Functional & Integrative Genomics
https://doi.org/10.1007/s10142-018-0644-6

ORIGINAL ARTICLE

Transcriptome-based mining and expression profiling of *Pythium* responsive transcription factors in *Zingiber* sp.

Vishnu Sukumari Nath[1,2] · Sayuj Koyyappurath[1,3] · Teena Elizabeth Alex[1] · Kiran Ayyanperumal Geetha[4] · Lesly Augustine[5] · Alka Nasser[1] · George Thomas[1]

Received: 16 January 2018 / Revised: 29 April 2018 / Accepted: 29 October 2018

Abstract

Transcription factors (TFs) fine-tune the host defense transcriptome in response to pathogen invasions. No information is available on *Zingiber zerumbet* (*Zz*) TFs involved in defense response against *Pythium myriotylum*. Here, we provide a global identification, characterization, and temporal expression profiling of *Zz* TFs following an incompatible interaction with *P. myriotylum* using a transcriptome sequencing approach. We identified a total of 903 TFs belonging to 96 families based on their conserved domains. Evolutionary analysis clustered the *Zz* TFs according to their phylogenetic affinity, providing glimpses of their functional diversities. High throughput expression array analysis highlighted a complex interplay between activating and repressing transcription factors in fine-tuning *Zz* defense response against *P. myriotylum*. The high differential modulation of TFs involved in cell wall fortification, lignin biosynthesis, and SA/JA hormone crosstalk allows us to envisage that this mechanism plays a central role in restricting *P. myriotylum* proliferation in *Zz*. This study lays a solid foundation and provides valuable resources for the investigation of the evolutionary history and biological functions of *Zz* TF genes involved in defense response.

Keywords Biotic stress · Defense response · Ginger · Plant-pathogen interaction · Transcriptional reprogramming

Introduction

Plants have developed a plethora of sophisticated defense mechanisms to ward off damages caused by pathogens. The host response to an invading pathogen involves a myriad of complex molecular, biochemical, and physiological processes that result in compatible or incompatible interaction (Jones and Dangl 2006). The coordinated activation of these processes involves the transcriptional reprogramming of several genes in a synchronized manner that is largely regulated by transcriptional factors (TFs). TFs act as transcriptional activators or repressors and play a central role in the regulation of developmental, metabolic processes, biotic and abiotic stresses. For example, WRKY TFs are involved in the plant immune system mediated by hormones (jasmonic acid and salicylic acid) that can respond to attacks by pathogens, such as bacteria, viruses, and fungi (Eulgem and Somssich 2007; Mukhtar et al. 2008). Members of the NAC TF family have been suggested to play important roles in the regulation of the transcriptional reprogramming associated with plant stress responses (Puranik et al. 2012). Considering the importance of TFs, research on TFs had gained momentum during recent years. Advances in transcriptome sequencing provide fast, cost-effective, and reliable approach to generate large expression datasets, especially suitable for non-model species to identify putative genes, key pathways, and regulatory mechanisms. Genome and transcriptome-wide approaches have allowed the identification and characterization of TF families regulating vital biological processes in several

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10142-018-0644-6) contains supplementary material, which is available to authorized users.

✉ George Thomas
gthomas@rgcb.res.in

[1] Plant Disease Biology and Biotechnology, Rajiv Gandhi Centre for Biotechnology, Thycaud P. O., Thiruvananthapuram 695014, India

[2] Present address: Department of Molecular Genetics, Institute of Plant Molecular Biology (IPMB), Branišovská 1160/31, České Budějovice 370 05, Czech Republic

[3] Present address: Department of Genome Structural Biology, Temasek Life Science Laboratory, 1-Research Link, National University of Singapore, Singapore 117604, Singapore

[4] Department of Biotechnology, Kerala Agricultural University, Vellayani, Thiruvananthapuram 695522, India

[5] Department of Botany, Sacred Heart College, Thevara, Ernakulam 682013, India

Published online: 11 November 2018

plant species such as the bHLH family in Apple (Yan et al. 2017); MYB, ADP-ribosylation factor, WRKY families in Foxtail Millet (Muthamilarasan et al. 2014, 2015, 2016), Rice (Smita et al. 2015); NAC family in durum wheat (Saidi et al. 2017); Dof family in cucumber (Wen et al. 2016); AP2/ERF family (Lata et al. 2014) and bZIP family in cassava (Hu et al. 2016).

As the largest group of plant fungal pathogens, the oomycete genus *Pythium* causes heavy crop losses worldwide. This necrotrophic, vascular pathogen causes soft rot disease in several plant species, including high-value spice crop ginger (*Zingiber officinale* Roscoe). Soft rot stands out to be a major production constraint for ginger cultivation globally. Besides its broad host range, the pathogen's ability to produce oospores able to endure in soil for many years makes *Pythium* extremely difficult to control (Yadeta and Thomma 2013). Therefore, plant genetic resistance is regarded as a promising key alternative to control crop diseases and pests. In contrast to the foliar diseases, the host response and physiological and molecular mechanisms underlying resistance to soil-borne necrotrophs are scanty (De Coninck et al. 2015). Characterization of the genetic components underlying disease resistance is a major research area in crop plants which is highly relevant for resistance breeding programs. However, sources of resistance against the soil-borne pathogen are rare in cultivated germplasm (Okubara and Paulitz 2005) and, ginger is not an exception (Kavitha and Thomas 2007). We have identified durable resistance in *Zingiber zerumbet* (L.) Smith (*Zz*), a wild relative of cultivated ginger against *Pythium* (Kavitha and Thomas 2008). Identifying and understanding the resistance mechanisms exhibited by the *Zz* could provide valuable clues to devising a sustainable and environment-friendly management option to mitigate soft rot disease in ginger. Because transcriptional reprogramming is a key event in plant's defense response to pathogens, it is extremely important to understand the role of transcription factors that play a pivotal role in fine-tuning this reprogramming. No information is available on TFs involved in the *Zz-Pythium* interaction. Here, we provide an extensive knowledge of *Zz* transcription factors involved in *P. myriotylum* incompatible interaction by a de novo transcriptome sequencing approach. We initially identified *Zz* TFs, classified them into various families based on their homology and investigated evolutionary relationships to that of model plants. Subsequently, we explored their expression signatures using a high throughput custom designed expression array. To the best of our knowledge, this study represents the first of its kind and provides illuminating insights into the incompatible *Zz-Pythium* interaction. The information generated in this study will serve as valuable resources for researchers who wish to explore this pathosystem in future.

Materials and methods

Plant material, pathogen inoculation, and tissue sampling

Zingiber zerumbet accession 2010-9, which is resistant to soft rot disease (Kavitha and Thomas 2008), was used for the study. Rhizomes harvested from mature plants were germinated in earthen pots in autoclaved red-earth:river-sand:leaf compost (1:1:1) mixture and the plants were maintained in an insect protected net-house under natural conditions at Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, India. Four-month-old plants with uniform growth were used for pathogen inoculation as previously described (Kavitha and Thomas 2008). The collar (intersection between the pseudostem and the rhizome) region of the plant was inoculated with four 10 mm diameter mycelial discs, excised from the growing margin of a 48-h-old highly aggressive *P. myriotylum* isolate (RGCB N14) grown on PDA. The authenticity of the isolate was confirmed by ITS amplification and sequencing using ITS1 and ITS4 universal primers (White et al. 1990). The inoculated region was covered with wet cotton to provide sustained humidity. Plants mock inoculated with plain PDA discs were used as controls. Total RNA was isolated from 1 in.-long collar region.

Library construction, transcriptome sequencing, and de novo assembly

Total RNA was isolated according to method of Salzman et al. (1999). RNA samples isolated from tissues sampled from control plants and from treated plants at 24 h post inoculation (hpi) with *P. myriotylum* were used for the library construction. We chose the RNA sample isolated at 24 hpi with *P. myriotylum* as the initiation of transcriptional reprogramming was reported previously in pathogen treated *Z. zerumbet* at this time period (Kavitha and Thomas 2008). The deep sequencing of the transcriptome was done at Genotypic Technology PVT. LTD., Bengaluru, India. Transcriptome libraries were constructed using an Illumina TruSeq RNA library Prep kit (Illumina, San Diego, California, USA). The libraries were separately sequenced using an Illumina GAIIx platform employing paired-end module with 73/72 base read length. Base-calls were performed with CASAVA 1.8.2 software programme and the raw reads were extracted in a FASTq format for further analysis. Quality of the raw reads was tested using the software NGS QC Toolkit v 2.2.3. The reads containing adaptor sequences and unknown nucleotides were removed and reads with $\geq 70\%$ bases with ≥ 20 phred score were filtered as high-quality reads.

High-quality reads obtained from control and treated samples were pooled and subjected to de novo assembly using a de Bruijn graph method (Zerbino and Birney 2008) as

implemented in the Trinity assembler-Version 20140717 (Grabherr et al. 2011) using default parameters. The quality of the Trinity assembled contigs was tested based on the number of reads aligned onto contigs (number of reads ≥ 10) and the FPKM (fragments per kilobase of exon per million fragments mapped) values (value ≥ 1) and the truly assembled contigs were filtered out. The truly assembled contigs were clustered at a 90% identity threshold using the software CD-HIT-EST (v 4.6.1) (Li and Godzik 2006) and filtered the longest contig (transcript) from each assembled Trinity locus.

Identification of *Z. zerumbet* transcription factors

To identify the *Zz* TFs, the transcriptome sequences were homology searched against the plant TF database (http://planttfdb.cbi.pku.edu.cn/) using the software iTAK v 1.7a (Zheng et al. 2016). This resulted in the classification of *Zz* transcripts in various TF families. To confirm the authenticity of this classification, the Hidden Markov Model (HMM) profile for each of the identified TF families was downloaded from the Pfam protein family database (http:// pfam.sanger.ac.uk/) and a HMMER search (http://hmmer.janelia.org/) was executed to identify the presence of conserved domains in identified putative TF sequences. All non-redundant sequences encoding complete TF domains were considered to be putative TF genes. Further, all the TF family sequences were double-checked using a batch CD-search in pfam protein database and NCBI conserved domain database (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Finally, a protein blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi) search was performed using the deduced amino acid sequences of the identified TFs to infer their homology with other similar sequences in the NCBI database.

Protein characterization, conserved motifs, and gene ontology identification

The TF proteins were further characterized by ProtParam tool available on Expert Protein Analysis System (ExPASy) proteomics server (http://web.expasy.org/protparam/) and isoelectric points (pI), molecular weight, aliphatic and instability index were computed. The subcellular localization of TFs was predicted using the CELLO online tool v 2.5 (http://cello.life.nctu.edu.tw/) (Yu et al. 2006).

