Bioinformatic Analysis of MYB Transcription Factors in *Fusarium graminearum*

Zhang Shengnan, Lan Jianwu, Li Ya

Key Laboratory of Biopesticides and Chemical Biology, Ministry of Education, Fujian Agriculture and Forestry University, Fuzhou, 350002, China

Corresponding author email: lina81@163.com


Received: 07 Jul., 2020
Accepted: 08 Jul., 2020
Published: 24 Jul., 2020

Abstract  The bioinformatic method was used to analyze the 17 *Fusarium graminearum* MYB transcription factors (TFs) included in the fungal transcription factor database. Analysis of gene sequence and protein physical and chemical properties showed that the 17 *F. graminearum* MYB TFs differ in gene length, protein size, theoretical isoelectric point, total number of atoms, and instability index; Motif analysis and protein secondary and tertiary structure prediction found that these transcription factors contain 2 characteristic motifs, and that the protein structures are diverse, but all contain the MYB TF characteristic structure (helix-turn-helix (HTH)); Phylogenetic tree analysis showed that 17 *F. graminearum* MYB TFs of and 12 reported MYB TF3 from other species have evolved into two large clades. Among them, FgMYB03, FgMYB04, FgMYB05, FgMYB06, FgMYB10, FgMYB12 and FgMYB13 have close relationship with known animal, plant and fungal MYB TFs including c-MYB, ATMYB2, MoMYb1, FlbD and Pol5, which implied that they possibly share similar biological functions. These results provide a reference for further research on the biological function of *F. graminearum* MYB TFs and the feature of other fungal MYB TF family.

Keywords  MYB transcription factors; *F. graminearum*; Bioinformatic analysis

MYB TFs are a family of genes containing 1 to 4 conserved MYB DNA binding domains. The typical characteristic of this domain is containing 50–53 amino acids, in which there is a conserved tryptophan every 18 or 19 amino acids (Ogata et al., 1996). In space, the MYB domain has a three helix structure, and the second and third helix turns into a certain angle, forming a helix-turn-helix (HTH) spatial configuration (Dubos et al., 2010). The MYB domain was first discovered in the oncogene (v-MYB) encoded by the avian myeloblastosis virus, and was subsequently reported in animals and plants. It is widespread in eukaryotes and has about 1 billion years of evolution (Klempnauer et al., 1982).

Up to now, MYB TFs has been widely studied in plants and animals. The animal MYB TF family is relatively small, with an average of only 4–5 members per species. They play an important role in regulating cell growth, tissue differentiation, and tumor formation (Golay et al., 1991). In comparison, there is a large number of MYB TFs in plants, with an average of 100 to 200 members in each species. They are involved in regulating plant cell morphogenesis, growth and development, primary and secondary metabolism, and abiotic and biological stress responses (Dubos et al., 2010). The study of fungal MYB TFs started late, to date they were only reported in the form of single gene in fungi such as yeast, *Aspergillus nidulans*, *Aspergillus fumigatus* and *Magnaporthe oryzae*. Preliminary studies have found that they are essential for fungal spore formation and fungal pathogenicity to plant (Arratia-Quijada et al., 2012; Dong et al., 2018; Sarikaya Bayram et al., 2018; Ramos-Sáenz et al., 2019).

*F. graminearum* is a plant pathogenic fungus, which can infect a variety of cereal crops such as wheat, barley, rice and oats, and cause diseases such as ear, stem, stem base and root rot. Its representative disease, wheat scab, is one of the most serious diseases in agricultural production (Chen et al., 2019). In this study, *F. graminearum* was selected as the research object. The sequence characteristics, protein structure and evolutionary relationship...
of MYB TFs were analyzed in the form of gene family. Our result provide clues for further research on the biological functions of *F. graminearum* MYB TFs and also provide reference for the research on MYB TF family in other plant pathogenic fungi.

