Molecular cloning and subcellular localization of *GLOBOSA* gene in *Nicotiana tabacum*

Cui Fangfang [1] Zheng Yun [1] Deng Lingfan [1] Yang Xiangfei [1] Zheng Jiuzhou [1] Meng Linfeng [1] Wang Jiange [2*] Liu Qiyuan [1]

1 The Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Ministry of Education, College of Agriculture, Jiangxi Agricultural University, Nanchang 330045;

2 College of Forestry, Jiangxi Agricultural University, Nanchang 330045

Corresponding author, wjgsd@126.com

Abstract Cytoplasmic male sterility is widely used for tobacco hybrid production. In order to study the mechanism of cytoplasmic male sterility (CMS) with carperlloid stamens, the *GLOBOSA* gene was cloned and sequenced from the flower buds of tobacco K326. The physicochemical properties ,functional domains and subcellular localization of its coding protein was predicted. The phylogenetic relationship was analyzed based protein sequences by maximum likelihood method. The recombinant expression vector of *GLOBOSA* gene and green fluorescent protein expression (GFP) was constructed, and transformed into *agrobacterium tumefaciens* and directly injected into tobacco leaves for transient expression. 36h later after the injection, the subcellular localization of *GLOBOSA* gene were observed by fluorescence microscope. The results showed that *GLOBOSA* was 768bp in length, coding 203aa. It was a hydrophobic protein with two domains of SRF-TF and K-box, and had no transmembrane region and signal peptide. Fluorescence microscopy showed that the protein was localized in cytoplasm and nucleus, and more was in the nucleus.

Keywords Tobacco; *GLOBOSA* genes; Cloning; Subcellular localization; Cytoplasmic male sterility

Cytoplasmic male sterility refers to the phenomenon in which plants cannot produce functional pollen but have normal pistils, exhibiting maternal inheritance and commonly occurring in flowering plant populations (Wang et al., 2019) . Cytoplasmic male sterility in many crops is a tool for utilizing heterosis . Phenotypically , in addition to pollen abortion, cytoplasmic male sterility presents a series of related changes:

* Corresponding author, wigsd@126.com

Funding: Supported by the National Natural Science Foundation of China (31960418)

pollen abortion, anther degeneration, stamen carpelization, and stamen petalization (Fan et al., 2016) . Studies have shown that environmental factors, natural variation, heterocytoplasm, and cytoplasmic fusion can all lead to cytoplasmic male sterility (Fang et al., 2019) , and sterility can be maintained or restored by nuclear-encoded genes (Wang et al., 2021) .

To date, cytoplasmic male sterility (CMS) has been studied in 150 species (Zhang, 2021) . Molecular biology research has found that mitochondrial DNA rearrangements are frequently observed in all sterile lines, producing new open reading frames (ORFs) or chimeric ORFs, leading to cytoplasmic male sterility (Wang et al., 2020) . However, not all of these variations are associated with the cytoplasmic male sterility phenotype, as some ORFs associated with CMS are hypothetical or chimeric. The petalization and carpelization of stamens in cytoplasmic male sterility are very similar to the flower developmental homeomorphism caused by nuclear genes, and are also referred to as cytoplasmic homeomorphism (Yao et al., 2021) . This type of cytoplasmic male sterility cannot form anthers and therefore cannot produce pollen. In cytoplasmic male sterile lines with carpel-like and petal-like stamens across multiple species, not only were mitochondrial DNA rearrangements detected (Liu, 2019) , but also downregulation of class B or C genes was observed (Chen et al., 2018; Linke et al., 2003; Rijpkema et al., 2006) . However, how cytoplasmic signaling retrogradely regulates cellular gene expression remains unclear.

Our research group, through years of continuous backcrossing, has cultivated the cytoplasmic male sterile line tobacco K326 with carpel-like stamens. Sequencing revealed downregulation of the class B gene *GLOBOSA* . Downregulation of class B genes is also present in cytoplasmic male sterile lines with carpel-like stamens in many species. Therefore, what impact would overexpression of class B genes, leading to restoration of the carpel-like phenotype, have on male sterility? Gene cloning and subcellular localization are prerequisites for studying gene function. Therefore, the objectives of this study are: 1) What are the physicochemical properties of the GLOBOSA protein? 2) Where is the encoded protein located in the cell? This research will provide data for elucidating the mechanism of cytoplasmic male sterility in tobacco.

