Overexpression of MpbHLH transcription factor, an encoding ICE-like protein, enhances Foc TR4-resistance of Cavendish banana

Haochen Li1,2#, Chunhua Hu1#, Aifeng Xie1,3#, Shaoping Wu1,4, Fangcheng Bi1, Tao Dong1, Chunyu Li1, Guiming Deng1, Weidi He1, Huijun Gao1, Ou Sheng1, Ganjun Yi1*, Qiao-Song Yang1,* and Tongxin Dou1,*

1 Key Laboratory of South Subtropical Fruit Biology and Genetic Resource Utilization (Ministry of Agriculture and Rural Affairs), Guangdong Key Laboratory of Tropical and Subtropical Fruit Tree Research, Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences, Guangzhou, China;
2 College of Plant Protection, South China Agricultural University, Guangzhou, China;
3 College of Life Sciences, South China Agricultural University, Guangzhou, China;
4 College of Horticulture, Hunan Agricultural University, Changsha, China.

*Correspondence:
Dr. Tongxin Dou
Email: doutongxin@gdaas.cn
or Dr. Qiaosong Yang
Email: soyang@hotmail.com
Or Dr. Ganjun Yi
Email: yiganjun@vip.163.com

#These authors contributed equally to this work.
Abstract

Background: Fusarium wilt is a worldwide problem that restrains the development of banana industry. Isolation and identification of resistance genes is not only significance for the selection of disease resistant varieties, but also important for banana molecular breeding.

Results: In this study, an ICE1-like transcription factor containing bHLH domain was cloned from Dajiao (Musa spp. ABB group) and named MpbHLH. We found that overexpression of MpbHLH in Cavendish banana (Musa spp. AAA group) could significantly enhance Fusarium wilt resistance which is caused by Foc TR4. Transcriptomics and proteomic analysis showed that phenylalanine ammonia lyase (PAL) and peroxidase (POD) were highly expressed in transgenic plants. Physiological experiments further revealed that transgenic plants, especially their roots, had high POD activity.

Conclusions: Overexpression of MpbHLH up-regulates the phenylpropanoid pathway-related genes and POD-associated genes in transgenic Cavendish banana. These changes may eventually lead to the improvement of disease resistance.

Keywords: Banana, Fusarium wilt, MpbHLH, PAL, POD, Transcription factor.
**Background**

Banana (*Musa* spp.) is among the most important food crops in many tropical and subtropical regions and is one of the most popular fruit crops worldwide [1,2]. Compared to other food crops, the yield and productivity of banana are adversely constrained by several key biotic and abiotic stresses, chief among these is Fusarium wilt [3,4]. The pathogen belongs to *Fusarium oxysporum* f.sp. *cubense* [5]. With the co-evolution of the pathogen and the host, the pathogen mainly evolves into four physiological races: *Foc* Race 1, *Foc* Race 2, *Foc* STR4, and *Foc* TR4. Among them, *Foc* TR4, which broke out in the 1990s, has the widest pathogenic range and the strongest pathogenicity, and is pathogenic to almost all types of bananas, increasing the vulnerability of the global banana industry [6,7].

To date, development of Fusarium wilt-resistant banana faces challenges of sexual reproduction barriers such as high sterility, complex genetic background, polyploid nature, and parthenogenesis, making it extremely difficult to develop new disease-resistant bananas via cross-breeding [8,9]. In recent years, researchers have invested a lot of efforts in cultivating new disease-resistant banana germplasm by the transformation of disease-resistant genes, and have made certain progress. For example, Mazizh et al. (2007) introduced soybean β-1,3glucanase gene into the banana, proving that the transgenic plant had a certain resistance to *Foc* Race 1 [10]; the overexpression of *PhDef1* and *PhDef2*, two antimicrobial peptide genes of Petunia, could significantly improve banana’s resistance to *Foc* TR4 [11]; Mohandasa et al. (2013) transferred *Ace-AMP1*, an antimicrobial peptide gene from onion seeds, into banana and conducted *Foc* Race 1 experiments, proving that the incidence rate of diseases in transgenic plants decreased significantly [12]; and the overexpression of diploid disease-resistant banana *RGA2* gene and nematode homologous gene *Ced9* could significantly enhance the resistance of Cavendish banana to *Foc* TR4 [3].

However, due to its constant evolution, the pathogen easily overcomes resistance barriers. Therefore, there is an urgent need to find disease-resistant genes through molecular breeding of *Foc*-resistant banana.