The additional conserved motifs of the TF proteins outside the family domain were statistically identified using the Multiple Expectation-maximization for Motif Elicitation (MEME) programme (http://meme-suite.org/tools/meme) with the following parameters: motif length set to 6–100 and motif sites to 2–120; maximum number of motifs: 10. The distribution of one single motif was "any number of repetitions" and the other search parameter was "search given strand only".

TFs are known to regulate multiple pathways. Therefore, to gain insights into the functional versatility of the identified TFs, Gene ontology (GO) annotation was executed using the Blast2GO tool v 4.1 (Conesa et al. 2005). The TF protein sequences were BLASTP searched against the Reference protein sequences (RefSeq) of NCBI under default parameters. The GO terms associated to hit sequences were obtained by mapping and finally, InterPro annotation was performed. The Blast2GO output provides information on molecular function, biological process and cellular component of each TF protein.

Phylogenetic analysis

To infer the phylogenetic relationships, multiple sequence alignments of Zz TFs along with reference TF proteins of *Arabidopsis thaliana, Glycine max*, and *Musa acuminata* were conducted using Clustal W (Thompson et al. 1994) with default settings. Subsequently, MEGA v 7.0 software (Kumar et al. 2016) was employed to build an unrooted phylogenetic tree based on the neighbor-joining (NJ) method with the following parameters: JTTmodel, pairwise gap deletion, and 1000 bootstraps to evaluate the significance of the nodes.

Identification of orthologous genes and protein interactions

The genes reported in other plant species that are closest (orthologous) to the *Zz* TFs were determined from well-characterized model species (*A. thaliana*, *G. max*, and *Oryza sativa indica*) and also from *M. acuminata*, the monocot that is phylogenetically closest to Zingiberaceae for which the whole genome is available. The orthologous genes were identified using eggnog-mapper (Huerta-Cepas et al. 2017) based on eggNOG 4.5 orthology data (Huerta-Cepas et al. 2016). To determine the interactions of TF proteins, protein-protein interaction (PPI) analysis was performed using STRING (http://string-db.org; (Franceschini et al. 2013) database in COG (Clusters of Orthologous Group) mode. The best-assigned COGs obtained based on most significant E-value using *Arabidopsis thaliana* as the organism was used to construct an interaction network. Those interactions with a confidence score of ≥ 0.7 and based on co-expression and experiment conditions were used to construct the network.

Microarray expression analysis

The custom gene expression microarray was performed at the University of Delhi South Campus Microarray Centre (UDSCMAC). The entire pool of 87,897 transcripts obtained by the de novo assembly of *Z. zerumbet* transcriptome data were subjected to probe design by using the software eArray (Agilent Technologies, Santa Clara, CA, USA; https://earray.chem.agilent.com/erray/). eArray treatment successfully

designed probes for 52,432 transcripts. The probes were printed onto a 1 × 3 in. glass slides using the SurePrint Inkjet technology (Agilent Technologies) in an 8 × 60 K format. Each array consisted of 58,476 probes, including probes designed from the de novo assembled *Z. zerumbet* transcripts (52,432 probes), non-plant negative controls (48 probes), plant-specific positive controls (2725 probes), and specific genes for probe replicates (3271 probes). There were eight separate arrays in the glass slides and each array contained 62,625 features and each feature had an array density of 15,000 to 100,000 probes. The high abundance for each probe in an array eliminates the possibility of probe saturation while using for one color gene expression analysis.

The RNA samples isolated from two biological replicates each from *Z. zerumbet* at 18 hpi, 36 hpi and 48 hpi with *P. myriotylum* and untreated control plants were used for expression analysis. Each biological replicate was sampled from an independent plant. The hybridization procedure was performed according to the Agilent Technologies' One-Color Microarray-Based Gene Expression Analysis Protocol (GA4140-90040) using a Gene Expression Hybridization kit (5188-5279). Hybridization was performed in a hybridization oven (Agilent Technologies, G2505-80085) at 60 °C for 17 h. The hybridized arrays were scanned using an Agilent Microarray scanner (G2565CA) at a resolution of 5 μm with scan control software according to the manufacturer's instructions (G2505-90020). The scanned images were analyzed using the Agilent Feature Extraction software v 10.5 by quantifying the pixel density of each hybridization spot. The software will automatically subtract background signals from the data and the feature extraction is received in .txt format.

Data from each array was extracted using the software Agilent Feature Extraction 10.5.1.1. The data analysis was performed using the software Genespring GX12 (Agilent Technologies). Signals were background corrected and baseline transformed to the median of all spots. The data was log2 transformed and normalized to 50th percentile using Loess normalization. The log2 ratios were averaged for replicate spots. The significance ($p < 0.05$) in the modulation of the signal intensity produced by RNA samples from treated samples (18 hpi, 36 hpi, 48 hpi) in relation to the untreated control samples was determined by unpaired Student's *t* test. The expression data was visualized as heat maps with clustering drawn with fold change values using the ClustVis online tool (Metsalu and Vilo 2015).

Results and discussion

Transcriptome sequencing and de novo assembly

More than 90% of the reads obtained in the two libraries were of high quality with an average of 98% of bases showing a phred score > 20. The NGS QC Toolkit analysis filtered 32.9 million and 31.8 million high-quality reads from control and treated samples of *Z. zerumbet* (GenBank Sequence Read Archive (SRA) accession numbers: control - SRA SRX959143; treated- SRX959144). The de novo assembly produced 129,905 contigs in *Z. zerumbet*. Quality of the de novo assembled contigs was further tested and filtered 104,349 truly assembled contigs in *Z. zerumbet*. CD-HIT-EST treatment of true contigs identified 87,897 distinct transcripts in *Z. zerumbet* with an average length of 900 bp and 44.5% GC content.

Identification of TFs from *Z. zerumbet* transcriptome

The iTAK and HMMER analysis of *Z. zerumbet* transcripts allowed the identification of 903 non-redundant TFs that were classified into 96 TF families (Table 1). For convenience, the identified TFs were named *Zz* followed by transcript number and the family name (e.g., Zz_c57373_g1_i1_WRKY). BLAST searches of the NCBI database revealed that the deduced amino acid sequences of majority TFs shared the highest similarity levels with *M. acuminata*. The sequence similarity ranged from 40 to 98% (Supplementary file: Table S1). More than 50% of the TFs belonged to well-characterized/described TF families viz. WRKY, AP2, bZIP, MYB, bHLH, and NAC which are known to play a crucial role in plant defense response (McGrath et al. 2005; Seo and Choi 2015; Tsuda and Somssich 2015). However, our analysis also identified some rare TFs that are recently described and less studied (Bai et al. 2011). Previous transcriptome-based studies have also highlighted the predominance of well-characterized TFs in defense response in *Arabidopsis* (Eulgem 2005), *Camellia sinensis* (Jayaswall et al. 2016) and *Chrysanthemum morifolium* (Song et al. 2016).

This study focuses on the characterization of 314 TFs belonging to WRKY, AP2, bZIP, MYB, bHLH, and NAC families.

Phylogenetic relationships of *Z. zerumbet* TFs

As evident from Fig. 1, all the *Zz* TF proteins could be efficiently grouped into their respective families with subgroups based on their phylogenetic affinity. Of the 3 main WRKY groups identified in *Arabidopsis*, the phylogenetic tree separated the 43 *Zz* WRKYs into Group IIb, IIc, and IIe, suggesting that all the members belonged to Group II WRKY (Eulgem et al. 2000). The Group IIc formed the major group with 26 members followed by Group IIb and Group IIe (Fig. 1b). A total of 12 groups have been described for the AP2/ERF family in *Arabidopsis* (Sakuma et al. 2002). In our study, the 32 *Zz* AP2/ERFs were grouped into 6 groups (A2, A5, B1, B4, AP2, and APETALA 2) with Group B1 forming the major clade with 10 members (Fig. 1a). Jakoby et al. (Jakoby et al. 2002) have described 8 groups for the bZIP family in

Table 1 List of *Z. zerumbet* transcription factor (TFs) families identified in this study. The TFs were identified using combination of iTAk and HMMER searches using the de novo assembled transcriptome sequences *Z. zerumbet* following *P. myriotylum* interaction

Transcription factor family	Count
Alfin	2
*AP2	32
ARID	1
AUXIAA	29
Auxin response factor	4
B3 DNA binding	17
BES 1	3
Bromo	1
BTB POZ	23
*bZIP	28
bZIP	2
bZIPc	2
CBFB NFYA	10
CCCT1	13
CCT2	3
CSD	1
CUPIN	2
DDT	2
DFRP	1
DNA POLY	1
DOG 1	3
DP	3
DUF 573	2
DUF 702	1
DUF 2431	1
E2FTDP	2
EIN3	4
ERCC3	1
FAR 1	6
FBOX	2
GAGA BIND	1
GATA	5
GLYCO HYDRO	1
GNAT	11
GRAS	43
HALZ	4
HAREHTH	1
*HLH	64
HOMEOBOX	31
HOMEOBOXKN	5
HSF DNA BIND	17
JMJC	3
KBOX	5
KNOX2	4
LIM	2
LINKERSTONE	1
LOB	7
MBF 1	1
MED 6	1
MED 7	1
MED 26	5
MED 31	1
MFMR ASSO	1
MFMR	1
mTEF	21
MYBCC	6
MYB DNA BIND 4	20
*MYB DNA BIND	111
NAM	49
NUF1P1	1
OVATE	12
PAH	3
PC4	3
PHD	12
PLATZ	7
POX	2
PP2C	4
PUB	1
RADICAL SAM	1
RCD1	6
RESPONSE REG	8
RUBISCO	2
RWPRK	1
S1FA	2
SANT	1
SBP	2
SET	20
SKP1 PO2	1
SNF2N1	10
SRFTF	15
SWIB	5
SWIRM	1
SWIRM ASSO	1
TCP	5
TCR	4
TIFFY	13
TUB	15
WHIRLY	3
*WRKY	63
YABBY	2
zFAN1	1
zfC2H2	7
zfC2H2 JAZ	2
zfCCH	27
zDOF	12
zfLSD	3
Total	903