1 Results and Analysis

1.1 Identification of *F. graminearum* MYB TF and analysis of protein physicochemical properties

A search of the fungal transcription factor database (FTFD) found that *F. graminearum* contains 17 MYB TF sequences, named as FgMYB01-013 for 13 of them. As the remaining 4 TFs have been previously studied (Kim et al., 2014; Son et al., 2014), the names MYT1, MYT2, MYT3, and FgFlbD designed when they were identified are used here (Table 1). Sequence analysis showed that the nucleic acid length of the 17 MYB TFs is between 6,915 bp and 678 bp. Among them, the longest and shortest ones are MYT3 and FgMYB05, respectively, their encoding amino acid length are 2,304 aa and 225 aa, respectively, and the corresponding molecular weights are 256,635.86 kD and 24,651.97 kD, respectively (Table 1). Analysis of physical and chemical properties found that the theoretical isoelectric point of the 17 MYB TFs is between 10.17 and 4.85, corresponding to the largest and smallest ones FgMYB09 and FgMYB05, respectively; the total number of atoms is between 35,416 and 3,452, which corresponds to the most And the smallest TFs are MYT3 and FgMYB05 respectively; the fat coefficient is between 75.84 and 43.11, corresponding to the largest and smallest TFs are FgMYB12 and FgMYB06; the instability coefficient is between 84.09 and 42.69, corresponding to the largest and smallest TFs are FgMYB01 and FgMYB03 (Table 1). These data indicate that the 17 MYB TFs of *F. graminearum* have differences in physicochemical properties such as nucleic acid sequence length, encoded protein size, and isoelectric point.

Table 1 Analysis of *F. graminearum* MYB TFs

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Accession number</th>
<th>Number of base pairs</th>
<th>Size (aa)</th>
<th>Molecular Weight (kD)</th>
<th>Theoretical pI</th>
<th>Total number of atoms</th>
<th>Aliphatic index</th>
<th>Instability index</th>
</tr>
</thead>
<tbody>
<tr>
<td>FgMYB01</td>
<td>XP_011317055.1</td>
<td>750</td>
<td>249</td>
<td>27,969.69</td>
<td>9.60</td>
<td>3908</td>
<td>71.00</td>
<td>84.09</td>
</tr>
<tr>
<td>FgMYB02</td>
<td>XP_011317263.1</td>
<td>2685</td>
<td>894</td>
<td>97,461.83</td>
<td>5.14</td>
<td>13477</td>
<td>67.24</td>
<td>61.61</td>
</tr>
<tr>
<td>FgMYB03</td>
<td>XP_011317776.1</td>
<td>2016</td>
<td>671</td>
<td>73,500.55</td>
<td>4.96</td>
<td>10159</td>
<td>66.59</td>
<td>42.69</td>
</tr>
<tr>
<td>FgMYB04</td>
<td>XP_011318475.1</td>
<td>858</td>
<td>285</td>
<td>31,339.18</td>
<td>6.48</td>
<td>4337</td>
<td>62.00</td>
<td>54.99</td>
</tr>
<tr>
<td>FgMYB05</td>
<td>XP_011318675.1</td>
<td>678</td>
<td>225</td>
<td>24,651.97</td>
<td>10.17</td>
<td>3452</td>
<td>44.00</td>
<td>79.86</td>
</tr>
<tr>
<td>FgMYB06</td>
<td>XP_011321075.1</td>
<td>861</td>
<td>286</td>
<td>30,403.95</td>
<td>5.71</td>
<td>4123</td>
<td>43.11</td>
<td>71.22</td>
</tr>
<tr>
<td>FgMYB07</td>
<td>XP_011324059.1</td>
<td>4890</td>
<td>1629</td>
<td>17,6914.21</td>
<td>6.37</td>
<td>24485</td>
<td>60.63</td>
<td>59.93</td>
</tr>
<tr>
<td>FgMYB08</td>
<td>XP_011326417.1</td>
<td>1770</td>
<td>589</td>
<td>64,429.22</td>
<td>5.99</td>
<td>8958</td>
<td>45.84</td>
<td>69.37</td>
</tr>
<tr>
<td>FgMYB09</td>
<td>XP_011327220.1</td>
<td>1371</td>
<td>456</td>
<td>51,435.63</td>
<td>4.85</td>
<td>7101</td>
<td>54.34</td>
<td>54.94</td>
</tr>
<tr>
<td>FgMYB10</td>
<td>XP_011320019.1</td>
<td>1671</td>
<td>556</td>
<td>62,175.85</td>
<td>8.24</td>
<td>8705</td>
<td>67.19</td>
<td>53.98</td>
</tr>
<tr>
<td>FgMYB11</td>
<td>XP_011327880.1</td>
<td>2337</td>
<td>778</td>
<td>86,759.78</td>
<td>7.11</td>
<td>12182</td>
<td>72.11</td>
<td>49.41</td>
</tr>
<tr>
<td>FgMYB12</td>
<td>XP_011319222.1</td>
<td>3345</td>
<td>1114</td>
<td>127,695.83</td>
<td>6.31</td>
<td>17909</td>
<td>75.84</td>
<td>44.18</td>
</tr>
<tr>
<td>FgMYB13</td>
<td>XP_011324238.1</td>
<td>1356</td>
<td>451</td>
<td>50,970.64</td>
<td>6.10</td>
<td>7097</td>
<td>69.31</td>
<td>55.22</td>
</tr>
<tr>
<td>MYT1</td>
<td>XP_011315967.1</td>
<td>885</td>
<td>294</td>
<td>32,503.95</td>
<td>9.87</td>
<td>4524</td>
<td>60.14</td>
<td>82.75</td>
</tr>
<tr>
<td>MYT2</td>
<td>XP_011327324.1</td>
<td>972</td>
<td>323</td>
<td>35,235.04</td>
<td>9.08</td>
<td>4824</td>
<td>55.08</td>
<td>59.23</td>
</tr>
<tr>
<td>MYT3</td>
<td>XP_011315973.1</td>
<td>6915</td>
<td>2304</td>
<td>256,635.86</td>
<td>8.81</td>
<td>35416</td>
<td>43.83</td>
<td>71.54</td>
</tr>
<tr>
<td>FgFlbD</td>
<td>XP_011317764.1</td>
<td>900</td>
<td>299</td>
<td>33,696.48</td>
<td>10.12</td>
<td>4665</td>
<td>60.67</td>
<td>74.34</td>
</tr>
</tbody>
</table>

