Fungal Genetics and Biology 161 (2022) 103712

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Transcription factor lineages in plant-pathogenic fungi, connecting diversity with fungal virulence

Evan John [a,*], Karam B. Singh [a,b], Richard P. Oliver [c], Kar-Chun Tan [a,*]

[a] Centre for Crop and Disease Management, Curtin University, Perth, Australia
[b] Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation, Perth, Australia
[c] School of Biosciences, University of Nottingham, United Kingdom

ARTICLE INFO

Keywords:
Transcription factor
Plant-pathogenic fungi
Virulence
Regulation
Orthogroup

ABSTRACT

Plant-pathogenic fungi span diverse taxonomic lineages. Their host-infection strategies are often specialised and require the coordinated regulation of molecular virulence factors. Transcription factors (TFs) are fundamental regulators of gene expression, yet relatively few virulence-specific regulators are characterised in detail and their evolutionary trajectories are not well understood. Hence, this study compared the full range of TFs across taxonomically-diverse fungal proteomes and classified their lineages through an orthology analysis. The primary aims were to characterise differences in the range and profile of TF lineages broadly linked to plant-host association or pathogenic lifestyles, and to better characterise the evolutionary origin and trajectory of experimentally-validated virulence regulators. We observed significantly fewer TFs among obligate, host-associated pathogens, largely attributed to contractions in several Zn2Cys6 TF-orthogroup lineages. We also present novel insight into the key virulence-regulating TFs Ste12, Pf2 and EBR1, providing evidence for their ancestral origins, expansion and/or loss. Ultimately, the analysis presented here provides both primary evidence for TF evolution in fungal phytopathogenicity, as well as a practical phylogenetic resource to guide further detailed investigation on the regulation of virulence within key pathogen lineages.

1. Introduction

Plant-pathogenic fungi have adapted to a wide range of ecological niches. It is now clear that pathogenic lifestyles, traditionally 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, phytotoxin-biosynthesis genes 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 coordinated regulation to maximise activity at the appropriate stage of infection (van der Does and Rep, 2017). Hence, the evolution of virulence-lifestyle 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 fungal taxa (Park et al., 2008; Todd et al., 2014; Shelest, 2017). TF gene duplications that are directly implicated in fungal virulence have been documented in several lineages of *Fusarium oxysporum*, the causal agent of vascular wilt on a wide range of plants (de Vega-Bartol et al., 2010; Niño-Sánchez et al., 2016; van der Does et al., 2016). On the other hand, broadly conserved regulators in plant-pathogenic fungi, such as the Velvet TFs, are absent in saprophytic yeasts (Bayram and Braus, 2012; Calvo et al., 2016). As such, it was hypothesised the range and profile of TF lineages in phytopathogenic fungi change in accordance with their host-

Abbreviations: TF, Transcription factor; DBD, DNA-binding domain; TFome, complete TF repertoire; ER, evolution rate; PC, principal component.

* Corresponding authors.
E-mail addresses: evan.john@curtin.edu.au (E. John), Kar-Chun.Tan@curtin.edu.au (K.-C. Tan).

https://doi.org/10.1016/j.fgb.2022.103712
Received 16 December 2021; Received in revised form 18 May 2022; Accepted 30 May 2022
Available online 3 June 2022

association & lifestyle.

Random mutagenesis or high-throughput gene knockout studies are established approaches that have been used to identify virulence-regulating TFs. For example, VdFTF1 in *Verticillium dahliae*, which regulates the expression of secreted virulence factors during infection on cotton, was identified by screening a random mutagenesis library (Zhang et al., 2018). On the other hand, a genome-wide TF knockout screen was undertaken for *Fusarium graminearum*, which identified several virulence regulators otherwise dispensable for saprophytic growth (Son et al., 2011). However, the generation and screening of fungal mutant libraries in this fashion can be labour intensive. Critically, the evolutionary lineages of such virulence regulators are often poorly understood through traditional database-similarity searches. It was therefore hypothesised that a comprehensive orthology-based analysis will reveal novel trajectories for the ancestral orthologues of experimentally-validated virulence regulators, and provide key insight into their evolution.

This study aimed to address these hypotheses in two distinct stages. First, the TF repertoires ("TFomes") were compiled from the annotated proteomes of 120 representative fungi to identify broad differences linked to plant host-association or pathogenic lifestyles. The TFs were then classified into orthogroups to characterise the individual lineages undergoing major expansion and/or contractions. The second stage made use of the TF-orthogroup phylogenies defined in the analysis to investigate the evolutionary origin and trajectories for several established fungal-virulence regulators to provide novel insight.

2. Methods

2.1. Compiling fungal TFomes

Non-redundant fungal proteomes were sourced for 100 plant pathogens, 10 saprophytes and 10 symbionts from UniProt (release 2020_05) (Bursteinas et al., 2016) or MycoCosm (Nordberg et al., 2014) and independent platforms where the proteomes were available with higher coverage assemblies (**Supplemental item 1**). Pathogens were selected from diverse fungal taxa of significant research interest (Pedro et al., 2016; Urban et al., 2020). The saprophytes and symbionts included both close and distant pathogen relatives across the taxa. Organisms were assigned a five letter organism ID (ORGID) for brevity (**Supplemental item 1**), adapted from their NCBI Taxonomy database naming convention (Schoch et al., 2020). General lifestyles (biotroph, hemibiotroph, necrotroph, saprophyte or symbiont) and plant host-associations (i.e. obligate, facultative or not associated) were assigned based on descriptions from the published proteome sources and/or literature covering the matter (Möller and Stukenbrock, 2017; Hane et al., 2020; Haridas et al., 2020). Interpro protein-domain annotations were also downloaded with the respective fungal proteome sources. Interproscan (release 82.0) was used to annotate proteomes from the independent servers where these were not available (Blum et al., 2020). TFs were selected from the annotated proteomes by cross-referencing a list of Interpro domain accessions that represent sequence-specific DNA-binding domains (DBDs) (**Supplemental item 2**). Unique protein identifiers were assigned to each TF following the format 'ProteinID | ORGID' where the 'ProteinID' was derived from the downloaded proteome source.

