Genome-wide characterization of WD40 protein family in *Monilinia fructigena*; insights in the evolutionary selective pressures acting on the WD40 repeats

Antonios Zambounis[1, *], Aliki Xanthopoulou[2]

[1]Institute of Plant Breeding and Genetic Resources, Department of Deciduous Fruit Trees, HAO-'Demeter', 59035, Naoussa, Greece. [2]Laboratory of Pomology, Department of Agriculture, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece.

Received: December 19, 2019; accepted: February 25, 2020.

The ascomycete fungus *Monilinia fructigena* is one of the most serious causal agents of brown rot in fruit trees. The genome of *M. fructigena* is available providing novel resources for the comprehensive characterization of particular gene families. WD40 proteins are scaffolding molecules depicting crucial roles in fundamental biological processes. In the present study, we identified 62 WD40 proteins in *M. fructigena* genome (MfWD40s). Based on their phylogenetic classification and domain architectures, MfWD40s were categorized into 5 clusters and 17 classes, respectively, which is indicative of their diverse expansion. Gene ontology analysis revealed that MfWD40s are involved in protein binding and various biological processes. RNA-seq data revealed that the highest number of *MfWD40s* genes showed stage-specific expression profiles with most of them being induced during germination of conidia. Furthermore, we accurately identified their WD40 repeats and assessed the evolutionary signatures acting upon them. The results postulate the existence of a purifying selection acting across their phylogenies. However, further analyses revealed that a number of amino acid residues is subjected to positively selection across the WD40 repeats. These findings would provide an important foundation towards deciphering the diverse functions of WD40 gene family in *M. fructigena* and the role of these proteins in pathogenicity during plant-microbe interactions.

Keywords: fungal diagnostics; *Monilinia fructigena*; WD40 repeats; evolutionary pressures; pathogens; plant-microbe interactions.

Abbreviations: MCL: markov cluster algorithm; NGS: next generation sequencing; PAML: phylogenetic analysis by maximum likelihood; RAxML: randomized axelerated maximum likelihood; PGGs: Paralogous Gene Groups.

[*]**Corresponding author:** Antonios Zambounis, Institute of Plant Breeding and Genetic Resources, Department of Deciduous Fruit Trees, HAO-'Demeter', 59035, Naoussa, Greece. Email: antonios.zamb@gmail.com.

Introduction

WD40 domain-containing proteins, also known as WD40-repeat proteins, are one of the largest protein families in eukaryotes and depict essential roles in fundamental cellular and biological processes [1, 2]. For example, in *Saccharomyces cerevisiae*, WD40 domain was ranked as the highest interacting domain [3]. These proteins typically contain repeated residues units of 40-60 residues [2] with conserved glycine-histidine (GH) and tryptophan-aspartate (WD) motifs [4-6]. Each WD40 repeat is folded into four anti-parallel *β*-strands with a characteristic propeller structure [2, 7, 8]. The number of these repeats in canonical WD40 domains is quite variable, while their sequences might reveal substantial diversity [2]. WD40 proteins are involved in specific cellular

processes serving as functional scaffolds for proteins interactions [2, 4, 8]. Thus, these proteins play pivotal roles in cell division, cytoskeleton assembly, DNA replication, RNA processing, apoptosis, and signal transduction [1]. The evolutionary expansion of WD40s proteins seems to be mainly oriented by intragenic duplication and recombination events following by an extensive sequence diversification [1, 7, 9]. Although genome-wide characterization of WD40 proteins has been conducted in many plant species, including foxtail millet [5], rice [9], cucumber, and *Arabidopsis* [10], little investigations have been reported in filamentous fungal species, where the putative functions of WD40 proteins remain unexplored so far.

Phytopathogenic fungi of the genus *Monilinia* (family *Sclerotiniaceae*) cause severe brown rot and blossom blight to stone and pome fruits all over the world [11]. Among them, *Monilinia fructigena* Honey *ex* Whetzel is one of the severe apothecial ascomycetes of this genus [12, 13]. The pathogen is under quarantine in many countries causing significant economic losses in plant species of the *Rosaceae* family, both in the field and mainly during postharvest storage [13]. Whole genome sequencing using NGS approaches of this important species (strain Mfrg269) is currently available [14]. The genome is assembled into 131 scaffolds accordingly to PacBio genomic data and harboring 9,960 predicted protein-coding genes. The availability of *M*. *fructigena* genome would definitely allow a further exploration of specific gene families and particularly of their structural and functional features.

In the present study, we carried out a genome-wide identification of the WD40 proteins in the genome of *M*. *fructigena* in order to shed light on structural characterization and the divergence in architecture of their functional domains, their transcriptome expression profiles in various developmental stages, as well as the evolutionary origins of their WD40 repeats. Our findings would undoubtedly facilitate the deeper understanding on the structure and function of the WD40s protein family of *M*. *fructigena*.

Materials and Methods

Sequences retrieval and genome-wide identification of MfWD40s proteins

For the genome-wide identification of WD40 proteins of *M. fructigena* (MfWD40s), the proteome of strain Mfrg269 was downloaded from NCBI genome database (https://www.ncbi.nlm.nih.gov/genome/genomes/66926; BioProject: PRJNA470675). Initially, the Hidden Markov Model (HMM) profile of the WD40 domain (PF00400) was downloaded from Pfam database (http://pfam.xfam.org) in order to pre-select the MfWD40s. The Hmmsearch program was employed from the HMMER3.0 software [15] under default parameters and using an *E*-value set at 1E-5.

Searches against the InterPro database (http://www.ebi.ac.uk/interpro/) were employed in order to ensure the integrity of MfWD40s upon InterProScan [16] results, and signatures matches of CDD and SMART enabled member databases. Finally, the WDSP software [17] was further adopted in a strict pipeline in order to screen the potentially reliable MfWD40s and to identify the final number of them based on average scores upon predictions of their WD40 repeats.

