Hu Y and Shu B (2021)
Notulae Botanicae Horti Agrobotanici Cluj-Napoca
Volume 49, Issue 2, Article number 12323
DOI:10.15835/nbha49212323
Research Article

Identifying strawberry *Whirly* family transcription factors and their expressions in response to crown rot

Yuyang HU, Bo SHU*

*Yangtze University, College of Horticulture and Gardening, 1 Nanhuan road, Jingzhou, China; 598944714@qq.com; bshbest@163.com (*corresponding author)*

Abstract

Crown rot is one of the most destructive diseases of cultivated strawberry. The correlation between *Whirly* family transcription factors, the one class of known resistance genes, and strawberry crown rot resistance has not been studied. In this study, the *Whirlys* of *Fragaria × ananassa, F. iinumae*, *F. vesca*, *F. viridis* and *F. nilgerrensis* were characterized by searching the strawberry genome database and analyzing the presence of Whirly domains. Five *Fa WHYs*, two *Fi WHYs*, three *Fn WHYs*, two *Fvi WHYs* and four *Fv WHYs* were identified from their respective genome. Two gene clusters with segmental duplications were obtained from the gene cluster analysis with two and three *Fa WHYs*, and three *Fa WHYs* showed syntenic relationships with *At WHYs of Arabidopsis thaliana*. *Fi WHY1, Fv WHY2* and *Fvi WHY1* showed syntenic relationships with *Fa WHY1* and *Fa WHY2*. At the same time, *Fi WHY2*, *Fv WHY3*, *Fvi WHY2* and *Fn WHY3* exhibited similar syntenic relationships with *Fa WHY4* and *Fa WHY5*. In addition, *Fn WHY1* and *Fn WHY2* corresponded to both *Fa WHY1* and *Fa WHY2*. Gene expression analysis revealed that five *Fa WHYs* were expressed in crowns, and the regulation of *Fa WHYs* was always consistent with the *cis*-elements in their promoters. All of them were downregulated by crown rot infected. Together, these results provided a basis for further functional studies of the FaWHYs proteins and their responses to crown rot.

Keywords: crown rot; gene structure; phylogenetic analysis; strawberry; transcriptional expression

Introduction

Strawberry, the small crop producing much appreciated fruits with unique flavour and high nutritious qualities, is of great importance throughout the world, but its productivity and quality are seriously limited by crown rot (Mangandi *et al.*, 2015; Anciro *et al.*, 2018). Crown rot occurs in the crown root neck, which is manifested as a short plant. After infection, the crown root neck produces red streaks, and then rapidly expands to dark, sunken spots, and finally the whole plant wilts and withers, which is a devastating disease of strawberry. Plants have a small family of single-stranded DNA (ssDNA) binding proteins called Whirly that are involved in the control of defence gene expression (Desveaux *et al.*, 2004; Isemer *et al.*, 2012). The Whirly transcription factor family can participate in plant resistance to adversity through the transduction of disease resistance signals and the regulation of hypersensitive responses (Yao *et al.*, 2008).

Whirly protein is a plant-specific protein, which is mainly distributed in chloroplasts, mitochondria and cell nuclei. In most species, this family contains only two members in most plant species, and three in

Received: 29 Mar 2021. Received in revised form: 23 Apr 2021. Accepted: 12 May 2021. Published online: 13 May 2021.
From Volume 49, Issue 1, 2021, Notulae Botanicae Horti Agrobotanici Cluj-Napoca journal uses article numbers in place of the traditional method of continuous pagination through the volume. The journal will continue to appear quarterly, as before, with four annual numbers.

Arabidopsis thaliana and a few species. There are currently many *Whirly* studies on *A. thaliana*. *AtWHY1* adapts to adversity and immune response through redox regulation of chloroplast components in retrograde signal (Lepage *et al.*, 2013; Foyer *et al.*, 2013). The *AtWHY1* genes are also involved in the salicylic acid (SA)-dependent disease resistance and SA-induced expression of the systemic acquired resistance response gene. *AtWHY1* is required for both full basal and specific disease resistance responses (Desveaux *et al.*, 2005). *WHYs* also function in response to microbe interactions, such as pathogen infection. A large set of defence genes with various biochemical functions, including pathogenesis-related (*PR*) genes, are activated or repressed in response to pathogen attack. Studies have found that the tagged *AtWHY1* is translocated from the plastid to the nucleus, which affects expression of target genes such as *PR1* (Isemer *et al.*, 2012). For instance, a transcriptional activator located in the nucleus of potato *PBF2* (*StWHY1*) was identified, which can combine with the elicitor response element on the promoter of the disease resistance gene *PR1*, activate the expression of *PR1*, and participate in the pathogen response process (Desveaux *et al.*, 2000). And overexpression of tomato *Whilry* gene in transgenic tobacco resulted in *Pseudomonas solanacearum* resistance (Zhao *et al.*, 2018). *Whirly* genes (*MeWHYs*) in cassava, *MeWRKY75* and *MeWHYs* confer improved disease resistance against cassava bacterial blight through forming an interacting complex of *MeWRKY75-MeWHY1/2/3* and transcriptional module of *MeWRKY75-MeWHY3* (Liu *et al.*, 2018).

