup.

The analysis of the OG0000017 orthogroup provided clarity for the relationship between Pf2-orthologues and their probable evolutionary origin from a lineage of carbohydrate/polysaccharide metabolic regulators. These processes are fundamental for plant pathogens that exhibit a necrotrophic phase of their lifecycle for the acquisition of

339 nutrients from their respective hosts. A question that remains will be to establish the unique
340 features that evolved for Pf2-orthologues which explain the distinct and conserved role
341 reported in fungal virulence.

342

343

Fig. 5 - Characterised Pf2-orthologues in relation to established carbohydrate regulators

A protein-sequence alignment is presented with a neighbour-joining tree depicting sequence divergence between Pf2-orthologues (highlighted in green) and the *A. nidulans* AmyR and *S. cerevisiae* MAL13 proteins that were identified in orthogroup OG0000017. The protein names correspond to those reported in the relevant literature followed by the five-letter ORGIDs used in this study for the respective fungi. The alignment presents the % of residue identity across the aligned loci with gaps being ignored. The regions corresponding to the Zn2Cys6 (IPR001138) and the 'Fungal transcription factor' (IPR007219) Interpro domains that were identified are presented above the alignment. Red boxes span the sequences where the respective domains are detected.

3.3.2. Evidence of domain architecture variation in Ste12 TFs

Ste12 TFs are downstream targets of mitogen-activated protein kinase (MAPK) signalling that control expansive hyphal growth (as opposed to yeast-like budding growth) and the pheromone response (sexual development) in *S. cerevisiae* (Rispaill and Di Pietro, 2010; Zhou et al., 2020). An analogous function has been attributed in fungal phytopathogens, where Ste12 TFs play a significant role in plant virulence that is manifested through the regulation of invasive hyphal growth (Wong Sak Hoi and Dumas, 2010; John et al., 2021). Interestingly, Ste12 orthologues in filamentous fungi also harbour C2H2 DBDs located in the C-terminal region of the protein (Wong Sak Hoi and Dumas, 2010; Gu et al., 2015; Sarmiento-Villamil et al., 2018).

An analysis of the corresponding OG0000022 protein-tree revealed that TFs harbouring the Ste12 DBD (IPR003120) are conserved even among the early-diverging fungal lineages such as the Chytridiomycota and Mucoromycota (**Supplementary item 10**). However, these were not detected among the plant-pathogenic fungi of the Ustilaginomycetes class in contrast to other Basidiomycetes such as *Puccinia striiformis*, where PstSTE12 is essential for pathogenicity on wheat (Zhu et al., 2018). In *Ustilago maydis*, a prominent maize pathogen of the Ustilaginomycetes, three other TFs were identified (Biz1, Mzr1 and Ztf1) that are reported to play related roles promoting infectious growth *in planta* (Flor-Parra et al., 2006; Zheng et al., 2008; Velez-Haro et al., 2020; de la Torre et al., 2020). These belong to a large clade of proteins restricted to the C2H2-DBD (**Supplementary item 10**) that included the BrIA conidiation regulators from the Aspergillaceae family and several regulators investigated in *M. oryzae* and *F. graminearum* with no obvious virulence-related function (Kwon et al., 2010; Son et al., 2011; M. Wang et al., 2015; Cao et al., 2016).

The common virulence-associated roles prompted a protein sequence alignment and phylogeny to visualise the relationship between characterised Ste12-orthologues and the *U.*

maydis C2H2-domain regulators Biz1, Mzr1 and Ztf1 (**Fig. 6**). The highest-degree of conservation was observed among the Ste12-orthologues, and was centred at the corresponding DBD which did not extend to Biz1, Mzr1 and Ztf1. Although the C2H2-domains were largely conserved across both clades, other regions of significant identity were not obvious (**Fig. 6**). Hence, it is somewhat paradoxical that the sexual cycle is intimately connected with host invasion in *U. maydis* (Vollmeister et al., 2012), since both of these roles are driven by Ste12 in other fungi and the Ste12-DBD has been entirely lost in the Ustilaginomycete lineage. It is plausible that the C2H2-DBDs share a common evolutionary origin in this orthogroup, but is difficult to assess due to high-sequence divergence beyond this domain. Nevertheless, the evidence suggests a regulatory shift from Ste12 has occurred in *U. maydis*, possibly involving Biz1, Mzr1 and Ztf1, which share analogous roles promoting invasive growth. This is supported by findings that Biz1 and Ste12 are both activated and inhibited by related MAPK and cyclin-dependent kinase pathways respectively (de la Torre et al., 2020). The Ustilaginomycete class is almost entirely composed of plant-pathogens (Q.-M. Wang et al., 2015), indicating that Ste12-DBD loss could be a direct adaptation to host-driven selective pressures. Consequently, the functional relationships identified in this orthogroup analysis have presented an important hypothetical-link to the evolution of virulence.

Fig. 6 - Characterised Ste12-orthologues relative to C2H2-only regulators of invasive growth

A protein-sequence alignment is presented with a neighbour-joining tree depicting sequence divergence between Ste12-orthologues (highlighted in green), the *Ustilago maydis* C2H2-domain regulators of invasive growth (blue) and the Aspergillaceae BrlA regulators of conidiation (red) that stem from putative common lineages in orthogroup OG0000022. The protein names correspond to those reported in the relevant literature followed by the five-letter ORGIDs used in this study for the respective fungi. The alignment presents the % of residue identity across the aligned loci with gaps being ignored. The regions corresponding to the Ste12 (IPR003120) and the C2H2 (IPR013087) Interpro domains that were identified are presented above or within the alignments. Red boxes span the sequences where the respective domains are detected.

3.3.3. Evidence for further expansion of the EBR1 TFs

The Zn2Cys6 TF EBR1 was first identified in *F. graminearum* through mutagenesis screening (Dufresne et al., 2008). The TF was subsequently investigated and shown to facilitate host invasion in *Fusarium* spp. by suppressing hyphal branching and promoting radial growth (Zhao et al., 2011). The large protein size and the internal positioning of the DBD, rather than the N-terminus, are relatively unusual features for the Zn2Cys6 TF family (MacPherson et al., 2006). In *F. oxysporum*, gene expansions are a characteristic of lineage-specific chromosomes (Coleman et al., 2009; Niño-Sánchez et al., 2016). Interestingly, this phenomenon has been well-documented for *EBR1*, where as many as nine paralogues have been reported (Jonkers et al., 2014; van der Does et al., 2016). An analysis here of the corresponding OG0000121 protein-tree highlights cases in the different *F. oxysporum* isolates assessed here (**Supplementary item 11**), but also reveals EBR1 expansion across other important Ascomycete plant pathogens. This included the *M. oryzae* species-isolates assessed and the Mycosphaerellaceae fungal family. A protein sequence alignment/phylogenetic tree was therefore produced to investigate these putative expansions in greater detail (**Fig. 7**).