Through nearly 10,000 years of evolution, plants have evolved a series of regulatory mechanisms that enable them to cope with environmental changes and
pathogen invasions [13-15]. When plants feel stress stimuli, signals are transmitted from cell membranes to the cytoplasm, and finally to the nucleus, which eventually result in physiological and metabolic changes through regulation of gene expressions and translation of proteins, thus improving the resistances of plants to different types of stress [16,17]. It is worth noting that transcription factors in the nucleus are critical in “response to stress”, and their functions are amplified through signal transduction cascades [13,18,19]. In recent years, a growing number of studies have found that many transcription factors, including WRKY, b-ZIP, AP2/ERF, and MYB, have superimposed biological functions on biotic and abiotic stresses in plants. For example, the rice overexpressing SNAC1- and OsNAC6/SNAC2 transcription factors are not only resistant to dehydration, salt, and cold treatment, but also show stronger resistance to Magnaporthe grisea [20]; Hénanff et al. (2013) found that Arabidopsis overexpressing VvNAC1 transcription factor of grape showed enhanced resistance to salt and cold treatment, and significantly improved resistance to staphylococcus griseus and Arabidopsis pathogens [22]; and the PacMYBA transcription factor overexpressing sweet cherry in Arabidopsis thaliana enhanced its resistance to salt, and also showed a certain resistance to pseudomonas syringae pv. infection [15]. In summary, transcription factors show superimposed biological functions in regulating biotic and abiotic stresses of plants, which may be a common rule. bHLH transcription factors have been widely studied in regulating abiotic stresses of plants, for example, the bHLH TFs MdCIbHLH1 of apple (Malus domestica.) was suggested to be involved in the response to cold stress [23]. A stress-responsive bHLH gene, RslICE1 from Raphanus sativus, could increase the cold tolerance of genetically modified rice [24]. Eleusine coracana bHLH57 is related to dehydration, salt, and oxidative tolerance, while bHLH122 plays an important role in dehydration and osmotic resistance in Arabidopsis [25]. OsbHLH148, a rice (Oryza sativa) bHLH gene, plays a role in dehydration tolerance as a component of the jasmonate signaling pathway [26]. Overexpression of PtrbHLH of Poncirus trifoliate increased cold tolerance in transgenic citrus [27]. However, bHLH transcription factors, especially bHLH transcription factors of banana, have so far not been reported in regulating disease resistance of banana.
In this study, a Dajiao candidate gene, \( MpbHLH \), encoding an ICE-like bHLH TF, was cloned and characterized. qPCR results revealed that \( MpbHLH \) transcript levels were significantly induced by abiotic stresses, such as cold treatment, salt, and dehydration. Overexpression of Dajiao \( MpbHLH \) led to enhanced superior resistance to Fusarium wilt of Cavendish banana transgenic lines. We used transcriptomics to analyze the gene expression changes. In addition, proteomic data were compared with transcriptome data. Our findings suggest that some stress-related genes, especially POD-associated genes, have a positive effect on banana Fusarium wilt.

Results

Isolation and structural analysis of \( MpbHLH \)

The full-length cDNA of \( MpbHLH \) was isolated from leaves of the Dajiao plantlets (Figure S1). The chromosome location of \( MpbHLH \) in the \( M. \) acuminata genome and \( M. \) balbisiana genome (http://banana-genome.cirad.fr/) shows that \( MpbHLH \) maps to chromosome 10, spanning 2998 bp and 2991 bp, respectively (Figure 1A). Analysis of its structural properties reveal that the predicted \( MpbHLH \) protein possesses features typical of MYC-like proteins, including a bHLH region (Figure 1B). The predicted three-dimensional structure of the \( MpbHLH \) protein is shown in Figure 1C. Amino acid sequence alignment shows that \( MpbHLH \) belongs to the bHLH TF family (Figure 2A). The \( MpbHLH \) amino acid sequence was compared with different species of monocotyledonous and dicotyledonous plants, and the phylogenetic tree analysis reveals that \( MpbHLH \) is clearly clustered with monocotyledonous plants (Figure 2B).

Expression profiles of \( MpbHLH \) in response to abiotic stresses

To gain insight into the biological functions of \( MpbHLH \), RT-qPCR was used to examine the expression profile of \( MpbHLH \) in Dajiao leaves in response to various abiotic stresses over time. Steady-state mRNA levels of \( MpbHLH \) were elevated under all tested stresses while the expression patterns varied. \( MpbHLH \) transcript levels increased progressively under cold stress until reaching the highest level at 6 hours (greater than 16-fold induction) before decreasing slightly after 24 hours (Figure 3A). Under extended salt stress, \( MpbHLH \) was induced by nearly 4.5-fold at 6 hours,
followed by reduction to a level slightly higher than that at the onset of the salt treatment (Figure 3B). Under dehydration stress, MpbHLH mRNA abundance underwent minor changes at 1 hour, increased to the highest level at 3 hours, and then reduced over the last three time points (Figure 3C). These results suggest that three abiotic stresses, including cold treatment, dehydration, and salt, could induce MpbHLH transcription.

Subcellular localization
In order to localize the expression of MpbHLH in cells, the GFP-MpbHLH fusion expression vector was constructed and transformed into rice protoplasts by PEG. The fluorescence was observed by laser confocal microscope after cultured at 28 °C under standard osmotic pressure for 12 hours. Overlay diagram showed that green fluorescence appeared in the nucleus of transformed cells (Figure 4), which indicated that MpbHLH is localized in the nucleus.

Transformation and characterization of transgenic Cavendish banana
Embryogenic cell suspensions (ECSs) of banana were genetically transformed using A. tumefaciens harboring a plant binary vector (pOx), to constitutively overexpress MpbHLH in transgenic banana plants (Figure S2A). Through co-cultures, solid medium screening, asexual propagation, and rooting (Figure S2B), 25 transformation lines were cultivated, to obtain the transformed plants; we applied hygromycin at each cultivation stage. Eleven out of 25 lines from the ECSs infected with Agrobacterium strain EHA105 harboring pOx-MpbHLH were positive for hygromycin, whereas it was absent in genomic DNA derived from untransformed WT plants (Figure S3A). Transgenic lines #DX11 and #DX13 with higher MpbHLH transcript levels were selected for further analysis along with the corresponding untransformed WT.