Phylogenetic, structural and collinear analyses

The phylogenetic relationships among the final dataset of MfWD40s were revealed by performing a MUSCLE alignment using 2 iterations, and then a tree reconstruction using

the RAxML program with 1,000 bootstrap replicates, a gamma model of rate heterogeneity and an ML estimate of the alpha parameter. The above analyses were performed in Geneious R7 platform [18].

In order to get a more precise view of the complex protein structures, the domain architectures of the MfWD40s were annotated based on the InterProScan, SMART and WDSP data. MfWD40s proteins with similar domain architectures were assigned to the same class. Diagrams for the representative domain architectures of MfWD40s were drawn employing the IBS software [19].

The MCScanX software [20] was used with typical parameter setting to identify collinear and segmentally duplicated regions across the MfWD40s in the genome. The scaffolds distributions and collinearity relationships of MfWD40s were visualized with the Circos software [21].

Gene ontology annotations

Gene Ontology (GO) analysis of MfWD40s was performed using Blast2GO software [22] with default settings. GO annotations were visualized using the WEGO software under a Pearson chi-square test (P<0.05).

Expression profiling using RNA-seq data

The RNA-seq data of *M. fructigena* strain Mfrg269 were used in order to elucidate the expression profiles of *MfWD40s* genes at 3 developmental stages. Pair-end reads that were corresponding to mycelium grown in the dark for 4 days (SRR6312183), mycelium grown in the dark for 2 days and then exposed to light for 2 days (SRR6312184), as well as in germinating conidia (SRR6312186), were all retrieved from SRA NCBI database. The *M. fructigena* annotated genes models (BioProject:PRJNA470675) were used to align the RNA-seq data to the reference genome employing the RNA STAR Gapped-read mapper software with default parameters [23]. The assignment of sequence reads to transcript features was performed using featureCounts program under default conditions, whereas the differentially expressed genes were retrieved using the DESeq2 R software using a model based on the negative binomial distribution [24]. The normalized transcripts counts of *MFWD40s* genes were used in order to perform an hierarchical clustering heatmap, that shows the expression profiles among each developmental stage based on average Euclidean distances employing ClustVis tool [25].

Identification of WD40 repeats

The annotation of WD40 repeats along with their accurate coordinates were accurately refined and confirmed based on WDSP predictions which provide highly precise WD40 repeats annotations [17]. The minimum number of repeats was 6 for assigning them in WD40 domains with known structures, while the repeat cut-off of the average score of WDSP was as high as 48 in order to reduce any false positive predictions and ensuring that the final WD40 repeats dataset was non-redundant [17, 26].

The reliable WD40 repeats of the MfWD40s that passed the above filters were used to assign their phylogenetic relationships employing the RAxML program with parameters as indicated above. The MEME Suite [27] was employed for discovery of the more reliable ungapped conserved motif across the identified WD40 repeats. The minimum motif width was set to 40 amino acids and the remaining parameters were set at default values.

Clustering and evolutionary analyses of WD40 repeats

An MCL-based approach was applied in order to assign unique PGGs and estimate the evolutionary signatures acting upon the identified WD40 repeats. All-against-all blast(p) searches were conducted with an *E*-value set at

1E-10. Similarity values of WD40 repeats up to 50% fed the MCL algorithm using an inflation value of "2" for clustering. The validity of the derived *Mf-WD40* PGGs was confirmed based on the overall amino acid identity (up to 50%) and the low existence of gaps among the aligned sequences in each PGG, eliminating by this way any gap-induced misalignments and elevating the divergence sequences in the following evolutionary analyses [28, 29].

Five PGGs, counting up to three WD40 repeats, were assigned to evolutionary analyses using the CODEML and CODEMLSITES programs from the PAML package [30]. The methodology for these evolutionary analyses were as reported by Zambounis *et al*. [28, 29, 31, 32], Lynn *et al*. [33]. Briefly, amino acid sequences MUSCLE alignments and neighbor joining (NJ) phylogenetic trees were constructed using the MEGA 5 software under default conditions [34]. In both CODEML and CODEMLSITES programs, log-likelihood calculations were performed under site-specific codon substitution models (such as M8, M7, M5, and M3) employing likelihood ratio tests (LRTs). The Bayesian method allowed the detection of codon residues subjected to positive selection by posterior probabilities. Thus, we estimated the variable selective pressures acting along the lineages and branches in the evolutionary trees, as well as the evolutionary signatures among amino acid residues using. Saturation effects of substitution were calculated among the aligned nucleotide sequences using MEGA 5 software [34].

Results and discussion

Identification of MfWD40 proteins

Specific protein families and signal transduction pathways are consistently recruited for the regulation of various developmental processes in phytopathogenic fungi [35]. Expansion of tandem repeats-containing gene families, such as the high repetitive WD40s, are well reported in plants [36], but little is known about these proteins in plant filamentous fungi. WD40 repeats are involved in numerous molecular processes, assembling themselves into a typical seven-bladed β-propeller fold [2, 8, 37]. In the present study, we systematically identified all WD40 proteins in *M. fructigena* genome, for a first time in a plant fungal pathogen, in order to gain valuable insights regarding their abundance, structural divergences, expression profiles and selective evolution pressures acting upon their WD40 repeats.

The hmmsearch program was applied to pre-select the WD40 proteins in predicted proteome using the HMM profile for the WD40 repeat (PF00400). An abundant number of 87 WD40 proteins were identified and the results were confirmed by InterProScan, SMART and CDD data. WDSP search-based domain analysis was further employed and entirely 62 non-redundant WD40s proteins (MfWD40s) were retained accordingly to the average confidence score employing by the WDSP software [17] in terms of accurately assigning WD40 repeats in a typical domain.

MfWD40s varied in length and physicochemical properties. The longest MfWD40 was RAL64015.1, while the shortest was RAL65436.1. The molecular weights varied from 854.99 kDa to 163,699.72 kDa, while isoelectric points ranged from 4.25 to 9.63. We also investigated *MfWD40s* gene structures and found that numbers of exons varied from 1 up to11.