1. Results and Analysis

1.1 Tobacco CLOBOSA Sequence Analysis

The tobacco *GLOBOSA* gene sequence is 768 bp, with a coding region of 635 bp, encoding a total of 209 amino acids (Figure 1). The molecule has a mass of approximately 24 kD , a theoretical isoelectric point (pI) of 8.94, and contains a total of 3,451 atoms ($C_{1060}H_{1729}N_{325}O_{328}S_{13)}$. Among the 20 amino acids that make up the protein, glutamic acid (Glu) has the highest proportion at 9.1%, while cysteine (Cys) has the lowest proportion at 0.5%. The instability index is 56.79, classifying it as an unstable protein, while the lipid index is 77.89, and the overall average hydrophilicity is -0.890, classifying it as a hydrophilic protein.

```
ATGGGAAGAGGAAAGATAGAGATCAAAAGAATAGAGAACTCAAGCAACAGGCAGGTCACT   60
 M  G  R  G  K  I  E  I  K  R  I  E  N  S  S  N  R  Q  V  T
TACTCAAAAAGAAGAAATGGGATCTTGAAAAAAGCTAAGGAAATCAGTGTTCTTTGTGAT  120
 Y  S  K  R  R  N  G  I  L  K  K  A  K  E  I  S  V  L  C  D
GCTCGTGTTTCTGTCATCTATTTTGCTAGTTCTGGCAAGATGCATGAGTTCTCCTCTACT  180
 A  R  V  S  V  I  I  F  A  S  S  G  K  M  H  E  F  S  S  T
TCGTTGGTTGATATTTTGGATCAATACCACAAGCTTACTGGAAGAAGATTGTGGGATGCT  240
 S  L  V  D  I  L  D  Q  Y  H  K  L  T  G  R  R  L  W  D  A
AAGCATGAGAACTTGGACAATGAAATCAACAAAGTCAAGAAAGACAATGACAACATGCAA  300
 K  H  E  N  L  D  N  E  I  N  K  V  K  K  D  N  D  N  M  Q
ATTGAACTCAGGCACCTAAAGGGTGAAGACATCACATCTTTGAACCACAGAGAGCTCATG  360
 I  E  L  R  H  L  K  G  E  D  I  T  S  L  N  H  R  E  L  M
ATGTTGGAAGATGCCCTTGATAATGGACTCACTAGTATCCGTAACAAGCAGAATGACCTT  420
 M  L  E  D  A  L  D  N  G  L  T  S  I  R  N  K  Q  N  D  L
CTGAGGATGATGAGGAAAAAGACTCAAAGTATGGAGGAGGAGCAAGACCAACTTAATTGG  480
 L  R  M  M  R  K  K  T  Q  S  M  E  E  E  Q  D  Q  L  N  W
CAATTGCGGCAACTAGAGATAGCAAGCATGAATAGGAATATGGGAGAAATAGGGGAAGTG  540
 Q  L  R  Q  L  E  I  A  S  M  N  R  N  M  G  E  I  G  E  V
TTTCACCAAAGGGAGAATGAATACCAAACTCAGATGCCTTTTGCCTTCCGAGTTCAGCCA  600
 F  H  Q  R  E  N  E  Y  Q  T  Q  M  P  F  A  F  R  V  Q  P
ATGCAGCCTAATTTGCAGGAGAGATTCTAAACACCTTGATTTATTTGGTGAATTATGACC  660
 M  Q  P  N  L  Q  E  R  F  *
TTTTATTAATGTTGTGCTATCAAAAAACTTGGTGTGTAATATCAATACTCATCTACCTTA  720
GCTTATGGTTTAAGTGACATTAATTATATATTTATCTATAACCATTAC               768
```