*Transcription factors explored in this communication

Arabidopsis. Here, the 28 *Zz* bZIPs could be assigned to 4 subgroups (Group A, C, H, and E), together with reference bZIPs used for the analysis. Group E formed the major group with 12 *Zz* bZIP members, whereas Group H was the smallest with only 5 members, suggesting the existence of a diversified bZIP family in *Zz* sp. (Fig. 1c). The *Arabidopsis* MYBs were extensively studied and classified into 22 groups by Stracke et al. (Stracke et al. 2001). The phylogram clustered the 111 *Zz* MYBs into 17 groups, which were further sub-grouped into smaller clades based on their phylogenetic affinity (Fig. 1f). The analysis revealed that the majority of *Zz* MYBs belong to the R2R3 group with only 11 TFs belonging to MYB-related group having a single MYB domain (Stracke et al. 2001). Among the 22 bHLH groups described in *Arabidopsis* (Toledo-Ortiz et al. 2003), the 64 *Zz* bHLH TFs could be categorized into seven groups (Fig. 1e), viz. Group 6, 8, 9, 11, 16, 17, and 18. Group 18 formed the major group with 13 members followed by Group 8 with 10 members. Zhu et al. (Zhu et al. 2012) have extensively described 21 groups in the NAC TF family. As illustrated in Fig. 1d, the 49 *Zz* NAC TFs

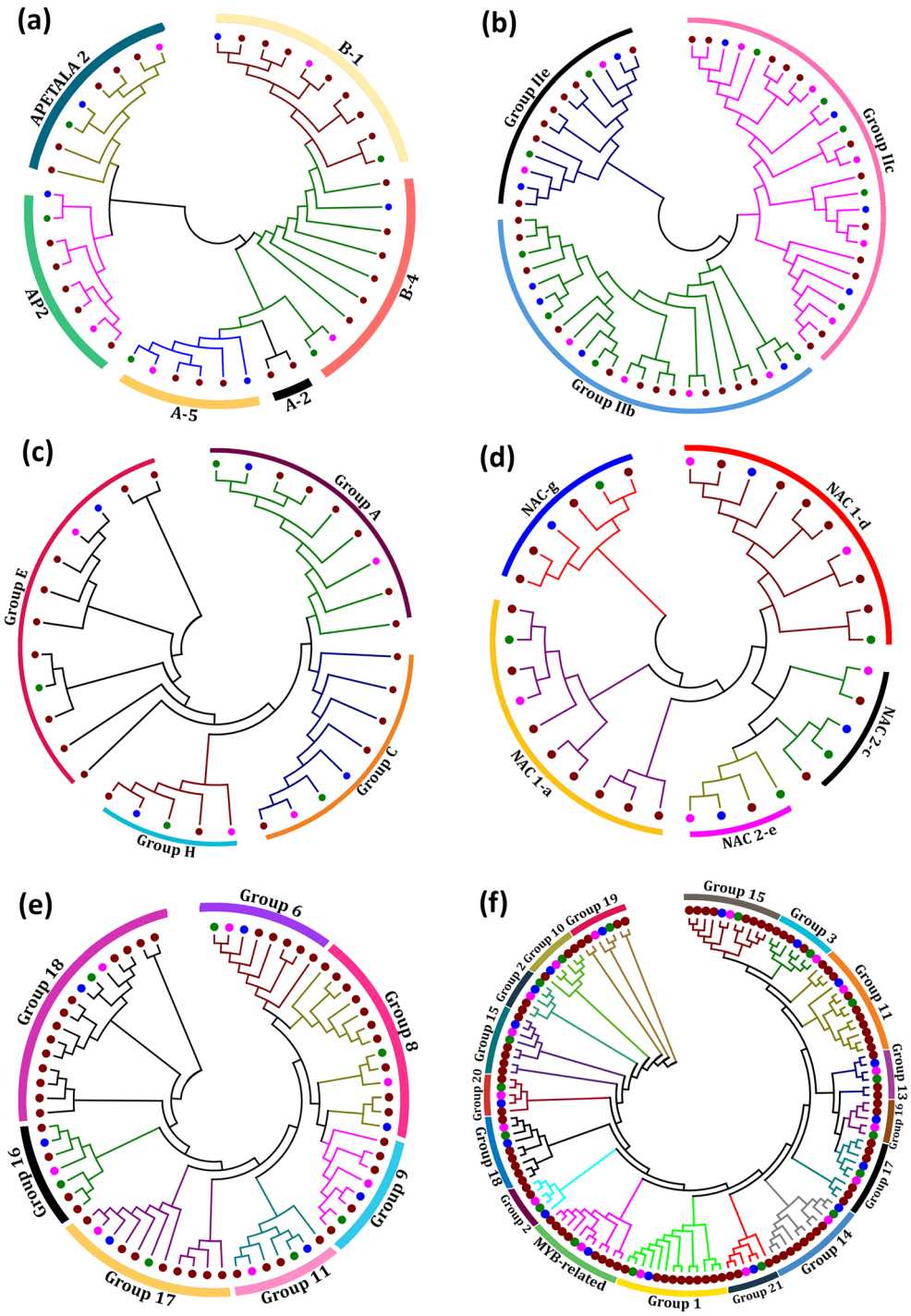


Fig. 1 Phylogenetic relationships and subfamily designations of *Z. zerumbet* TF families along with model species *A. thaliana*, *G. max*, and *M. acuminata*. The deduced amino acid sequences were aligned using Clustal W and the phylogenetic tree was constructed using the neighbor-joining method implemented in MEGA 7 software. The colored branch indicates the different subfamilies (**a**) AP2, (**b**) WRKY, (**c**) bZIP, (**d**) NAC, (**e**) bHLH, and (**f**) MYB

were distributed into 5 groups (Nac 1-a, Nac 1-d, Nac 2-c, Nac 2-e, and Nac-g) (Zhu et al. 2012) with Nac 1-a and Nac 1-d forming the major clades with 10 members each. Intriguingly, we observed that few *Zz* TFs (e.g., AP2 and bZIP family TFs) that belonged to the same subfamily were clustered in different clades in the phylogenetic tree (Fig. 1a, c). This may be due to the occurrence of duplication and divergent events in *Zz* TF genes. A similar observation was reported by Muthamilarasan et al. (Lata et al. 2014) where Foxtail millet MYBs failed to cluster in the same clade according to their subfamilies.

Protein characterization, conserved motif, and gene ontology identification

The ExPASy analysis of *Zz* TFs showed a range of variation in their molecular weight (7.5–69.3 kDa) and isoelectric point (4.2–12.01). The protein characteristics of *Zz* TFs belonging

to six studied families are presented in Supplementary Table S2. The average value of the aliphatic index was 75.9 which was in agreement with previous reports (Hu et al. 2016; Song et al. 2016; Wen et al. 2016). More than 85% of TFs was calculated to be unstable proteins with an instability index value of more than 40. We envisage that these variations could be attributed to the presence of putative novel variants, which needs to be validated by further research.

Subcellular location analysis based on the CELLO (http://cello.life.nctu.edu.tw/) showed that most *Zz* TFs proteins were located in the nucleus followed by mitochondria/plasma membrane (Additional file: Table S2).

To investigate the structural diversity of motif compositions in *Zz* TFs, a total of 10 conserved motifs for each family were captured by MEME software. The conserved motifs were named motif 1 through motif 10 for the individual family and are represented in their relative location within the protein (Supplementary Figs. 1 and 2). All TFs genes possessed their highly conserved core domain characteristic of the TF family. For the WRKY family, Motif 1 was present in all the sequences irrespective of the subgroups and was confirmed to be the highly conserved WRKYGQK domain. The majority of the *Zz* TFs in the same group shared similar motifs, suggesting that these conserved motifs play crucial roles in group-specific functions and reflects functional similarities. (Pinheiro et al. 2009) However, considerable divergence in their structures was found between the different groups. Few subgroups had common motifs in all the members, whereas other subgroups possessed special motifs. For example, WRKY Group IIb contained mostly similar motifs, whereas Group IIC had special motifs (Additional file: Fig. 1), reflecting the complex nature of the function of TF proteins in *Z. zerumbet*. The motif distribution indicated that the genes containing the same motifs were likely produced via gene expansion within the same groups. Remarkably, we also observed that despite the presence of a highly conserved motif or DNA-binding domain characteristic of TF family, the sequence similarity of other regions was relatively low in most TF genes. One reason for this observation is that the conserved motif or DNA-binding domains of transcription factor genes are critical for their function and thus appear to have diverged at a slow pace (Liu et al. 1999). Additionally, nucleotide substitution might have played a central role in the evolution of conserved regions, whereas substitutions and small insertions/deletions contributed to variable region diversification (Purugganan et al. 1995).

The GO analysis using Blast2GO software distributed the *Zz* TFs into three main functional biological categories: biological process, cellular component, and molecular function (Supplementary Fig. 3). In the biological processes category, a predominant portion of TFs was found to be involved in cellular, metabolic and biosynthetic processes. In the molecular function category, the highest proportion of TF proteins corresponded to transcription factor activity, heterocyclic compound binding, organic cyclic compound binding, and protein binding. The cellular component analysis revealed that most of the *Zz* TF proteins are nuclear localized followed by localization in plastid and mitochondrion which complemented the CELLO results. TF localization and its nucleo-cytoplasmic traffic constitute an important regulatory checkpoint in fine-tuning of gene expression and the control of gene expression (Vandromme et al. 1996). Nuclear, mitochondrial and chloroplast signaling have key roles in governing host defense response against pathogens (Kangasjärvi et al. 2012).

Identification of orthologous genes and protein interactions

Comparative genomic analyses across different taxa allow the transfer of functional information from a well-characterized taxon, such as the model plant *Arabidopsis*, to another less-studied taxon, like *Z. zerumbet*. To this end, orthologous genes remain the best candidates for functional annotation transfer. Here, we surveyed orthologous genes from well-characterized model plants viz. *A. thaliana*, *G. max*, *O. sativa* and *M. acuminata* which facilitated in predicting the putative functions of *Zz* TF proteins (Supplementary Table S3). However, for genes belonging to multigene families, such as plant transcription factors, the definition of orthologous relationships is highly challenging (Conte et al. 2008). Accordingly, it is worth commenting that several *Zz* TF genes that showed a strong differential modulation in our study could not be attributed to any known function. Given the fact that the functional characterization of TFs involved in plant-pathogen interaction is still in its infancy with only a handful of studies elucidating the functional role of TFs involved in plant-pathogen interactions (particularly necrotrophic), this is justified (Eulgem 2005; Buscaill and Rivas 2014; Seo and Choi 2015).

Structural protein-protein interaction (PPI) network provides a comprehensive understanding of interactions between proteins. The PPI network highlighted several protein functional groups interacting with each other (Fig. 2), suggesting its role in diverse biological processes. The interaction network clearly portrays the interconnected network of diverse TF families (e.g., AP2 and bZIP; Fig. 2a) in response to biotic stress. Most hub genes of this network were involved in the defense response against pathogens. The detailed description of key genes involved in the interaction is presented in the Supplementary Table S4.