The distribution of *MfWD40s* genes was widespread and uneven on 38 scaffolds, whereas QKRW01000002.1 scaffold contained the largest number of *MfWD40s* genes (5 genes, 8.06%). In plants the distribution of the *WD40* genes on

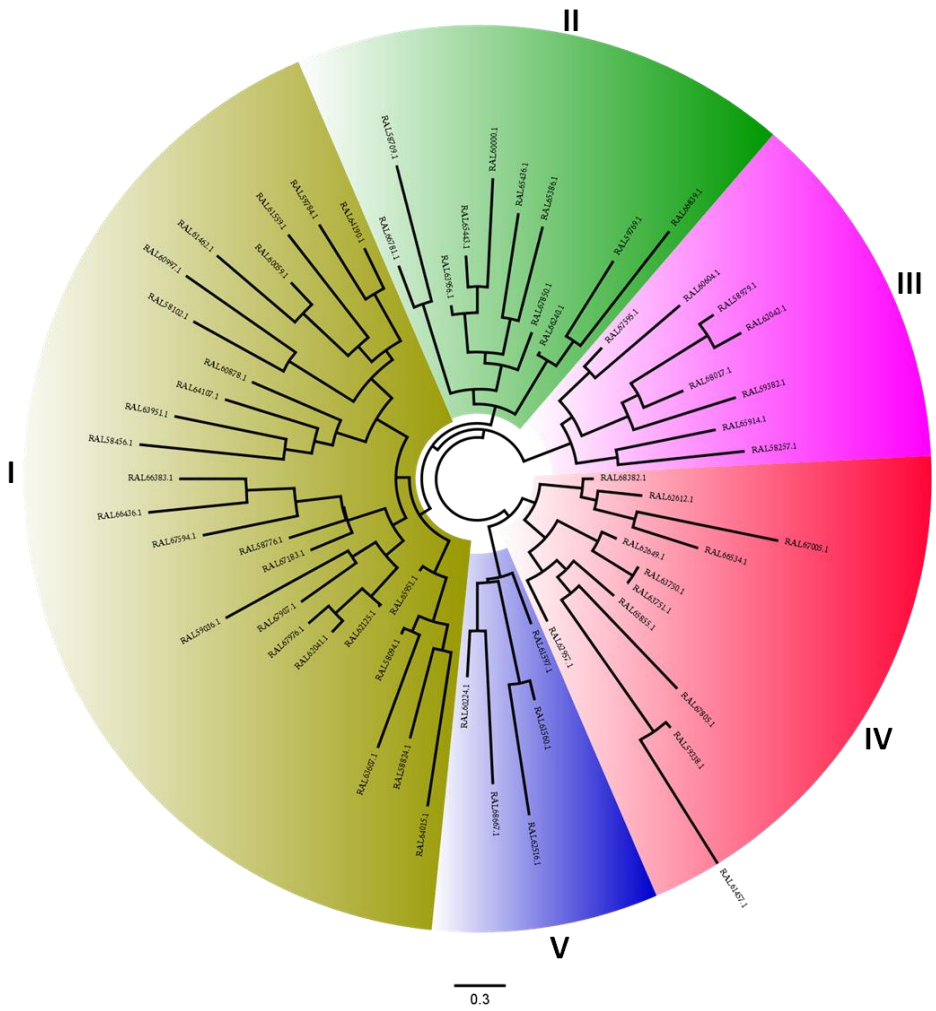

Figure 1. Phylogenetic classification of MfWD40s proteins. The tree was divided into five main distinct clusters (I–V).

chromosomes were also extensive and uneven [5, 9, 10], as well as in silkworm [1]. In general, segmental and tandem duplication might be responsible for expansions of WD40-containing genes families during their evolution [1, 9].

Phylogenetic classification, domain structural architectures and gene duplication analysis of the MfWD40 family

In order to gain insight into the expansion and phylogenetic relationships of the WD40 family in *M. fructigena*, we used the 62 identified MfWD40s to construct a RAxML phylogenetical tree. MfWD40s were classified into 5 main distinct clusters (I-V) containing 26, 11, 8, 12, and 5 proteins, respectively (Figure 1). The backbone topology of the tree was efficiently resolved containing 122 branches and reflecting a series of putative gene duplications up to the most recent bursts of expansion at terminal branches (Figure 1). The overall pair-wise identity was 6.6%, implying a rather high divergence among them.

The annotation results of InterProScan and SMART predictions were enabled in order to determine the domain architectures and decipher the complex domain structures of MfWD40s. Particularly, the WD40 repeats annotations replaced with WDSP predictions,