1.2 Motif and gene structure analysis of *F. graminearum* MYB TF

Motif analysis of 17 MYB TFs of *F. graminearum* by MEME showed that different TFs contained different motif numbers (Figure 1). Among them, FgMYB12 contains the largest Motif number, 8 motifs were included; FgMYB01 contains the smallest Motif number 2 motifs were included; the remaining MYB TFs each contain 2-8 motifs (Figure 1). Of the 20 predicted Motifs, Motif 1 and 2 have the highest frequency, presuming that they are characteristic motifs of the *F. graminearum* MYB TF.
Using SOPMA to predict the secondary structure of the MYB TF of *F. graminearum*, we found that the secondary structure of these TFs is mainly the α helix (blue line) and random curl (purple line), and also contains a few β sheets (Green line) (Figure 2); Furthermore, in the secondary structure of each protein, the α-helix and β-sheet are arranged at intervals (Figure 2), which allows that a helix-turn-helix (HTH) conformation is formed in space. This prediction result is consistent with the structural characteristics of MYB protein (Dubos et al., 2010).

**Figure 1 Motif prediction of *F. graminearum* MYB TFs**

Using SWISS-MODEL to predict the tertiary structure and construct model of the *F. graminearum* MYB TFs, we found that, consistent with the prediction of the secondary structure, these proteins have structures such as α helix, random coil and β sheet, and obvious helix-turn-helix (HTH) structure can be found in the model diagram (Figure 3). However, different MYB TFs also have differences in protein structure. For example, different proteins contain different numbers of helix-turn-helix (HTH), and the spatial configuration also exhibits diversity (Figure 3).

Domain analysis of the 17 MYB TF of *F. graminearum* by SMART revealed that, except for MYT3 and FgMYB01, other MYB proteins contain at least 1 SANT domain. Among them, FgMYB04, FgFlbD, FgMYB10, FgMYB11 and FgMYB12 contain the most 2 SNAT domains (Figure 4). As reported, the SNAT domain was highly similar to MYB domain in structure, and is generally considered to be the characteristic domain of the MYB TF (Grüne et al., 2003). In addition, besides having a SANT domain, FgMYB03 and FgMYB13 contain a ZNF ZZ zinc finger domain, and FgMYB12 contains DEXDc and HELICc domains (Figure 4), presuming that they have other functions besides the function of MYB TF. Analysis of the distribution of SNAT domain in MYB protein revealed that its position is not fixed, present at both the N-terminal, C-terminal, and intermediate positions of full amino acid sequence, while the SNAT domain in animal and plant MYB TFs is mainly gathered at the N-terminal.

**1.3 Evolutionary analysis of MYB TF of *F. graminearum***

Using MEGA 7.0, the homologous alignment and evolutionary tree construction of 17 MYB TFs of *F. graminearum* and 12 MYB TFs of other animals, plants and fungi were performed. It was found that these 29 MYB TFs have evolved into two large branches (I and II). Branch I is larger and mixed with 21 MYB genes
from both animals, plants and fungi; branch II is smaller and contains 8 MYB genes only from fungi, 7 from *F. graminearum* and 1 from *Saccharomyces cerevisiae* (Figure 5).