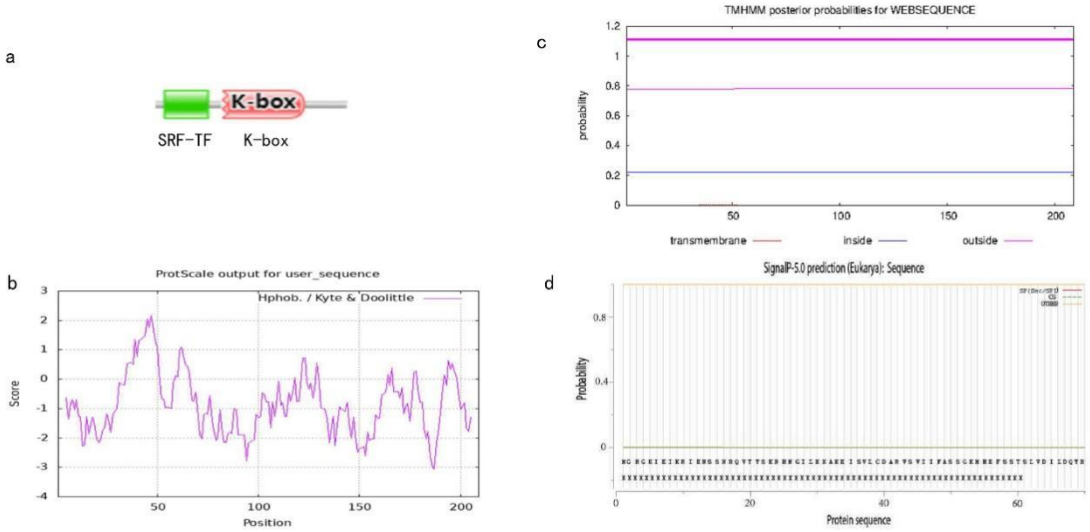
Figure 1. Sequence of tobacco GLOBOSA gene and its encoded protein

Figure 1 Nucleotide sequence and its coding protein sequence of *CLOBOSA* gene in *Nicotiana tabacum*

1.2 Analysis of the structure of the tobacco GLOBOSA gene

Structural analysis of the tobacco *GLOBOSA gene* (Figure 1) . It can be seen that the tobacco *GLOBOSA* gene has two domains: SRF-TF (PF00319) and K-box (PF01486). The SRF-TF domain is located at positions 10–57 aa, and the K-box domain is located at positions 75–164 aa. The SRF-TF domain protein has two antiparallel facultative α-helices, which can dimerize the protein and determine the binding of transcription factors to specific DNA (Ratcliffe et al., 2001). The K-box domain protein sequence has three α-helices, forming affinity helices within the protein dimer, which may be related to protein-protein

interactions (Melzer et al., 2010). MADS-box genes are widely distributed in animals, plants, and fungi, participating in the regulation of growth and development processes; the *GLOBOSA* gene is one such member. In the four-factor model of flower development, the ability to form polymers is fundamental to floral allotropy and determines floral organ specificity (Hanano et al., 2011). *The GLOBOSA* gene structure also suggests that these polymers may bind to target genes during flower development, participating in floral organ formation. (This information was obtained through ExPASy...) ProtScale results indicate the absence of hydrophobic regions in the encoded protein (Figure 1), consistent with TMHMM analysis. No transmembrane regions were found in the Figure 1). SignalP-5.0 results show the absence of a signal peptide in the encoded protein (Figure 1), and Cell -PLoc 2.0 prediction indicates its subcellular localization to the nucleus. Therefore, *GLOBOSA* is a MIKC $^{C-}$ type MADS-box gene with SRF-TF and K-box domains, capable of dimerizing to form a complex that binds to specific DNA and participates in gene regulation. The encoded protein lacks hydrophobic regions and a signal peptide, and is located in the nucleus.



a:预测的GLOBOSA结构域；b：预测的GLOBOSA疏水性；c：预测的GLOBOSA跨膜域；d：预测的GLOBOSA信号肽。

a:Predicted motifs of GLOBOSA;b:Predicted hydrophobicity of GLOBOSA;c:Predicted transmembrane domain of GLOBOSA;d:Predicted dicted signal of GLOBOSA.