2.2. TFome regression analysis

The total occurrences of each annotated TF domain across the annotated fungal proteomes were counted using a custom bash script (which accounted for cases where multiple domains exist within a single protein). 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 were 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. Orthology analysis and species-tree inference

Two orthology analyses were separately undertaken using OrthoFinder (release 2.4.0) (Emms and Kelly, 2019). The entire set of fungal proteomes was analysed to define proteome-orthogroups and the annotated TFomes were analysed in isolation to define TF-orthogroups. The default OrthoFinder parameters were used with the exception that MMseqs2 (Steinegger and Söding, 2017) was invoked for the sequence search stage. The annotated TFs among the 64,299 proteome-orthogroups were calculated by extracting the number of member-proteins with at least one DBD from the annotated list in **Supplemental item 2**. A 'species tree' was also produced from 502 of the universally conserved proteome-orthogroups as part of the OrthoFinder analysis with STAG (Emms and Kelly, 2018), which was used to present phylogenetic taxonomic distances. The species tree was manually rooted at the most distantly related lineage (Chytridiomycota), 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 TF-orthogroup 'protein trees'. The protein trees were built on the basis of the DendroBLAST distance measure (Kelly and Maini, 2013) as that is built in to 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).

2.4. TF-orthogroup size comparative analyses

Fungal TF-orthogroup datasets were ordered as per the leaf-node order derived from the species-tree (beginning with the root at SYNEN) for subsequent visualisation and analyses. TF evolution-rate parameters were calculated using COUNT, which implements a phylogenetic birth-and-death model (Csűös, 2010). These were calculated for all internal and extant nodes on the species-tree using a table of the 855 TF-orthogroup counts across each of the fungi assessed. The default gain-loss-duplication model was used in COUNT with the convergence threshold of 0.1 and the default "same gain/loss ratio in all lineages" to produce a single TF evolution rate parameter at each species-tree node. A principal component analysis was then undertaken to identify components underpinning fungal differences based on rlog normalised TF-orthogroup counts to reduce biases associated with proportionally-low-overall orthogroup sizes (Love et al., 2014). A hierarchical clustering analysis was also conducted to group the TF orthogroups based on normalised 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. Data formatting, analysis and the production of plots and heatmaps were undertaken in R (version 4) using the phylogram, complex heatmap, ggplot2, ggpubr and the rstats packages (R Core Team, 2020).

2.5. Independent analysis of established fungal-virulence regulators

Orthogroups harbouring established virulence regulators were determined by cross-referencing the respective ProteinIDs with experimentally-characterised TFs from the scientific literature (John et al., 2021). The related 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. Insights on the range and profile of TFs across 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 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 non-specific single-stranded nucleic-acid binding proteins that did not fit the TF definition (IPR012340 and IPR001878) (**Supplemental item 2**). Hence, the TFomes were assembled from a refined fungal TF-annotation criteria, which can also be updated in future studies when DBD-models are further improved. The 100 plant-pathogenic fungi analysed encompassed a broad taxonomic set of pathogens, based on research interest and available genomic resources, while accounting for genomic variability in major pathogen lineages such as *Fusarium*, *Bipolaris* and *Puccinia* (Ma et al., 2010; Duplessis et al., 2011; Baroncelli et al., 2017). Along with the 20 saprophytes and symbiotic fungi included, this provided a practical resource to investigate broad trends in the range and profile of TFs relevant to virulence on plants (**Supplemental item 1**).

3.1.1. Reduction of TFs is correlated with obligate host-associated fungi

The relative TF content (TFome size) across each of the 120 proteomes was first assessed to gauge any broader trends in 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

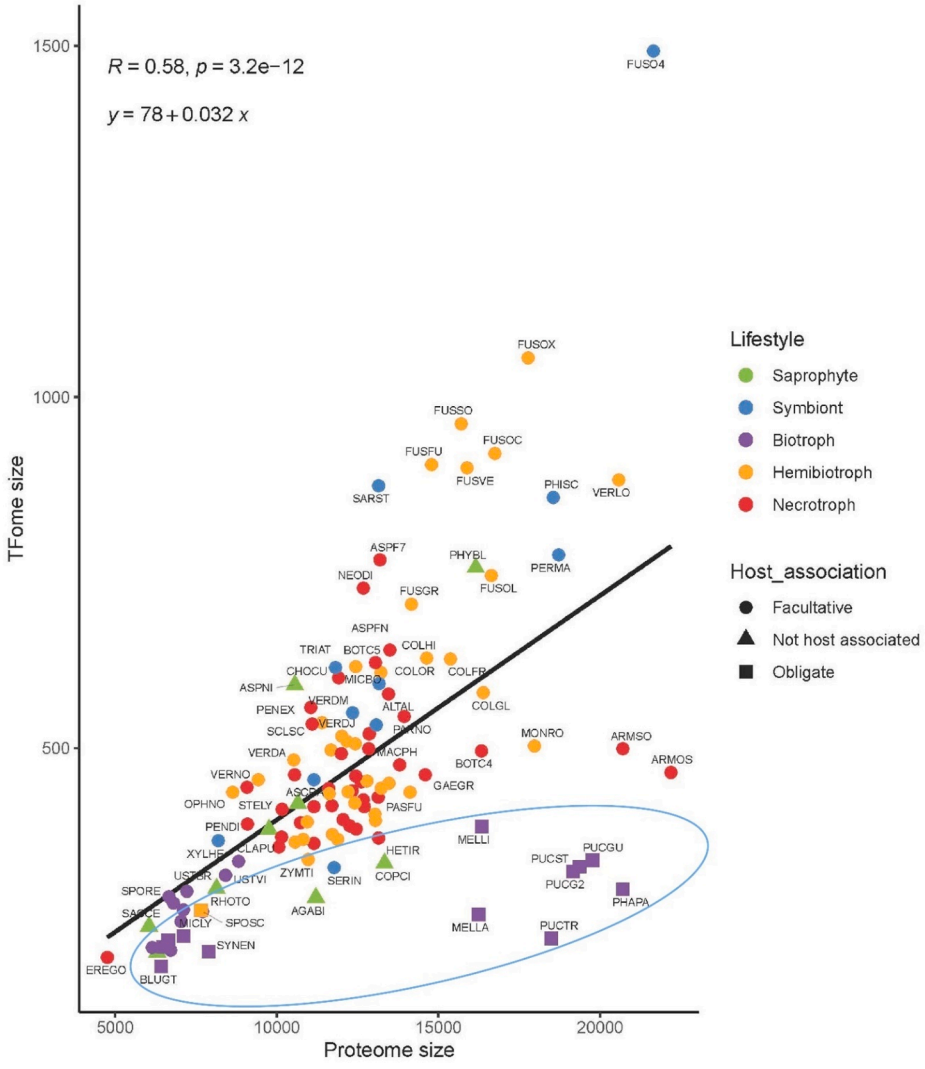

Fig. 1. Fungal transcription factor content (TFome) relative to proteome size. **TFome vs proteome sizes for the 120 fungi used in this analysis. The five-letter fungal ORGIDs are used to label the respective points which provide an indication of the respective lifestyle and host-association (Supplemental item 1). 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 blue circle encompasses the obligate host-associated pathogens, which exhibited a significant reduction relative to proteome sizes (detailed in Supplemental item 3).** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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 variation was higher among the fungi assessed in this study.