Overexpression of MpbHLH increases Foc TR4-resistance in transgenic Cavendish banana
The above two transgenic lines (#DX11 and #DX13) and WT plants were inoculated
with *Foc* TR4 in a greenhouse, to assess the disease resistance of transgenic banana plants. Fourteen days after the inoculations it was found that the transgenic plants grew well, while most leaves of the WT plants turned yellow and the whole plants died, manifesting the typical symptoms of Fusarium wilt (Figure 5A). In addition, the disease index of transgenic banana plants was significantly lower than that of WT plants (Figure 5B). The infection and localization of *Foc* TR4 in the roots and bulbs of banana were regularly observed using a laser confocal microscope after inoculation. As shown in Figure 5C, after 24 hours of inoculation, it was observed that *Foc* TR4 invaded the root tip surface of WT plants and left brown spots around them. Seven days after inoculation, fluorescence signals of *Foc* TR4 were observed in the root cross section of WT plants. These fluorescence signals mainly concentrated around the vascular bundles and produced large areas of brown spots. Brown spots were also observed in transgenic plants, but no obvious fluorescence signal was detected. After 14 days of inoculation, the fluorescence signals around the vascular bundles of WT plants were enhanced, indicating that more *Foc* TR4 colonized. Several scattered brown spots were observed in the cross-section of transgenic plants. However, only a small amount of fluorescence signal generated by *Foc* TR4 was detected. These results indicate that banana plants overexpressing MpbHLH could significantly inhibit the infection and localization rates of *Foc* TR4 from the root tip to vascular bundles, which may be the direct reason for transgenic banana plants to be highly resistant to Fusarium wilt.

**Overexpression of MpbHLH in Cavendish banana leads to dramatic transcriptomic alterations**

As a TF, MpbHLH might regulate an array of downstream target genes that may account for the noticeable enhancement of disease resistance in the Cavendish banana overexpression lines. To verify this hypothesis, we carried out transcriptomic analysis to compare the expression profiles of WT banana and one single-copy overexpression line (#DX11) before and after *Foc* TR4 treatment at BGI Genomics, BGI-Shenzhen, Shenzhen, Guangdong, P.R. China, using 6 Gb of clean sequencing data per sample. Using 2-fold change as a selection threshold, a total of 2114 and 422 genes were up-
and down-regulated, respectively in #DX11, compared with the WT without Foc TR4 treatment (Additional File 1). However, 478 and 123 genes were up- and down-regulated, respectively in line #DX11, compared with the WT after 7 days of Foc TR4 treatment (Additional File 2). As for the 14-day Foc TR4 treatment, there were 1515 up-regulated genes and 1816 down-regulated genes in line #11 compared with the WT (Additional File 3), indicating a noticeable alteration of the gene expression profile in the transgenic plant. Interestingly, 119 genes showed significant changes at the same times (0, 7, and 14 days) (Figure 6A and Additional File 4). A hierarchical heat map image was generated to analyze these 119 differentially expressed genes (DEGs) in response to different Foc TR4 treatment time (Figure 6B).

To further validate the results from this transcriptomic analysis, the expression of 6 candidate genes was analyzed using RT-qPCR (Figure 7 and Additional File 5). The results show that the RT-qPCR expression patterns of the selected genes were largely consistent with those obtained from the transcriptomic data despite the difference in the absolute fold change between the two methods, suggesting that results from this preliminary microarray analysis are reliable.

Gene Ontology (GO) catalogue analysis of the 119 DEGs in the transgenic line showed that GO terms under the category of “molecular functions” before and after Foc TR4 treatment were primarily related to “catalytic activity”, “transporter activity”, and “binding”. Biological processes of the DEGs were mainly related to “metabolic process”, “localization”, “response to stimulus”, and “cellular process” before and after Foc TR4 treatment. Cellular components primarily included the GO terms “cell”, “membrane part”, and “membrane”, with or without Foc TR4 treatment (Figure S4A).

The results of fluorescence showed that Foc TR4 could invade the plant and induce immune response within 7 days after inoculation. We enriched 270 DEGs differentially expressed within 7 days after inoculation by KEGG pathway. The results showed that 14 DEGs were enriched in the MAPK signal pathway and 13 DEGs were enriched in the phenylpropanoid biosynthesis pathway (Figure 8A). Further analysis of phenylpropanoid biosynthesis pathway showed that the gene expression of PAL, mannitol dehydrogenase, and POD 52 were up-regulated in transgenic plants compared with WT plants within 7 days after inoculation (Figure
Overexpression of MpbHLH in Cavendish banana leads to differences in protein expression

To further understand the effect of MpbHLH expression on banana plants, proteomics was adopted for protein-level verifications of transgenic (#DX11) and wild-type plants before *Foc* TR4 inoculation. The analysis results show that 211 proteins were up-regulated and 158 proteins were down-regulated compared to WT plants (Additional File 6). As shown in Figure S5, Go classification analysis results show that the biological process consists of 13 components that are mainly concentrated in the response to stimulus, metabolic process, and cellular process. The cellular component consists of 9 components, mainly concentrated in the cell, organelle, and membrane. The cellular component consists of 10 components, mainly focusing on catalytic activity, binding activity, and transporter activity. The differentially expressed proteins were enriched by protein domain. Further analysis shows that the differentially expressed proteins were mainly located in the glutathione-S-transferase domain, peroxidase domain, cupin domain, and disease-related protein family domain (Figure 8C). After searching and analyzing the proteins contained in the above domains, we found that POD, germin-like protein, and disease related protein were all up-regulated (Table 1). In addition, two PAL were identified and up-regulated about 1.3 times.