Figure 1 Tobacco *GLOBOSA* gene analysis

Figure 2. An analysis of *GLOBOSA* gene i in *Nicotiana tabacum*

1.3 System Evolution Analysis

Using the tobacco *GLOBOSA* sequence, protein data from six other species were scanned. Under the criteria of a threshold of 1e-10 and a score greater than 120, ten homologous protein sequences were obtained. After Hmmer identification and removal of sequences containing only SRF-TF but no K-box, eight *GLOBOSA* homologous sequences were finally obtained. These sequences, along with the *GLOBOSA* sequences, were used to construct a phylogenetic tree based on maximum likelihood using IQTREE, as Figure 2shown in Figure 3. It can be seen that the basal group *Cinnamomum camphora* is at the outermost end, followed by the monocotyledonous rice , then the eudicotyledonous * Arabidopsis thaliana*, with potato and tomato at the other end. In the phylogenetic tree, Solanaceae plants cluster together, with *Nicotiana benthamiana* and *Tobacco Bungei* initially clustering together, while potato and tomato are at the top of the evolutionary hierarchy. Although the phylogenetic tree was constructed using *GLOBOSA* homologous protein sequences, the relationships reflected are consistent with the evolutionary positions of the species. The figure also shows that rice and * Nicotiana benthamiana* each have two homologous protein sequences, reflecting gene duplication events during evolution.

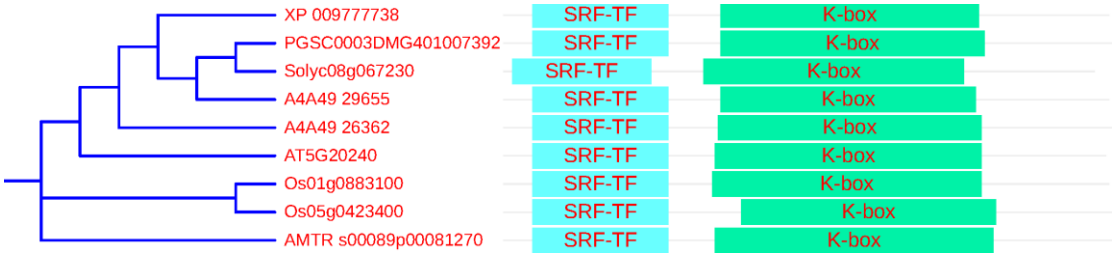


Figure 2 *GLOBOSA* phylogenetic tree

Figure 3 Phylogenetic tree of *GLOBOSA*

1.4 Construction of tobacco GLOBOSA vector

The expression vector pCAMBIA1302 is an expression vector carrying the GFP gene. It has a strong 35S promoter before its multiple cloning site and carries hygromycin (HygR) and kanamycin (Kan) resistance genes, making it a good vector for gene expression studies. The *GLOBOSA* gene was inserted between the NcoI and SpeI restriction sites to construct a recombinant expression vector carrying the 35S: *GLOBOSA* :GFP gene (Figure 3) . During vector construction, the stop codon needs to be removed, and protective bases are added. Therefore, the target gene on the vector cannot be cleaved by enzymes due to the

altered enzyme recognition site at the *GLOBOSA gene insertion position. Thus, PCR amplification was used to obtain the target gene. After successful vector construction,* the vector was verified by double digestion with Figure 3NcoI / BsiWI . The verification results are shown in Figure 4b. It can be seen that the recombinant expression vector initially showed one band, and after enzyme digestion, two bands appeared, indicating successful construction of the recombinant expression vector. Before transient expression, the recombinant expression vector needs to be transformed into Agrobacterium. Positive clones are selected and PCR is used to verify whether the target gene has been successfully transformed into Agrobacterium. The results Figure 3are shown in Figure 4c. As can be seen from the figure, the target gene band is as expected, indicating that the recombinant expression vector carrying the target gene has been successfully transformed into Agrobacterium.