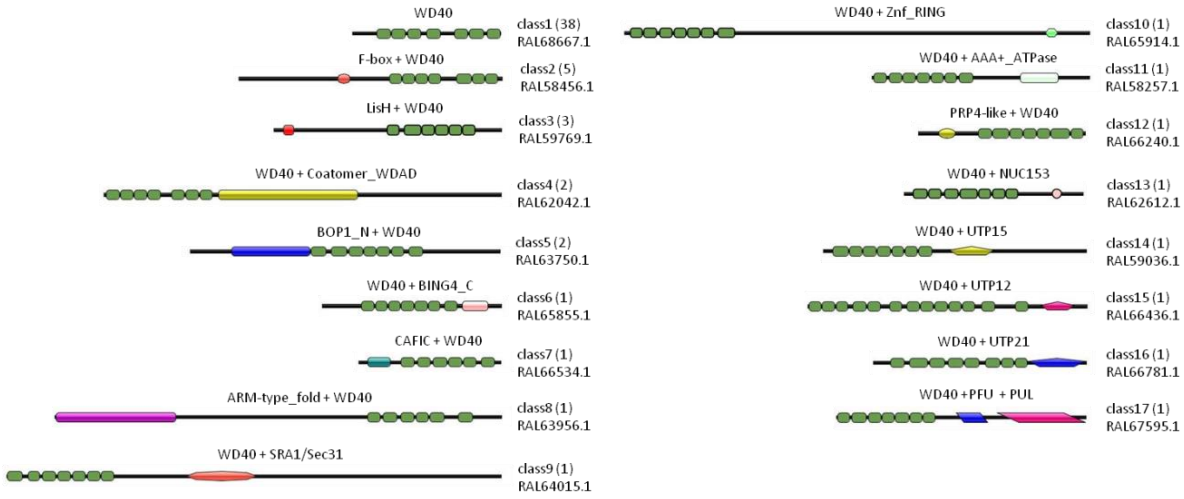

Figure 2. Domain architectures of MfWD40s proteins which are organized into 17 classes. The WD40 repeats are colored in dark green, and other domains are filled in with other shapes and colors separately. The number of members in each class is shown in parentheses. The name of the displayed protein is given below.

since WDSP can provide more complete and precise WD40 repeats annotations [17, 38]. Based on the domain annotations, proteins with similar domain architectures were assigned to the same class. Thus, MfWD40s were categorized into 17 classes indicating that MfWD40s were functionally a rather diverse family (Figure 2). A total of 38 MfWD40s were grouped into class 1. The other MfWD40s with additional functional domains were categorized into classes as follows: 5 MfWD40s containing the F-box domain were grouped into class 2; 3 MfWD40s comprising the LisH domain were identified as class 3; 2 MfWD40s with the Coatomer WD associated region were categorized into class 4; 2 MfWD40s with the BOP1 N domain were identified as class 5; MfWD40s with each of the BING4, CAFIC, ARM-type fold, SRA1/Sec31, Znf RING, AAA+ATPase, PRP4-like, NUC153, UTP15, UTP11, UTP21 domains were categorized into classes 6-16, respectively. Finally, an MfWD40 protein comprised from a fusion of two structurally related domains, named PFU and PUL, it was categorized into class 17.

Phylogenetic taxonomy revealed that many members of class 1 are separately grouped along with MfWD40s members of other classes with multi-domains (Figure 1). Only a few MfWD40s from the same classes were grouped together into the same sub-clade, as some members of the class 2 (WD40 proteins containing the F-box domain) which was also reported in WD40s proteins in *Bombyx mori* [1]. Consequently, these findings imply that this multi-domain MfWD40s architecture may have evolved from genes with only the WD40 domain within the same sub-clade. The domain recombination events might have happened mainly after gene duplication events in the evolution of the multi-domain *MfWD40s* genes.

As it was also proposed by He *et al*. [1], we speculate that both tandem and segmental duplication events might partially contribute to the expansion of *MfWD40s* gene family resulting at the diverse domain architectures. The gene duplication events among the *MfWD40s* genes, which were distributed at 38 scaffolds, were analyzed with MCScanX software. Based on our collinearity analysis, we found that 7 genes