Temporal expression signatures of *Zz* TFs in response to *P. myriotylum*

The activation of the defense transcriptome is a complex and multidimensional process. Analyzing temporal gene

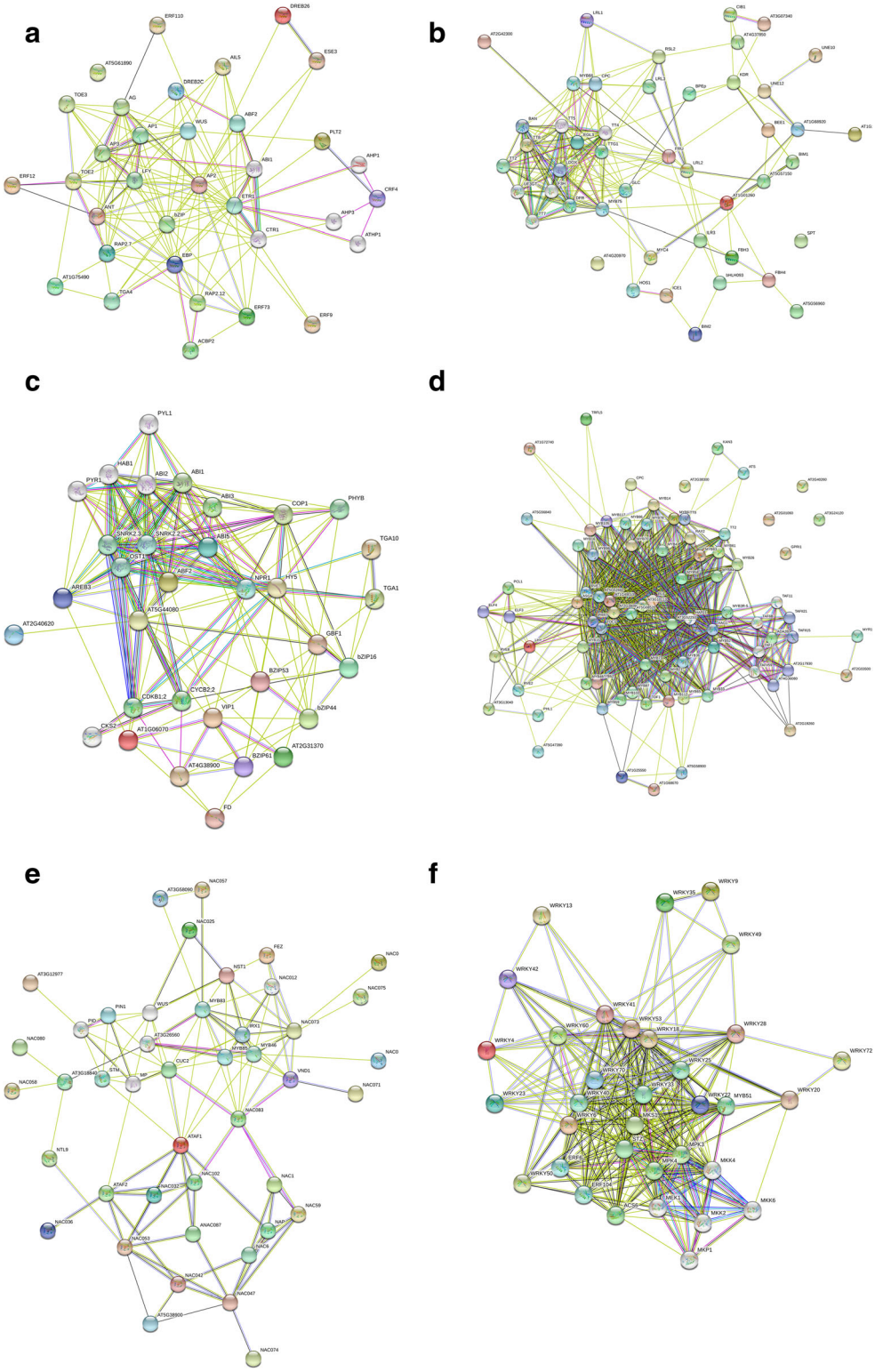


Fig. 2 Network analysis of *Zz* TFs in COG mode using STRING 10.0 database showing all connections of TF genes with a confidence score > 0.4. The connection colors represent the types of evidence for inferring association: recurring neighborhood in different genomes (green line), events of gene fusion (red), co-occurrence of those genes in the same organisms (dark blue), co-expression (black), experimental protein–protein interaction data (pink), pathway described by other databases (light blue), literature text-mining (yellow), and homology (purple lines) (**a**) AP2 family, (**b**) bHLH family, (**c**) bZIP family, (**d**) MYB, (**e**) NAC, and (**f**) WRKY

expression patterns may provide important clues for gene functions. Here, we conducted a custom designed high throughput expression array analysis of the probes designed to the *Zz* transcripts obtained by the whole genome transcriptome sequencing and de novo assembly. From this dataset, expression signatures of *Zz* TFs were filtered out and 314 TF genes holding a significant expression ($P \leq 0.05$) at least in one time frame as compared to control were shortlisted (Supplementary Table S5). Modulation varied in different time frames, some TFs showed a steady upregulation

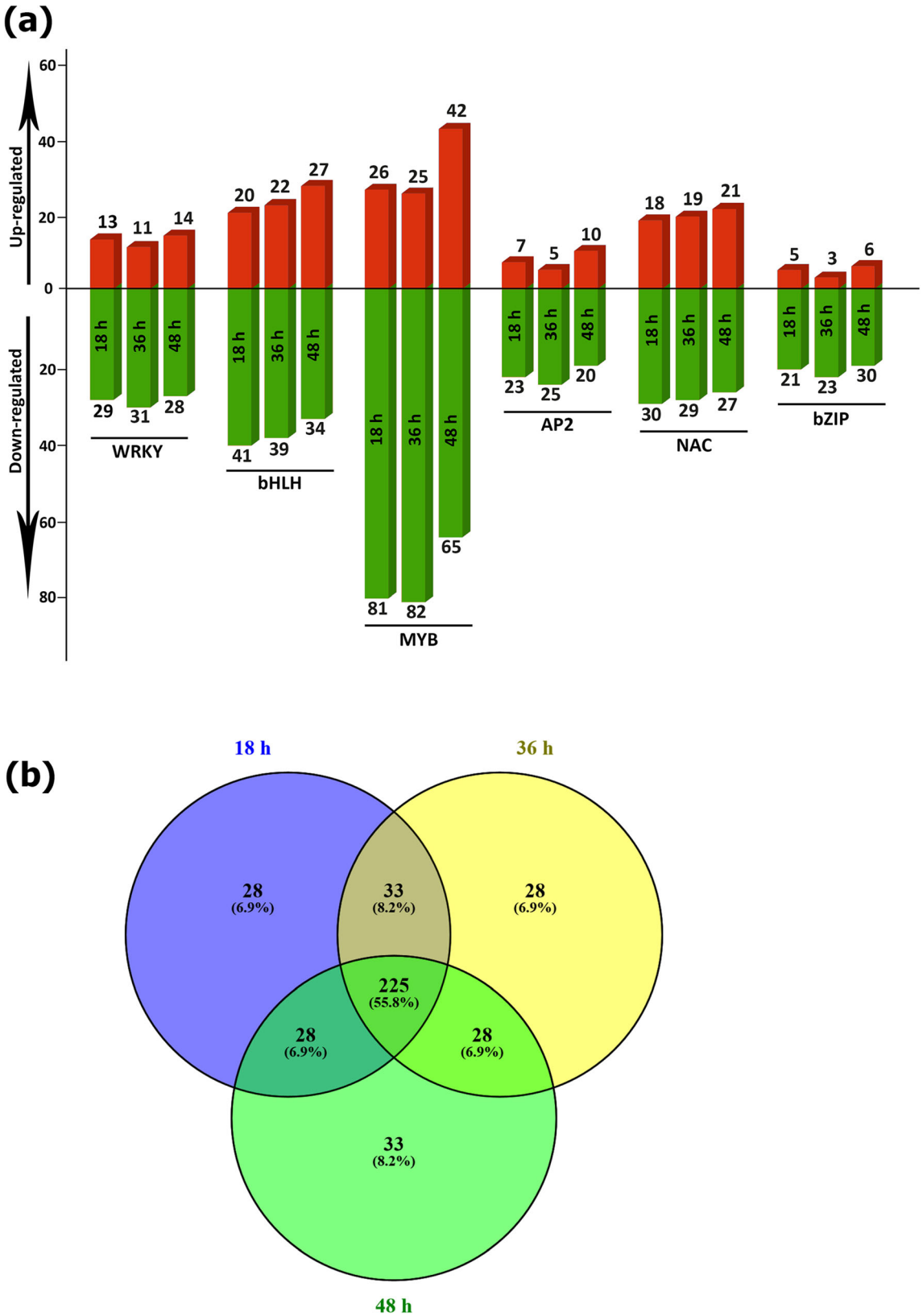


Fig. 3 Summary statistics of *Zz* TFs modulated in response to *P. myriotylum* infection. **a** Number of TFs upregulated and downregulated at 18, 36, and 48 hpi, ($P \leq 0.05$). The red bar represents upregulation while green bar represents downregulation. **b** Venn diagram representing the complex modulation of *Zz* TFs

and downregulation, while others depicted variable expression profiles (Fig. 3b). Of the total of 314 TFs, a total of 89, 85 and 120 TFs were upregulated at 18, 36, and 48 hpi while 225, 229, and 194 TFs were downregulated at 18, 36, and 48 hpi respectively (Fig. 3a) in this study reflecting their concerted activator or repressor roles during biotic stress response (Schenk et al. 2000; Singh et al. 2002; Windram et al. 2012; Seo and Choi 2015). In comparison to other TF families, the WRKY and MYB families were found to be predominantly participating in the transcriptional response to *P. myriotylum* as evidenced by their high differential modulation. A total of 13 (20%) WRKY and 34 (30%) MYB TFs were highly upregulated, whereas 30 (55%) WRKY and 72 (68%) MYB TFs were found to be downregulated at least in one of the time points (Fig. 4e, f). The WRKY and MYB TFs are known to play pivotal roles in regulatory networks controlling plant development, secondary metabolism, hormone signal transduction, disease resistance and abiotic stress tolerance (Eulgem et al. 2000; Dubos et al. 2010). We also observed at least 15 AP2 and 14 bHLH TFs strongly induced during the study, suggesting their crucial role in *Zz* defense response (Fig. 4a, b). The majority (60%) of the differential modulation identified in the NAC and bZIP TFs were of downregulation with 8 bZIP and 14 NAC TFs showing a strong induction (Fig. 4c, d). Not surprisingly, a predominant portion (80%) of the TFs depicted a downregulation in our study. It is well established that when a plant perceives a pathogen it switches its cellular machinery, recruiting its resources from normal cellular metabolism towards defense responses (Scheideler et al. 2002; Windram et al. 2012). In *Arabidopsis*, transcript levels of photosynthesis, carbon reduction cycle and pigment synthesis genes decreased in response to biotic stress, while genes coding for the synthesis of jasmonic acid (JA), salicylic acid (SA), and ethylene (ET) were upregulated (Bilgin et al. 2010). It has been suggested that the downregulation of photosynthetic gene expression enables plant's nitrogen resources to be reallocated for the synthesis of new defense proteins (Rojas et al. 2014). Therefore the role of repressor-type TFs observed in our study may be to regulate the activation of plant defense responses in a finely controlled manner in coordination with other cellular processes. Nevertheless, more research is needed to understand the exact mechanisms behind this observation.