The higher variation in the TFome ~ proteome relationship prompted a comparison based on fungal sub-groups, defined by their lifestyle (i.e. biotrophic, hemibiotrophic, necrotrophic, saprophytic or symbiotic) or their association with plant hosts (i.e. facultative, obligate or not host-associated), to identify any broad trends relevant to phytopathogenicity. Linear relationships were maintained within each division. Interestingly, the TFome as a proportion of proteome size was significantly reduced for obligate host-associated fungi relative to facultative pathogens and non-host saprophytes (**Supplemental item 3**). 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. Although we did not identify significant differences based on fungal-lifestyle definitions, a large overlap clearly exists where most obligates are also biotrophs. In facultative-biotroph lineages (e.g. Clavicipitaceae, Ustilaginales, Mixiomycetes and Taphrinomycetes) the TFome sizes, but also proteome sizes, were all lower than the median (437.5 and 12,296.5 respectively; **Supplemental item 1**). These observations provide evidence that a restricted TF 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.

3.1.2. *TF lineages evolve in plant-pathogenic fungi*

Differences in size of the fungal TFomes, particularly the reduced range in obligate/biotrophic pathogens, motivated a deeper analysis into the specific profile of the TF lineages undergoing expansion or contraction. Previous studies compared fungal TFs by DBD-family classifications and revealed the Zn2Cys6 and C2H2 zinc finger domains are the most frequent (Todd et al., 2014; Shelest, 2017), which paralleled this study. Of the 56,754 TFs compiled from the 120 fungal proteomes 32,023 Zn2Cys6, 11,436 C2H2, 4,777 homeodomain, 3,168 bZIP and 2,932 CCCH-type DBDs represented the top families overall (**Supplemental item 2**). Their large sizes indicated an orthology analysis would better resolve distinct TF lineages to specify differences in TF evolution across pathogenic taxa.

The first approach encompassed the entire set of 120 fungal proteomes which produced 64,299 orthogroups (**Supplemental item 4**). There were 2,743 proteome-orthogroups with at least one annotated TF. Most of the TFs were identified in orthogroups where the majority of the proteins also harboured DBDs (Fig. 2**A**). Specifically, 89% of the 56,754 annotated TFs were classed in orthogroups where at least half of the proteins were considered TFs (**Supplemental item 5**). This provided confidence that DBDs are a conserved factor underpinning proteome-orthogroup assignment. However, a large proportion of proteins that did not harbour DBD annotations had been incorporated in the remaining proteome-orthogroups (Fig. 2**A;** Fig. 2**B**). These could represent real TFs where the DBD model excludes the domain variant (false negatives), ancestral TFs where the DBD has mutated or been lost (true negatives), unrelated proteins where a protein-rearrangement incorporated a single DBD to an orthogroup-member (true negatives) or proteins that are homologous to real TFs in regions outside the DBD (true negatives).

Since most annotated TFs had grouped in common lineages using the first approach, a second orthology analysis was undertaken based solely on annotated fungal TFomes to remove the true negatives. The result was 855 TF-orthogroup lineages that encompassed 56,418 of the 56,754 annotated TFs (Fig. 2**B; Supplemental item 6**). While likely some real

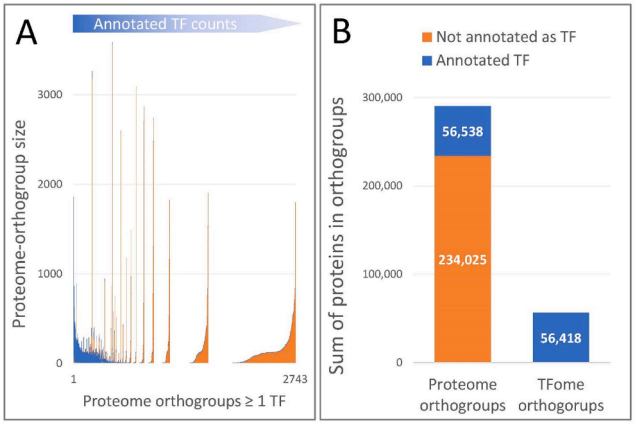

Fig. 2. Annotated transcription factor (TF) composition of orthogroups. **Panel A) depicts the annotated-TF distribution across the 2,743 (of 64,299) proteome-orthogroups containing at least one annotated TF. The Y-axis corresponds to the respective orthogroup sizes. The stacked-bars distributed along the X-axis represent the proportion of annotated TFs (blue) in the 2,743 orthogroups. These were distributed from left to right by decreasing size of their annotated-TF counts. Panel B) compares the annotated-TF orthogroup composition from the two orthology analyses conducted (full-proteome based vs TFome based). The first bar represents the total composition of the 2,743 proteome-orthogroups that possessed at least one annotated TF. The second bar represents the total composition from the 'high-confidence' set of 855 TF-orthogroups built solely from fungal TFomes. Accordingly, proteins not annotated as TFs had been excluded prior to TF-orthogroup assignment and are absent in the second bar.** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

TFs are excluded with this approach, these presented a high-confidence set that excluded most non-TFs, and with which expansion and contraction could be measured in specific TF lineages across plant-pathogenic fungi.

The rates of TF expansion and contraction at the phylogenetic species-tree nodes, their 'evolution-rate' parameters (Csűös, 2010), were calculated from the relative sizes of the 855 TF-orthogroups across the fungal taxa (**Supplemental item 7**). The species-tree topology aligned well with the fungal taxonomic architecture (Fig. 3) and was investigated to identify the higher TF evolution-rates (ER) relative to phylogenetic distance at pathogen lineages. The highest rate (ER = 1.235) was observed at the Mucoromycota node (N3), an early-diverging fungal lineage. Since the early-diverging phyla (Mucoromycota and Chytridiomycota) do not represent major pathogen lineages, it is difficult to attribute any significant trends. Within the remaining fungal lineages, ERs were highest overall at the terminal fungal nodes, suggesting the TF profiles generally do not become fixed and continue to change (**Supplemental item 7**). The highest ER observed at an internal ancestral node was the Pezizomycotina division (N15: ER = 0.346), which represents filamentous Ascomycete fungi. Interestingly, the subsequent highest internal-node ERs were identified at the Clavicipitaceae (N94: ER = 0.248), the Erysiphales (N38: ER = 0.248) and the Pucciniales (N11: ER = 0.207). These biotrophic pathogen lineages also encompassed a large fraction of the obligate host-associated fungi with a reduced overall TFome size, likely through contraction (Fig. 3). This suggested the reductions associated with these pathogenic lifestyles could be attributed to specific TF-orthogroup lineages.