Overexpression of *MpbHLH* improves the POD activity of transgenic plants

To investigate whether POD activity contributes to the improved resistance to Fusarium wilt of the transgenic banana plants, POD activity was evaluated before and after *Foc* TR4 treatment. As shown in Figure 9, POD activity of WT plants was significantly lower than that of the transgenic lines, both in roots and leaves. POD activity of WT plant was 2050 (U/g) before *Foc* TR4 treatment, however, the POD activity of #DX11 and #DX13 were 2760 (U/g) and 3980 (U/g), respectively; after *Foc* TR4 treatment, the WT had a POD activity ranging from 1950-2280 (U/g), while the POD activity of the two transgenic lines were significantly higher, 2690-3150
(U/g) for #DX11 and 3080-3420 (U/g) for DX13 (Figure 6B). These results indicate that the resistance of transgenic banana to Fusarium wilt may have improved due to increased POD activity.

**Discussion**

bHLH proteins constitute one of the largest TF families in plants; 177 and 167 bHLH TFs have been identified from the genomes of rice and Arabidopsis [28,29]. Numerous studies have shown that plant bHLH-type proteins are implicated in a number of biological processes such as cell elongation [30], cell specification [31], flowering [32], pollen development [33], hormone signaling response [34], and secondary metabolism [35]. Additionally, increasing evidence indicates that bHLH TFs play critical roles in plant responses to various abiotic stresses, including dehydration stress [25], salt stress [36], and cold stress [37-39]. However, little is known about the roles of bHLH homologs in Dajiao, a very special disease-resistant banana variety in China. Thus, characterization of the function of bHLH genes from Dajiao is pivotal in providing valuable gene candidates for genetic manipulation. In the current study, one putative bHLH-ortholog, MpbHLH from Dajiao, was cloned and functionally analyzed. We observed that overexpression of MpbHLH in Cavendish banana resulted in significant increase in resistance to Fusarium wilt (Figure 5).

Expression profiles further suggest that MpbHLH transcript level is induced by abiotic stresses, including cold treatment, salt, and dehydration (Figure 3). Previous studies described the general rule that transcription factors play superimposed biological roles in regulating biological and abiotic stresses of plants [15, 20-22], which prompted us to investigate their potential function in resistance to Fusarium wilt by generating transgenic plants. Thus, MpbHLH was transformed into Cavendish banana, a Foc TR4-sensitive perennial plant with important agronomic value. Overexpression of MpbHLH in banana markedly reduced the disease index and resulted in better plant phenotypes than that of WT plants after Foc TR4 treatment (Figure 5). Note that growth retardation has been frequently observed in banana plants overexpressing stress-responsive TFs such as MusabZIP53 [43]. However, MpbHLH
overexpression resulted in no obvious phenotypic changes from the WT under normal
growth conditions implying that MpbHLH may hold great potential for genetic
engineering, to improve disease resistance in banana.