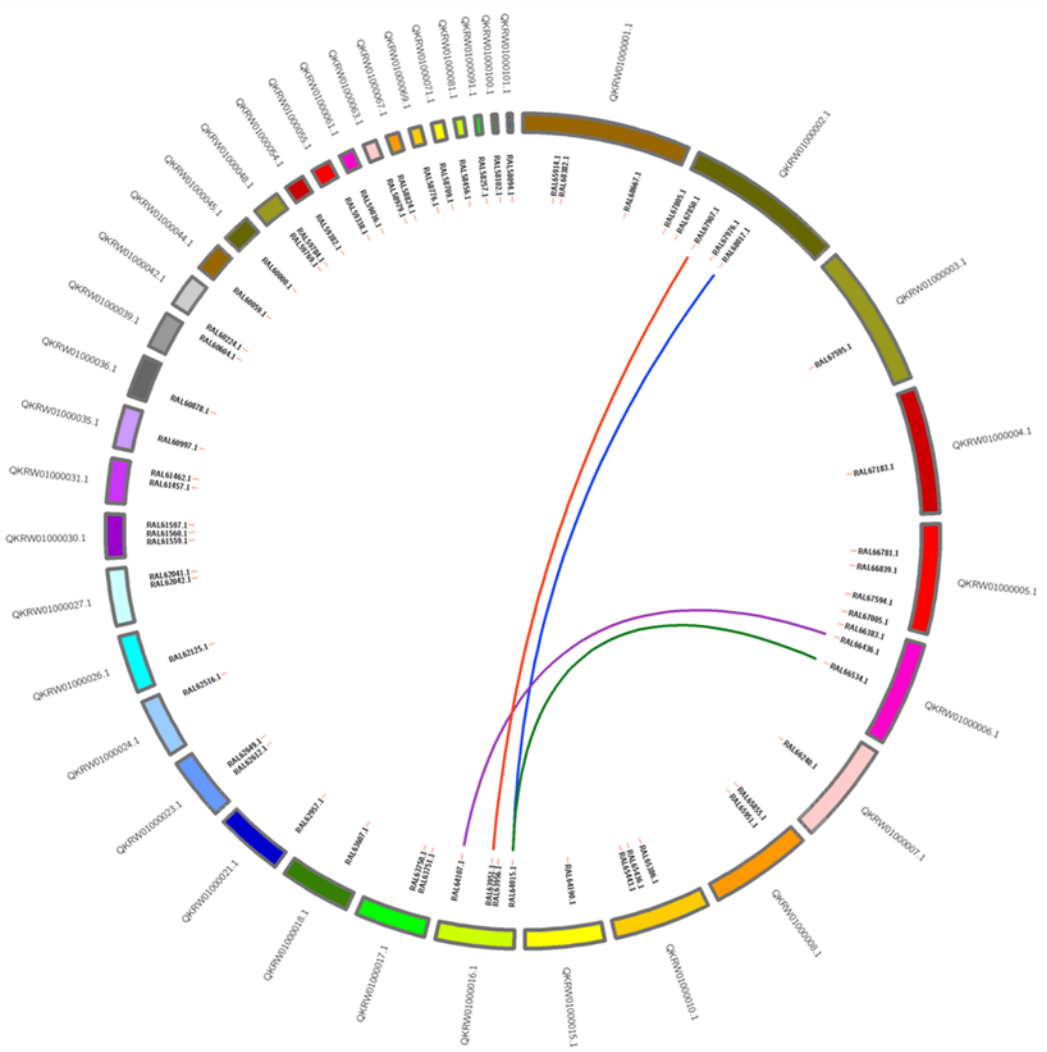

Figure 3. Scaffold locations of *MfWD40s* genes and segmentally duplicated gene pairs in *Monilinia fructigena* genome. The distribution of each *MfWD40* gene is marked on the circle with a short red line. The collinear blocks are shown with colored lines which denote the syntenic regions between genes.

(~11%) were involved in 4 segmental duplication events across 3 scaffolds (Figure 3). These results indicate that segmental duplication events only partially contribute to the expansion of *MfWD40s* genes.

Gene ontology analysis of the MfWD40s

GO analysis revealed that all MfWD40s participate in protein binding and are involved in diverse biological processes (Figure 4). The results revealed also that the MfWD40s of all 5 distinct phylogenetical clusters (I-V) had a variety of biological functions with a high degree of functional specificity.

Expression profiles of the MfWD40s

The stage-specific expression profiles of *MfWD40* genes was assessed using the RNA-seq data across 3 experimental conditions [14]. *MfWD40s* genes showed stage-specific expression patterns and most of them were highly expressed during germination of conidia implying that they are directly involved in infection processes (Figure 5). Among them, there were induced all *MfWD40s* genes containing the F-box domain, except

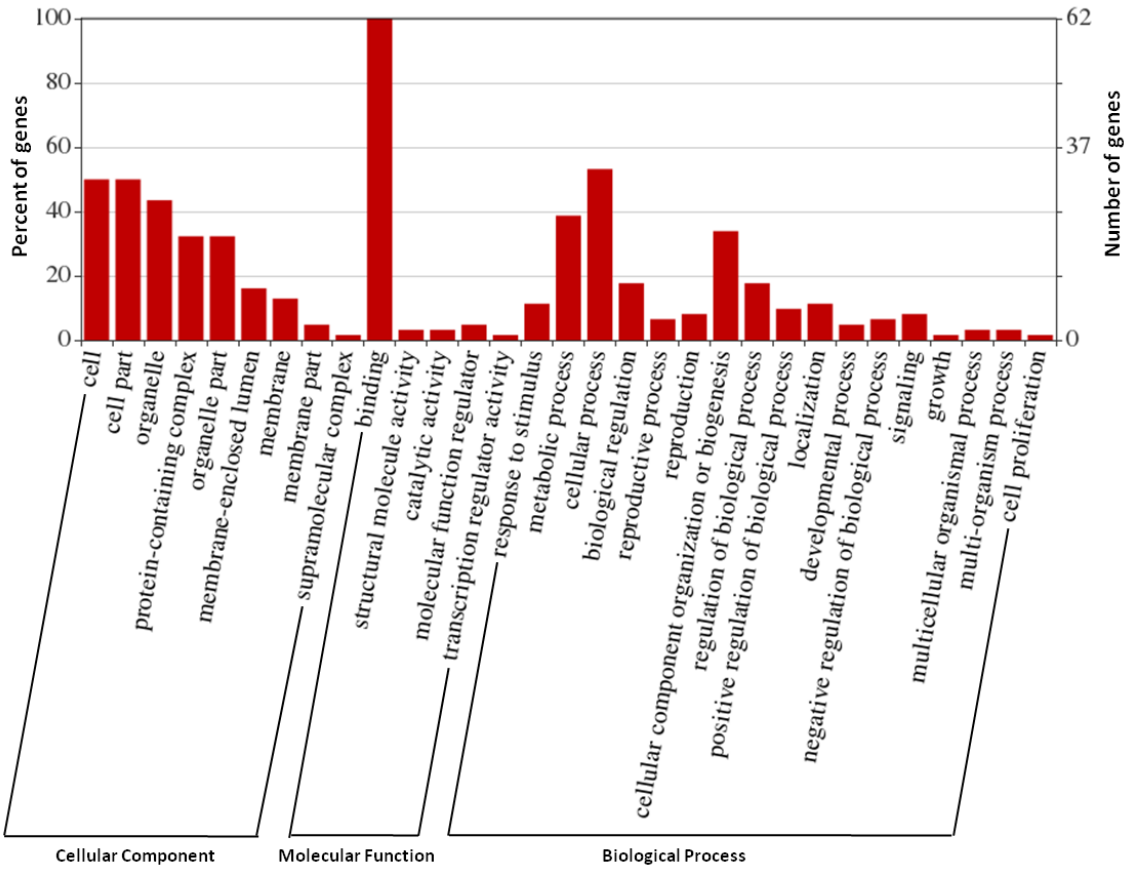

Figure 4. Gene Ontology (GO) categories of the 62 MfWD40s proteins.

DID88_005160, as well as those comprising the LisH domain, the Coatomer WD associated region, and the BOP1_N domain. Only 15 genes, with most of them belonging to class 1 of *MfWD40s*, were specifically less induced in the stage of germinating conidia and exhibited higher expression in the mycelium growing stages, either in dark or in light. We speculate that the specific expression of these genes might contribute to the light-dependent homeostasis in culture stages of *M. fructigena*. Overall, the diverse expression profiles of *MfWD40s* genes imply that certain genes might participate in specific growing and developmental stages.

WD40 repeats identification, phylogenetics analysis, and assignment of PGGs

The precise prediction of the number and coordinates of WD40 repeats in a WD40 protein is quite a challenging task [17, 26]. In our study, the WD40 repeats were predicted upon an average confidence score employing the WDSP software [17]. Using these thresholds, we retained 468 WD40 repeats among the 62 MfWD40s proteins.