A principal component (PC) analysis of normalised TF-orthogroup counts across the fungal species/strains provided insight into some of the major ER differences linked to ancestral fungal nodes. A major separation was observed for PC1, which was the largest PC and explained 38.75% of the variance, as opposed to PC2 which only accounted for 6.33% (Fig. 4). The filamentous Ascomycota (Pezizomycotina) formed a group distinct from other fungi with the exception that

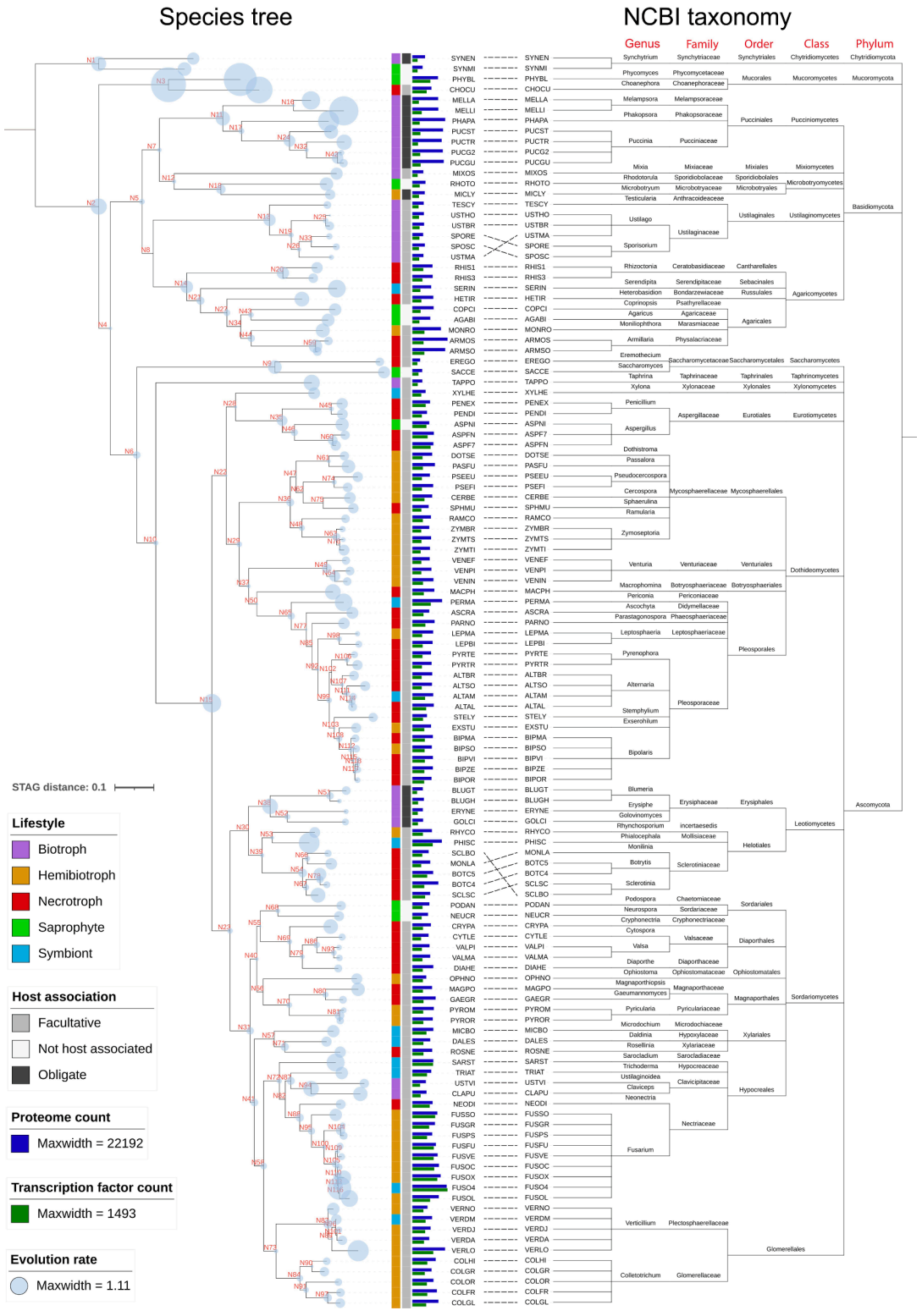

Fig. 3. Fungal phylogenetic species-tree, taxonomy and relative transcription factor (TF) evolution rates. **The species tree represents fungal evolutionary distances, which was built from the 502 proteins that were universally conserved across proteome-orthogroups using STAG (Emms and Kelly, 2018). The internal nodes, representing putative ancestral lineages are labelled with node IDs N1 – N118. Clade architectures derived from NCBI taxonomic are presented in the right tree with the terminal nodes aligned. The five-letter fungal ORGIDs are used with the respective lifestyle and host-association described in the corresponding legends (Supplemental item 1). Fungal proteome and TFome sizes are represented by the width of the blue and green bars respectively; the largest bar (maxwidth) corresponds to the respective counts indicated in the legend. The blue bubbles represent the square-root values for the COUNT (Csűös, 2010) TF evolution-rate (ER) parameter which measures relative TF-orthogroup expansion/contraction. Higher ER values are represented by larger bubble diameters up to the largest value (maxwidth) indicated in the legend. A pdf version is provided for enhanced viewing as Supplemental item 8.** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