As is well known, the response of stress is an overly complex process that is
regulated by multiple signaling pathways [9]. To identify the molecular mechanism of
the MpbHLH to regulate banana’s resistance to Fusarium wilt, transcriptional profiles
were compared between WT and the transgenic line (#DX11) before and after Foc
TR4 treatment. Notably, MpbHLH both induces and suppresses the transcript levels
of many genes, implying both positive and negative impacts on the expression atlas.
Such a phenomenon is not an exception as extensive transcriptional reprogramming
has been demonstrated in many reports comparing global transcript profiles between
transgenic plants overexpressing TF and their WT counterparts [27,44]. As shown in
Additional File 2 and File 3, the number of genes with altered expression levels after
14 days of Foc TR4 treatment (1515 up-regulated and 1816 down-regulated) is
remarkably larger than that for 7 days of Foc TR4 treatment (478 and 123,
respectively). We speculate that the transgenic plant undergoes certain unidentified
modifications upon exposure to biotic stress that result in the expression or
suppression of a set of stress-responsive genes involved in the plant stress response
signaling network. Fusarium wilt of banana is a typical soil-borne vascular disease
[25,45,46]. It was observed by laser confocal microscopy that the hyphae of the
treated plants broke through the cell wall barrier of WT plants and filled the root
vascular bundles. However, the infection of Foc TR4 in transgenic plants was
significantly delayed (Figure 5C). According to transcriptomics and proteomic
analysis, we believe that the up-regulation of phenylpropanoid synthesis-related genes
and proteins in transgenic plants plays a key role in this process. Phenylpropane
metabolism is an important pathway of plant secondary metabolism, which directly or
indirectly produces all the phenylpropanoid skeleton materials in plants, such as lignin,
phytoalexin, flavonoids, and so on [47,48]. PAL is the first enzyme in the
phenylpropanoid metabolic pathway [49]. It can catalyze the deamination of L-Phe to
trans-cinnamic acid, which is the direct or indirect precursor of many secondary
metabolites [50]. Barros et al. (2016) found that PAL was involved in nearly half of
the lignin synthesis through its Tyrosine ammonia lyase (TAL) activity, and lignin is an important part of the vascular bundle [51]. Shadle et al. (2003) found that overexpression of PAL not only increased lignin content, but also reduced growth rate and delayed flowering [52]. Although fewer Foc fluorescence signals were observed in the root cross section of transgenic plants than in WT plants, a large number of brown spots still appeared in transgenic plants after Foc treatment 14 days (Figure 5C). These blotches may indicate immune response and ROS production. Liu et al. (2020) showed that as a kind of semi-saprophytic fungus, the fusaric acid secreted by Foc could be transported by plant vascular bundles before colonization, which caused plant immune response and ROS outbreak and further led to banana cell apoptosis [25]. Our transcriptomic data indicates that the transgenic plants have several peroxidase genes up-regulated before and at the early stage after Foc TR4 inoculation compared with wild-type plants. This was further confirmed by proteomic analysis results (Table 1). A large number of studies have shown that POD can remove excessive ROS from cells and maintain intracellular ROS levels. This is an important defense mechanism for plants to adapt to biotic and abiotic stresses [53-55]. The following physiological experiments showed that the POD activity of transgenic plants, especially in roots, was higher than that of wild-type plants before inoculation, 7 days and 14 days after inoculation (Figure 9). Therefore, we speculate that the transgenic plants overexpressing MpHLH may enhance the ROS scavenging ability of transgenic banana roots by increasing POD activity, thus enhancing the resistance of transgenic banana to Foc TR4. We also noticed that several germinatin-like proteins and disease-related proteins were up-regulated in transgenic plants before Foc infection (Table 1). Disease-related proteins have been proved to be related to plant disease resistance in many crops [56,57]. The study of germinating protein showed that there are antimicrobial protein sub-families in germinating protein. The expression of this protein can be induced by pathogen attack as well as exogenous hydrogen peroxide, and these proteins can not only regulate host resistance to pathogens but also possess superoxide dismutase (SOD) activity in vitro [58]. In addition, overexpression of rice germinatin-like protein gene in tobacco showed tolerance to high concentration of H₂O₂ and enhanced resistance to Fusarium and
Our data demonstrates that MpbHLH is a novel transcription factor from Dajiao that plays a role in positively regulating banana’s resistance to Fusarium wilt. Overexpression of MpbHLH in Foc TR4-sensitive banana can improve the resistance of transgenic banana to Fusarium wilt by regulating the expression of multiple resistance-related genes including phenylpropanoid pathway-related genes and oxidative stress-related genes. These findings not only reveal the fact that at a molecular level MpbHLH transcription factor positively regulates banana’s resistance to Fusarium wilt, but also guides the design of strategies to improve disease resistance in banana and possibly other agricultural crops and is of important theoretical and practical significance.

**Materials and methods**

**Plant materials, growth conditions and stress treatment**

Dajiao (*Musa* spp. ‘Dajiao’; ABB group; Accession No. LCDJ_01) was collected from the national banana germplasm repository, Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences. For cold treatment, the potted plantlets were grown up to the five-leaf-stage in a growth chamber at 30/28 °C (day/night) under a photon flux density of 240 μmol m⁻² s⁻¹ throughout a 12 h photoperiod and at a relative humidity of 60–80%. Plants with a uniform growth status were then transferred to a chamber set at 10 °C for 0, 1, 3, 6, 24, and 48 h in the dark. For dehydration stress, one-month-old rooted plantlets were put on filter paper above a workbench and dried at 25 °C (with a relative humidity of 44.0 %) for 0, 1, 3, 6, 24, and 48 h in an ambient environment with natural daylight. Salt treatment was applied by placing the one-month-old rooted plantlets in 200 mM NaCl solution for 0, 1, 3, 6, 24, and 48 h in an ambient environment with natural daylight. Plants without any treatment were regarded as controls (0 h). The fully expanded leaves on the top of the plants/plantlets were collected at the indicated time points for each treatment, immediately frozen in liquid nitrogen, and stored at −80 °C for further analysis. All samples from each time point were collected in triplicate.
Gene isolation and sequence analysis

Total RNA derived from Dajiao leaves was isolated using a plant RNA extraction kit (Code No. 9767, TaKaRa, Dalian, China), and first strand cDNA was synthesized from 1 µg of total RNA using a PrimeScript RT Reagent Kit, according to the manufacturer’s instructions (TaKaRa). Based on the cDNA sequences of *MpbHLH* acquired from the Dajiao transcriptome data [61], partial Dajiao bHLH cDNA fragments were amplified by PCR using primers containing SpeI and BamHI restriction sites (GSP1, Table S1). Chromosomal location prediction was performed by the BLAT server (http://banana-genome.cirad.fr/blast) on the Genoscope Genome Browser. The molecular weights (kDa) and isoelectric points (pI) of the corresponding proteins were predicted using the ProParam tool (http://www.expasy.ch/tools/pi_tool.html). Homology searches were conducted using the NCBI BLAST server (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Construction of the molecular model of MpbHLH was performed by the SWISS-MODEL server (http://swissmodel.expasy.org). Sequence alignment was carried out using ClustalW, and the phylogenetic tree was constructed using MEGA 5.1 software.

Quantitative real-time (RT-qPCR) PCR analysis

The primer pairs used for RT-qPCR are listed in Table S1 (GSP2, GSP7-GSP12). The PCR reaction consisted of 10 µL of 2xSYBR Green PCR Master Mix (TaKaRa), 200 nM primers, and 2 µL of 1:40-diluted template cDNA in a total volume of 20 µL. The RT-qPCR program included an initial denaturation step at 95 °C for 10 s, followed by 40 cycles of 5 s at 95 °C, and annealing and extension at 60 °C for 20 s. The *MaACT1* gene (GSP3, Table S1) was selected as a reference gene, as per our previous report [9]. The relative expression levels of the amplified products were calculated based on the comparative threshold cycle method [9]. Transcript abundance was normalized against the reference gene.