In the *M. oryzae* reference isolate causing rice blast (PYROR) both MoCod2 (MGG_09263) and Cnf2 (MGG_15023) are TFs required for pathogenic development (Chung et al., 2013; Lu et al., 2014). The alignment revealed sequence conservation at the Zn2Cys6 DBD and several unannotated regions towards the C-terminus (**Fig. 7**). This strongly indicates both proteins stem from the EBR1 lineage, but from the fungi assessed here Cnf2 is unique to *M. oryzae* and has acquired an independent function in the virulence of this pathogen. The alignment also reveals that in the *M. oryzae* millet pathovar included in this study (PYROM), three MoCod2 paralogues exist (**Fig. 7**). The unique/redundant roles for these paralogues should be explored, as well as any functional interactions between MoCod2 and Cnf2 in *M. oryzae* to better understand the evolution and regulation of virulence by this TF lineage.

444

445 Several EBR1 duplication events were evident across the Mycosphaerellaceae
446 family, which encompasses several damaging pathogens (**Supplementary item 11**).
447 Inspection of the protein alignment (**Fig. 7**) suggests this is accompanied by the loss of a
448 major proportion of the C-terminal conserved region in a sub-lineage which included *Z. tritici*
449 Mycgr3g111569 (AlmA). This TF was characterised through overexpression which affected
450 hyphal development during axenic growth (Cairns et al., 2015) and suggests a functional
451 connection to other EBR1 orthologues is still maintained. Other TFs in the
452 Mycosphaerellaceae-family lineage retain the conserved C-terminal region, but a functional
453 role is yet to be reported.

454

455 It was striking that the majority of pathogens where EBR1 expansion events are
456 observed are considered hemibiotrophs. It is plausible that a consequence of EBR1
457 duplication is the up-regulation of genes underpinning radial growth, thereby promoting
458 asymptomatic colonisation of the host typical of this lifestyle. The novel evidence for gene
459 expansions in key pathogen lineages, in particular the Mycosphaerellaceae, may warrant
460 further exploration to establish their individual and/or connected roles during host infection.

461

Fig. 7 - EBR1 expansion events identified in fungal pathogen lineages

A protein-sequence alignment is presented with a neighbour-joining tree depicting sequence divergence between the EBR1-orthologues identified in orthogroup OG000121. The respective lineages (and their taxonomic ranking) are highlighted according to the figure legend, while yellow indicates those within each lineage that are characterised through functional investigation. The protein names correspond to those reported in the relevant literature followed by the five-letter ORGIDs used in this study for the respective fungi. The alignment presents the % of residue identity across the aligned loci with gaps being ignored. The region corresponding to the Zn2Cys6 (IPR001138) and Interpro domain that was identified is presented above the alignment. Red boxes span the sequences where the respective domains are detected.

3.3.4. Tracing phytotoxic SM-gene cluster regulators

Pathway-specific regulators are typically encoded as part of fungal SM-gene clusters with most of those identified belonging to the Zn2Cys6 TF family (Keller, 2019; Lyu et al., 2020). A prominent feature of these clusters is diversifying selection and their frequent loss from fungal genomes, as well as horizontal transfer across taxa (Lind et al., 2017; de Jonge et al., 2018; Rokas et al., 2018). While a TF based analysis does not replace full SM-cluster phylogenies and functional investigation, a close inspection of the TF protein-tree clades revealed some surprising and significant connections for regulators of phytotoxic-SM biosynthesis.

An interesting case presented for BcBot6, a Zn2Cys6 TF assigned to orthogroup OG0000292 that regulates the biosynthesis of the phytotoxin botrydial in *Botrytis cinerea* (Porquier et al., 2016) (**Supplementary item 12**). A putative orthologue of BcBot6 was identified here in the phylogenetically-distant sunflower pathogen *Diaporthe helianthi* (DHEL01_v212770), and a sequence alignment with the adjacent branch on the protein-tree revealed little conservation relative to other fungal TFs (**Fig. 8A**). Although *D. helianthi* is reported to secrete several SMs that are phytotoxic (Ruocco et al., 2018), it is not known how these SMs are regulated, or if any are structurally analogous to botrydial. Moving forward, some interesting research questions that could be addressed are A) to explore the role of DHEL01_v212770 in *D. helianthi* or B) to cross-complement DHEL01_v212770 into the *bcbot6* deletion mutants of *B. cinerea* to determine if it is functionally conserved. The sequence alignment also suggested the putative BcBot6 orthologue annotated in the T4 isolate of *B. cinerea* (BOTC4) was truncated relative to BcBot6 in *B. cinerea* B05.10 (BOTC5) (**Fig. 8A**). Nevertheless, botrydial is an important virulence factor produced by T4 (Pinedo et al., 2008; Porquier et al., 2016) and further analysis will be required to define translational or regulatory differences attributable to this difference.

Another *B. cinerea* pathway-specific phytotoxin regulator is BcBOA13, which controls botcinic acid biosynthesis and is present in the B05.10 isolate but missing in the T4 isolate (Dalmais et al., 2011; Porquier et al., 2019). Interestingly, a homologous botcinic acid biosynthetic gene-cluster was previously identified in the genome of *Sclerotinia sclerotiorum*, a broad-host range pathogen that is closely related to *B. cinerea* (Porquier et al., 2019). However, inspection of the BcBOA13 clade in the OG0000000 orthogroup protein-tree revealed the homologous protein sscle_15g106440 was absent (**Supplementary item 12**). A closer inspection and sequence alignment that included sscle_15g106440 (Uniprot accession A0A1D9QLR3) suggested a replacement has occurred spanning the N-terminal DBD (**Fig. 8B**). It will be important to ascertain whether the sscle_15g106440 protein-annotation is accurate, or whether expression of the SM-gene cluster in *S. sclerotiorum* has been affected due to this mutation. The nearest branches in the BcBOA13 clade correspond to TF homologues that regulate the biosynthesis of aspirochlorine, a mycotoxin produced in *Aspergillus* spp. with no structural or functional relationship to botcinic acid having been described (Chankhamjon et al., 2014; Uka et al., 2020). Hence, it is unlikely that a BcBOA13 biosynthetic cluster was horizontally transferred across these fungal genomes. If a common ancestral protein existed, the sequence alignment indicates significant divergence has occurred relative to BcBOA13 outside the Zn2Cys6-DBD and some C-terminal residues (**Fig. 8B**).