The *WHY* family has been characterized in several plants, including potato (Desveaux *et al.*, 2000), *A. thaliana* (Desveaux *et al.*, 2004), wheat (Chitnis *et al.*, 2014), tomato (Zhao *et al.*, 2018), cassava (Liu *et al.*, 2018), tobacco (Zhao *et al.*, 2018), chili (Lu *et al.*, 2019), soybeans (Li *et al.*, 2019), and rice (He *et al.*, 2020). *WHYs* play important roles in biotic stress and may function in the strawberry biotic stress response. However, strawberry-specific *WHY* studies are lacking. In the present study, the strawberry *Whirly* family transcription factors members were identified via bioinformatics tools, and their expression patterns in response to biotic stress were characterized. This study provides basic information on the protein structures, subfamily divisions, chromosome localization in the strawberry genome, and expression patterns of the Whirly proteins response to crown rot.

Materials and Methods

Fa WHY searching and characteristics

The complete genome assembly of strawberry *Fragaria × ananassa* 'Camarosa', *F. iinumae*, *F. vesca*, *F. viridis* and *F. nilgerrensis* were downloaded from the strawberry Genome Database (*https://www.rosaceae.org*). Three *A. thaliana Whirly* protein sequences (AtWHYs) were obtained from the *A. thaliana* Information Resource (TAIR) (*https://www.arabidopsis.org/*) (Cappadocia *et al.*, 2013). *F. ananassa* Whirly (FaWHYs), *F. iinumae* Whirly (FiWHYs), *F. vesca* Whirly (FvWHYs), *F. viridis* Whirly (FviWHYs) and *F. nilgerrensis* Whirly (FnWHYs) sequences were selected by comparison to the *A. thaliana* query sequences *via* BLASTP, respectively. Redundant proteins were manually deleted based on their E-values. The molecular weight (MW) and isoelectric point (pI) of the candidate protein sequences were determined by ExPASy (https://web.expasy.org/compute/pi) (Gasteiger *et al.*, 2003). Conserved Whirly domains were verified for all potential Whirly proteins using the NCBI Batch CD-Search program (*https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi*).

Sequence alignment and phylogenetic analysis

The full-length *Whirly* protein sequences from *A. thaliana* and strawberries were aligned *via* muscle in MEGA version 7.0, with default parameters (Edgar, 2004; Kumar *et al.*, 2013). A neighbour-joining (NJ) tree was also generated with bootstrapping (1000 replicates). The phylogenetic relationships among the five kinds of strawberries *WHYs* and *A. thaliana WHYs* were estimated.

Conserved motifs and gene structure analysis

Motif analysis was conducted on the MEME website (*http://meme-suite.org/tools/meme*) to identify conserved motifs with the following optimized parameters: zero or one occurrence per sequence, a maximum of 10 motifs and an optimum motif width between 6 and 50 residues. The default settings were used for all other parameters. Comparing the coding sequence with the corresponding genome sequence, the structure of *Fa WHYs* was determined using TBtools (Chen *et al.*, 2020).

The Strawberries Generic Feature Format (GFF) files were downloaded from the strawberry Genome Database and used to elucidate the structure information of the *Whirly* gene. An illustration of the FaWHYs protein motifs, conserved domain, gene structures and a phylogenetic tree was also constructed in TBtools (Chen *et al.*, 2020).

Chromosomal distribution, gene duplication and collinearity

The chromosome locations of the candidate strawberry *Whirly* genes were analyzed from the GFF information and visualized by TBtools (Chen *et al.*, 2020). Gene duplication events of the *Fa WHYs* and collinearity between the *A. thaliana* Whirly protein sequences and five kinds of strawberries Whirly protein sequences were investigated by MCScanX (Wang *et al.*, 2012). The results were visualized in TBtools (Chen *et al.*, 2020).

Fa WHY expression in response to biotic stress

A single factor experiment was performed using different treatments causing inoculation with *Colletotrichum siamense SCR-7*. The healthy and consistent strawberry seedlings were divided into treatment group (JZ) and control group (CK) with 12 pots each. The JZ treatment group was inoculated with *C. siamense SCR-7* using a sterilized needle as described in Li *et al.* (2014), while the CK group was inoculated with non-toxic medium using the same method. The seedlings were grown with or without *C. siamense SCR-7* and 0 or 6 days after vaccination, resulting in four treatment groups: inoculation with no pathogens after 0 day (0DCK), inoculation with no pathogens after 6 days (6DCK), inoculation with pathogens after 0 day (0DJZ) and inoculation with pathogens after 6 days (6DJZ). Each treatment had three biological replicates. The photo of strawberry inoculation with pathogens after 6 days and CK was shown in Figure S1.

Transcriptomic data of seedling crowns from the four treatments were analysed as described by Shu *et al.* (2016). Twelve libraries of seedling crowns were sequenced using the Illumina HiSeq 2000 system. Reads that contained adapters, more than 10% unknown nucleotides, and more than 50% bases with a quality value ≤5 were removed to obtain uncontaminated sequences based on the raw data. Uncontaminated sequences were mapped to the genome of *F. ananassa* 'Camarosa' (v1.0.a1) for annotation. The transcriptomic data were uploaded to the NCBI Sequence Read Archive as PRJNA715088. Gene expression was analysed based on the transcriptomic data, where the transcriptional abundance of *Fa WHY* was calculated as fragments per kilobase of exon model per million mapped reads (FPKM) using the Cufflinks package cuffdiff version 2.2.1. The FPKM value of 0DCK was considered the relevant control. Heat maps were created using TBtools software based on the transformed data of $\log_2$ (FPKM+1) values (Chen *et al.* 2020).