Sub-cellular localization of MpbHLH protein

The full-length cDNA of MpbHLH was isolated as described above and subcloned into the pMD18-T vector (TaKaRa). To determine the subcellular localization of
MpbHLH, plasmid pMD18-T-containing *MpbHLH* was amplified using primers GSP4 (Table S1) containing *SalI* and *ClaI* restriction sites. The amplicon was digested with these enzymes and subcloned into the pUC19-GFP vector, containing the GFP reporter gene, to produce the fusion construct pUC19-MpbHLH-GFP under the control of the cauliflower mosaic virus 35S promoter (CaMV 35S). The fusion construct and the control vector (pUC19-GFP) were separately transformed into rice protoplasts, with reference to the previous reports [62]. Images were collected from transiently transformed rice protoplast cells grown at 28 °C using a confocal laser-scanning microscope (LAM510, Carl Zeiss GmbH, Jena, Germany). Captured images were analyzed by Image-Pro software.

**Generation of transgenic plants by *Agrobacterium tumefaciens*-mediated transformation**

*MpbHLH* cDNA was amplified with specific primers (GSP1) containing *SpeI* or *BamHI* restriction sites. The PCR product was confirmed, digested with *SpeI* and *BamHI*, and then ligated into the *SpeI/BamHI*-linearized binary vector pOx driven by the Zea mays polyubiquitin promoter (Figure S2A). The constructed binary vector is denoted as pOx-MpbHLH. The newly constructed pOx-MpbHLH plant expression vector was transferred into *A. tumefaciens* strain EHA105 by heat shock (Dou et al. 2016). As shown in Figure S2B, the overexpression vector was used for transformation of Cavendish banana (*Musa* spp. Cavendish; AAA Group; Accession No. BX_01) embryogenic cell suspensions (ECSs) using the protocol, as described previously [9, 63]. All the biochemical agents used in the tissue culture process are listed in Table S2. Hygromycin-resistant plants were identified by PCR using special primers (as shown in Table S2, GSP5). Positive banana plants were multiplied vegetatively using meristems of in *vitro* plantlets. The rooted plantlets were hardened in the greenhouse and used for further assays.

**Disease resistance test of the transgenic plants**

The *Foc* TR4-GFP was provided by professor Zheng Shijun (Yunnan Academy of Agricultural Sciences). The *Foc* TR4-GFP isolate was grown on a PDA medium for 7
days at 28 °C. A small agar portion containing the mycelial growth was inoculated in PDB and allowed to grow for 2 days at 28 °C. The fungal mycelium was sieved out using four layers of cheese cloth. Conidiospores present in the filtrate were centrifuged and washed with sterile distilled water, and subsequently, the spore density was adjusted to $2 \times 10^5$ spores/mL [64]. To evaluate the resistance of transgenic banana plants to pathogen *Foc* TR4, two-month-old transgenic and WT plants (20 plants for each line) were treated with root injury and inoculated in the above-mentioned *Foc* TR4 suspensions with GFP markers, and then inoculated and planted in the nutrition cups. The plants followed normal fertilizer and water management, and disease incidence rate was observed and summarized every day using the method stipulated in our previous report [64]. After 14 days, the plants were cut longitudinally along the root bases to evaluate the disease index of the diseased plants based on the areas of disease spots. The grading criteria of the index of disease is as follows: grade 1 (browning area less than 5%); grade 2 (browning area between 5% and 25%); grade 3 (browning area between 25% and 50%); grade 4 (browning area between 50% and 75%); and grade 5 (browning area more than 75%) [46]. The experiment was repeated three times, and the representative phenotypes were photographed.

**Physiological measurement and laser confocal observation**

Two transgenic lines (#DX11 and #DX13) and WT plants were inoculated at a density of $2 \times 10^5$ spores/mL *Foc* TR4-GFP. POD activity of banana roots and leaves were measured before and after *Foc* TR4-GFP treatment (three top leaves of each plant were used, with five plants for the WT or transgenic lines). POD activity was tested using the protocol as previously described [65-66], which was then slightly modified. To observe the infection and localization of pathogenic bacteria in the root system, the roots/bulbs of wild-type and transgenic plants were observed using a laser confocal microscope (Zeiss LSM710) 0 days, 3 days, 7 days, 10 days, and 14 days after inoculation. The experiment was repeated three times, and the representative phenotypes were photographed.
**RNA sequence**

Transcriptional profiling of WT banana and transgenic line #DX11 was carried out by RNA-Seq analysis at BGI Genomics, BGI-Shenzhen, Shenzhen, Guangdong, P.R. China. Three biological replicates were used for each genotype, the WT, and #DX11, under normal growth conditions, while three biological replicates were used for each genotype before and after *Foc* TR4 treatment (7d and 14d). RNA preparation, library construction, and sequencing on the BGISEQ-500 platform of PE 100 with 30 million reads per sample were performed at BGI (www.genomics.org.cn, BGI, Shenzhen, China). Gene expression levels were quantified using the RSEM software package [67]. To further validate the reliability of this transcriptomic data, 6 candidate genes were assayed by RT-qPCR with specific primers (GSP-11), as shown in Table S1.