Cercosporin and its structural homologues are potent light-activated phytotoxic SMs produced by *Cercospora* spp. (Daub et al., 2005; Newman and Townsend, 2016). Evidence of horizontal transfer, lineage diversification and/or gene loss has recently been detailed for the cercosporin biosynthetic-gene cluster in several Ascomycete lineages (de Jonge et al., 2018). *CbCTB8* encodes a cercosporin pathway-specific Zn2Cys6 regulator in *Cercospora beticola* and several cluster-orthologues were identified here in the protein-tree clade (**Supplementary item 13**). This included *Parastagonospora nodorum* EclR, a TF that controls the production of the phytotoxic structural homologue of cercosporin, elsinichrome C

(Chooi et al., 2017). A sequence alignment was therefore undertaken to explore the conserved protein regions in CbCTB8, EclR and the other predicted cercosporin-regulators that were identified in *M. oryzae* pathovars and the cucurbit pathogen *Colletotrichum orbiculare* (**Fig. 8C**). The alignment revealed that, in addition to the N-terminal Zn2Cys6 DBD, the AfIR domain associated with regulators of aflatoxin biosynthesis (Yu et al., 1996; Liu and Chu, 1998) is co-conserved. This was also observed in the adjacent protein-tree clade that included TFs from the symbiotic endophytes *Xylona heveae* and *Periconia macrospinoso*, suggesting the domains are not strictly-linked to the control of phytotoxin production (**Fig. 8C**). The distinguishing features of CbCTB8 orthologues that target genes controlling phytotoxin biosynthesis must therefore be determined, yet their common evolutionary lineage was predicted in line with the more comprehensive gene-cluster analysis conducted previously (de Jonge et al., 2018).

Fig. 8 - Phytotoxic secondary-metabolite cluster regulators across fungal pathogen lineages

Protein-sequence alignments are presented with a neighbour-joining tree depicting sequence divergence between A) the *B. cinerea* BcBot6 regulator of botrydial biosynthesis, B) the BcBoa13 regulator of botcinic acid biosynthesis and C) the CTB8 regulator of cercosporin biosynthesis. Putative homologous-regulators are highlighted in blue, while yellow indicates those within each lineage that are characterised through functional investigation. The protein names correspond to those reported in the relevant literature followed by the five-letter ORGIDs used in this study for the respective fungi. The alignment presents the % of residue identity across the aligned loci with gaps being ignored. The region corresponding to the Zn2Cys6 (IPR001138) and (in panel C only) the AfIR (IPR013700) Interpro domains that were identified are presented above the alignment. Red boxes span the sequences where the respective domains are detected. Homologous proteins that are yet to be functionally verified are indicated with by the asterix* and the double asterisk** indicates adjacent clades in the respective protein-trees (**Supplementary items 12 & 13**) that are not predicted to encode functional homologues.

4. Conclusions

The aims of the analysis presented here were to explore the conservation, expansion and loss of TFs associated with pathogenic lifestyles, as well as provide a novel resource to effectively trace the evolutionary trajectory of regulators that are correlated with virulence. A taxonomically-diverse range of fungi were assessed and a significantly-reduced TF content was observed in obligate host-associated pathogens. This suggested that, in general, the limited ecological-niche of these fungi lends itself to regulation by relatively few TFs. This study then resolved the TF classification system from 70 DBD-based families to an informative set of 855 orthogroups. Recent approaches have used large genomic datasets to demonstrate that the content of other distinct gene families, such as carbohydrate-acting enzymes, can both predict and refine our definitions of plant-pathogenic lifestyles (Hane et al., 2020; Haridas et al., 2020). These definitions more accurately reflect the molecular factors that underpin virulence. Here we demonstrated that many TF orthogroups were over/underrepresented in distinct pathogen lineages (**Fig. 4**). This suggests accurate information on both the type and content of fungal TFs should complement these approaches and enhance our understanding of the factors that contribute to virulence in plant-pathogenic fungi.

For TFs that have been under strong diversifying selection, precisely inferring their true orthologues will require detailed datasets that can incorporate a large number of closely related species. Nevertheless, as demonstrated in the virulence regulators discussed, the exploration of specific clades within the orthogroup protein-trees has provided novel insights into their evolutionary origins and trajectories. Importantly, the orthogroups presented in this analysis are a useful resource for the future exploration of virulence-regulating TFs. It is envisaged this will occur using different and complementary approaches:

- The orthogroup lineages (protein-trees) of functionally characterised TFs can be explored to define taxonomic orthologues and search for cases of gene duplication or loss in a pathogen of interest.
- Investigate orthogroups that exhibit expansion in a pathogen lineage. These can be identified from the heatmap in **Fig. 4 & Supplementary item 7** which correspond to columnar layout of the orthogroup counts in **Supplementary item 6**. The protein-trees for the expanded orthogroups can then be explored in detail to guide further phylogenetic and functional analyses that investigate any lifestyle-associated selection pressures.
- Search for cases of horizontal-gene transfer events for TFs implicated in the acquisition of fungal virulence. The orthogroup clade of a known virulence regulator can be examined in the corresponding protein-tree to identify anomalous TF neighbours (i.e. those from taxonomically distant fungi). This is particularly relevant for regulators of phytotoxic SM-gene clusters where genetic mobility is common (Lind et al., 2017; de Jonge et al., 2018; Rokas et al., 2018).

Combined, these are powerful means to assist the characterisation of the fundamental virulence regulators in plant-pathogenic fungi. This will provide insight into the distinct factors underpinning plant infection to assist the design of targeted mitigation strategies. This could include the direct inhibition of the TFs themselves, or their upstream/downstream signalling components.

Acknowledgements

We would like to thank Dr. Mark Derbyshire, Dr. Robert Syme, Dr. Darcy Jones and Dr. James Hane for their advice and support in undertaking the computational analysis.

Funding

This study was supported by the Centre for Crop and Disease Management, a joint initiative of Curtin University (<https://www.curtin.edu.au/>) and the Grains Research and Development Corporation (<https://grdc.com.au/>) [research grant CUR00023]. Evan John was supported by an Australian Government Research Training Program Scholarship (<https://www.dese.gov.au/>) administered through Curtin University. This work was also supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary items

Supplementary item 1 - Organism and TF annotation metadata

A table providing the fungal species/strains used in the analysis [columns 1-2], their assigned five letter ORGIDs [3], alternative names used for each organism in the scientific literature [4], their lifestyle (either necrotroph, biotroph, hemibiotroph, symbiont or saprophyte) [5], host association (either facultative or obligate and not applicable for saprophytes) [6], TF count relative to the proteome size [7-9], details on the proteome sources [10-11], references for the associated genome publication [12-13] and the NCBI taxonomic identifiers with phylogenetic rankings for each organism [14-20].

Supplementary item 2 - Assessment of total TF DBDs compiled

A table listing the respective Interpro IDs [column 1], Pfam/Superfam IDs [2], DBD/TF family description [3], the total count across species analysed in this study [4] and an indication whether the DBD was in the list adapted for this study [5] (Shelest, 2017).

Supplementary item 3 - Orthogroups defined for fungal TFs

Spreadsheets detailing individual TFs assigned to orthogroups (spreadsheet 1) or the unassigned TFs (spreadsheet 2). Orthogroups are listed by size in descending order. TF identifiers follow the format: 'ProteinID | ORGID' where the ProteinID was derived from the downloaded proteome source.