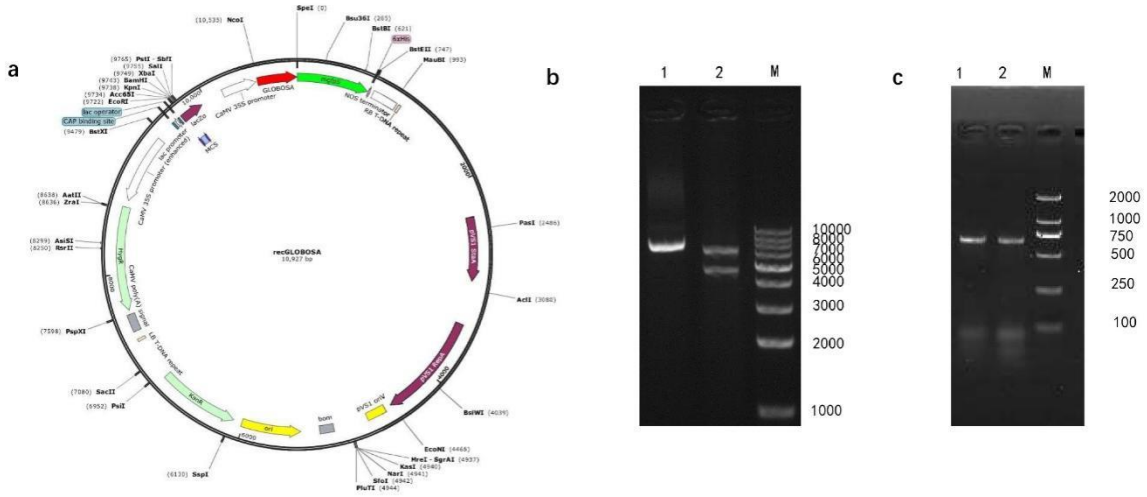


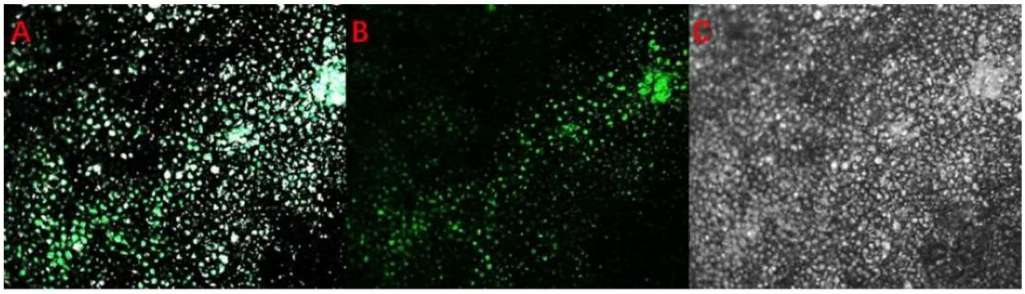
Figure 3 A: Recombinant expression vector containing *GLOBOSA* ; B: Recombinant expression vector enzyme digestion verification ; C: Recombinant expression vector PCR verification.

Figure 4 A: Recombinant expression vector containing *GLOBOSA* ; B: Validation of recombinant expression vector by restriction enzyme digestion ; C: Validation of recombinant expression vector by PCR

1.5 Transient expression and subcellular localization of the target gene

The function of a gene is closely related to the subcellular localization of its encoded product. In the recipient, the reporter gene GFP expressed on the recombinant expression vector emits green fluorescence, which can be located using fluorescence microscopy to determine the subcellular localization of the target gene. When tobacco K326 reached the 6-leaf stage, Agrobacterium tumefaciens with the successfully

introduced gene recombinant expression vector was injected into the tobacco leaves. After 36 days... Transient expression results can be observed after h (picture 4) . In Cell- PLoc prediction, the *GLOBOSA* gene is located in the cell nucleus. As can be seen, there is strong fluorescence in the cell nucleus, indicating that the target gene is subcellularly located within the cell nucleus, consistent with Cell- PLoc prediction.



A：合并图像；B：荧光图像；C：明场。
a: Merged image; b: Fluorescence image; C: Bright-field image

picture 4 Subcellular localization of *GLOBOSA*

Figure5 Subcellular localization of *GLOBOSA*

2 Discussion

investigated the *GLOBOSA gene* in the K326 tobacco line, a male-sterile line with carpel-derived cytoplasm . The gene contains a total of 768... bp, encoding 209 amino acids, the encoded protein is mainly composed of two domains, SRF-TF and K-box, has no transmembrane region, no signal peptide, and is subcellularly located on the cell nucleus.