The DEGseq method was used to screen differentially expressed genes (DEGs) between groups, as previously described [68]. Gene Ontology (GO) and pathway annotation and enrichment analyses were based on the GO Database (http://www.geneontology.org/) and the KEGG pathway database (http://www.genome.jp/kegg/), respectively. Statistical analysis was performed and DEGs were selected with the criteria of a fold change ≥ 2 and adjusted p-value ≤ 0.001, as described previously [68].

**Proteomic analysis**

To determine the constitutive expression of transgenic plants at the protein level under normal growth conditions, banana leaves of two-month-old transgenic line (#DX11) and WT plants were ground into powder and transferred to 5 ml centrifuge tubes, added with 3 ml of lysis buffer (containing 40 mM Tris, 5M urea, 2M thiourea, 50 mM dithiothreitol, 1% sodium dodecyl sulfate, 2% CHAPS, 2% amphoteric carrier, at pH 3-10), and then centrifuged at 12,000 rpm for 15 min at 4 °C. The supernatant was removed, and the remaining precipitates were precipitated with 20% TCA at -20 °C for 2h. After centrifugation at 12,000 rpm at 4 °C for 10 min, the supernatant was removed, and the remaining precipitates were rinsed with cold acetone three
times. Finally, 1 mg of protein sample was added along with 1.0 ml of 1% SDS, to
determine the protein concentration after vacuum concentration and extraction of
acetone. The above protein solution was reduced with 5 mM dithiothreitol at 56 °C
for 30 minutes, and then alkylated with 11 mM iodoacetamide in the dark for 15
minutes. After that, the protein sample was diluted to 100 mM TEAB with urea
concentration lower than 2M, and then added with trypsin at the mass ratio of 1:50
trypsin/protein for the first digestion overnight. Finally, trypsin with the mass ratio of
1:100 trypsin/protein was added for the second time for 4h. The digested protein was
labeled with tandem mass spectrometry labeling kit (Thermo, No. 90062), and liquid
chromatography and mass spectrometry were conducted using the methods proposed
by Jiang et al. (2017) [69].

Statistical analysis
Experiments were performed according to a complete randomized design (CRD). The
data reported in the figures are means of the values with standard error (SE) and were
examined statistically by ANOVA. The least significant difference at the 5% level was
analyzed by DPS software (version 3.01; Zhejiang University, Hangzhou, China).

List of abbreviations
bHLH, Basic helix loop helic; Cef, Cefotaxime; DEGs, differentially expressed genes;
ECS, Embryogenic cell suspension; GFP, Green fluorescent protein; hpt, Hygromycin;
IAA, Indole-3-acetic acid; MW, Molecular weight; NAA, Naphthylacetic; PCR,
Polymerase chain reaction; PAL, Phenylalanine ammonia lyase; POD, Peroxidase; PI,
Isoelectric point; qPCR, Real-time quantitative; ROS, Reactive oxygen species;
TAL : Tyrosine ammonia lyase; TF, Transcription factor; WT, Wild type; 2, 4-D, 2,
4-Dichlorophenoxyacetic acid; 6-BA, 6-benzyl adenine

Ethics approval and consent to participate
Not applicable.
Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interests.

Funding

This work was supported by the Natural Science Foundation of Guangdong Province, No. 2019A1515012120; National Natural Science Foundation of China, Nos. 31801843 and 31872939; Guangdong Academy of Agricultural Sciences Foundation, No. 201815; Tutor system project of Guangdong Academy of Agricultural Sciences, No. R2018QD-018.

Authors' contributions

H-C.L. and T-X.D conceived the study and analyzed the data. T-X.D., S-P.W., C-H.H., A-F.X., O.S., F-C.B., G-M.D., C-Y.L., T.D., H-J.G. and W-D.H performed the experiments. T-X.D drafted the manuscript. G-J.Y and Q-S.Y revised the manuscript.

Acknowledgements

The authors heartily appreciate Shijun Zheng for his generous offer of Pathogen with fluorescent label.

Availability of data and materials

The mass transcriptome data have been deposited to the interactive reporting system (Website: https://dataview.ncbi.nlm.nih.gov/object/PRJNA681292).
Username: doutongxin555; Password: Kk123456
Accession Number: PRJNA681292

The mass proteomic data have been deposited to the interactive reporting system (Website: http://www.ebi.ac.uk/pride).
Username: reviewer_pxd023313@ebi.ac.uk; Password: lgpbzBHA

Project accession: PXD023313

**Reference**


Figure Legends

Figure 1. Gene structure and locus of MpbHLH. (A) Gene structure and locus. The chromosomal localization of MpbHLH predicted in M. acuminata and M. balbisiana. Compared with the M. acuminata genome, MpbHLH contains four exons spanning 2998 bp and has a predicted Open reading frame (ORF) of 1677 bp. MpbHLH is mapped to Chromosome 10 at position 21218882-21221879. Compared with the M. balbisiana genome, MpbHLH also contains four exons spanning 2991 bp and has a predicted ORF of 1680 bp. MpbHLH is mapped to Chromosome 10 at position 14056264-14059254. Nucleotide numbers are indicated under the gene structure. (B) Schematic protein structures of MpbHLH with the N-terminal domain containing a bHLH domain. Codon numbers are indicated under the protein structures. (C) The predicted tertiary structure of MpbHLH evaluated and obtained using the SWISS-MODEL server.