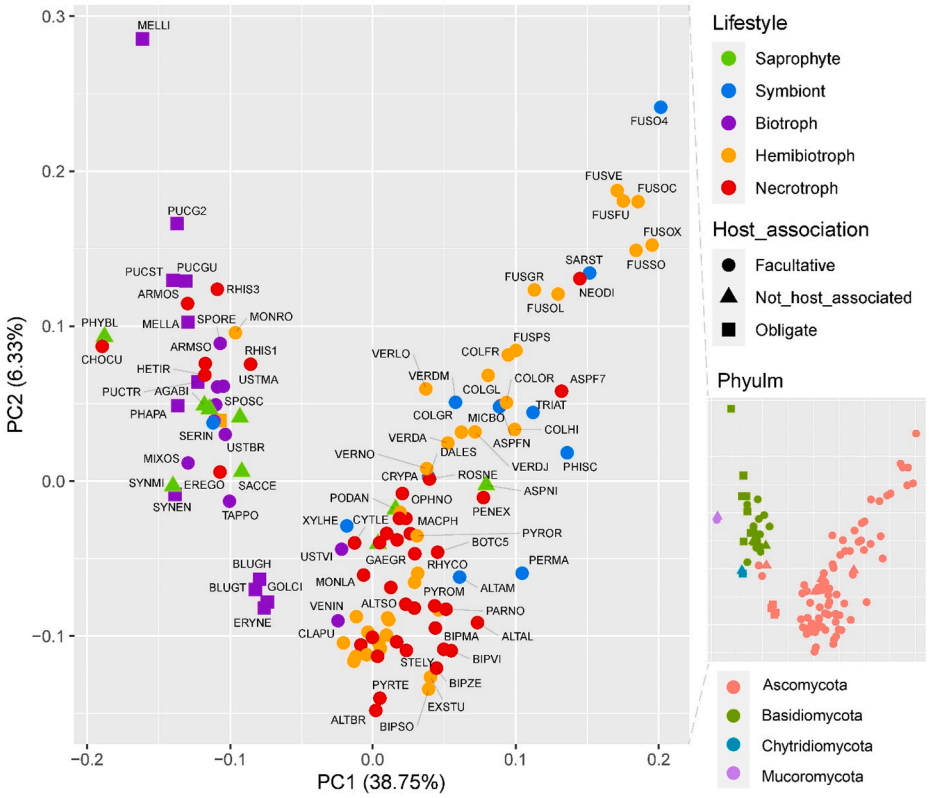

Fig. 4. Principal components (PC) by transcription factor (TF) orthogroup counts. **A PC-analysis plot separating fungi based on normalised counts in the 855 high-confidence annotated TF-orthogroup lineages. The greatest TF-orthogroup variance is explained by PC1 on the X-axis. The five-letter fungal ORGIDs are used to label the respective points which provide an indication of the respective lifestyle and host-association (Supplemental item 1). The mini-plot represents the same plot with the divisions coloured by taxonomic Phyla. Within and across these divisions, the obligate-biotrophic pathogens trend to separate towards negative values across PC1. The TF-orthogroup loadings underlying PC1 and PC2 are documented alongside the orthogroup DNA-binding domain annotations in Supplemental item 9.**

the biotrophic lineages within this group (Clavicipitaceae and the Erysiphales) grouped closer to the other taxonomies. The Basidiomycota, which encompassed a greater proportion of biotrophic fungi overall, formed the other major taxonomic grouping separated from the Pezizomycotina across PC1, the obligately-pathogenic Pucciniales in particular (Fig. 4). By examining the top loadings values for PC1 (**Supplemental item 9**), 16 of the top 20 TF-orthogroups consisted primarily of TFs harbouring a Zn2Cys6 DBD. This suggested many ancestral Zn2Cys6 TF orthologues have undergone major expansions in the Pezizomycotina, which subsequently contracted in the biotrophic lineages. It is therefore likely that most Zn2Cys6 TFs are dispensable for the biotrophic lifestyle, even more so if an obligate host-association evolves.

A clustering analysis was then undertaken on normalised TF-orthogroup counts across the 120 fungi to uncover any other distinct trends (**Supplemental item 10**). There were no obvious orthogroup clusters revealing TF expansion/contraction associated with the broadly defined pathogenic lifestyle or host-associations. However, several isolated fungal lineages exhibited orthogroup clusters well above (Z-score > 2) and/or below (Z-score < -2) the average TF counts. This indicated TF-lineages have evolved, but that identifying changes correlated with the regulation of virulence will require more precise lifestyle-definitions, or an investigation assessing a greater number of pathogens vs non-pathogens from closely related taxa. Nevertheless, this analysis provides a useful resource to place TF lineage-composition in the broader context of fungi at large.

3.2. *Regulators of virulence in conserved and recently evolved TFs*

It was hypothesised that the TF-orthogroup phylogenies (their 'protein trees') would provide insight into the evolutionary trajectories of ancestral TF orthologues. The protein trees of several experimentally-validated virulence regulators were investigated to document variation in plant-pathogen lineages. These were chosen on the basis of a recent review (John et al., 2021) which suggested they underpin virulence but not viability, share conserved roles in the pathogens experimentally-investigated, and that their prevalence across fungi remained unclear which this orthology-analysis could better resolve. This included Pf2, Ste12 and EBR1 but should not be considered exhaustive. Hence, the respective protein-trees for all TF-orthogroups (OG0000000 > OG0000854) are provided (**Supplemental item 11**) to facilitate future investigation on the evolutionary origin and trajectories of other TFs that underpin fungal virulence.

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

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) were identified in the TF-orthogroup OG0000017. These were traced as direct taxonomic orthologues that form a distinct clade with *Neurospora crassa* Col-26 (**Supplemental item 12**). 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), processes highly relevant to phytopathogenicity.

The OG0000017 comprised 370 TFs with annotated Zn2Cys6 DBDs (IPR001138) and 'Fungal transcription factor' domains (IPR007219) spanning the Ascomycota. Interestingly, it also includes MAL13, ZNF1, MAL33 and YGR288W, TFs 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 indicates these four TFs are paralogues in the context of this orthogroup (**Supplemental item 12**). 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 hydrolysis (Tani et al., 2001; Nakamura et al., 2006). 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 for which there were no other functionally defined

orthologues identified. A protein sequence alignment/phylogenetic tree comparing MAL13, AmyR and the experimentally-characterised Pf2-orthologues highlighted common and distinguishing features (Fig. 5). A relatively high degree of sequence conservation was observed overall at the Zn2Cys6 DBD and the Fungal transcription factor domains. However, a region unique to the Pf2-orthologues was evident at the C-terminus of the Zn2Cys6 DBD (Fig. 5), suggesting a possible differentiation from both AmyR and MAL13 at DNA-binding residues.

Outside the Pf2-orthologue clade in OG0000017 there were no other TFs identified with experimentally-validated virulence roles. This encompassed 12 other TFs that were previously screened as part of Zn2Cys6 gene-deletion-mutant libraries in both *M. oryzae* and *F. graminearum* (Son et al., 2011; Lu et al., 2014). Hence, this TF-orthogroup analysis established an evolutionary origin for a distinct Pf2-orthologue clade from a lineage of carbohydrate/polysaccharide metabolic regulators. This clade, based on current experimental investigation, has evolved a fundamental regulatory role in the respective fungi, particularly relevant to the plant pathogens that exhibit a necrotrophic phase in their lifecycle.

3.2.2. *Ste12 lineage loss in Ustilaginomycetes*

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* (Rispail 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).