![Figure 2 Protein secondary structure prediction of *F. graminearum* MYB TFs](image)

![Figure 3 Protein tertiary structure prediction of *F. graminearum* MYB TFs](image)
Figure 4 Domain prediction of *F. graminearum* MYB TFs

Figure 5 Phylogenetic analysis of *F. graminearum* MYB TFs

Note: AC: *Acremonium chrysogenum*; AF: *Aspergillus fumigatus*; AMV: *Avian myeloblastosis virus*; AN: *Aspergillus nidulans*; AT: *Arabidopsis thaliana*; FG: *Fusarium graminearum*; HS: *Homo sapiens*; MO: *Magnaporthe oryzae*; SC: *Saccharomyces cerevisiae*; ZM: *Zea mays*

In branch I, *F. graminearum* FgMYB12, FgMYB13, FgFlbD and FgMYB06 locate in the same small branch with *Arabidopsis thaliana* ATMYB2 (Dubos et al., 2010), *Zea mays* ZmybC1 (Dubos et al., 2010), *Homo sapiens* c-MYB (Zhao et al., 2019), *Avian myeloblastosis virus* v-myb (Klempnauer et al., 1982), *Magnaporthe oryzae* MoMyb1 (Dong et al., 2018) and *Aspergillus fumigatus* FlbD (Arratia-Quijada et al., 2012) (blue area in Figure 5), suggesting that they have a close relationship. In addition, *F. graminearum* FgMYB04, FgMYB05,
MYT1 and MYT3 locate in another small branch with *Aspergillus nidulans* BasR (Fischer et al., 2018) and *Acrocnium chrysogenum* AcMybA (Wang et al., 2018) (Figure 5 red Area), which means that they have a close relationship.

In branch II, *F. graminearum* FgMYB03 and FgMYB10 are in the same small branch with *S. cerevisiae* Pol5 (Ramos-Sáenz et al., 2019) (the gray area in Figure 5), which indicates that they have a close relationship. All the remaining 5 MYB genes FgMYB01, FgMYB02, FgMYB08, FgMYB09 and FgMYB11 are from *F. graminearum* and locate in another small branch (Figure 5 purple area), indicating that they have a close relationship. The above results indicate that the *F. graminearum* MYB TF has sequence homology and evolutionary affinity with the known MYB TFs.

## 2 Discussions

This study found that *F. graminearum* contains 17 members of the MYB TF family. Further search for other fungal MYB TF families in the fungal transcription factor database (http://ftfd.snu.ac.kr/tf.php?a= dv_list&id=8) showed that the number of MYB TFs of a single fungal species is mostly between 10 and 50, only a few species have the number of MYB TF number less than 10 or more than 50, and no species have more than 60 MYB TFs. In comparison, the number of reported MYB TFs for individual animals and plants is between 4-5 and 100-200, respectively (Prouse and Campbell, 2012), indicating that the size of the fungal MYB TF family is between animals and plants. It is speculated that MYB TFs of different species have different amplification ratios during evolution.

According to the number of MYB domain repeats (R), animal and plant MYB TFs can be divided into four types: 1R, 2R, 3R, and 4R, representing MYB domains containing 1, 2, 3, and 4 repeats, respectively (Dubos et al., 2010). In this study, domain prediction of the 17 *F. graminearum* MYB TFs showed that except for 5 TFs including 2 MYB domains (belonging to type 2R), the remaining 10 contained only 1 MYB domain (belonging to type 1R), no type 3R and 4R MYB TFs were discovered, the types are not as rich as animal and plant MYB TFs. In addition, although the fungal transcription factor database contains 17 MYB TFs in *F. graminearum*, domain prediction by the SMART database showed that two of them MYT3 and FgMYB01 do not contain MYB domain. We speculated that it may be caused by the different prediction basis and judgment threshold of different databases.

Through the analysis of protein physical and chemical properties and protein structure prediction, it is found that, although 17 *F. graminearum* MY TFs all have conservative MYB structural characteristics, different MYB TFs showed big differences in nucleic acid length, protein size, theoretical isoelectric point, total number of atoms, and instability coefficient. Especially the difference between the minimum and maximum values is more obvious, which is not as balanced as plant TFs (Liu X.T. et al., 2109). These indicate that 17 *F. graminearum* MYB TFs have diverse characteristics in biochemical or biological functions. Phylogenetic studies showed that *F. graminearum* MYB TFs FgMYB03, FgMYB04, FgMYB05, FgMYB06, FgMYB10, FgMYB12 and FgMYB13 have close relationship to the reported animal, plant and fungal MYB TFs c-MYB, ATMYB2, MoMyb1, FlBD and Pol5. In view of the role of the reported TFs in regulating animal cell differentiation, plant stress response, fungal sporulation and pathogenicity (Arratia-Quijada et al., 2012; Dong et al., 2018; Zhao L.R. et al., 2019; Dubos et al., 2010), the MYB TFs in *F. graminearum* may have similar biological functions.