Regulation of *Zz* TFs with putative defense functions

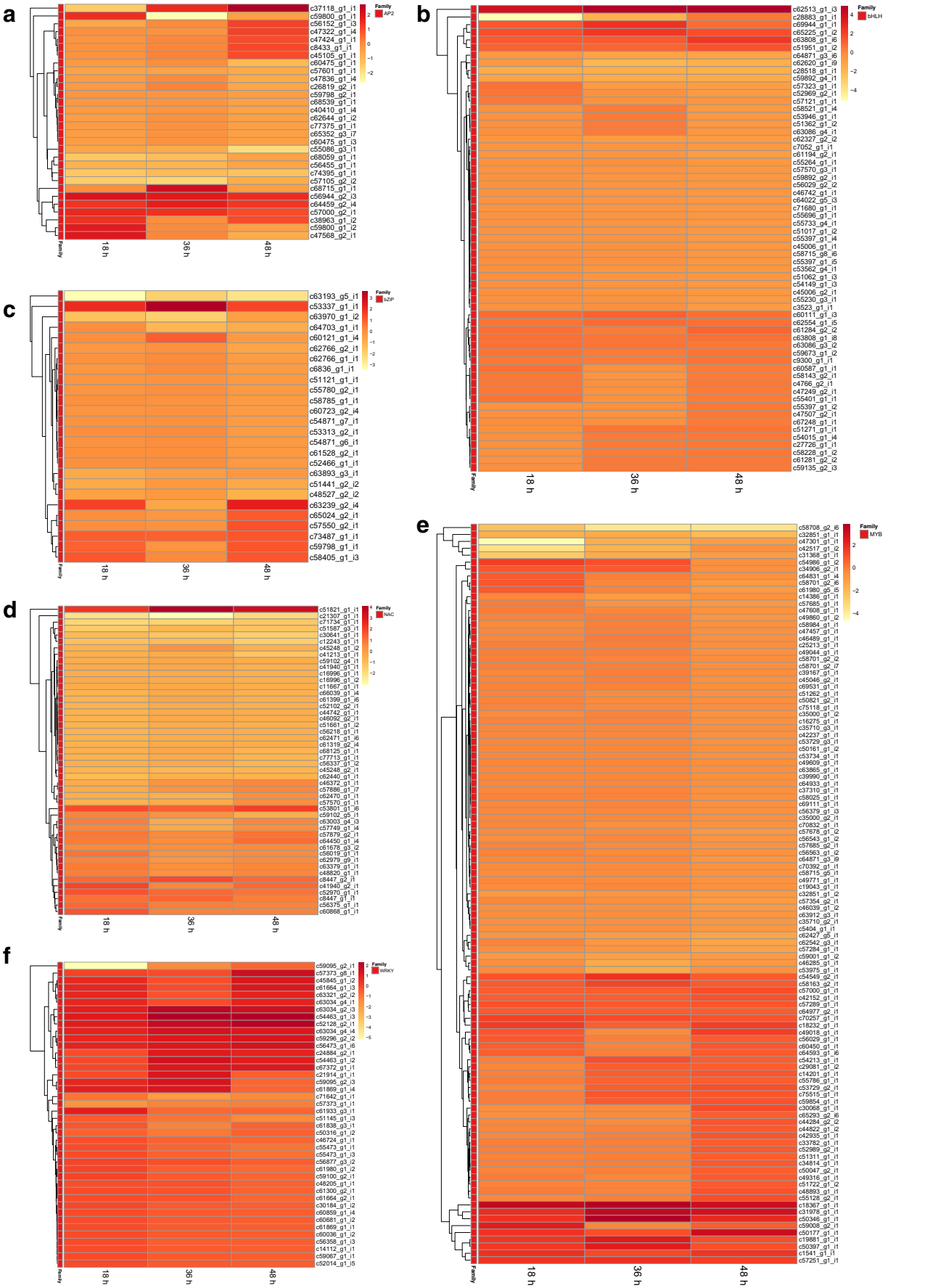
Defense response

The ability to perceive and mount a rapid response to pathogen attack is critical for plant survival. A study by Windram et al. (Windram et al. 2012) reported altered expression levels of at least 53 *Arabidopsis* AP2/ERF TFs during *B. cinerea* infection, suggesting its predominant role in the defense response. In our study, the expression of ERF1 (ortholog of Zz_c8433_g1_i1_AP2) was highly abundant at 48 hpi. AtERF1 is known to induce defense gene expression and plant resistance against necrotrophs (Lorenzo 2004). Another TF Zz_63086_g4_i1_HLH, an ortholog of BIM1 displayed a strong induction at 18 hpi. BIMs are known to regulate Brassinosteroids (BZR1/BES1) genes (growth-promoting steroid hormones in plants) that are involved in various developmental processes including plant defense (Kim and Wang 2010). The diverse roles played by NAC TFs in response to biotic and abiotic stresses are established beyond doubt (Nuruzzaman et al. 2013). Here, ATAF1 (ortholog of Zz_c51821_g1_i1_NAC and Zz_c62470_g1_i1_NAC) was strongly induced at 36 hpi. ATAF subfamily clearly appears to have a conserved, but non-redundant function in regulating the plant defense responses to different pathogens. Wu et al. (Wu et al. 2009) has reported an increase in the transcript abundance of ATAF1 in response to abiotic stress, wounding and *B. cinerea* infection, suggesting that NAC TFs play complex roles in plant defense. The Zz_c75515_g1_i1_MYB, an ortholog of *Arabidopsis* MYB protein AtMYB30 displayed high induction during 36 and 48 hpi. AtMYB30 is a key regulator of plant defenses and one of the best characterized MYB regulators directing defense-related transcriptional responses (Raffaele et al. 2008; Froidure et al. 2010).

Fig. 4 Heat map representing temporal expression profiles of *Zz* TFs following interaction with *P. myriotylum*. The samples were collected at 18, 36, and 48 hpi. **a** AP2 family, **b** bHLH family, **c** bZIP family, **d** NAC family, **e** MYB, and **f** WRKY

Hormone signaling

Signal transduction by hormone signaling is a key component of basal plant immunity. Among the hormones, SA, JA, and ethylene are regarded as core immune phytohormones. Extensive crosstalk between hormone signaling pathways is well established (Derksen et al. 2013) enabling plants to fine-tune its defense response against specific pathogens. These crosstalk between signaling pathways can be synergistic or antagonistic (Pieterse et al. 2009) and was also evident in our expression array data. WRKY TFs are a large family of regulators involved in various developmental and physiological processes, specifically in coping with diverse biotic and abiotic stresses (Eulgem et al. 2000). In this study, *Arabidopsis* WRKY33, an ortholog of Zz_c59296_g2_i2_WRKY TF was strongly induced at all time frames. This TF plays important role in pathogen resistance and was shown to be critical for resistance against the necrotrophic fungal pathogen *B. cinerea* (Zheng et al. 2006) and *Alternaria brassicicola* (Birkenbihl et al. 2012). It is believed that WRKY33 acts as a node of convergence for integrating SA and JA defense signaling. (Li 2004) Similarly, Zz_c67372_g1_i1_WRKY (ortholog of