All TFs harbouring an Ste12-DBD (IPR003120) classed within orthogroup OG0000022. An analysis of the corresponding protein-tree revealed these TFs are conserved, even among the early-diverging fungal lineages such as the Chytridiomycota and Mucoromycota (**Supplemental item 13**). However, they 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*, the model pathogen of the Ustilaginomycetes, three other TFs were identified outside the Ste12 clade in OG0000022 (Biz1, Mzr1 and Ztf1). Interestingly, they are each reported to regulate infectious growth *in planta* (Flor-Parra et al., 2006; Zheng et al., 2008; Velez-Haro et al., 2020; de la Torre et al., 2020), in contrast to any other experimentally-characterised regulators in OG0000022 that lack Ste12 DBDs (**Supplemental item 13**) (Kwon et al., 2010; Son et al., 2011; Wang et al., 2015; Cao et al., 2016).

This prompted a protein sequence alignment to identify any relationships 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 centred at the corresponding DBD, but 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 orthologues in other fungi. Although Biz1 and Ste12 TFs are activated and inhibited by related MAPK and cyclin-dependent kinase pathways respectively (de la Torre et al., 2020), there was little evidence for a direct evolutionary link beyond the C2H2 domain. Consequently, while Ste12 domain loss is a major distinction identified in the Ustilaginomycete pathogen-lineage, the causes and functional significance will require deeper investigation.

3.2.3. *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) and 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 such cases in several different *F. oxysporum* isolates, but also provides novel evidence of EBR1 expansion in other important pathogen lineages within the Ascomycota (**Supplemental item 14**). A protein sequence alignment/phylogenetic tree was therefore produced to investigate this in greater detail (Fig. 7).

In the *M. oryzae* reference isolate causing rice blast (PYROR) both

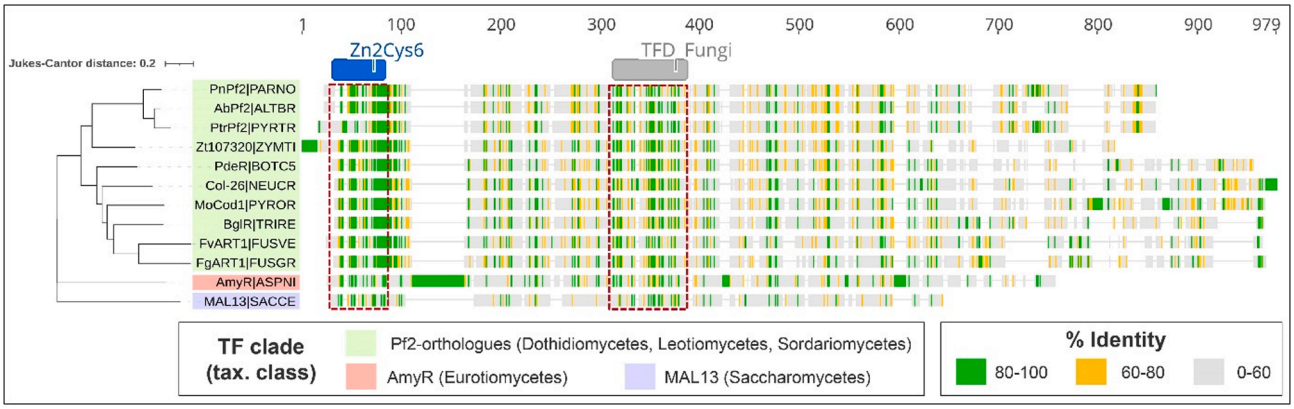

Fig. 5. Characterised Pf2-orthologues in relation to established carbohydrate regulators. **A protein-sequence alignment is presented with a neighbour-joining tree depicting sequence divergence between experimentally-validated Pf2-orthologues (highlighted in green) and the *A. nidulans* AmyR and *S. cerevisiae* MAL13 proteins that were identified in external clades of 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' (TFD_Fungi - IPR007219) TF domains defined in this study are presented above the alignment. Red boxes span the sequences where the respective domains are detected. A corresponding protein tree encompassing the full OG0000017 orthogroup complements this figure to provide further context (Supplemental item 12).** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

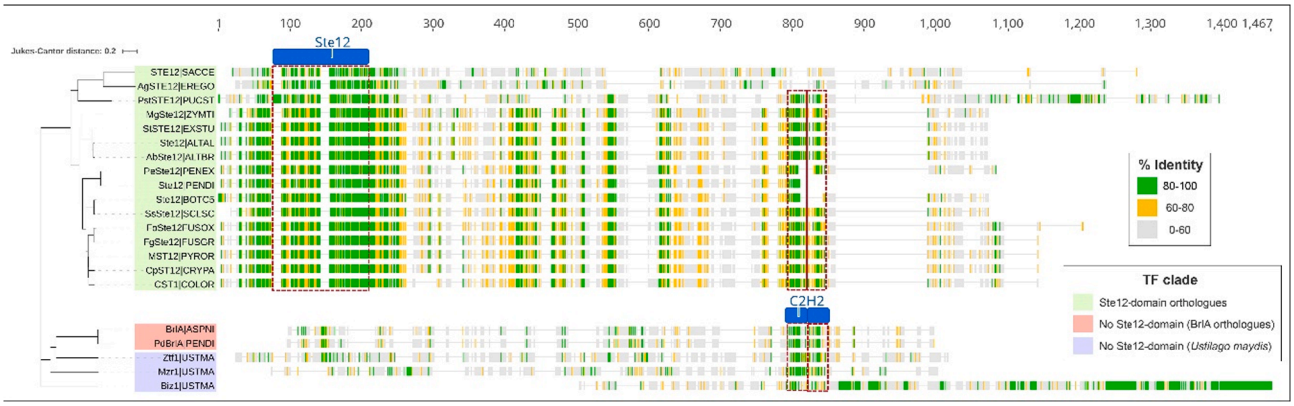

Fig. 6. Characterised Ste12-orthologues relative to related C2H2-domain lineages. **A protein-sequence alignment is presented with a neighbour-joining tree depicting sequence divergence between experimentally-characterised Ste12-orthologues (highlighted in green) and other characterised C2H2 transcription factors (TFs) from the common TF-orthogroup OG0000022. These included the *Ustilago maydis* regulators of invasive growth (blue) and the Aspergillaceae BrlA regulators of conidiation (red), where little identify is shared with Ste12 TFs beyind the C2H2 domain. 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) TF domains defined in this tudy are presented above or within the alignments. Red boxes span the sequences where the respective domains are detected. A corresponding protein tree encompassing the full OG0000022 orthogroup complements this figure to provide further context (Supplemental item 13).** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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. The yellow labels indicate experimentally-characterised transcription factors. 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) TF domain defined in this study is presented above the alignment. Red boxes span the sequences where the respective domains are detected. A corresponding protein tree encompassing the full OG0000121 orthogroup complements this figure to provide further context (Supplemental item 14).** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