The formation mechanism of cytoplasmic male sterility is a question worthy of exploration. Cytoplasmic male sterility refers to the inability to produce functional pollen. Although the manifestations differ, they are all maternally inherited, indicating that they are influenced by cytoplasmic genes. Mitochondrial gene rearrangements have been found in all types of cytoplasmic male sterility (Yan et al. , 2022) , which supports this view. Among cytoplasmic male sterility, pollen abortion and anther abortion are types of cytoplasmic male sterility in many crops , while cytoplasmic male sterility caused by stamen degeneration, stamen petalization, and stamen carpelization is another type of sterility caused by the inability to form male organs . Cytoplasmic male sterility lines with stamen carpelization and stamen petalization not

only show mitochondrial DNA rearrangements but also show downregulation of class B or C genes. According to the flower ABC model, the interaction between class B genes and class A genes determines petal development, while the interaction between class B genes and class C genes determines stamen development (Yang et al., 2017) . In Arabidopsis , B functional genes also act as transcriptional activators. The GLO/DEF heterodimer directly binds to the CAR -ArG motif, leading to promoter activation (Mahajan and Yadav, 2014) . In tomato, there are four B genes: TAP3 and TM6 from the AP3 branch, and SLGLO1 and TPI from the PI branch. Different B genes have different functions, resulting in different loss-of-function phenotypes. Loss of TM6 function causes stamen defects, while silencing TPI leads to flower homeomorphism defects (Cao et al., 2019) .

From the perspective of origin, heterocytoplasmic male sterility can be caused by heterocytoplasmic male sterility, cytoplasmic fusion, and natural mutations. Considering heterocytoplasmic male sterility alone, the same cytoplasm under different nuclear backgrounds , exhibiting stable performance across different years and environments, suggests that nuclear gene function is intact. A reasonable explanation is that cytoplasmic genes influence nuclear gene expression. The question then becomes: how do mitochondrial genes affect nuclear gene expression?

Nucleocytoplasmic coordination is a prerequisite for correct gene expression, and cytoplasmic male sterility is one of the consequences of nucleocytoplasmic imbalance. Cytoplasmic genes can only encode a small portion of the proteins they need; most of the required proteins are encoded by nuclear genes. Nucleocytoplasmic imbalance likely occurs when the nuclear genes providing proteins to mitochondria do not encode the proteins required by the mitochondria. The mitochondria then send feedback signals affecting the nuclear genes. Simultaneously, the accumulation of these nuclear proteins becomes toxic to the cell, triggering mechanisms to degrade them. These cellular activities ultimately result in cytoplasmic male sterility. In cytoplasmic male sterility, flower development genes also change, affecting male organogenesis. If overexpression of class B or class C genes restores normal male organogenesis, what will their fertility be like?

In the tobacco carpel-derived cytoplasmic male sterile line K326, transcriptome sequencing revealed not only mitochondrial genome rearrangement but also downregulation of the class B gene *GLOBOSA* . *This study aims to observe* phenotypic and fertility changes in the sterile line K326 by overexpressing the *GLOBOSA gene; the final results await confirmation* by transgenic studies.

3 Materials and Methods

3.1 Test Materials

The tobacco maintainer line K326 used in the experiment was planted in the experimental field of Jiangxi Agricultural University Science and Technology Park. Field management methods for high-quality tobacco production were applied. At the bud stage, K326 flower buds were immediately immersed in liquid nitrogen and stored at -80℃ for RNA extraction. The transient expression material was tobacco K326, which was used for Agrobacterium-mediated transient expression 60 days after sowing, when 5-6 true leaves had emerged .

3.2 Cloning of the tobacco *GLOBOSA* gene

The GLOBOSA gene cloning method followed the method described by Qiu et al. (2021) . TaKaRa The MiniBEST Plant RNA Extraction Kit , PrimeScript™ 1st Strand cDNA Synthesis Kit , and pGME -T Easy vector were purchased from TAKARA. RNA extraction and reverse transcription were performed according to the kit instructions.