Cis-acting elements of the Fa WHYs

The 2,000 bp sequences upstream of the transcription initiation site of the candidate genes were extracted from the strawberry genome sequences. The PlantCARE software (*http://bioinformatics.psb.ugent.be/webtools/plantcare/html*) was used to search for *cis*-acting elements (Rombauts *et al.*, 1999), and the results were visualized in TBtools (Chen *et al.*, 2020)

Results

Identification, characteristics, and chromosomal distribution of the WHYs in strawberry genome

Five *FaWHYs* were identified in the *F. ananassa* genome after searching for Whirly domain sequences. The *FaWHYs* were named according to their positions on each chromosome. The FaWHY protein lengths ranged from 164 aa (FaWHY3) to 298 aa (FaWHY5), the pI was ranged from 8.53 (FaWHY3) to 9.59 (FaWHY1), and the molecular weight ranged from 18.46 kDa (FaWHY3) to 33.10 kDa (FaWHY5). For the other four diploid strawberry, the Whirly protein lengths ranged from 112 aa (FnWHY1 and FnWHY2) to 273 aa (FiWHY2), the pI was ranged from 9.24 (FiWHY1) to 9.96 (FvWHY1), and the molecular weight ranged from 12.41 kDa (FnWHY1 and FnWHY2) to 30.08kDa (FiWHY2) (Table 1).

The *FaWHYs* were mapped on the strawberry chromosomes, and five *FaWHYs* were located on chromosomes Fvb1-4, Fvb2-2, Fvb4-2, Fvb4-3, and Fvb4-4, with one gene on each chromosome (Figure 1). The *WHYs* of the other diploid strawberries were all distributed on second and fourth chromosome, respectively (Figure S2). Gene duplication and divergence are important in gene family expansion and in the evolution of novel functions. Two gene clusters with segmental duplications were obtained from the gene cluster analysis. One cluster contained two genes (*FaWHY1* and *2*), whereas the other cluster had three genes (*FaWHY3*, *4* and *5*). There were no tandem duplications in the *FaWHY* genes (Figure 1).

Table 1. Basic information on Whirly family transcription factors identified in the strawberry genome database

Gene name	Original ID	Length (aa)	pI	MW (kDa)
FaWHY1	maker-Fvb1-4-augustus-gene-207.32-mRNA-1	200	9.59	25.41
FaWHY2	snap_masked-Fvb2-2-processed-gene-230.10-mRNA-1	261	9.51	28.72
FaWHY3	maker-Fvb4-2-augustus-gene-26.47-mRNA-1	164	8.53	18.46
FaWHY4	maker-Fvb4-3-augustus-gene-29.63-mRNA-1	271	9.53	29.77
FaWHY5	maker-Fvb4-4-augustus-gene-27.61-mRNA-1	298	9.44	33.10
FiWHY1	evm.model.scaf_51.28	242	9.24	26.77
FiWHY2	evm.model.scaf_10.712	273	9.57	30.08
FnWHY1	evm.model.ctg76.9	112	9.74	12.41
FnWHY2	evm.model.ctg77.22	112	9.74	12.41
FnWHY3	evm.model.ctg49.187	271	9.53	29.85
FviWHY1	evm.model.ctg86.54	159	9.30	17.55
FviWHY2	evm.model.ctg68.71	271	9.61	30.00
FvWHY1	FvH4_2g00250.t2	177	9.96	19.57
FvWHY2	FvH4_2g00250.t1	231	9.66	25.38
FvWHY3	FvH4_4g31850.t1	271	9.53	29.77
FvWHY4	FvH4_4g31850.t2	243	9.67	26.41

Notes: aa - amino acids; pI - isoelectric point; MW - molecular weight

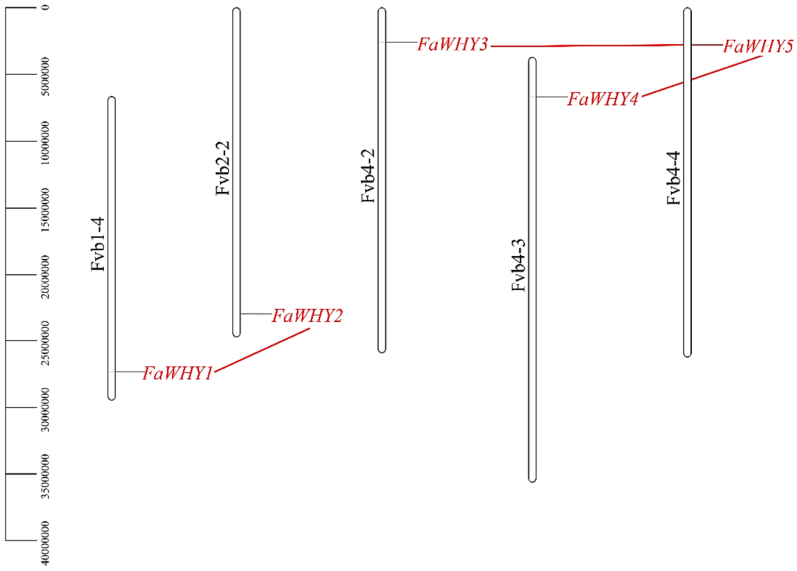

Figure 1. Chromosomal distribution of the *Fa WHY* genes
Chromosome numbers are provided at the top of each chromosome together with the approximate size. The *Fa WHYs* were named *Fa WHY1* to *Fa WHY5* based on their order on the chromosomes. Red lines mark the gene clusters with tandem duplications