Figure 2. Characterization of MpbHLH. (A) A protein sequence alignment of MpbHLH and other bHLH transcription factors from different species. Accession numbers are as follows: AtICE2 (BAC42644), CbICE1 (AAS79350), PtICE1 (ABN58427), PsICE1 (ABF48720), MdICE1 (ABS50251), GmICE1 (ACJ39211), GmICE2 (ACJ39212), AtICE1 (AAP14668), GmICE3 (ACJ39213), GmICE4 (ACJ39214), RcICE1 (EEF51703), OsICE1 (BAD88163), TaICE41 (ACB69501), TaICE87 (ACB69502), ZmICE2 (ACG46593), OrbHLH2 (ABAB93991), and MpbHLH (KM379133). Identical residues are shown in black, conserved residues in dark gray, and similar residues in light gray. The highly conserved bHLH and zipper domains are marked. (B) Phylogenetic analysis of MpbHLH. The phylogenetic tree was constructed using the DNAMAN program. The deduced amino acid sequence, with GenBank accession numbers listed in parentheses is as follows: AtICE2 (BAC42644), CbICE1 (AAS79350), PtICE1 (ABN58427), PsICE1 (ABF48720), MdICE1 (ABS50251), GmICE1 (ACJ39211), GmICE2 (ACJ39212), AtICE1 (AAP14668),
GmICE3 (ACJ39213), GmICE4 (ACJ39214), RcICE1 (EEF51703), OsICE1 (BAD88163), TaICE41 (ACB69501), TaICE87 (ACB69502), ZmICE2 (ACG46593), OrbHLH2 (ABA93991), and MpbHLH (KM379133).

**Figure 3.** Time-course expression analysis of *MpbHLH* under various stress treatments. (A) Expression of *MpbHLH* in Dajiao under cold stress for 0, 1, 3, 6, 24, and 48 h. (B) Expression of *MpbHLH* in Dajiao under salt stress for 0, 1, 3, 6, 24, and 48 h. (C) Expression of *MpbHLH* in Dajiao under dehydration stress (25 °C and relative room humidity of 44%) for 0, 1, 3, 6, 24, and 48 h. Error bars for qPCR analysis stand for SD based on three replicates.

**Figure 4.** Subcellular localization analysis of MpbHLH. The MpbHLH-GFP fusion protein and GFP (used as a control), driven by the CaMV 35S promoter, were separately transformed into rice protoplasts and visualized by fluorescence microscopy. Images were taken of representative cells expressing GFP. From left to right: cell morphology in the bright field, spontaneous fluorescence map of chloroplasts, dark field figure of GFP fusion proteins, and a combined figure of GFP fusion proteins and spontaneous fluorescence of chloroplasts (bar:10 µm).

**Figure 5.** *Foc* TR4-resistance assay of transgenic banana plants. (A) Phenotypes of transgenic banana plants (#DX11, and #DX13) and the WT before and after *Foc* TR4 treatment (14 d). (B) Disease index distribution of WT banana and transgenic plants analyzed after *Foc* TR4 treatment for 14 d. (C) Laser confocal observation of transgenic banana plants (#11 and #13) and the WT before and after *Foc* TR4 treatment (24 h, 7 d and 14 d). R, Root.

**Figure 6.** Transcriptomic analysis of WT banana and a transgenic line (#DX11). Three independent biological replicates were collected for each of the genotypes, WT, or #DX11, before and after *Foc* TR4 treatment. (A) DEGs identified among WT and transgenic line #DX11 before and after *Foc* TR4 treatment (7 and 14 d). (B) Thermogram of 119 differentially expressed genes of WT and transgenic line #DX11 before and after *Foc* TR4 treatment (7 and 14 d). The color key represents (FPKM+1) normalized log2 transformed counts. Red and green indicate high- and low-level gene expression, respectively. Each column represents a treatment and each row represents a gene.
Figure 7. Real-time quantitative RT-PCR analysis of the expression of six differentially expressed genes before and after *Foc* TR4 treatment (7 and 14 d).

Figure 8. Differential genes and differential protein analysis. (A) The differentially expressed genes were enriched according to KEGG classification. (B) Phenylpropanoid biosynthesis differentially expressed genes cluster analysis. (C) The differentially expressed proteins were enriched according to the protein domain.

Figure 9. POD activity of transgenic banana plants (#DX11 and #DX13) and WT before and after *Foc* TR4 treatment (7 and 14 d).

Table 1. Up regulated protein expression.

Supplemental Table S1. Primers used in this study.

Supplemental Table S2. Composition of various medium used for regeneration and transformation experiments.

Supplemental Figure S1. Open reading frame (ORF) sequence of *MpHLH* along with its predicted protein sequence.

Supplemental Figure S2. Generation of transgenic Cavendish banana plants using *A. tumefaciens*-mediated transformation.

Supplemental Figure S3. Generation and molecular identification of transgenic banana plants overexpressing *MpHLH*.

Supplemental Figure S4. Transcriptomic analysis of WT and a transgenic line (#DX11).

Supplemental Figure S5. Proteomics analysis of WT and a transgenic line (#DX11).


Additional File 4: Genes with significant differences in WT and a transgenic line
(#DX11) before and after *Foc* TR4 treatment (7 and 14 d).

**Additional File 5:** Six candidate genes selected to verify the reliability of transcriptome data using the qRT-PCR method

**Additional File 6:** Up and down-regulated proteins in the proteomic data.