GLOBOSA gene amplification were: forward primer : 5′-ATGGGAAGAGGAAAGAT-3′, reverse primer : 5′-TCACTTAAACCATAAGCTAA-3′. The reaction mixture was 2×T5 Super PCR Mix for PAGE 22. μL , forward primer 1 μL , reverse primer 1 μL , template 1 μL . The PCR reaction program was as follows: 94℃ for 5 min pre-denaturation, 94℃ for 30 s denaturation, 58.7℃ for 30 s, 72℃ for 50 s, for a total of 35 cycles, with a final 72℃ for 5 min followed by incubation at 4℃. The amplified products were verified by agarose gel electrophoresis , and the target band was recovered via a low-melting-point agarose gel . The recovered product was ligated into the pGME -T Easy vector , transformed into *E. coli* DH5α , and cultured on LB agar plates containing kanamycin . After 24 h , clones were selected for PCR verification. Positive clones were sent to Kexin Technology Co., Ltd. for sequencing.

3.3 Bioinformatics Analysis

Analyzing the physicochemical properties of gene-encoded proteins can help study gene function and its location. Bioinformatics analysis methods were performed according to those of Qiu et al. (2021) . To understand the physicochemical properties and subcellular localization of the protein encoded by the *GLOBOSA gene,* ProtParam (http://web.expasy.org/protparam/) was used. The physicochemical properties of the protein were inferred ; the domains of the encoded protein were analyzed using the Pfam database

(http://pfam.xfam.org/) , and the results were further validated using hmmer (https://www.ebi.ac.uk/Tools/hmmer) ; the hydrophobicity of the encoded protein was inferred using ExPASy (https://web.expasy.org/protscale/) ; the transmembrane region was predicted using TMHMM (https://dtu.biolib.com/DeepTMHMM) ; the signal peptide was determined using SignalP 5.0 (http://www.cbs.dtu.dk/services/) ; and the subcellular localization prediction of the protein was based on Cell- Ploc (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/) .

3.4 Evolutionary Analysis

To investigate the evolutionary relationships of *GLOBOSA* genes among different species, a phylogenetic tree was constructed based on *GLOBOSA homologous protein sequences* , following the method described by Qiu Shanshan (2021) . When studying interspecies relationships, wild tobacco, potato, and tomato, all belonging to the Solanaceae family, were considered first. The model species Arabidopsis and rice were used as references. Furthermore, to trace earlier origins, the earliest angiosperm, *Cinnamomum camphora*, was included. Genomic data were downloaded from NCBI (https://blast.ncbi.nlm.nih.gov/) and stored locally. A blast search was then performed on the local library using tobacco *GLOBOSA* protein sequences, with a threshold of 1e-10 and a scroe count greater than 200. The obtained sequences were further identified using hmmer (https://www.ebi.ac.uk/Tools/hmmer). A phylogenetic tree was constructed using the maximum likelihood method with IQTREE software, with 1000 bootstrap iterations. Results were presented using Evolview (https://www.evolgenius.info/evolview) .

3.5 Construction of expression vectors

The expression vector was constructed according to the method described by Qiu Shanshan. Plasmid extraction kits, expression vector pCAMBIA1302, and *Nco* I enzyme were purchased from Qingke Biotechnology Co., Ltd. DH5α containing the expression vector was cultured in LB liquid medium (Tryptone 10 g/L, Yeast extract 5 g/L, NaCl 10 g/L) . Six hours later, plasmids were extracted using a kit. The cloning vector containing the target fragment pGME -T Easy was digested with *Nco* I , and the target fragment was recovered . Simultaneously, the expression vector pCAMBIA1302 was digested with *Nco* I , and the target fragment was then ligated into the *Nco* I- *Spe* I multiple cloning site of the expression vector. After transformation into DH5α , the vector was cultured and amplified in LB medium . After 24 hours , the recombinant expression vector was extracted and the ligation of the recombinant vector was verified by

digestion with *Nco* I / *BsiWI* .