In summary, this study was the first to use bioinformatics in fungi to analyze the characteristics the MYB TF family in *F. graminearum*. The results obtained provide a reference for the study of the single gene function of this family member in *F. graminearum* and for the bioinformatics analysis of other fungal MYB TF families.
3 Materials and Methods

3.1 Screening and sequence acquisition of MYB TF of F. graminearum

The 17 MYB TFs of *F. graminearum* were obtained from the fungal transcription factors database (http://ffid.snu.ac.kr/index.php?a=view). Use the keyword "Myb" to search and then select "Fusarium graminearum" in the species list, 17 MYB TF records and corresponding amino acid sequences could be generated by the browser. By blast in the NCBI database with the amino acid sequence (https://www.ncbi.nlm.nih.gov/), the nucleotide sequence of each MYB TF can be obtained.

3.2 Analysis of physical and chemical properties of MYB TF of F. graminearum

Using Expasy (https://web.expasy.org/protparam/) to analyze the physical and chemical properties of *F. graminearum* MYB TF. Upload the protein sequence of the family members to the Expasy server and calculate, the protein molecular weight, theoretical isoelectric point, total number of atoms, fat coefficient and instability coefficient and other values could be generated.

3.3 Motif and protein structure analysis of MYB TF of F. graminearum

Use MEME (http://meme-suite.org/tools/meme) to predict the motif of MYB TF protein sequence (Motif). The search number of motif is set as 20, and other parameters are set as the default values (Ma et al., 2019). Use SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) to predict the secondary structure of the MYB protein sequence. The Swiss-model (https://swissmodel.expasy.org/interactive) was used to predict the tertiary structure of the MYB protein sequence based on the homology modeling method. The SMART (http://smart.embl-heidelberg.de/) was used to predict the conserved domain of the MYB protein sequence.

3.4 MYB TF evolution analysis

The phylogenetic tree was constructed using the MEGA7.0 software based on the amino acid sequences of 17 MYB TFs of *F. graminearum* and 12 MYB TFs of other species. First, Clustal is used for sequence alignment; then the model is calculated to be JTT+G by MEGA7.0; finally, the maximum likelihood (ML) method is used to build the tree, and the Bootstrap is set as 1000.

Authors’ contributions

Zhang Shengnan and Lan Jianwu are the experimental designers and executors of this research; Zhang Shengnan completed data analysis and wrote the manuscript; Lan Jianwu participated in the experimental design and analyzed the experimental results; Li Ya is the planner and the person in charge of the project, directing experimental design, data analysis, and paper writing. All authors read and approved the final manuscript.

Acknowledgments

This research was jointly funded by the National Natural Science Foundation of China (31871914) and the Fujian Natural Science Foundation (2018J01615).

References

Arratia-Quijada J., Sánchez O., Scorzecchio C., and Aguirre J., 2012, FlbD, a Myb transcription factor of *Aspergillus* nidulans, is uniquely involved in both asexual and sexual differentiation, Eukaryot Cell, 11(9): 1132-1142

https://doi.org/10.1128/EC.00101-12

PMid:22798393 PMCid:PMC3445977


https://doi.org/10.1146/annurev-phyto-082718-100318

PMid:30893009


https://doi.org/10.1186/s12866-018-1263-z

PMid:30285616 PMCid:PMC6171296
https://doi.org/10.1016/j.tplants.2010.06.005
PMid:20674465

https://doi.org/10.7554/elife.40969
PMid:30311191
PMCID:PMC6234034

https://doi.org/10.1182/blood.V77.1.149.bloodjournal771149
PMid:1984793

https://doi.org/10.1016/S1097-2765(03)00273-9

https://doi.org/10.1371/journal.pone.0094359
PMid:24722578
PMCID:PMC3983115

https://doi.org/10.1016/0092-8674(82)90138-6


https://doi.org/10.1038/nstr8746-178
PMid:8564545

https://doi.org/10.1016/j.bbadgr.2011.10.010
PMid:22067744

https://doi.org/10.1261/rna.072116.119
PMid:31413149

https://doi.org/10.1007/s00294-017-0740-5
PMid:28840304

https://doi.org/10.1007/s12275-014-4384-6
PMid:25277408

https://doi.org/10.1016/j.fgb.2018.05.006
PMid:29870835


8