a
Family
AP2
c37118_g1_i1
c59800_g1_i1
c56152_g1_i3
c47322_g1_i4
c47424_g1_i1
c8433_g1_i1
c45105_g1_i1
c60475_g1_i1
c57601_g1_i1
c47836_g1_i4
c26819_g2_i1
c59798_g2_i1
c68539_g1_i1
c40410_g1_i4
c62644_g1_i2
c77375_g1_i1
c65352_g3_i7
c60475_g1_i3
c55086_g3_i1
c68059_g1_i1
c56455_g1_i1
c74395_g1_i1
c57105_g2_i2
c68715_g1_i1
c56944_g2_i3
c64459_g2_i4
c57000_g2_i1
c38963_g1_i2
c59800_g1_i2
c47568_g2_i1
18 h
36 h
48 h
b
Family
bHLH
c62513_g1_i3
c28883_g1_i1
c69944_g1_i1
c65225_g1_i2
c63808_g1_i6
c51951_g1_i2
c64871_g3_i6
c62620_g1_i9
c28518_g1_i1
c59892_g4_i1
c57323_g1_i1
c52969_g2_i1
c57121_g1_i1
c58521_g1_i4
c53946_g1_i1
c51362_g1_i2
c63086_g4_i1
c62327_g2_i2
c7052_g1_i1
c61194_g2_i1
c55264_g1_i1
c57570_g3_i1
c59892_g2_i1
c56029_g2_i2
c46742_g1_i1
c64022_g5_i3
c71680_g1_i1
c55696_g1_i1
c55733_g4_i1
c51017_g1_i2
c55397_g1_i4
c45006_g1_i1
c58715_g8_i6
c55397_g1_i5
c53562_g4_i1
c51062_g1_i3
c54149_g1_i3
c45006_g2_i1
c55230_g3_i1
c3523_g1_i1
c60111_g1_i3
c62554_g1_i5
c61284_g2_i2
c63808_g1_i8
c63086_g3_i2
c59673_g1_i2
c9300_g1_i1
c60587_g1_i1
c58143_g2_i1
c4766_g2_i1
c47249_g2_i1
c55401_g1_i1
c55397_g1_i2
c47507_g2_i1
c67248_g1_i1
c51271_g1_i1
c54015_g1_i4
c27726_g1_i1
c58228_g1_i2
c61281_g2_i2
c59135_g2_i3
18 h
36 h
48 h
c
Family
bZIP
c63193_g5_i1
c53337_g1_i1
c63970_g1_i2
c64703_g1_i1
c60121_g1_i4
c62766_g2_i1
c62766_g1_i1
c6836_g1_i1
c51121_g1_i1
c55780_g2_i1
c58785_g1_i1
c60723_g2_i4
c54871_g7_i1
c53313_g2_i1
c54871_g6_i1
c61528_g2_i1
c52466_g1_i1
c63893_g3_i1
c51441_g2_i2
c48527_g2_i2
c63239_g2_i4
c65024_g2_i1
c57550_g2_i1
c73487_g1_i1
c59798_g1_i1
c58405_g1_i3
18 h
36 h
48 h
d
Family
NAC
c51821_g1_i1
c21307_g1_i1
c71734_g1_i1
c51587_g3_i1
c30641_g1_i1
c12243_g1_i1
c45248_g1_i2
c41213_g1_i1
c59102_g4_i1
c41940_g1_i1
c16996_g1_i1
c16996_g1_i2
c11667_g1_i1
c66039_g1_i4
c61399_g1_i6
c52102_g2_i1
c44742_g1_i1
c46092_g2_i1
c51661_g1_i2
c56218_g1_i1
c62471_g1_i6
c61319_g2_i4
c68125_g1_i1
c77713_g1_i1
c56337_g1_i2
c45246_g2_i1
c62440_g1_i1
c46372_g1_i1
c57886_g1_i7
c62470_g1_i1
c57570_g1_i1
c53801_g1_i6
c59102_g5_i1
c63003_g4_i3
c57749_g1_i4
c57879_g2_i1
c64450_g1_i4
c61678_g3_i2
c56019_g1_i1
c62979_g9_i1
c63379_g1_i1
c48820_g1_i1
c8447_g2_i1
c41940_g2_i1
c52970_g1_i1
c8447_g1_i1
c56375_g1_i1
c60868_g1_i1
18 h
36 h
48 h
e
Family
MYB
c58708_g2_i6
c32851_g1_i1
c47301_g1_i1
c42517_g1_i2
c31368_g1_i1
c54986_g1_i2
c34906_g2_i1
c64831_g1_i4
c58701_g2_i6
c61980_g5_i6
c14386_g1_i1
c57685_g1_i1
c47608_g1_i1
c49860_g1_i2
c58984_g1_i1
c47457_g1_i1
c46489_g1_i1
c25213_g1_i1
c49044_g1_i1
c58701_g2_i2
c58701_g2_i7
c39167_g1_i1
c45046_g2_i1
c69531_g1_i1
c51262_g1_i1
c50821_g2_i1
c75118_g1_i1
c35000_g1_i2
c16275_g1_i1
c35710_g3_i1
c42237_g1_i1
c53729_g3_i1
c50161_g1_i2
c53734_g1_i1
c49609_g1_i1
c53865_g1_i1
c39990_g1_i1
c64933_g1_i1
c37310_g1_i1
c58025_g1_i1
c69111_g1_i1
c56379_g1_i3
c35000_g2_i1
c70832_g1_i1
c57678_g1_i2
c56543_g1_i2
c57685_g2_i1
c56563_g1_i2
c64871_g3_i9
c70392_g1_i1
c58715_g5_i1
c49771_g1_i1
c19043_g1_i1
c32851_g1_i2
c57354_g2_i1
c46039_g1_i2
c63912_g3_i1
c35710_g2_i1
c5404_g1_i1
c62427_g5_i1
c62542_g3_i1
c57284_g1_i1
c59001_g1_i2
c46285_g1_i1
c53975_g1_i1
c54549_g2_i1
c58163_g2_i1
c57000_g1_i1
c42152_g1_i1
c57289_g1_i1
c64977_g2_i1
c70257_g1_i1
c18232_g1_i1
c49018_g1_i1
c56029_g1_i1
c60450_g1_i1
c64593_g1_i6
c54213_g1_i1
c29081_g1_i2
c14201_g1_i1
c55786_g1_i1
c53729_g2_i1
c75515_g1_i1
c59854_g1_i1
c30068_g1_i1
c65293_g2_i6
c44284_g2_i2
c44822_g1_i2
c42935_g1_i1
c33782_g1_i1
c52989_g2_i1
c51311_g2_i1
c34814_g1_i1
c50047_g2_i1
c49316_g1_i1
c51722_g1_i2
c48893_g1_i1
c55128_g2_i1
c18367_g1_i1
c31978_g1_i1
c50346_g1_i1
c59008_g2_i1
c50177_g1_i1
c19881_g1_i1
c50397_g1_i1
c1541_g1_i1
c57251_g1_i1
18 h
36 h
48 h
f
Family
WRKY
c59095_g2_i1
c57373_g6_i1
c45845_g1_i2
c61664_g1_i3
c63321_g2_i2
c63034_g4_i1
c63034_g2_i3
c54463_g1_i3
c52128_g2_i1
c63034_g4_i4
c59296_g2_i2
c56473_g1_i6
c24884_g2_i1
c54463_g1_i2
c67372_g1_i1
c21914_g1_i1
c59095_g2_i3
c61869_g1_i4
c71642_g1_i1
c57373_g1_i1
c61933_g3_i1
c51145_g1_i3
c61838_g3_i1
c50316_g1_i2
c46724_g1_i1
c55473_g1_i1
c55473_g1_i3
c56877_g3_i2
c61980_g1_i2
c59100_g2_i1
c48205_g1_i1
c61300_g2_i1
c61664_g2_i1
c30184_g1_i2
c60859_g1_i4
c60681_g1_i2
c61869_g1_i1
c60036_g1_i2
c56358_g1_i3
c14112_g1_i1
c59067_g1_i1
c52014_g1_i5
18 h
36 h
48 h

AtWRKY22); Zz_c54463_g1_i3_WRKY and Zz_c63321_g2_i2_WRKY (orthologs of WRKY70) displayed high upregulation in this study. These TFs are involved in crosstalk between SA and JA-dependent defense pathways against bacterial and fungal pathogens in *A. thaliana.* (Li et al. 2006; Phukan et al. 2016) The MYC subfamily of bHLH is widely known for its role in plant defense. In our study, MYC2 (an ortholog of Zz_c67248_g1_i1_HLH) was found to be highly expressed at 48 hpi. This is particularly significant as the MYC2 has been demonstrated to act as a master regulator in the crosstalk between SA-JA-ET signaling networks in response to necrotrophs leading to resistance (Pieterse et al. 2009; Kazan and Manners 2013). The members of TGA family are best-known bZIP TFs involved in plant defense (Alves et al. 2013). Here, the members of the TGA family (orthologs of Zz_c53337_g1_i1_bZIP and Zz_c63239_g2_i4_bZIP) were found to be significantly upregulated during the entire course of the study. Several TGA-interacting proteins are known to modulate their activity during plant defense response and they are involved in hormonal crosstalk, particularly which connect the SA and JA/ET pathway (Zander et al. 2010; Gatz 2012). An increasing body of research suggests the role of AP2/ERF TFs in defense responses against multiple pathogens (Seo and Choi 2015). In our study, AtERF9 and RAP2-2 orthologs of Zz_c45105_g1_i1_AP2 and Zz_c56944_g2_i3_AP2 respectively showed high induction. In *A. thaliana*, AtERF9 and RAP2-2 are known to play a pivotal role in resistance to the necrotrophic fungus *Botrytis cinerea* by modulating ethylene/JA signaling pathway (Zhao et al. 2012; Maruyama et al. 2013). MYB TFs are also known to be involved in hormone signal transduction. The expression of Zz_c18232_g1_i1_MYB (ortholog of AtMYB 72) was strongly induced at 18 and 36 hpi. AtMYB 72 is known to play a central role in JA/ET-dependent induced systemic resistance (ISR) that is effective against a broad range of pathogens (Van der Ent et al. 2008). From these results, we envisage that hormonal signaling crosstalk particularly that involves SA and JA plays a crucial role in providing resistance to *Zz* against *P. myriotylum*.

Cell wall fortification

Besides acting as a physical barrier, plant cell wall also serves as a defense barrier against pathogen attack. Modification of cell wall architecture is an essential part of plant response to invading pathogens whereby they restrict the pathogen spread. This is generally associated with the deposition of polymers such as lignin, a byproduct of phenylpropanoid secondary metabolite pathway that originates from phenylalanine. These compounds have established roles in plant growth and development, more importantly in defense against biotic stress (Liu et al. 2015). MYB TFs are regarded as the major regulator playing a key role in regulating sets of enzymatic genes involved in the phenylpropanoid biosynthetic pathway (Liu et al. 2015). Remarkably, we observed a steep increase in the proportion of upregulated MYB TF genes towards 48 hpi (Fig. 4e), suggesting the concerted role of MYB TFs in cell wall fortification events in *Zz*. Corroborating to this observation, AtMYB6, an ortholog of Zz_c1541_g1_i1_MYB_DNA_BINDING was strongly overexpressed in our study. AtMYB6 plays a crucial role in cell wall thickening, lignin and phytoalexins biosynthesis (Atkinson and Urwin 2012). In addition, Zz_c57289_g1_i1_MYB_DNA_BINDING (ortholog of MYB36) was abundantly expressed throughout the time period, which regulates the gene expression of both casparian strip genes (*CASP*1, *PER*64, and *ESB*1) and lignin polymerization in the cell wall (Kamiya et al. 2015). Accordingly, it is tempting to speculate that Zz_c57289_g1_i1_MYB_DNA_BINDING and Zz_c1541_g1_i1_MYB_DNA_BINDING might be playing central roles in restricting *P. myriotylum* invasion in *Zz* and should be interesting candidates for future research.

Circadian regulator mediated defense

Circadian regulators play crucial roles in the temporal control of the defense genes in plants. Wang et al. (Wang et al. 2011) has reported that several defense genes are under circadian control by CCA1, allowing plants to 'anticipate' infection at dawn when the pathogen normally disperses the spores and time immune responses according to the perception of different pathogenic signals upon infection. Consistent with this report, in this study, *Arabidopsis* circadian regulator, CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1), orthologs of Zz_c18367_g1_i1_MYB; Zz_c59008_g2_i1_MYB and Zz_c31978_g1_i1_MYB was found to be significantly upregulated ($P \leq 0.05$) at all the time frames implying its crucial role in *Zz-Pythium* interaction leading to resistance.