Synteny analysis of the Whirlys in A. thaliana and five strawberries

To further investigate the phylogenetic patterns of the *Fa WHYs*, a comparative syntenic map of five strawberries and *A. thaliana* was constructed. *Fa WHY2 and Fa WHY5* showed syntenic relationships with the *At WHY2*, which is located on the *Fvb2-2 and Fvb4-3* chromosome. While *Fa WHY4* showed syntenic relationships with both *At WHY1* and *At WHY3*, indicating that it may have played an important role in the evolution of the *Whirly* family (Figure 2). *F. iinumae*, *F. vesca*, *F. viridis* and *F. nilgerrensis* showed similar collinearity with *F. ananassa*. *F. iinumae*, *F. vesca* and *F. viridis* on the second chromosome: chr2, fvb2, fvir2, which have *Fi WHY1, Fv WHY2, Fvi WHY1* showed syntenic relationships with *Fa WHY1* and *Fa WHY2*. At the same time, the genes on four fourth chromosome: *Fi WHY2*, *Fv WHY3*, *Fvi WHY2 Fn WHY3* had similar syntenic relationships with *Fa WHY4* and *Fa WHY5*. In addition, genes *Fn WHY1* and *Fn WHY2* on the second chromosome of *F. nilgerrensis* corresponded to both *Fa WHY1* and *Fa WHY2*. This showed that in the process of strawberry evolution, genes have been duplicated, and the *Whirly* gene is highly conserved (Table S1).

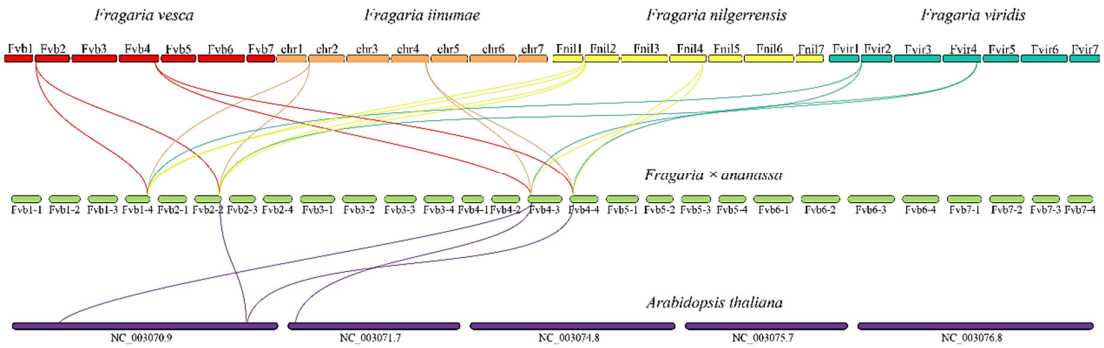

Figure 2 Synteny analysis of *Fa WHYs* between *Arabidopsis thaliana* and five kind of strawberries
Purple lines in the background indicate the collinear blocks within *Fragaria× ananassa* and *A. thaliana* genomes. And the red, orange, yellow and green lines indicate the collinear blocks within four diploid strawberries and *F. ananassa* genomes.

Phylogenetic, exon-intron structure, conservative domains and motifs analysis of the Fa WHYs

The phylogenetic relationships of five different kinds of strawberries and AtWHYs were analysed by a phylogenetic tree of the protein sequence alignment. As shown in Figure 3, the strawberry Whirlys and AtWHYs clustered into three major groups. Groups I to III have 9, 2 and 8 members, respectively, and differences were observed between *A. thaliana* and strawberries. The number of strawberries and *A. thaliana* Whirlys in groups I and III was nearly equal and groups II only have two AtWHYs (Figure 3).

The *Fa WHYs* in different groups were characterized according to their Whirly domain numbers and exon–intron structures. Motifs 1 and 2 composed the whirly domain, and all 5 *Fa WHYs* had one characteristic domains. The number of introns varied from 4 (*Fa WHY3*) to 8 (*Fa WHY2*). *Fa WHYs* in group I possessed motifs 1, 2, 3, 4, 5 and 8. Group II contained no *Fa WHY* member. *Fa WHY4* and *Fa WHY5* in groups III contained 8 same motifs, and the differences were in the exon-intron structure, while the *Fa WHY3* lacked untranslated region (5' UTR) (Figure 4; Figure S3).