MoCod2 (MGG_09263) and Cnf2 (MGG_15023) are TFs simultaneously required for pathogenic development (Chung et al., 2013; Lu et al., 2014). In addition to these individual roles, *MoCod2* was also required for conidiation and *Cnf2* deletion led to overproduction of conidia. This analysis presents novel evidence that both are orthologous to EBR1 but have therefore undergone duplication and neofunctionalization in *M. oryzae*, with additional duplication events evident in the *M. oryzae* millet pathovar (PYROM) included in this study (Fig. 7). The unique/redundant roles for these paralogues will need to be investigated further, 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. Putative EBR1 duplication events were also evident across the fungal Mycosphaerellaceae family (**Supplemental item 14**), which encompasses several damaging pathogens. The protein alignment (Fig. 7) suggests the C-terminal region was lost in a sub-lineage that included the *Z. tritici* TF AlmA (Mycgr3g111569). AlmA modulates hyphal development (Cairns et al., 2015), suggesting a functional connection to other EBR1 orthologues, yet the remaining TFs in the expanded Mycosphaerellaceae lineage are yet-to-be experimentally characterised. It was striking that most pathogens where EBR1

expansion events are observed are considered hemibiotrophs. It is plausible that EBR1 duplication up-regulates genes underpinning radial growth, thereby promoting asymptomatic colonisation of the host typical of this lifestyle. The novel evidence presented for EBR1 expansion beyond the *Fusarium* lineage warrants further exploration to establish the regulatory consequences during host infection.

4. Conclusions

At the onset of this study, it was hypothesised that the range and profile of TF lineages in phytopathogenic fungi change in accordance with their host-association and lifestyle. Therefore, we proposed that a comprehensive orthology-based analysis would reveal novel trajectories of important virulence regulators, and thereby further our understanding of the factors controlling plant pathogenicity. In the first stage of the analysis, a reduction in the overall TF content of obligate-biotrophic pathogens was identified. Through a TF-orthology analysis, this was largely attributed to contractions in several Zn2Cys6-DBD TF-lineages. Beyond this, the TF analysis supports the view that the strict-tripartite division of fungal pathogenicity into biotroph, hemibiotroph and necrotrophy is not well supported by the data. Recent studies have used large genomic datasets to demonstrate that the content of other distinct gene families, such as carbohydrate-acting enzymes, are predictive of plant-pathogenic lifestyles and can provide more precise definitions (Hane et al., 2020; Haridas et al., 2020; Wu and Cox, 2021). These definitions more accurately reflect the molecular factors that underpin virulence, but often only emerge within more recently evolved fungal taxa. Incorporating phylogenetic approaches may well offer the best way to analyse the evolution of phytopathogenicity. Here we demonstrated TFs evolve in different fungal lineages at different rates (Fig. 3). Deeper insight into the evolutionary pressures underpinning these changes in plant-pathogenic fungi will likely be gained using more precise lifestyle-definitions, or by assessing the TF-profile in a greater number of pathogens vs non-pathogens from closely related taxa.

The second stage of this analysis demonstrated TF-orthogroup lineages were useful for defining the evolutionary-trajectories of several key virulence regulators. For example, a Pf2-orthologue clade was distinguished and placed within carbohydrate metabolic-regulators. Ste12 domain loss was also documented in the Ustilaginomycetes and novel EBR1 expansions were then presented in *M. oryzae* and the Mycosphaerellaceae lineages. The orthogroups presented in this analysis, such as these specific examples highlighted, are a resource for future studies to investigate the evolutionary origins and trajectories of virulence-regulating TFs in plant-pathogenic fungi. This will provide insight into the distinct regulators underpinning plant infection to assist the design of targeted mitigation strategies, that could include the direct inhibition of the TFs themselves or their upstream/downstream signalling components.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

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

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.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fgb.2022.103712.

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), 1–15. https://doi.org/10.5598/imafungus.2017.08.01.01.

Baroncelli, R., Talhinhas, P., Pensec, F., Sukno, S.A., Le Floch, G., Thon, M.R., 2017. The *Colletotrichum acutatum* Species Complex as a Model System to Study Evolution and Host Specialization in Plant Pathogens. Frontiers in Microbiology 8. https://doi.org/10.3389/fmicb.2017.02001.

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), 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., 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res. 49 (D1), D344–D354. 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.o., 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 (3), 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).

Charoensawan, V., Wilson, D., Teichmann, S.A., 2010a. Lineage-specific expansion of DNA-binding transcription factor families. Trends Genet. 26 (9), 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 (3), 498–514. https://doi.org/10.1111/tpj.12217.

Choi, JaeJin, Kim, S.-H., 2017. A genome tree of life for the fungi kingdom. PNAS 114 (35), 9391–9396. https://doi.org/10.1073/pnas.1711939114.

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.-J., 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., Madhani, H.D., 2009. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene

Expansion. PLoS Genet. 5 (8), e1000618. https://doi.org/10.1371/journal.pgen.1000618.
Collemare, J., Lebrun, M.-H., 2011. In: Effectors in Plant–Microbe Interactions. Wiley, pp. 377–400. https://doi.org/10.1002/9781119949138.ch15.
Csuos, M., 2010. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics 26 (15), 1910–1912. https://doi.org/10.1093/bioinformatics/btq315.
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 (48), 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 (4), 470–479. https://doi.org/10.1094/PHYTO-09-10-0252.
Dufresne, M., Lee, T.V.D., M'Barek, S.B., Xu, X., Zhang, X.u., 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 (12), 1552–1561. https://doi.org/10.1016/j.fgb.2008.09.004.
Duplessis, S., Cuomo, C.A., Lin, Y.-C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D.L., Hacquard, S., Amselem, J., Cantarel, B.L., Chiu, R., Coutinho, P.M., Feau, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M.G., Kodira, C.D., Kohler, A., Kües, U., Lindquist, E.A., Lucas, S.M., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J.L., Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Salamov, A.A., Schmutz, J., Selles, B., Shapiro, H., Tanguay, P., Tuskan, G.A., Henrissat, B., Van de Peer, Y., Rouzé, P., Ellis, J.G., Dodds, P.N., Schein, J.E., Zhong, S., Hamelin, R.C., Grigoriev, I.V., Szabo, L.J., Martin, F., 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc. Natl. Acad. Sci. 108 (22), 9166–9171. https://doi.org/10.1073/pnas.1019315108.
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. 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), 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 (1), 124–138. https://doi.org/10.1111/mpp.12886.
Han, J.W., Kim, D.Y., Lee, Y.J., Choi, Y.R., Kim, B., Choi, G.J., Han, S.-W., Kim, H., 2020. Transcription factor PdeR is involved in fungal development, metabolic change, and pathogenesis of gray mold *Botrytis cinerea*. J. Agric. Food Chem. 68, 9171–9179. https://doi.org/10.1021/acs.jafc.0c02420.
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 (3), 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 (7), 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 (7), 1982–2003. https://doi.org/10.1111/1462-2920.12331.
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*. Mol. Microbiol. 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, J., Liu, Q., Li, J., 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*. Biotechnol. Biofuels 14, 33. https://doi.org/10.1186/s13068-021-01877-2.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8.
Lu, J., Cao, H., Zhang, L., Huang, P., Lin, F., Xu, J.-R., 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 (10), e1004432. https://doi.org/10.1371/journal.ppat.1004432.
Ma, L.-J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P.M., Kang, S., Shim, W.-B., Woloshuk, C., Xie, X., Xu, J.-R., Antoniw, J., Baker, S.E., Bluhm, B.H., Breakspear, A., Brown, D.W., Butchko, R.A.E., Chapman, S., Coulson, R., Coutinho, P.M., Danchin, E.G.J., Diener, A., Gale, L.R., Gardiner, D.M., Goff, S., Hammond-Kosack, K.E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C.D., Koehrsen, M., Kumar, L., Lee, Y.-H., Li, L., Manners, J.M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S.-Y., Proctor, R.H., Regev, A., Ruiz-Roldan, M.C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D.C., Turgeon, B.G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C.A., Kistler, H.C., Rep, M., 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464 (7287), 367–373. https://doi.org/10.1038/nature08850.
MacPherson, S., Larochelle, M., Turcotte, B., 2006. A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol. Mol. Biol. Rev. 70 (3), 583–604. https://doi.org/10.1128/MMBR.00015-06.
Möller, M., Stukenbrock, E.H., 2017. Evolution and genome architecture in fungal plant pathogens. Nat. Rev. Microbiol. 15 (12), 756–771. https://doi.org/10.1038/nrmicro.2017.76.
Mukaka, M., 2012. A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 24, 69–71.
Nakamura, T., Maeda, Y., Tanoue, N., Makita, T., Kato, M., Kobayashi, T., 2006. Expression profile of amylolytic genes in *Aspergillus nidulans*. Biosci. Biotechnol. Biochem. 70 (10), 2363–2370. https://doi.org/10.1271/bbb.50694.
Niño-Sánchez, J., Casado-Del Castillo, V., Tello, V., De Vega-Bartol, J.J., Ramos, B., Sukno, S.A., Díaz Mínguez, J.M., 2016. The *FTF* gene family regulates virulence and expression of SIX effectors in *Fusarium oxysporum*. Mol. Plant Pathol 17 (7), 1124–1139. https://doi.org/10.1111/mpp.12373.
Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 42 (D1), D26–D31. https://doi.org/10.1093/nar/gkt1069.
Oh, M., Son, H., Choi, G.J., Lee, C., Kim, J.-C., Kim, H., Lee, Y.-W., 2016. Transcription factor ART1 mediates starch hydrolysis and mycotoxin production in *Fusarium graminearum* and *F. verticillioides*. Mol. Plant Pathol 17, 755–768. https://doi.org/10.1111/mpp.12328.
Park, J., Park, J., Jang, S., Kim, S., Kong, S., Choi, J., Ahn, K., Kim, J., Lee, S., Kim, S., Park, B., Jung, K., Kim, S., Kang, S., Lee, Y.-H., 2008. FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. Bioinformatics 24 (7), 1024–1025. https://doi.org/10.1093/bioinformatics/btn058.
Pedro, H., Maheswari, U., Urban, M., Irvine, A.G., Cuzick, A., McDowall, M.D., Staines, D.M., Kulesha, E., Hammond-Kosack, K.E., Kersey, P.J., 2016. PhytoPath: an integrative resource for plant pathogen genomics. Nucleic Acids Res. 44 (D1), D688–D693. https://doi.org/10.1093/nar/gkv1052.
Plissonneau, C., Benevenuto, J., Mohd-Assaad, N., Fouché, S., Hartmann, F.E., Croll, D., 2017. Using population and comparative genomics to understand the genetic basis of effector-driven fungal pathogen evolution. Front. Plant Sci. 8 https://doi.org/10.3389/fpls.2017.00119.
Pusztahelyi, T., Holb, I.J., Pócsi, I., 2016. Plant-Fungal Interactions: Special Secondary Metabolites of the Biotrophic, Necrotrophic, and Other Specific Interactions. In: 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 (4), 327–332.
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 (4), 986–1004. https://doi.org/10.1111/mpp.12584.
Schoch, C.L., Ciufo, 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 Sci. 27 (1), 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., Xu, J.-R., 2011. A phenome-based functional analysis of transcription factors in the cereal head blight fungus, Fusarium graminearum. PLOS Pathogens 7 (10), 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. Nat. Biotechnol. 35 (11), 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 Res. 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 (1), 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 (1), 214. https://doi.org/10.1186/1471-2164-15-214.

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., Stukenbrock, E.H., 2016. Transcription factors encoded on core and accessory chromosomes of *Fusarium oxysporum* induce expression of effector genes. PLoS Genet. 12 (11), 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. Annu. Rev. Phytopathol. 55 (1), 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 (1), 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 Microbiol. Rev. 36 (1), 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. PdbrlA, PdabaA and PdwetA control distinct stages of conidiogenesis in *Penicillium digitatum*. Res. Microbiol. 166 (1), 56–65. https://doi.org/10.1016/j.resmic.2014.12.003.

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.

Wu, B., Cox, M.P., 2021. Comparative genomics reveals a core gene toolbox for lifestyle transitions in Hypocreales fungi. Environ. Microbiol. 23 (6), 3251–3264. https://doi.org/10.1111/1462-2920.15554.

Xiong, Y.i., 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., Swaminathan, K., 2017. A fungal transcription factor essential for starch degradation affects integration of carbon and nitrogen metabolism. PLoS Genet. 13 (5), e1006737. https://doi.org/10.1371/journal.pgen.1006737.

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 (4), 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 (12), 1407–1418. https://doi.org/10.1094/MPMI-06-11-0158.

Zheng, Y., Kief, J., Auffarth, K., Farfsing, J.W., Mahlert, 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. Mol. Microbiol. 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 (4), 961–974. https://doi.org/10.1111/mpp.12582.