Hypersensitive response

Hypersensitive response (HR) triggered by pathogens leads to rapid cell death in the vicinity of infection and thus prevents microbes from spreading. Production of reactive oxygen species (ROS) and localized host programmed cell death (PCD) are crucial mechanisms through which plants respond to pathogen attack. ROS may contribute to resistance by strengthening the host cell wall, thereby confining the pathogen in the infected site and further orchestrating HR-mediated defense gene activation (Torres et al. 2006; Lehmann et al. 2015). In our study, Zz_c8447_g2_i1_NAC (ortholog of NAC87) was

highly induced at 36 hpi. Recently, Yan et al. (2017) has shown that NAC87 in *Brassica napus* L. modulates ROS and PCD accompanied by typical changes at the morphological and cellular levels in response to multiple stresses. Similarly, two of our bZIP TFs, Zz_c63239_g2_i4_bZIP and Zz_c53337_g1_i1_bZIP, orthologous to the TGA sub family of basic domain/leucine zipper was highly upregulated at 18 and 48 hpi, 36 hpi respectively. The members of the TGA family are known to regulate ROS scavenger, glutathione-S-transferase and pathogenesis-related (PR) genes (Ndamukong et al. 2007). Additionally, AtMYB30 (ortholog of Zz_c75515_g1_i1_MYB), highly expressed at 36 and 48 hpi has been shown to be involved in pathogen-induced hypersensitive responses (HR) and cell death (Raffaele et al. 2008). This phenomenon regulates multiple physiological processes, including terminal differentiation, senescence, and disease resistance (Pajerowska-Mukhtar and Dong 2009).

Conclusion

Presently our picture of the transcriptional networks underlying plant immunity is extremely fragmentary and mostly relies on isolated cases of TFs impacting selective signaling pathways, and the expression of a limited number of individual target genes. This study has provided the comprehensive identification, characterization, and temporal expression profiling of *Zz* TF genes with a special emphasis on TFs involved in *Zz-P. myriotylum* incompatible interaction. Our results highlighted a complex interplay between activating and repressing transcription factors from multiple families in fine-tuning defense response against *P. myriotylum* attack. Several TFs could be assigned putative functions in defense responses based on their well-characterized orthologous genes from model plants. In particular, the strong induction of TFs involved in cell wall fortification, lignin biosynthesis, and SA/JA hormone crosstalk allows us to envisage that this mechanism plays a central role in restricting *P. myriotylum* proliferation in *Zz*. Functional analysis of these TFs and delineating their signaling pathways could be the next step of this research. We have already initiated studies in this regard and are currently in the process of silencing key TFs using Virus-Induced Gene Silencing (VIGS) technique.

Acknowledgments VSN and LA gratefully acknowledge Kerala State Council for Science, Technology and Environment (KSCSTE), Government of Kerala Post-Doctoral Fellowship (Award No: 001-11/PDF/2016/KSCSTE) and research fellowship (Award No. 010-55/FHSP/2010/CSTE), respectively. SK thanks Science and Engineering Research Board (SERB), Government of India for National Post-Doctoral Fellowship (File No. PDF/2017/002022). KG gratefully acknowledges Council for Scientific and Industrial Research (CSIR), Government of India for research fellowship (F. No. 9/716(0103)/2008-EMR-I). GT acknowledges CSIR, Government of India for a research grant (No. 38(1397)/14EMR-II). We thank the anonymous reviewers for their valuable comments in improving the manuscript.

Authors' contributions VSN executed the bioinformatics pipelines, performed data analysis, and drafted the manuscript. SK contributed to the bioinformatics analysis and drafting the manuscript. TEA contributed to the data analysis. KAG performed transcriptome sequencing and assembly. LA performed microarray analysis. AN contributed to data analysis. GT conceived and supervised the study. All authors read, commented on, and approved the final version of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Alves SM, Dadalto PS, Gonçalves BA et al (2013) Plant bZIP transcription factors responsive to pathogens: a review. Int J Mol Sci 14: 7815–7828

Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. J Exp Bot 63:3523–3543. https://doi.org/10.1093/jxb/ers100

Bai Y, Meng Y, Huang D, Qi Y, Chen M (2011) Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. Genomics 98:128–136. https://doi.org/10.1016/j.ygeno.2011.05.002

BILGIN DD, ZAVALA JA, ZHU JIN et al (2010) Biotic stress globally downregulates photosynthesis genes. Plant Cell Environ 33:1597–1613. https://doi.org/10.1111/j.1365-3040.2010.02167.x

Birkenbihl RP, Diezel C, Somssich IE (2012) Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses toward Botrytis cinerea infection. Plant Physiol 159:266–285. https://doi.org/10.1104/pp.111.192641

Buscaill P, Rivas S (2014) Transcriptional control of plant defence responses. Curr Opin Plant Biol 20:35–46. https://doi.org/10.1016/j.pbi.2014.04.004

Conesa A, Götz S, García-Gómez JM et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Conte MG, Gaillard S, Lanau N, Rouard M, Perin C (2008) GreenPhylDB: a database for plant comparative genomics. Nucleic Acids Res 36:D991–D998. https://doi.org/10.1093/nar/gkm934

De Coninck B, Timmermans P, Vos C et al (2015) What lies beneath: belowground defense strategies in plants. Trends Plant Sci 20:91–101. https://doi.org/10.1016/j.tplants.2014.09.007

Derksen H, Rampitsch C, Daayf F (2013) Signaling cross-talk in plant disease resistance. Plant Sci 207:79–87. https://doi.org/10.1016/j.plantsci.2013.03.004

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. Trends Plant Sci 15:573–581. https://doi.org/10.1016/j.tplants.2010.06.005

Eulgem T (2005) Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci 10:71–78. https://doi.org/10.1016/j.tplants.2004.12.006

Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol 10:366–371. https://doi.org/10.1016/j.pbi.2007.04.020

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Plant Pathol J 26:110–114. https://doi.org/10.5423/PPJ.2010.26.2.110

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41:808–815. https://doi.org/10.1093/nar/gks1094

Froidure S, Canonne J, Daniel X, Jauneau A, Briere C, Roby D, Rivas S (2010) AtsPLA2-α nuclear relocalization by the Arabidopsis transcription factor AtMYB30 leads to repression of the plant defense response. Proc Natl Acad Sci 107:15281–15286. https://doi.org/10.1073/pnas.1009056107

Gatz C (2012) From pioneers to team players: TGA transcription factors provide a molecular link between different stress pathways. Mol Plant-Microbe Interact 26:151–159. https://doi.org/10.1094/MPMI-04-12-0078-IA

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol 29:644–652. https://doi.org/10.1038/nbt.1883

Hu W, Yang H, Yan Y, Wei Y, Tie W, Ding Z, Zuo J, Peng M, Li K (2016) Genome-wide characterization and analysis of bZIP transcription factor gene family related to abiotic stress in cassava. Sci Rep 6. https://doi.org/10.1038/srep22783

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P (2016) EGGNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293. https://doi.org/10.1093/nar/gkv1248

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P (2017) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol 34:2115–2122. https://doi.org/10.1093/molbev/msx148

Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F, bZIP Research Group (2002) bZIP transcription factors in Arabidopsis. Trends Plant Sci 7:106–111. https://doi.org/10.1016/S1360-1385(01)02223-3

Jayaswall K, Mahajan P, Singh G, Parmar R, Seth R, Raina A, Swarnkar MK, Singh AK, Shankar R, Sharma RK (2016) Transcriptome analysis reveals candidate genes involved in blister blight defense in tea (Camellia sinensis (L) Kuntze). Sci Rep 6:1–14. https://doi.org/10.1038/srep30412

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323–329. https://doi.org/10.1038/nature05286

Kamiya T, Borghi M, Wang P, Danku JMC, Kalmbach L, Hosmani PS, Naseer S, Fujiwara T, Geldner N, Salt DE (2015) The MYB36 transcription factor orchestrates Casparian strip formation. Proc Natl Acad Sci 112:10533–10538. https://doi.org/10.1073/pnas.1507691112

Kangasjärvi S, Neukermans J, Li S et al (2012) Photosynthesis, photorespiration, and light signalling in defence responses. J Exp Bot 63:1619–1636. https://doi.org/10.1093/jxb/err402

Kavitha PG, Thomas G (2007) Evaluation of Zingiberaceae for resistance to ginger soft rot caused by Pythium aphanidermatum (Edson) fitzp. Plant Genet Resour Newslett 152:54–57

Kavitha PG, Thomas G (2008) Defence transcriptome profiling of Zingiber zerumbet (L.) Smith by mRNA differential display. J Biosci 33:81–90. https://doi.org/10.1007/s12038-008-0002-2

Kazan K, Manners JM (2013) MYC2: the master in action. Mol Plant 6:686–703. https://doi.org/10.1093/mp/sss128

Kim T-W, Wang Z-Y (2010) Brassinosteroid signal transduction from receptor kinases to transcription factors. Annu Rev Plant Biol 61:681–704. https://doi.org/10.1146/annurev.arplant.043008.092057

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054

Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Prasad M (2014) Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (Setaria italica L.). PLoS One 9:1–14. https://doi.org/10.1371/journal.pone.0113092

Lehmann S, Serrano M, L'Haridon F, Tjamos SE, Metraux JP (2015) Reactive oxygen species and plant resistance to fungal pathogens. Phytochemistry 112:54–62. https://doi.org/10.1016/j.phytochem.2014.08.027

Li J (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell Online 16:319–331. https://doi.org/10.1105/tpc.016980

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

Li J, Brader G, Kariola T, Tapio Palva E (2006) WRKY70 modulates the selection of signaling pathways in plant defense. Plant J 46:477–491. https://doi.org/10.1111/j.1365-313X.2006.02712.x

Liu L, White MJ, Macrae TH (1999) Functional domains, evolution and regulation. Eur J Biochem 257:247–257. https://doi.org/10.1046/j.1432-1327.1999.00349.x

Liu J, Osbourn A, Ma P (2015) MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Mol Plant 8:689–708. https://doi.org/10.1016/j.molp.2015.03.012

Lorenzo O (2004) JASMONATE-INSENSITIVE1 encodes a MYC transcription factor essential to discriminate between different Jasmonate-regulated defense responses in Arabidopsis. Plant Cell Online 16:1938–1950. https://doi.org/10.1105/tpc.022319

Maruyama Y, Yamoto N, Suzuki Y, Chiba Y, Yamazaki KI, Sato T, Yamaguchi J (2013) The Arabidopsis transcriptional repressor ERF9 participates in resistance against necrotrophic fungi. Plant Sci 213:79–87. https://doi.org/10.1016/j.plantsci.2013.08.008

McGrath KC, Dombrecht B, Manners JM, Schenk PM, Edgar CI, Maclean DJ, Scheible WR, Udvardi MK, Kazan K (2005) Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. Plant Physiol 139:949–959. https://doi.org/10.1104/pp.105.068544

Metsalu T, Vilo J (2015) ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res 43:W566–W570. https://doi.org/10.1093/nar/gkv468

Mukhtar MS, Deslandes L, Auriac MC, Marco Y, Somssich IE (2008) The Arabidopsis transcription factor WRKY27 influences wilt disease symptom development caused by Ralstonia solanacearum. Plant J 56(6):935–947. https://doi.org/10.1111/j.1365-313X.2008.03651.x.