An RAxML phylogenetical tree was constructed and the 468 WD40 repeats were classified into 6 (I-VI) distinct clusters (Figure 6). WD40 repeats are thought to diversify from intragenic duplication and recombination events during evolution [7]. It was therefore thought that the sequences and structures of WD40 repeats in the same protein are more similar to each other than those from different proteins [4]. In our dataset, WD40 repeats were classified independently

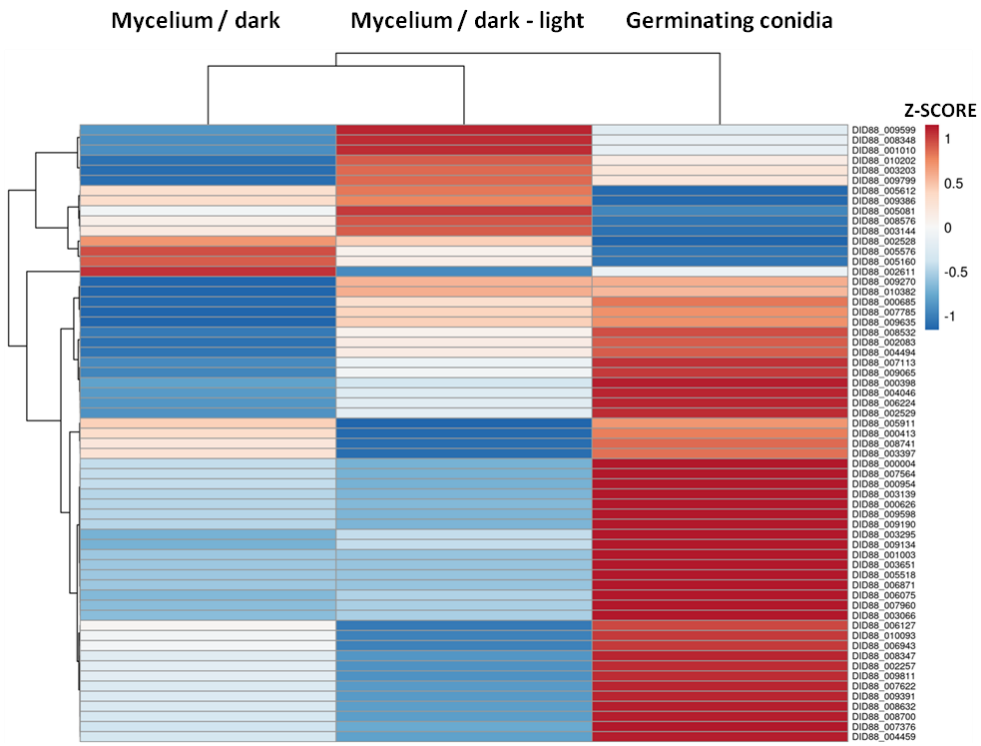

Figure 5. Hierarchical clustering heatmap of *MfWD40* genes based on expression profiles from RNAseq data. The colors in the heatmap represent the numbers of DESeq-normalized transcript counts (log10) in the three growing and developmental stages of *Monilinia fructigena*.

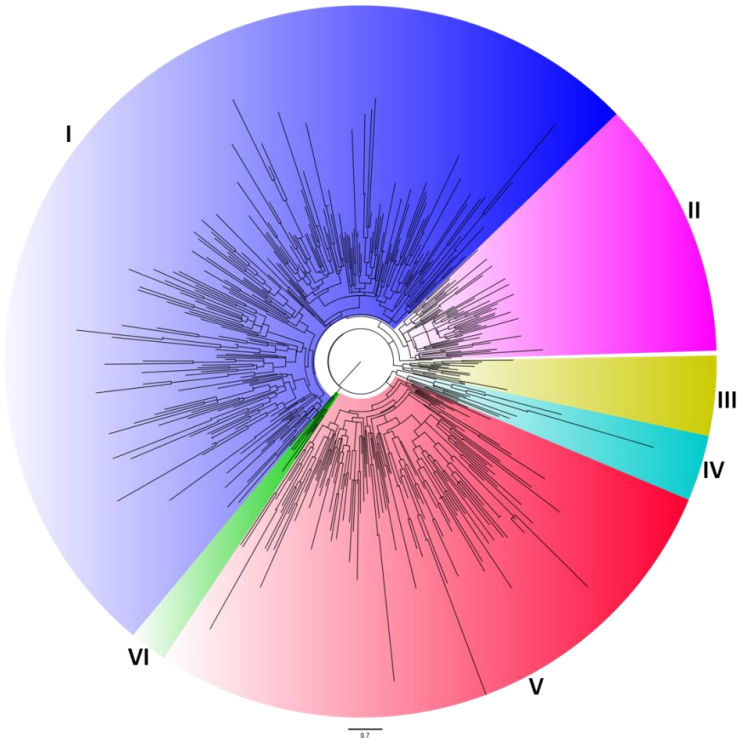

Figure 6. RAxML phylogeny of amino acid sequences from the 468 WD40 repeats. The tree was divided into six main distinct clusters (I-VI).