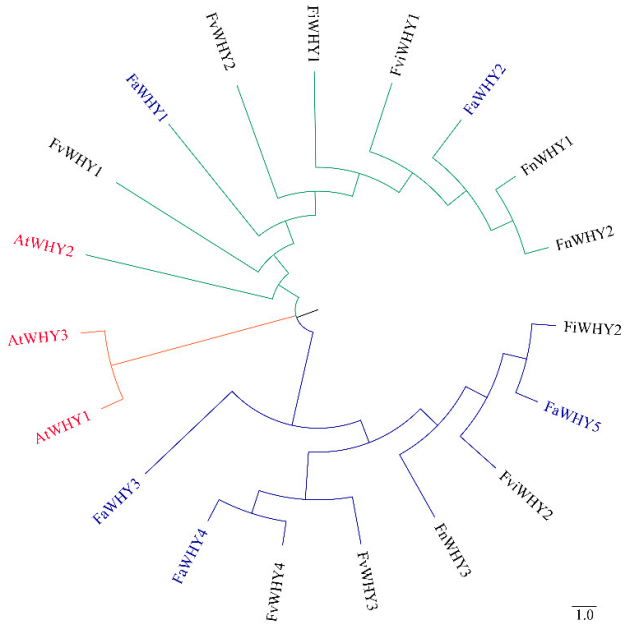

Figure 3 Phylogenetic tree of the Whirlys based on an alignment of strawberries and *Arabidopsis thaliana* proteins

The phylogenetic tree was constructed using the neighbour-joining method implemented in MEGA 7.0. Reliability of the predicted tree was tested using bootstrapping with 1,000 replicates. Branch lines with different colours represent different Whirly groups.

Analysis of Fa WHY expression and the cis-elements in Fa WHY promoters

Gene expression analysis revealed that 5 *Fa WHYs* were variably expressed in the crown roots and all of them were downregulated by crown rot infected (Figure 5). Several *cis*-elements, including 'hormone-responsive', were identified in the upstream regulatory regions (promoters) of the *Fa WHYs*. The *cis*-elements for 'defense and stress responsive' and 'salicylic acid responsive' are responsible for the plant response to pathogen infection. 'Defense and stress responsive' *cis*-elements were identified in promoters of four *Fa WHYs* (*Fa WHY1*, *3*, *4* and *5*) and one 'salicylic acid responsive' *cis*-elements were identified in promoters of *Fa WHY1*. There was only one 'defence and stress response' *cis* elements of the same species found in *Fa WHY1*, *3*, *4*, and *5* indicates that they have the same ability to regulate stress. The promoters of *Fa WHY1* contained 'defence and stress responsive' and 'salicylic acid responsive' *cis*-elements, suggesting that this gene may be regulated by SA-dependent disease resistance responsive and microbial interactions (Figure 6).

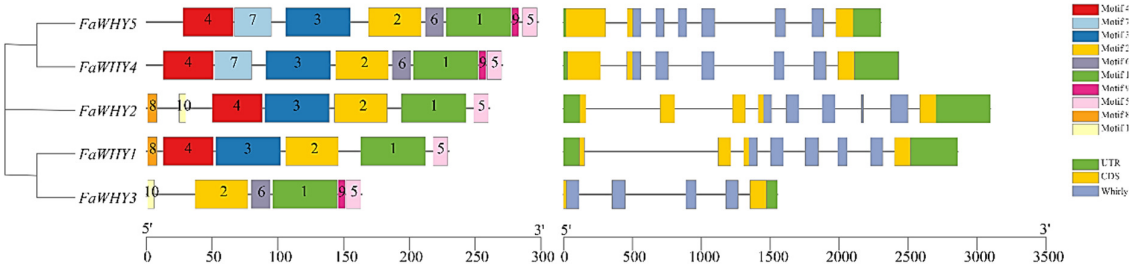

Figure 4 Phylogenetic tree of deduced FaWHY proteins associated with the motif composition and exon-intron composition of *Fa WHY* genes

The phylogenetic tree was constructed using the neighbour-joining method (left-hand side of the figure). Reliability of the predicted tree was tested using bootstrapping with 1,000 replicates. The motif composition related to each FaWHY protein is displayed in the middle of the figure. The motifs, numbered 1-10, are displayed in different colored boxes. The information for each motif is provided in Figure S3.

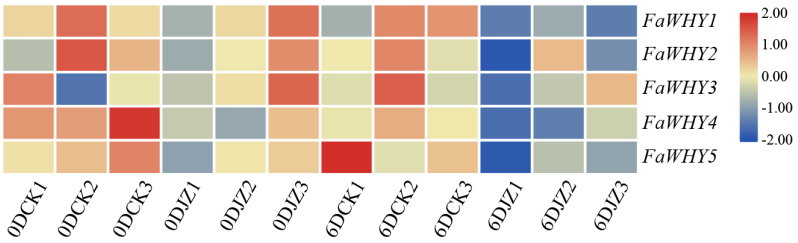

Figure 5. Expression profiles of the *Fa WHY* genes responding to crown rot

Red and blue indicate up- and downregulated genes compared to the relevant control.