Supplementary item 4 - TF-distance based species-tree

The species-tree built from orthogroup TF distances (Emms and Kelly, 2018) and rooted at *Synchytrium endobioticum* (SYNEN) which was used for orthologue inferences. The full fungal species/strain IDs are included alongside the respective ORGIDs at the tree leaves.

Supplementary item 5 - NCBI taxonomy tree

The NCBI taxonomy tree for the fungal species/strains used in the analysis. Node points indicate the respective taxonomic divisions listed in **Supplementary item 6**. The full fungal species/strain IDs are included alongside the respective ORGIDs at the tree leaves.

Supplementary item 6 - TF orthogroup counts

A spreadsheet detailing the orthogroup counts for each organism [column 1], which are listed using their ORGIDs in the order of the leaf nodes in the TF species-tree. The respective NCBI taxonomic rankings for species, genus, family, order, class and phylum are listed [2-7]. The listed TFs for 855 orthogroups [8-862] are ordered by the hierarchical clustering depicted in **Fig. 4**.

Supplementary item 7 - Clustered TF-orthogroup heatmap (high resolution)

A high-resolution heatmap of normalised counts (Z-scores) for the 855 hierarchically-clustered orthogroups as depicted in **Fig. 4**. Fungal ORGIDs are listed on the Y-axis and TF-orthogroup IDs are listed on the X-axis which correspond to the layout for the TF counts presented in **Supplementary item 6**.

Supplementary item 8 - TF orthogroup protein-trees

Accessible via: <https://doi.org/10.6084/m9.figshare.17211737>

The TF orthogroup protein-trees generated in this study (for the orthogroups OG0000000 > OG0000591). These are in Newick format and can be visualised in the reader's chosen tree-viewing platform such as iTOL (Letunic and Bork, 2019). The branch lengths are defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

Supplementary item 9 - OG0000017 protein-tree

The full OG0000017 (Pf2-orthologue) protein-tree harbouring the TFs explored in **Fig. 5**. Highlighted are the *S. cerevisiae* MAL13 outgroup in blue, *A. nidulans* AmyR in orange and

the Pf2 taxonomic clade (green) with the functionally defined orthologues in yellow (for the fungi FUSVE, FUSGR, NEUCR, PYROR, BOTC5, ALTBR, PYRTR, PARNO and ZYMTI respectively). The branch length is defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

Supplementary item 10 - OG0000022 protein-tree

The full OG0000022 (Ste12) protein-tree harbouring the TFs explored in **Fig. 6**. Highlighted are the Ste12 orthologues in green with the functionally defined TFs in yellow (for the fungi ZYMTI, ALTBR, ALTAL, EXSTU, PENDI, PENEX, SCLSC, BOTC5, FUSGR, FUSOX, PYROR, CRYPA, VERDA, COLOR and PUCST respectively) and the Ste12 TFs that lack C-terminal C2H2 domains in blue (for SACCE and EREGO). The *U. maydis* C2H2-only TFs Biz1, Mrz1 and Ztf1 are highlighted in purple and the BrIA orthologues (PENDI and ASPNI) in orange. The branch length is defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

Supplementary item 11 - OG0000121 protein-tree

The full OG0000121 (EBR1) protein-tree harbouring the TFs explored in **Fig. 7**. Highlighted are the functionally defined EBR1 orthologues in yellow (for the fungi FUSOL, FUSGR, ZYMTI and PYROR respectively) and the putative expansions in the *Fusarium*, *M. oryzae* and the Mycosphaerellaceae lineages. The branch length is defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

Supplementary item 12 - OG0000292 (BcBot6) and OG0000000 (node 6240 - BcBOA13) protein-trees

Panel A highlights the BcBot6 transcription factor (TF) for *B. cinerea* T4 (BOTC4) and B05.10 (BOTC5) isolates in yellow alongside the putative *D. helianthi* (DIAHE) homologue in blue, indicating a shared ancestral origin despite relatively large species distances. Panel B highlights the BcBOA13 TF from BOTC5 in yellow and the aspirochlorine TF regulators from

Aspergillus spp. (corresponding to ORGIDs in blue: ASPNI, ASPFN and ASPF7) that also suggest a shared ancestral origin despite distant fungal taxonomies. The branch length is defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

Supplementary item 13 - OG0000000 protein-tree (node 6486 - CbCTB8)

A tree representing the relationship between transcription factors regulating the biosynthesis of Cercosporin (i.e. CbCTB8) or its structural homologues. The functionally defined CbCTB8 orthologues in the literature (represented by the fungal ORGIDs for CERBE and PARNO) are highlighted in yellow. Other putative homologous regulators reported for the cercosporin secondary-metabolite-gene clusters (de Jonge et al., 2018) were also grouped in a neighbouring orthogroup-clade and are highlighted in blue (represented by the fungal ORGIDs for PYROM, PYROR and COLOR). The branch length is defined by the DendroBLAST distance measure (Kelly and Maini, 2013).

References

- Akache, B., Wu, K., Turcotte, B., 2001. Phenotypic analysis of genes encoding yeast zinc cluster proteins. *Nucleic Acids Res* 29, 2181–2190. <https://doi.org/10.1093/nar/29.10.2181>
- Aylward, J., Steenkamp, E.T., Dreyer, L.L., Roets, F., Wingfield, B.D., Wingfield, M.J., 2017. A plant pathology perspective of fungal genome sequencing. *IMA Fungus* 8, 1–15. <https://doi.org/10.5598/ima fungus.2017.08.01.01>
- Bayram, Ö., Braus, G.H., 2012. Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. *FEMS Microbiol Rev* 36, 1–24. <https://doi.org/10.1111/j.1574-6976.2011.00285.x>
- Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G.A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D.H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Orengo, C.A., Pandurangan, A.P., Rivoire, C., Sigrist, C.J.A., Sillitoe, I., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Bateman, A., Finn, R.D., 2020. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkaa977>
- Bursteinas, B., Britto, R., Bely, B., Auchincloss, A., Rivoire, C., Redaschi, N., O'Donovan, C., Martin, M.J., 2016. Minimizing proteome redundancy in the UniProt Knowledgebase. *Database*: J Biol Databases Curation 2016. <https://doi.org/10.1093/database/baw139>
- Cairns, T.C., Sidhu, Y.S., Chaudhari, Y.K., Talbot, N.J., Studholme, D.J., Haynes, K., 2015. Construction and high-throughput phenotypic screening of *Zymoseptoria tritici* over-expression strains. *Fungal Genet Biol* 79, 110–117. <https://doi.org/10.1016/j.fgb.2015.04.013>
- Calvo, A.M., Lohmar, J.M., Ibarra, B., Satterlee, T., 2016. Velvet regulation of fungal development, in: *Growth, Differentiation and Sexuality, The Mycota*. Springer International Publishing, Cham, pp. 475–497. https://doi.org/10.1007/978-3-319-25844-7_18
- Cao, H., Huang, P., Zhang, L., Shi, Y., Sun, D., Yan, Y., Liu, X., Dong, B., Chen, G., Snyder, J.H., Lin, F., Lu, J., 2016. Characterization of 47 Cys2-His2 zinc finger proteins required for the development and pathogenicity of the rice blast fungus *Magnaporthe oryzae*. *New Phytol* 211, 1035–1051. <https://doi.org/10.1111/nph.13948>
- Chamberlain, S., Szoecs, E., Foster, Z., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., O'Donnell, J., 2020. taxize: Taxonomic information from around the web [WWW Document]. URL <https://github.com/ropensci/taxize> (accessed 1.1.21).
- Chankhamjon, P., Boettger-Schmidt, D., Scherlach, K., Urbansky, B., Lackner, G., Kalb, D., Dahse, H.-M., Hoffmeister, D., Hertweck, C., 2014. Biosynthesis of the halogenated mycotoxin aspirochlorine in Koji old involves a cryptic amino acid conversion. *Angew Chem Int Ed* 53, 13409–13413. <https://doi.org/10.1002/anie.201407624>
- Charoensawan, V., Wilson, D., Teichmann, S.A., 2010a. Lineage-specific expansion of DNA-binding transcription factor families. *Trends Genet.* 26, 388–393. <https://doi.org/10.1016/j.tig.2010.06.004>
- Charoensawan, V., Wilson, D., Teichmann, S.A., 2010b. Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Res* 38, 7364–7377. <https://doi.org/10.1093/nar/gkq617>
- Cho, Y., Ohm, R.A., Grigoriev, I.V., Srivastava, A., 2013. Fungal-specific transcription factor *AbPf2* activates pathogenicity in *Alternaria brassicicola*. *The Plant J* 75, 498–514. <https://doi.org/10.1111/tpj.12217>
- Choi, J., Kim, S.-H., 2017. A genome tree of life for the fungi kingdom. *PNAS* 114, 9391–9396. <https://doi.org/10.1073/pnas.1711939114>
- Chooi, Y.-H., Zhang, G., Hu, J., Muria-Gonzalez, M.J., Tran, P.N., Pettitt, A., Maier, A.G., Barrow, R.A., Solomon, P.S., 2017. Functional genomics-guided discovery of a light-activated phytotoxin in the wheat pathogen *Parastagonospora nodorum* via pathway