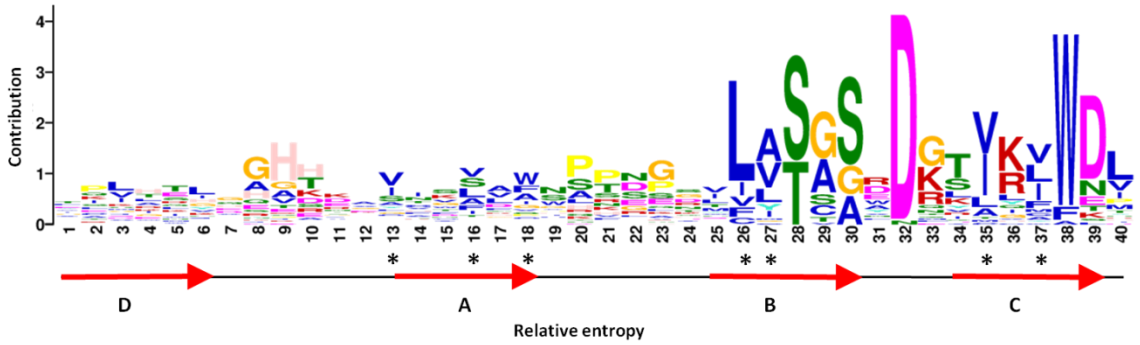

Figure 7. HMM motif logo of the 468 WD40 repeats. The residues marked by stars are conserved hydrophobic residues involved in WD40 β-propeller stabilization.

upon their origin domain which indicates that their expansion was happened at different evolution timescales among their proteins' domains. Taking into account that these 468 WD40 repeats show a rather low average identity (8.9%), we speculate that WD40 repeats with low sequence similarity arose at an early evolution stage. The HMM-based logo of the 468 WD40 repeats is shown in Figure 7.

Based on a MCL clustering approach these WD40 repeats were clustered in 35 PGGs, whereas 26 out of these repeats were clustered in 5 PGGs with each one containing from 4 up to 8 repeats. These 5 PGGs (*Mf*-WD40 PGGs) were further proceeded for evolutionary analysis (Table 1).

Evolutionary signatures using maximum likelihood approaches

In order to investigate the selective evolutionary signatures upon WD40 repeats, which might also contribute to the overall divergence of MfWD40s, the ratios of non-synonymous (Ka) versus synonymous (Ks) substitution rate were estimated for all sequences of WD40 repeats in each *Mf*-WD40 PGG. All the WD40 repeats in each PGG had a valid structure in means of an amino acid length up to 40 residues. In all sequence pairs no saturation effects based on nucleotide substitutions were observed in terms of the average codon-based evolutionary divergences of synonymous (dS) mutations per synonymous sites.

The selective evolutionary pressures acting on all the 5 *Mf*-WD40 PGGs lineages were initially investigated using the CODEML program. Statistically significant evidence of positive selection was not detected among the branches of the PGGs evolutionary trees, since the ω (Ka/Ks) values were < 1 in all the *Mf*-WD40 PGGs datasets. We hypothesize that these WD40 repeats might have undergone a purifying selection leading to a limited functional divergence, putatively after the duplication of the *MfWD40s* genes as it was also reported in the foxtail millet [5]. Furthermore, our results imply that recent episodes of purifying selection were acting on these WD40 repeats, overlapping ancient intra-gene duplications events during the expansion of WD40 domains in *M. fructigena*. This is in agreement with findings by Hu *et al*. [2] that the repertoire of WD40 domains might contain both ancient and recent duplication events, whereas sequence identity between repeats within the same domain can be related to the evolution time of intra-gene duplication events.

Table 1. CODEMLSITES analysis of *Mf*-WD40 PGGs.

Mf-WD40 PGGs	Number of clustered WD40 repeats	2ΔL (models M7 /M8)	Number of sites	Number of positively selected sites	Positively selected sites with posterior Bayesian probabilities
Mf-WD40-1	8	0	111	-	-
Mf-WD40-2	6	0.82	114	10	1 K, 2 N, 4 R, 7 P, 10 V, 12 A,15 F, 17 P, 19 G, 37 V
Mf-WD40-3	4	0.26	123	9	2 I, 4 V, 6 S, 7 Q, 18 R, 20 N, 21 Q, 23 Y, 35 S
Mf-WD40-4	4	0	117	18	3 M, 4 Q, 5 H, 7 K, 10 T, 14 L, 17 A, 19 S, 20 F, 22 G, 23 M, 24 S, 26 F, 29 G, 30 M, 35 V. 36 V, 39 Q
Mf-WD40-5	4	0.24	117	-	-

The CODEMLSITES program was also employed in order to evaluate the selective pressures acting among amino acid residues across the *Mf*-WD40 PGGs (Table 1). Various models were tested across the *Mf*-WD40 PGGs dataset alignments using LRTs. Extensive signs of positive selection with high posterior Bayesian probabilities were found in 3 out of the 5 PGGs acting widely upon amino acid residues across the respective tree branches. The 37 predicted positively selected sites along with the relevant CODEMLSITES statistics are showed in Table 1. Several of these positively selected sites were similar with the predicted hotspot residues based on the WDSP repeats predictions, as well as with the conserved hydrophobic amino acids involved in WD40 β-propeller stabilization (Figure 7). These amino acids sites on the surface of WD40 repeats extensively take part in assembling proteins or nucleic acids into functional complexes serving as scaffolds in various cellular networks [4, 8].

Conclusion

WD40 proteins play key roles in various biological processes. Although these proteins were extensively characterized in other organisms, little investigations have been reported in filamentous fungal species. Taking advantage of the recently available genome sequences of *M. fructigena*, we performed a genome-wide analysis of these proteins. Sixty-two WD40 proteins were identified and accordingly to their domain architectures these proteins were structurally classified in 17 classes. Segmental duplications of these genes seem not to play a major role for the expansion of this gene family in *M. fructigena*. We further explored the potential biological roles of these genes by decipher their RNA-seq expression profiles. In parallel, we systematically predicted their WD40 repeats and we comprehensively investigated the selective evolutionary signatures acting upon them. These findings increase our knowledge of the *M. fructigena* WD40 protein family and decipher its landscape for the first time in phytopathogenic fungi.

Acknowledgments

This work has been supported by Hellenic Agricultural Organization (HAO) - Demeter, Greece.

References

1. He S, Tong X, Han M, Hu H, Dai F. 2018. Genome-wide identification and characterization of WD40 protein genes in the Silkworm, *Bombyx mori*. Int J Mol Sci. 19:527.
2. Hu XJ, Li T, Wang Y, Xiong Y, Wu XH, Zhang DL, Ye ZQ, Wu YD. 2017. Prokaryotic and highly repetitive WD40 proteins: a systematic study. Sci Rep. 7:10585.