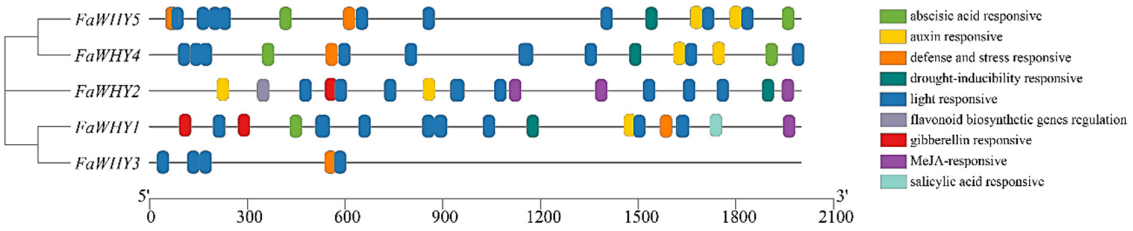

Figure 6 Predicted *cis*-elements in the promoters of the *Fa WHYs*

The 'defense and stress responsive' and 'salicylic acid responsive' *cis*-elements are indicated by orange and nattier blue, respectively.

Discussion

Plant *Whirlys* are a multigene family; there are 3 *Whirlys* in *A. thaliana*, 2 *Whirlys* in potato (Maréchal *et al.*, 2008), 2 *Whirlys* in tomato (Akbudak *et al.*, 2019), in addition, the amino acid sequence of the protein can also be found in dozens of plants such as soybean, wheat, rice, corn, and lily (Kong *et al.*, 2012). Although the size of *F. ananassa* genome (780 Mb) (*https://www.rosaceae.org*) was larger than the genome of *A. thaliana* (125 Mb) (*https://www.arabidopsis.org*), the total number of *Fa WHY* genes was similar to the number of these genes in *A. thaliana*. Five *Fa WHYs* were identified on five chromosomes in the *F. ananassa* genome. Two gene clusters with segmental duplications were obtained from the gene cluster analysis with two and three *Fa WHYs*, these clusters likely arose from segmental duplications, suggesting gene family expansion during evolution. Among them, three *Fa WHYs* (*Fa WHY2*, *4* and *5*) showed syntenic relationships with the

AtWHYs. It can be speculated that *FaWHY1* and *FaWHY3* may be new genes produced during plant evolution. The syntenic relationships of *whirly* between other four kinds of strawberries and the *F. ananassa* showed that the *F. ananassa* had gene duplication during the evolution process. However, *FaWHY3* did not find collinearity in the four diploid strawberries, but *FaWHY3* had segmental duplications between *FaWHY4* and *FaWHY5*, which may be caused by chromosomal variation during evolution. These results provided insights that would assist in the prediction of the evolution of *FaWHYs*. *F. ananassa* is a common allopolyploid, and its parental ancestors still exist. Recently, the possibility of ancestral parents of octoploid strawberry were researched (Edger *et al.*, 2019; Liston *et al.*, 2020; Edger *et al.*, 2020), through the syntenic relationship between octoploid strawberry and its possible ancestor parents to confirm its evolution and genetic characteristics at the early stage of formation. In our study, we found that *FaWHY1* and *FaWHY2* have syntenic relationship with *FvWHY1*, *FiWHY1* and *FviWHY1*, but *FnWHY1* and *FnWHY2* have syntenic relationship with both *FaWHY1* and *FaWHY2*. The *FnWHYs* have not increased exponentially but has decreased exponentially in the syntenic relationship analysis, which provided evidence for other researchers' study that *F. nilgerrensis* is not the ancestor of *F. ananassa* (Feng *et al.*, 2021).

The whirly domain was highly conserved during the evolution process, which provided information for the prediction of the structure and function of the *FaWHYs* gene. The whirly protein (WHYs) have three domains: Whirl*y* domain, an N-terminal domain and C-terminal variable region. The Whirly domain is the most important domain, which has the ability to bind to ssDNA and the KGKAAL, YDW and K amino acid residues in this region may play a role as important sites of WHY protein (Desveaux *et al.*, 2005). The N-terminal domain may have chloroplast or mitochondrial signal peptides and transcription activation regions; the C-terminal variable region has a self-regulating region, which can regulate ssDNA binding activity (Desveaux *et al.*, 2002). All of the FaWHYs clustered into two major groups, with distinct protein domains, motifs and sequences. Whirly domain is the most conserved region in whirly protein, Motifs 1 and 2 composed the whirly domain, and all 5 FaWHYs had one characteristic domain.

The phylogenetic tree generated from the protein sequence alignment of strawberries and *A. thaliana* segregated the 5 FaWHYs into two large groups. Group members shared similar protein sequence lengths, motif compositions and exon–intron structures, suggesting a close relationship. Thus, FaWHY1, FaWHY2 and their homolog AtWHY2 in the same branch may play similar roles in plant–microbe interactions and biotic stress responses. AtWHY2 clustered with FaWHY1, FaWHY2 speculated that they may be located in mitochondrial cells and participate in the transmission of disease resistance signals (Cappadocia *et al.*, 2012). The phylogenetic tree predicted that the FaWHYs are involved in pathogen infection interactions, but this hypothesis requires verification in future studies. In addition, gene expression analysis revealed that 5 *FaWHYs* were expressed in the crown roots, with identical expression patterns. All of them were down-regulated by crown rot infected. Furthermore, 'Defence and stress responsive' *cis*-elements were identified in the promoters of *FaWHYs1*, *3*, *4*, and *5*. The regulation of *FaWHYs* expression was similar to the *cis*-elements in the promoters, it can be speculated that they have similar functions. This result suggested that all of the *FaWHYs* have the ability to regulate pathogen infection stress.