activation. *Environ Microbiol* 19, 1975–1986. <https://doi.org/10.1111/1462-2920.13711>

Chung, H., Choi, J., Park, S.-Y., Jeon, J., Lee, Y.-H., 2013. Two conidiation-related Zn(II)2Cys6 transcription factor genes in the rice blast fungus. *Fungal Genet Biol* 61, 133–141. <https://doi.org/10.1016/j.fgb.2013.10.004>

Coleman, J.J., Rounsley, S.D., Rodriguez-Carres, M., Kuo, A., Wasmann, C.C., Grimwood, J., Schmutz, J., Taga, M., White, G.J., Zhou, S., Schwartz, D.C., Freitag, M., Ma, L., Danchin, E.G.J., Henrissat, B., Coutinho, P.M., Nelson, D.R., Straney, D., Napoli, C.A., Barker, B.M., Gribskov, M., Rep, M., Kroken, S., Molnár, I., Rensing, C., Kennell, J.C., Zamora, J., Farman, M.L., Selker, E.U., Salamov, A., Shapiro, H., Pangilinan, J., Lindquist, E., Lamers, C., Grigoriev, I.V., Geiser, D.M., Covert, S.F., Temporini, E., VanEtten, H.D., 2009. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene Expansion. *PLOS Genetics* 5, e1000618. <https://doi.org/10.1371/journal.pgen.1000618>

Collemare, J., Lebrun, M.-H., 2011. Fungal secondary metabolites: ancient toxins and novel effectors in plant–microbe interactions, in: *Effectors in Plant–Microbe Interactions*. John Wiley & Sons, Ltd, pp. 377–400. <https://doi.org/10.1002/9781119949138.ch15>

Dalmaís, B., Schumacher, J., Moraga, J., Pêcheur, P.L., Tudzynski, B., Collado, I.G., Viaud, M., 2011. The *Botrytis cinerea* phytotoxin botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with botrydial. *Mol Plant Pathol* 12, 564–579. <https://doi.org/10.1111/j.1364-3703.2010.00692.x>

Daub, M.E., Herrero, S., Chung, K.-R., 2005. Photoactivated perylenequinone toxins in fungal pathogenesis of plants. *FEMS Microbiology Letters* 252, 197–206. <https://doi.org/10.1016/j.femsle.2005.08.033>

de Jonge, R., Ebert, M.K., Huitt-Roehl, C.R., Pal, P., Suttle, J.C., Spanner, R.E., Neubauer, J.D., Jurick, W.M., Stott, K.A., Secor, G.A., Thomma, B.P.H.J., Van de Peer, Y., Townsend, C.A., Bolton, M.D., 2018. Gene cluster conservation provides insight into cercosporin biosynthesis and extends production to the genus *Colletotrichum*. *PNAS* 115, E5459–E5466. <https://doi.org/10.1073/pnas.1712798115>

de la Torre, A., Castanheira, S., Pérez-Martín, J., 2020. Incompatibility between proliferation and plant invasion is mediated by a regulator of appressorium formation in the corn smut fungus *Ustilago maydis*. *PNAS* 117, 30599–30609. <https://doi.org/10.1073/pnas.2006909117>

de Vega-Bartol, J.J., Martín-Dominguez, R., Ramos, B., García-Sánchez, M.-A., Díaz-Mínguez, J.M., 2010. New virulence groups in *Fusarium oxysporum* f. sp. *phaseoli*: the expression of the gene coding for the transcription factor Ftf1 correlates with virulence. *Phytopathology* 101, 470–479. <https://doi.org/10.1094/PHYTO-09-10-0252>

Dufresne, M., van der Lee, T., Ben M'barek, S., Xu, X., Zhang, X., Liu, T., Waalwijk, C., Zhang, W., Kema, G.H.J., Daboussi, M.-J., 2008. Transposon-tagging identifies novel pathogenicity genes in *Fusarium graminearum*. *Fungal Genet Biol* 45, 1552–1561. <https://doi.org/10.1016/j.fgb.2008.09.004>

Emms, D.M., Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *bioRxiv* 466201. <https://doi.org/10.1101/466201>

Emms, D.M., Kelly, S., 2018. STAG: species tree inference from all genes. *bioRxiv* 267914. <https://doi.org/10.1101/267914>

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. *The Plant Cell* 18, 2369–2387. <https://doi.org/10.1105/tpc.106.042754>

Gu, Q., Zhang, C., Liu, X., Ma, Z., 2015. A transcription factor FgSte12 is required for pathogenicity in *Fusarium graminearum*. *Mol Plant Pathol* 16, 1–13. <https://doi.org/10.1111/mpp.12155>