Muthamilarasan M, Khandelwal R, Yadav CB, Bonthala VS, Khan Y, Prasad M (2014) Identification and molecular characterization of MYB transcription factor superfamily in C4 model plant foxtail millet (Setaria italica L.). PLoS One 9:e109920. https://doi.org/10.1371/journal.pone.0109920

Muthamilarasan M, Bonthala VS, Khandelwal R, Jaishankar J, Shweta S, Nawaz K, Prasad M (2015) Global analysis of WRKY transcription factor superfamily in Setaria identifies potential candidates involved

in abiotic stress signaling. Front Plant Sci 6:1–15. https://doi.org/10.3389/fpls.2015.00910

Muthamilarasan M, Mangu VR, Zandkarimi H, Prasad M, Baisakh N (2016) Structure, organization and evolution of ADP-ribosylation factors in rice and foxtail millet, and their expression in rice. Sci Rep 6:1–13. https://doi.org/10.1038/srep24008

Ndamukong I, Al AA, Thurow C et al (2007) SA-inducible Arabidopsis glutaredoxin interacts with TGA factors and suppresses JA-responsive PDF1.2 transcription. Plant J 50:128–139. https://doi.org/10.1111/j.1365-313X.2007.03039.x

Nuruzzaman M, Sharoni AM, Kikuchi S (2013) Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front Microbiol 4:1–16. https://doi.org/10.3389/fmicb.2013.00248

Okubara PA, Paulitz TC (2005) Root defense responses to fungal pathogens: a molecular perspective. Plant Soil 274:215–226. https://doi.org/10.1007/s11104-004-7328-9

Pajerowska-Mukhtar K, Dong X (2009) A kiss of death—proteasome-mediated membrane fusion and programmed cell death in plant defense against bacterial infection. Genes Dev 23:2449–2454. https://doi.org/10.1101/gad.1861609

Phukan UJ, Jeena GS, Shukla RK (2016) WRKY transcription factors: molecular regulation and stress responses in plants. Front Plant Sci 7:1–14. https://doi.org/10.3389/fpls.2016.00760

Pieterse CMJ, Leon-Reyes A, Van Der Ent S, Van Wees SCM (2009) Networking by small-molecule hormones in plant immunity. Nat Chem Biol 5:308–316. https://doi.org/10.1038/nchembio.164

Pinheiro GL, Marques CS, Costa MDBL, Reis PAB, Alves MS, Carvalho CM, Fietto LG, Fontes EPB (2009) Complete inventory of soybean NAC transcription factors: sequence conservation and expression analysis uncover their distinct roles in stress response. Gene 444:10–23. https://doi.org/10.1016/j.gene.2009.05.012

Puranik S, Sahu PP, Srivastava PS, Prasad M (2012) NAC proteins: regulation and role in stress tolerance. Trends Plant Sci 17:369–381. https://doi.org/10.1016/j.tplants.2012.02.004

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics 140:345–356. https://doi.org/10.1016/S0169-5347(99)01816-9

Raffaele S, Vailleau F, Léger A et al (2008) A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in Arabidopsis. Plant Cell 20:752–767. https://doi.org/10.1105/tpc.107.054858

Rojas CM, Senthil-Kumar M, Tzin V, Mysore KS (2014) Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. Front Plant Sci 5:1–12. https://doi.org/10.3389/fpls.2014.00017

Saidi MN, Mergby D, Brini F (2017) Identification and expression analysis of the {NAC} transcription factor family in durum wheat (Triticum turgidum L. ssp. durum). Plant Physiol Biochem 112:117–128. https://doi.org/10.1016/j.plaphy.2016.12.028

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem Biophys Res Commun 290:998–1009. https://doi.org/10.1006/bbrc.2001.6299

Salzman RA, Fujita T, Hasegawa PM (1999) An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. Plant Mol Biol 17:11–17. https://doi.org/10.1023/a:1007520314478

Scheideler M, Schlaich NL, Fellenberg K, Beissbarth T, Hauser NC, Vingron M, Slusarenko AJ, Hoheisel JD (2002) Monitoring the switch from housekeeping to pathogen defense metabolism in Arabidopsis thaliana using cDNA arrays. J Biol Chem 277:10555–10561. https://doi.org/10.1074/jbc.M104863200

Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proc Natl Acad Sci U S A 97:11655–11660. https://doi.org/10.1073/pnas.97.21.11655

Seo E, Choi D (2015) Functional studies of transcription factors involved in plant defenses in the genomics era. Brief Funct Genomics 14:260–267. https://doi.org/10.1093/bfgp/elv011

Singh KB, Foley RC, Oñate-Sánchez L (2002) Transcription factors in plant defense and stress responses. Curr Opin Plant Biol 5:430–436. https://doi.org/10.1016/S1369-5266(02)00289-3

Smita S, Katiyar A, Chinnusamy V, Pandey DM, Bansal KC (2015) Transcriptional regulatory network analysis of MYB transcription factor family genes in rice. Front Plant Sci 6:1–19. https://doi.org/10.3389/fpls.2015.01157

Song A, Gao T, Li P, Chen S, Guan Z, Wu D, Xin J, Fan Q, Zhao K, Chen F (2016) Transcriptome-wide identification and expression profiling of the DOF transcription factor gene family in Chrysanthemum morifolium. Front Plant Sci 7. https://doi.org/10.3389/fpls.2016.00199

Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol 4:447–456. https://doi.org/10.1016/S1369-5266(00)00199-0

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680. https://doi.org/10.1093/nar/22.22.4673

Toledo-Ortiz G, Huq E, Quail PH (2003) The Arabidopsis basic / helix-loop-helix transcription factor family. Plant Cell 15:1749–1770. https://doi.org/10.1105/tpc.013839.et

Torres MA, Jones JDG, Dangl JL (2006) Reactive oxygen species signaling in response to pathogens. Plant Physiol 141:373–378. https://doi.org/10.1104/pp.106.079467.conjunction

Tsuda K, Somssich IE (2015) Transcriptional networks in plant immunity. New Phytol 206:932–947. https://doi.org/10.1111/nph.13286

Van der Ent S, Verhagen BWM, Van Doorn R et al (2008) MYB72 is required in early signaling steps of Rhizobacteria-induced systemic resistance in Arabidopsis. Plant Physiol 146:1293–1304. https://doi.org/10.1104/pp.107.113829

Vandromme M, Gauthier-Rouviere C, Lamb N, Fernandez A (1996) Regulation of transcription factor localization: fine tuning of gene expression. Trends Biochem Sci 21:59–64

Wang W, Barnaby JY, Tada Y, Li H, Tör M, Caldelari D, Lee DU, Fu XD, Dong X (2011) Timing of plant immune responses by a central circadian regulator. Nature 470:110–115. https://doi.org/10.1038/nature09766

Wen CL, Cheng Q, Zhao L, Mao A, Yang J, Yu S, Weng Y, Xu Y (2016) Identification and characterisation of Dof transcription factors in the cucumber genome. Sci Rep 6:1–11. https://doi.org/10.1038/srep23072

White TJ, Bruns TD, Lee SB, Taylor JW (1990) Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. In: PCR Protocols, a Guide to Methods and Applications, pp. 315-322.

Windram O, Madhou P, McHattie S, Hill C, Hickman R, Cooke E, Jenkins DJ, Penfold CA, Baxter L, Breeze E, Kiddle SJ, Rhodes J, Atwell S, Kliebenstein DJ, Kim YS, Stegle O, Borgwardt K, Zhang C, Tabrett A, Legaie R, Moore J, Finkenstadt B, Wild DL, Mead A, Rand D, Beynon J, Ott S, Buchanan-Wollaston V, Denby KJ (2012) Arabidopsis defense against Botrytis cinerea: chronology and regulation deciphered by high-resolution temporal transcriptomic analysis. Plant Cell 24:3530–3557. https://doi.org/10.1105/tpc.112.102046

Wu Y, Deng Z, Lai J, Zhang Y, Yang C, Yin B, Zhao Q, Zhang L, Li Y, Yang C, Xie Q (2009) Dual function of Arabidopsis ATAF1 in abiotic and biotic stress responses. Cell Res 19:1279–1290

Yadeta KA, J Thomma BP (2013) The xylem as battleground for plant hosts and vascular wilt pathogens. Front Plant Sci 4:1–12. https://doi.org/10.3389/fpls.2013.00097

Yan J, Tong T, Li X, Chen Q, Dai M, Niu F, Yang M, Deyholos MK, Yang B, Jiang YQ (2017) A novel NAC-type transcription factor, NAC87, from oilseed rape modulates reactive oxygen species accumulation and cell death. Plant Cell Physiol 0:1–14. https://doi.org/10.1093/pcp/pcx184

Yu C-S, Chen Y-C, Lu C-H, Hwang J-K (2006) Prediction of protein subcellular localization. Proteins Struct Funct Bioinf 64:643–651. https://doi.org/10.1002/prot.21018

Zander M, La Camera S, Lamotte O et al (2010) Arabidopsis thaliana class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses. Plant J 61:200–210. https://doi.org/10.1111/j.1365-313X.2009.04044.x

Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107

Zhao Y, Wei T, Yin KQ, Chen Z, Gu H, Qu LJ, Qin G (2012) Arabidopsis RAP2.2 plays an important role in plant resistance to Botrytis cinerea and ethylene responses. New Phytol 195:450–460. https://doi.org/10.1111/j.1469-8137.2012.04160.x

Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J 48:592–605. https://doi.org/10.1111/j.1365-313X.2006.02901.x

Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, Zhao PX, Rhee SY, Fei Z (2016) iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. Mol Plant 9:1667–1670. https://doi.org/10.1016/j.molp.2016.09.014

Zhu T, Nevo E, Sun D, Peng J (2012) Phylogenetic analyses unravel the evolutionary history of nac proteins in plants. Evolution (N Y) 66:1833–1848. https://doi.org/10.1111/j.1558-5646.2011.01553.x