Conclusions

Strawberry crown rot occurs all over the world, and the correlation between its disease resistance and the disease resistance gene *Whirly* is still unclear. In our current study, we identified five *FaWHYs*, two *FiWHYs*, three *FnWHYs*, two *FviWHYs* and four *FvWHYs* in the *F. ananassa, F. iinumae*, *F. nilgerrensis*, *F. viridis*, and *F. vesca* genome, respectively. In the syntenic relationship analysis with *A. thaliana*, it was found that *F. ananassa* produced a new genome (*FaWHY1*) during the evolution process. In the syntenic relationship analysis with four diploid strawberries, *FiWHY1*, *FvWHY2*, and *FviWHY1* showed syntenic relationships with *FaWHY1* and *FaWHY2*. At the same time, *FiWHY2*, *FvWHY3*, *FviWHY2*, and *FnWHY3* have

similar syntenic relationships with *Fa WHY4* and *Fa WHY5*, and *Fn WHY1* and *Fn WHY2* corresponded to both *Fa WHY1* and *Fa WHY2*. It showed that *F. ananassa* may have chromosomal variation during the evolution process, which also proved it is highly conserved during whirly evolution. It was revealed that these genes are simultaneously down-regulated in the process of disease resistance. The analysis of phylogenetic tree and *cis*-elements in promoters indicated that the genes may have the ability to regulate the pressure of pathogen infection. However, the study of Whirly's mechanism of action is still not thorough, so it is necessary to further study the signal pathways involved to further study the specific mechanism of action. Collectively, the results of this study provided a basis for future functional studies of the strawberry Whirly and their responses to crown rot.

Authors' Contributions

SB conceived and designed the experiments, supervised and revised the manuscript. YH conducted the experiments and wrote the original manuscript.

Both authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Scientific Research Foundation for Doctor of Yangtze University, grant number 802100270303.

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References

Akbudak MA, Filiz E (2019). *Whirly* (*Why*) transcription factors in tomato (*Solanum lycopersicum*): genome-wide identification and transcriptional profiling under drought and salt stresses. Molecular Biology Report 46(4):4139-4150. *https://doi.org/10.1007/s11033-019-04863-y*

Anciro A, Mangandi J, Verma S, Peres N, Whitaker VM, Lee S (2018). *FaRCg1*: a quantitative trait locus conferring resistance to Colletotrichum crown rot caused by *Colletotrichum gloeosporioides* in octoploid strawberry. Theoretical and Applied Genetics 131(10):2167-2177. *https://doi.org/10.1007/s00122-018-3145-z*

Cappadocia L, Parent JS, Sygusch J, Brisson N (2013). A family portrait: structural comparison of the *Whirly* proteins from *Arabidopsis thaliana* and *Solanum tuberosum*. Acta Crystallographica. Section F Structural Biology Crystallization Communications 69(11):1207-1211. *https://doi.org/10.1107/S1744309113028698*

Cappadocia L, Parent JS, Zampini, E, Lepage E, Sygusch J, Brisson N (2012). A conserved lysine residue of plant *Whirly* proteins is necessary for higher order protein assembly and protection against DNA damage. Nucleic Acids Research 40(1):258-269. *https://doi.org/10.1093/nar/gkr740*

Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. Molecular Plant 13(8):1194-1202. *https://doi.org/10.1101/289660*

Chitnis VR, Gao F, Yao Z, Jordan MC, Park S, Ayele BT (2014). After-ripening induced transcriptional changes of hormonal genes in wheat seeds: the cases of brassinosteroids, ethylene, cytokinin and salicylic acid. PloS One 9(1):e87543. *https://doi.org/10.1371/journal.pone.0087543*

Desveaux D, Allard J, Brisson N, Sygusch J (2002). A new family of plant transcription factors displays a novel ssDNA-binding surface. Nature Structural Biology 9(7):512-517. *https://doi.org/10.1038/nsb814*

Desveaux D, Després C, Joyeux A, Subramaniam R, Brisson N (2000). *PBF-2* is a novel single-stranded DNA binding factor implicated in *PR-10a* gene activation in potato. The Plant Cell 12(8):1477-1489. *https://doi.org/10.2307/3871144*

Desveaux D, Maréchal A, Brisson N (2005). *Whirly* transcription factors: defense gene regulation and beyond. Trends in Plant Science 10(2):95-102. *https://doi.org/10.1016/j.tplants.2004.12.008*

Desveaux D, Subramanian R, Després C, Mess J, Lévesque C, Fobert P, Dangl J (2004). A *WHIRLY* transcription factor is required for salicylic acid dependent disease resistance in *Arabidopsis*. Developmental Cell 6(2):229-240. *https://doi.org/10.1016/s1534-5807(04)00028-0*

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5):1792-1797. *https://doi.org/10.1093/nar/gkh340*

Edger PP, McKain MR, Yocca AE, Knapp SJ, Qiao Q, Zhang TC (2020). Reply to: Revisiting the origin of octoploid strawberry. Nature Genetics 52(1):5-7. *https://doi.org/10.1038/s41588-019-0544-2*