Habig, M., Bahena-Garrido, S.M., Barkmann, F., Haueisen, J., Stukenbrock, E.H., 2020. The transcription factor Zt107320 affects the dimorphic switch, growth and virulence of the fungal wheat pathogen *Zymoseptoria tritici*. *Mol Plant Pathol* 21, 124–138. <https://doi.org/10.1111/mpp.12886>

Hane, J.K., Paxman, J., Jones, D.A.B., Oliver, R.P., de Wit, P., 2020. "CATAStrophy," a genome-informed trophic classification of filamentous plant pathogens – how many different types of filamentous plant pathogens are there? *Front Microbiol* 10. <https://doi.org/10.3389/fmicb.2019.03088>

Haridas, S., Albert, R., Binder, M., Bloem, J., LaButti, K., Salamov, A., Andreopoulos, B., Baker, S.E., Barry, K., Bills, G., Bluhm, B.H., Cannon, C., Castanera, R., Culley, D.E., Daum, C., Ezra, D., González, J.B., Henrissat, B., Kuo, A., Liang, C., Lipzen, A., Lutzoni, F., Magnuson, J., Mondo, S.J., Nolan, M., Ohm, R.A., Pangilinan, J., Park, H.-J., Ramírez, L., Alfaro, M., Sun, H., Tritt, A., Yoshinaga, Y., Zwiers, L.-H., Turgeon, B.G., Goodwin, S.B., Spatafora, J.W., Crous, P.W., Grigoriev, I.V., 2020. 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. *Stud Mycol* 96, 141–153. <https://doi.org/10.1016/j.simyco.2020.01.003>

Ikeda, K., Park, P., Nakayashiki, H., 2019. Cell biology in phytopathogenic fungi during host infection: commonalities and differences. *J Gen Plant Pathol* 85, 163–173. <https://doi.org/10.1007/s10327-019-00846-w>

John, E., Singh, K.B., Oliver, R.P., Tan, K.-C., 2021. Transcription factor control of virulence in phytopathogenic fungi. *Mol Plant Pathol* 22, 858–881. <https://doi.org/10.1111/mpp.13056>

Jones, D.A.B., John, E., Rybak, K., Phan, H.T.T., Singh, K.B., Lin, S.-Y., Solomon, P.S., Oliver, R.P., Tan, K.-C., 2019. A specific fungal transcription factor controls effector gene expression and orchestrates the establishment of the necrotrophic pathogen lifestyle on wheat. *Sci Rep* 9, 1–13. <https://doi.org/10.1038/s41598-019-52444-7>

Jonkers, W., Xayamongkhon, H., Haas, M., Olivain, C., van der Does, H.C., Broz, K., Rep, M., Alabouvette, C., Steinberg, C., Kistler, H.C., 2014. EBR1 genomic expansion and its role in virulence of *Fusarium* species. *Environ Microbiol* 16, 1982–2003. <https://doi.org/10.1111/1462-2920.12331>

Keller, N.P., 2019. Fungal secondary metabolism: regulation, function and drug discovery. *Nat Rev Microbiol* 17, 167–180. <https://doi.org/10.1038/s41579-018-0121-1>

Kelly, S., Maini, P.K., 2013. DendroBLAST: approximate phylogenetic trees in the absence of multiple sequence alignments. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0058537>

Kwon, N.-J., Garzia, A., Espeso, E.A., Ugalde, U., Yu, J.-H., 2010. FlbC is a putative nuclear C2H2 transcription factor regulating development in *Aspergillus nidulans*. *Molecular Microbiology* 77, 1203–1219. <https://doi.org/10.1111/j.1365-2958.2010.07282.x>

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47, W256–W259. <https://doi.org/10.1093/nar/gkz239>

Li, Jinyang, Liu, Q., Li, Jingen, Lin, L., Li, X., Zhang, Y., Tian, C., 2021. RCO-3 and COL-26 form an external-to-internal module that regulates the dual-affinity glucose transport system in *Neurospora crassa*. *Biotechnology for Biofuels* 14, 33. <https://doi.org/10.1186/s13068-021-01877-2>

Lind, A.L., Wisecaver, J.H., Lameiras, C., Wiemann, P., Palmer, J.M., Keller, N.P., Rodrigues, F., Goldman, G.H., Rokas, A., 2017. Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. *PLoS Biol* 15. <https://doi.org/10.1371/journal.pbio.2003583>

Liu, B.-H., Chu, F.S., 1998. Regulation of aflR and its product, AflR, associated with Aflatoxin biosynthesis. *Appl Environ Microbiol* 64, 3718–3723. <https://doi.org/10.1128/AEM.64.10.3718-3723.1998>

Lu, J., Cao, H., Zhang, L., Huang, P., Lin, F., 2014. Systematic analysis of Zn2Cys6 transcription factors required for development and pathogenicity by high-throughput gene knockout in the rice blast fungus. *PLoS Pathog* 10. <https://doi.org/10.1371/journal.ppat.1004432>

Lyu, H.-N., Liu, H.-W., Keller, N.P., Yin, W.-B., 2020. Harnessing diverse transcriptional regulators for natural product discovery in fungi. *Nat. Prod. Rep.* 37, 6–16.