3. Stirnimann CU, Petsalaki E, Russell RB, Müller CW. 2010. WD40 proteins propel cellular networks. Trends Biochem Sci. 35:565–574.
4. Xu C, Min J. 2011. Structure and function of WD40 domain proteins. Protein Cell. 2:202–214.
5. Mishra AK, Muthamilarasan M, Khan Y, Parida SK, Prasad M. 2014. Genome-wide investigation and expression of WD40 protein family in the model plant foxtail millet (*Setaria italica* L.). PLoS One. 9:e86852.
6. Schapira M, Tyers M, Torrent M, Arrowsmith C. 2017. WD40 repeat domain proteins: A novel target class? Nat Rev Drug Discov. 16:10.1038.
7. Andrade MA, Perez-Iratxeta C, Ponting CP. 2001. Protein repeats: Structures, functions and evolution. J Struct Biol. 134:117–131.
8. Smith TF. 2008. Diversity of WD-repeat proteins. Subcell Biochem. 48:20–30.
9. Ouyang Y, Huang X, Lu Z, Yao J. 2012. Genomic survey, expression profile and co-expression network analysis of OsWD40 family in rice. BMC Genom. 13:100.
10. Li Q, Zhao P, Li J, Zhang C, Wang L, Ren Z. 2014. Genome-wide analysis of the WD-repeat protein family in cucumber and *Arabidopsis*. Mol Genet Genom. 289:103–124.
11. Rivera Y, Zeller K, Srivastava S, Sutherland J, Galvez M, Nakhla M. Poniatowska A, Schnabel G, Sundin G, Abad ZG. 2018. Draft genome resources for the phytopathogenic fungi *Monilinia fructicola*, *M. fructigena*, *M. polystroma*, and *M. laxa*, the causal agents of brown rot. Phytopathology. 108:1141-1142.
12. Naranjo-Ortíz MA, Rodríguez-Píres S, Torres R, De Cal A, Usall J, Gabaldón T. 2018. Genome sequence of the brown rot fungal pathogen *Monilinia laxa*. Genome Announc. 6:e00214-18.
13. Ortega SF, Del Pilar Bustos López M, Nari L, Boonham N, Gullino ML, Spadaro D. 2019. Rapid detection of *Monilinia fructicola* and *Monilinia laxa* on peach and nectarine using loop-mediated isothermal amplification. Plant Dis. 103:2305-2314.
14. Landi L, De Miccolis Angelini RM, Pollastro S, Abate D, Faretra F, Romanazzi G. 2018. Genome sequence of the brown rot fungal pathogen *Monilinia fructigena*. BMC Res Notes. 11:758.
15. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Res. 39:29-37.
16. Jones P, Binns D, Chang HY, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30:1236-1240.
17. Wang Y, Hu XJ, Zou XD, Wu XH, Ye ZQ, Wu YD. 2015. WDSPdb: A database for WD40-repeat proteins. Nucleic Acids Res. 43:339-344.
18. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28:1647-1649.
19. Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, Lahrmann U, Zhao Q, Zheng Y, Zhao Y, Xue Y, Ren J. 2015. IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics. 31:3359-3361.
20. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40:e49.
21. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639-1645.
22. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36:3420-3435.
23. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29:15-21.
24. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome Biol. 11:R106.
25. Metsalu T, Vilo J. 2015. Clustvis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 43:566-570.
26. Wang Y, Jiang F, Zhuo Z, Wu XH, Wu YD. 2013. A method for WD40 repeat detection and secondary structure prediction. PLoS One. 8:e65705.
27. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37:W202-W208.
28. Zambounis A, Ganopoulos I, Kalivas A, Tsaftaris A, Madesis P. 2016a. Identification and evidence of positive selection upon resistance gene analogs in cotton (*Gossypium hirsutum* L.). Physiol Mol Biol Plants. 22:415-421.
29. Zambounis A, Psomopoulos F, Ganopoulos I, Avramidou E, Aravanopoulos FA, Tsaftaris A, Madesis P. 2016b. In silico analysis of the LRR receptor-like serine threonine kinases subfamily in *Morus notabilis*. Plant Omics. 9:319-326.
30. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586-1591.
31. Zambounis A, Ganopoulos I, Avramidou E, Aravanopoulos FA, Tsaftaris A, Madesis P. 2016. Evidence of extensive positive selection acting on cherry (*Prunus avium* L.) resistance gene analogs (RGAs). Aust J Crop Sci. 10:1324-1329.
32. Zambounis A, Elias M, Sterck L, Maumus F, Gachon CM. 2012. Highly dynamic exon shuffling in candidate pathogen receptors… What if brown algae were capable of adaptive immunity? Mol Biol Evol. 29:1263-1276.
33. Lynn DJ, Lloyd AT, Fares MA, O'Farrelly C. 2004. Evidence of positively selected sites in mammalian α-defensins. Mol Biol Evol. 21:819-827.
34. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 28:2731-2739.
35. Shor E, Chauhan NA. 2015. Case for two-component signaling systems as antifungal drug targets. PLoS Pathog. 11:e1004632.
36. Sharma M, Pandey G. 2016. Expansion and function of repeat domain proteins during stress and development in plants. Front Plant Sci. 6:10.3389.
37. Bian S, Li X, Mainali H, Chen L, Dhaubhadel S. 2017. Genome-wide analysis of DWD proteins in soybean (*Glycine max*): Significance of Gm08DWD and GmMYB176 interaction in isoflavonoid biosynthesis. PLoS One. 12:e0178947.
38. Zou XD, Hu XJ, Ma J, Li T, Ye ZQ, Wu YD. 2016. Genome-wide analysis of WD40 protein family in human. Sci Rep. 6:39262.