Edger PP, Poorten TJ, VanBuren R, Hardigan MA, Colle M, McKain MR, ... Knapp PS (2019) Origin and evolution of the octoploid strawberry genome. Nature Genetics 51:541-547. *https://doi.org/10.1038/s41588-019-0356-4*

Feng C, Wang J, Harris AJ, Folta KM, Zhao M, Kang M (2021). Tracing the diploid ancestry of the cultivated octoploid strawberry. Molecular Biology and Evolution 38(2):478-485. *https://doi.org/10.1093/molbev/msaa238*

Foyer CH, Karpinska B, Krupinska K (2013). The functions of *WHIRLY1* and REDOX-RESPONSIVE TRANSCRIPTION FACTOR1 in cross tolerance responses in plants: a hypothesis. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 369(1640):20130226. *https://doi.org/10.1098/rstb.2013.0226*

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003). Expasy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research 31(13):3784-3788. *https://doi.org/10.1093/nar/gkg563*

He YT, Ju CM, Chen JG (2020). Cloning of OsWHY1 gene in rice and construction of overexpression and RNAi vectors. School of Life Sciences, Hubei University 42:390-397.

Isemer R, Mulisch M, Schäfer A, Kirchner S, Koop HU, Krupinska K (2012). Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. FEBS Letters 586(1):85-88. *https://doi.org/10.1016/j.febslet.2011.11.029*

Kong FY, Deng YS, Zhou B, Meng QW (2012). Research advancement of Whirly transcription factors. Plant Physiology Journal 48(7):643-653. *https://doi.org/10.13592/j.cnki.ppj.2012.07.014*

Kumar S, Stecher G, Tamura K (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33(7):1870-1874. *https://doi.org/10.1093/molbev/msw054*

Lepage É, Zampini É, Brisson N (2013). Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signalling in *Arabidopsis*. Plant Physiology 163(2):867-881. *https://doi.org/10.1104/pp.113.223560*

Li Q, Wang H, Guo LQ, Luo QL (2019). Identification and expression analysis of the *Whirly* gene family in soybean. Soybean Science 38:204-211.

Li Z, Wang YT, Gao L, Wang F, Ye JL, Li GH (2014). Biochemical changes and defence responses during the development of peach gummosis caused by *Lasiodiplodia theobromae*. European Journal of Plant Pathology 138(1):195-207. *https://doi.org/10.1007/s10658-013-0322-4*

Liston A, Wei N, Tennessen JA, Li JM, Dong M, Ashman TL (2020) Revisiting the origin of octoploid strawberry. Nature Genetics 52(1):2-4. *https://doi.org/10.1038/s41588-019-0543-3*

Liu W, Yan Y, Zeng HQ, Li XL, Wei YX, Liu GY, ... Shi HT (2018). Functional characterization of *WHY-WRKY75* transcriptional module in plant response to cassava bacterial blight. Tree Physiology 38(10):1502-1512. *https://doi.org/10.1093/treephys/tpy053*

Lu YY, Wei XC, Zhao YY, Yuan YX, Wang ZY, Yang SJ, ... Zhang XW (2019). Identification and expression analysis of *Whirly* gene family in *Capsicum annuum* L. Acta Agriculturae Boreali-Sinica 34:72-78.

Mangandi J, Peres NA, Whitaker VM (2015). Identifying resistance to crown rot caused by Colletotrichum gloeosporioides in strawberry. Plant Disease 99(7):954-961. *https://doi.org/10.1094/pdis-09-14-0907-re*

Maréchal A, Parent JS, Sabar M, Véronneau-Lafortune F, Abou-Rached C, Brisson N (2008). Overexpression of mtDNA-associated *AtWhy2* compromises mitochondrial function. BMC Plant Biology 8(1):42. *https://doi.org/10.1186/1471-2229-8-42*

Rombauts S, Déhais P, Montagu MV, Rouzé P (1999). PlantCARE, a plant *cis*-acting regulatory element database. Nucleic Acids Research 27(1):295-296. *https://doi.org/10.1093/nar/27.1.295*

Shu B, Li WC, Liu LQ, Wei YZ, Shi SY (2016). Transcriptomes of arbuscular mycorrhizal fungi and litchi host interaction after tree girdling. Frontiers in Microbiology 7:408. *https://doi.org/10.3389/fmicb.2016.00408*

Wang YP, Tang HB, Debarry JD, Tan X, Li JP, Wang XY, ... Paterson AH (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Research 40(7):e49. *https://doi.org/10.1093/nar/gkr1293*

Yao QT, Zhang WW, Liu L, Han R, Jian GL, Qi FJ (2008). Negative regulation of rice *Whirly* transcription factor for the hypersensitive response induced by non-host pathogen bacterium. Journal of Agriculture Science and Technology 10:53-58.

Zhao SY, Wang GD, Zhao WY, Zhang S, Kong FY, Dong XC, Meng QW (2018). Overexpression of tomato WHIRLY protein enhances tolerance to drought stress and resistance to *Pseudomonas solanacearum* in transgenic tobacco. Biologia Plantarum 62:55-68. *https://doi.org/10.1007/s10535-017-0714-y*