- 883 <https://doi.org/10.1039/C8NP00027A>
- 884 MacPherson, S., Larochelle, M., Turcotte, B., 2006. A fungal family of transcriptional
- 885 regulators: the zinc cluster proteins. *Microbiol Mol Biol Rev* 70, 583–604.
- 886 <https://doi.org/10.1128/MMBR.00015-06>
- 887 Möller, M., Stukenbrock, E.H., 2017. Evolution and genome architecture in fungal plant
- 888 pathogens. *Nat. Rev. Microbiol.* 15, 756–771. <https://doi.org/10.1038/nrmicro.2017.76>
- 889 Mukaka, M., 2012. A guide to appropriate use of Correlation coefficient in medical research.
- 890 *Malawi Med J* 24, 69–71.
- 891 Nakamura, T., Maeda, Y., Tanque, N., Makita, T., Kato, M., Kobayashi, T., 2006. Expression
- 892 profile of amyolytic genes in *Aspergillus nidulans*. *Bioscience, Biotechnology, and*
- 893 *Biochemistry* 70, 2363–2370. <https://doi.org/10.1271/bbb.50694>
- 894 Newman, A.G., Townsend, C.A., 2016. Molecular characterization of the cercosporin
- 895 biosynthetic pathway in the fungal plant pathogen *Cercospora nicotianae*. *J Am*
- 896 *Chem Soc* 138, 4219–4228. <https://doi.org/10.1021/jacs.6b00633>
- 897 Niño Sánchez, J., Casado Del Castillo, V., Tello, V., De Vega Bartol, J.J., Ramos, B.,
- 898 Sukno, S.A., Díaz Mínguez, J.M., 2016. The FTF gene family regulates virulence and
- 899 expression of SIX effectors in *Fusarium oxysporum*. *Mol Plant Pathol* 17, 1124–1139.
- 900 <https://doi.org/10.1111/mpp.12373>
- 901 Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T.,
- 902 Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy
- 903 Joint Genome Institute: 2014 updates. *Nucleic Acids Res* 42, D26–31.
- 904 <https://doi.org/10.1093/nar/gkt1069>
- 905 Oh, M., Son, H., Choi, G.J., Lee, C., Kim, J.-C., Kim, H., Lee, Y.-W., 2016. Transcription
- 906 factor ART1 mediates starch hydrolysis and mycotoxin production in *Fusarium*
- 907 *graminearum* and *F. verticillioides*. *Mol Plant Pathol* 17, 755–768.
- 908 <https://doi.org/10.1111/mpp.12328>
- 909 Park, Jongsun, Park, Jaejin, Jang, S., Kim, Seryun, Kong, S., Choi, J., Ahn, K., Kim, J., Lee,
- 910 S., Kim, Sunggon, Park, B., Jung, K., Kim, Soonok, Kang, S., Lee, Y.-H., 2008. FTFD:
- 911 an informatics pipeline supporting phylogenomic analysis of fungal transcription
- 912 factors. *Bioinformatics* 24, 1024–1025. <https://doi.org/10.1093/bioinformatics/btn058>
- 913 Pedro, H., Maheswari, U., Urban, M., Irvine, A.G., Cuzick, A., McDowall, M.D., Staines, D.M.,
- 914 Kulesha, E., Hammond-Kosack, K.E., Kersey, P.J., 2016. PhytoPath: an integrative
- 915 resource for plant pathogen genomics. *Nucleic Acids Res* 44, D688–D693.
- 916 <https://doi.org/10.1093/nar/gkv1052>
- 917 Pinedo, C., Wang, C.-M., Pradier, J.-M., Dalmais, B., Choquer, M., Pêcheur, P.L., Morgant,
- 918 G., Collado, I.G., Cane, D.E., Viaud, M., 2008. The sesquiterpene synthase from the
- 919 botrydial biosynthetic gene cluster of the phytopathogen *Botrytis cinerea*. *ACS Chem*
- 920 *Biol* 3, 791–801. <https://doi.org/10.1021/cb800225v>
- 921 Plissonneau, C., Benevenuto, J., Mohd-Assaad, N., Fouché, S., Hartmann, F.E., Croll, D.,
- 922 2017. Using population and comparative genomics to understand the genetic basis of
- 923 effector-driven fungal pathogen evolution. *Front Plant Sci* 8.
- 924 <https://doi.org/10.3389/fpls.2017.00119>
- 925 Porquier, A., Moraga, J., Morgant, G., Dalmais, B., Simon, A., Sghyer, H., Collado, I.G.,
- 926 Viaud, M., 2019. Botcinic acid biosynthesis in *Botrytis cinerea* relies on a
- 927 subtelomeric gene cluster surrounded by relics of transposons and is regulated by the
- 928 Zn2Cys6 transcription factor BcBoa13. *Curr Genet* 65, 965–980.
- 929 <https://doi.org/10.1007/s00294-019-00952-4>
- 930 Porquier, A., Morgant, G., Moraga, J., Dalmais, B., Luyten, I., Simon, A., Pradier, J.-M.,
- 931 Amselem, J., Collado, I.G., Viaud, M., 2016. The botrydial biosynthetic gene cluster of
- 932 *Botrytis cinerea* displays a bipartite genomic structure and is positively regulated by
- 933 the putative Zn(II)2Cys6 transcription factor BcBot6. *Fungal Genet Biol* 96, 33–46.
- 934 <https://doi.org/10.1016/j.fgb.2016.10.003>
- 935 Pusztahelyi, T., Holb, I.J., Pócsi, I., 2016. Plant-Fungal Interactions: Special Secondary
- 936 Metabolites of the Biotrophic, Necrotrophic, and Other Specific Interactions, in:
- 937 Mérillon, J.-M., Ramawat, K.G. (Eds.), *Fungal Metabolites, Reference Series in*

Phytochemistry. Springer International Publishing, Cham, pp. 1–58.
https://doi.org/10.1007/978-3-319-19456-1_39-1

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rispail, N., Di Pietro, A., 2010. The homeodomain transcription factor Ste12. *Commun Integr Biol* 3, 327–332.

Rokas, A., Wisecaver, J.H., Lind, A.L., 2018. The birth, evolution and death of metabolic gene clusters in fungi. *Nature Reviews Microbiology* 16, 731–744.
<https://doi.org/10.1038/s41579-018-0075-3>

Ruocco, M., Baroncelli, R., Cacciola, S.O., Pane, C., Monti, M.M., Firrao, G., Vergara, M., Magnano di San Lio, G., Vannacci, G., Scala, F., 2018. Polyketide synthases of *Diaporthe helianthi* and involvement of DhPKS1 in virulence on sunflower. *BMC Genomics* 19, 27. <https://doi.org/10.1186/s12864-017-4405-z>

Sarmiento-Villamil, J.L., Prieto, P., Klosterman, S.J., García-Pedrajas, M.D., 2018. Characterization of two homeodomain transcription factors with critical but distinct roles in virulence in the vascular pathogen *Verticillium dahliae*. *Mol Plant Pathol* 19, 986–1004. <https://doi.org/10.1111/mpp.12584>

Schoch, C.L., Ciuffo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., McVeigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J.P., Sun, L., Turner, S., Karsch-Mizrachi, I., 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*: J Biol Databases Curation 2020. <https://doi.org/10.1093/database/baaa062>

Shelest, E., 2017. Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. *Front. Genet.* 8. <https://doi.org/10.3389/fgene.2017.00053>

Shelest, E., 2008. Transcription factors in fungi. *FEMS Microbiol. Lett.* 286, 145–151. <https://doi.org/10.1111/j.1574-6968.2008.01293.x>

Sievers, F., Higgins, D.G., 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Science* 27, 135–145. <https://doi.org/10.1002/pro.3290>

Son, H., Seo, Y.-S., Min, K., Park, A.R., Lee, J., Jin, J.-M., Lin, Y., Cao, P., Hong, S.-Y., Kim, E.-K., Lee, S.-H., Cho, A., Lee, S., Kim, M.-G., Kim, Y., Kim, J.-E., Kim, J.-C., Choi, G.J., Yun, S.-H., Lim, J.Y., Kim, M., Lee, Y.-H., Choi, Y.-D., Lee, Y.-W., 2011. A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. *PLOS Pathogens* 7, e1002310. <https://doi.org/10.1371/journal.ppat.1002310>

Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>

Tan, K.-C., Oliver, R.P., 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. *PLOS Pathogens* 13, e1006241. <https://doi.org/10.1371/journal.ppat.1006241>

Tangsombatvichit, P., Semkiv, M.V., Sibirny, A.A., Jensen, L.T., Ratanakhanokchai, K., Soontorngun, N., 2015. Zinc cluster protein Znf1, a novel transcription factor of non-fermentative metabolism in *Saccharomyces cerevisiae*. *FEMS Yeast Research* 15. <https://doi.org/10.1093/femsyr/fou002>

Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K., Kobayashi, T., Tsukagoshi, N., 2001. Characterization of the amyR gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. *Curr Genet* 39, 10–15. <https://doi.org/10.1007/s002940000175>

Todd, R.B., Zhou, M., Ohm, R.A., Leeggangers, H.A., Visser, L., de Vries, R.P., 2014. Prevalence of transcription factors in ascomycete and basidiomycete fungi. *BMC Genomics* 15, 214. <https://doi.org/10.1186/1471-2164-15-214>

Uka, V., Cary, J.W., Lebar, M.D., Puel, O., Saeger, S.D., Mavungu, J.D.D., 2020. Chemical repertoire and biosynthetic machinery of the *Aspergillus flavus* secondary metabolome: A review. *Comprehensive Reviews in Food Science and Food Safety* 19, 2797–2842. <https://doi.org/10.1111/1541-4337.12638>

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S.Y., De Silva, N.,
Martinez, M.C., Pedro, H., Yates, A.D., Hassani-Pak, K., Hammond-Kosack, K.E.,
2020. PHI-base: the pathogen–host interactions database. *Nucleic Acids Res* 48,
D613–D620. <https://doi.org/10.1093/nar/gkz904>

van der Does, H.C., Fokkens, L., Yang, A., Schmidt, S.M., Langereis, L., Lukasiewicz, J.M.,
Hughes, T.R., Rep, M., 2016. Transcription factors encoded on core and accessory
chromosomes of *Fusarium oxysporum* induce expression of effector genes. *PLOS*
Genetics 12, e1006401. <https://doi.org/10.1371/journal.pgen.1006401>

van der Does, H.C., Rep, M., 2017. Adaptation to the host environment by plant-pathogenic
fungi. *Annual Review of Phytopathology* 55, 427–450.
<https://doi.org/10.1146/annurev-phyto-080516-035551>

Velez-Haro, J.M., Martínez-Soto, D., Guevara-Olvera, L., Ruiz-Herrera, J., 2020. Ztf1, an
Ustilago maydis transcription factor involved in virulence. *Eur J Plant Pathol* 156,
189–200. <https://doi.org/10.1007/s10658-019-01877-x>

Vollmeister, E., Schipper, K., Baumann, S., Haag, C., Pohlmann, T., Stock, J., Feldbrügge,
M., 2012. Fungal development of the plant pathogen *Ustilago maydis*. *FEMS*
Microbiology Reviews 36, 59–77. <https://doi.org/10.1111/j.1574-6976.2011.00296.x>

Wang, M., Sun, X., Zhu, C., Xu, Q., Ruan, R., Yu, D., Li, H., 2015. Pdb1A, PdabA and
PdwtA control distinct stages of conidiogenesis in *Penicillium digitatum*. *Res*
Microbiol 166, 56–65. <https://doi.org/10.1016/j.resmic.2014.12.003>

Wang, Q.-M., Begerow, D., Groenewald, M., Liu, X.-Z., Theelen, B., Bai, F.-Y., Boekhout, T.,
2015. Multigene phylogeny and taxonomic revision of yeasts and related fungi in the
Ustilaginomycotina. *Stud Mycol* 81, 55–83.
<https://doi.org/10.1016/j.simyco.2015.10.004>

Wong Sak Hoi, J., Dumas, B., 2010. Ste12 and Ste12-like proteins, fungal transcription
factors regulating development and pathogenicity. *Eukaryotic Cell* 9, 480–485.
<https://doi.org/10.1128/EC.00333-09>

Xiong, Y., Wu, V.W., Lubbe, A., Qin, L., Deng, S., Kennedy, M., Bauer, D., Singan, V.R.,
Barry, K., Northen, T.R., Grigoriev, I.V., Glass, N.L., 2017. A fungal transcription
factor essential for starch degradation affects integration of carbon and nitrogen
metabolism. *PLOS Genetics* 13, e1006737.
<https://doi.org/10.1371/journal.pgen.1006737>

Yu, J.-H., Butchko, R.A.E., Fernandes, M., Keller, N.P., Leonard, T.J., Adams, T.H., 1996.
Conservation of structure and function of the aflatoxin regulatory gene aflR
from *Aspergillus nidulans* and *A. flavus*. *Curr Genet* 29, 549–555.
<https://doi.org/10.1007/BF02426959>

Zhang, W.-Q., Gui, Y.-J., Short, D.P.G., Li, T.-G., Zhang, D.-D., Zhou, L., Liu, C., Bao, Y.-M.,
Subbarao, K.V., Chen, J.-Y., Dai, X.-F., 2018. *Verticillium dahliae* transcription factor
VdFTF1 regulates the expression of multiple secreted virulence factors and is
required for full virulence in cotton. *Mol Plant Pathol* 19, 841–857.
<https://doi.org/10.1111/mpp.12569>

Zhao, C., Waalwijk, C., de Wit, P.J.G.M., van der Lee, T., Tang, D., 2011. EBR1, a novel
Zn(2)Cys(6) transcription factor, affects virulence and apical dominance of the hyphal
tip in *Fusarium graminearum*. *Mol Plant Microbe Interact* 24, 1407–1418.
<https://doi.org/10.1094/MPMI-06-11-0158>

Zheng, Y., Kief, J., Auffarth, K., Farfing, J.W., Mahler, M., Nieto, F., Basse, C.W., 2008.
The *Ustilago maydis* Cys2His2-type zinc finger transcription factor Mzr1 regulates
fungal gene expression during the biotrophic growth stage. *Molecular Microbiology*
68, 1450–1470. <https://doi.org/10.1111/j.1365-2958.2008.06244.x>

Zhou, W., Dorrity, M.W., Bubb, K.L., Queitsch, C., Fields, S., 2020. Binding and regulation of
transcription by yeast Ste12 variants to drive mating and invasion phenotypes.
Genetics 214, 397–407. <https://doi.org/10.1534/genetics.119.302929>

Zhu, X., Liu, W., Chu, X., Sun, Q., Tan, C., Yang, Q., Jiao, M., Guo, J., Kang, Z., 2018. The
transcription factor PstSTE12 is required for virulence of *Puccinia striiformis* f. sp.
tritici. *Mol Plant Pathol* 19, 961–974. <https://doi.org/10.1111/mpp.12582>

1048

Lifestyle

Host association

Fungal division	ANCOVA	Regression line groupings
Lifestyle	Pval = 0.561	Biotroph ^A , Hemibiotroph ^A , Necrotroph ^A , Saprophyte ^A , Symbiont ^A
Host dependence	Pval = 9.13E-5*	Facultative ^A , Not host-associated ^A , Obligate ^B

NCBI taxonomy tree

TF-distance species tree

Lifestyle

- Biotroph
- Hemibiotroph
- Necrotroph
- Saprophyte
- Symbiont

Host association

- Facultative
- Not host associated
- Obligate

Clustered fungal transcription factor orthogroups

Jukes-Cantor distance: 0.2