3.6 Subcellular localization

The subcellular localization method for GLOBOSA was performed according to the method of Qiu et al. (2021) . Competent Agrobacterium GV3101 was purchased from Shanghai Weidi Biotechnology Co., Ltd., and the method was performed according to the reagent instructions . The expression vector was transformed into competent Agrobacterium GV3101 using the freeze-thaw method. After shaking culture at 28 °C for 2–3 h, the Agrobacterium was further cultured on LB solid medium (containing kanamycin and hygromycin resistance). Positive clones were selected and cultured on LB liquid medium (containing kanamycin and hygromycin resistance) at 28 °C with shaking . When $_{\text{the OD600}}$ *reached* 1.0, the bacterial solution was directly injected into tobacco K326 leaves . After 36 h , Olympus fluorescence microscopy was used for detection, examining 6 leaves with 3 fields of view per leaf.

Author's Contribution

Cui Fangfang was the main executor of the experimental research in this study and completed the data processing and writing of the first draft of the paper; Zheng Yun, Deng Lingfan, Yang Xiangfei, Zheng Jiuzhou , and Meng Linfeng participated in some of the experiments; Liu Qiyuan and Wang Jiange were the conceivers and leaders of the project, and guided the experimental design, data analysis, paper writing and revision. All authors agreed on the final text.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (31960418).

References

Cao X., Liu 2125-2135 .

Chen G. , Ye

Fang C., Li L., He R., Wang D., Wang M., Hu Q., Ma Q., Qin K., Feng X., Zhang G., Fu X., and Liu Z., 2019, Identification of S23 causing both interspecific hybrid male sterility and environment-conditioned male sterility in rice, Rice (NY), 12(1):10. doi: 10.1186/s12284-019-0271-4 .

Yj , Wang Y., Liu QY, Zhou QM, and Zhou W., 2016, Advances in Cytoplasmic Male Sterility in Plant, Zhongguo Nongxue Tongbao (Chinese Agricultural Science Bulletin) , 32 (18) : 70-75 .

Liu ZH, 2019, Establishment of Abortion Regulatory Network and Function Analysis of *TaUGP1-6A* in Isonuclear Alloplasmic Male Sterility Wheat (*Triticum Aestivum* L.), Dissertation for Ph.D. , Northwest A&F University.

Linke B., Nothnagel T., and Borner T., 2003, Flower development in carrot CMS plants: mitochondria affect the expression of MADS box genes homologous to GLOBOSA and DEFICIENS, Plant J, 34(1): 27-37 . doi: 10.1046/j.1365-313x.2003.01703.x.

Mahajan M., and Yadav SK, 2014, Gain of function mutation in tobacco MADS box promoter switch on the expression of flowering class B genes converting sepals to petals., Molecular Biology Reports, 41(2) 705-12 . doi: 10.1007/s11033-013-2909-4.

Qiu SS, Deng LF, Liu MM, Zhang JQ, Wang YB, Liu QY and Wang JG, 2021, Molecular cloning and subcellular localization of *SRK* Gene in *Sinomanglietia glauca* , Acta Agriculturae Universitatis Jiangxiensis, 43(6):1339-1347. (邱珊珊，邓凌帆，刘苗苗，张建强，王炎波，刘齐元，王江革, Cloning and subcellular localization of the *SRK gene of* the rare plant Magnolia grandiflora , Journal of Jiangxi Agricultural University 2021, 43(6): 1339-1347.)

Rijpkema AS, Royaert S., Zethof J., van der Weerden G., Gerats T., and Vandenbussche M., 2006, Analysis of the Petunia TM6 MADS box gene reveals functional divergence within the DEF/AP3 lineage, Plant Cell, 18(8): 1819-1832 .

BM , Chen YZ, Li HB, Mo H., Huang LB, 2019, Progress of Mechanism of Male Sterility of Plants and Its Application, Henan Nongye Kexue (Journal of Henan Agricultural Sciences) , 48 (05): 1-9 .

Wang Q., Zhang X., Lin S., Yang S., Yan X., Bendahmane M., Bao M., and Fu X., 2020, Mapping a double flower phenotype-associated gene DcAP2L in Dianthus chinensis, J Exp Bot, 71(6): 1915-1927.

Wang T., He T., Ding X., Zhang Q., Yang L., Nie Z., Zhao T., Gai J., and Yang S., 2021, Confirmation of GmPPR576 as a fertility restorer gene of cytoplasmic male sterility in soybean, J Exp Bot, 72(22): 7729-7742 .

an HX, Lin CJ, Fan YJ, Zhang CB, 2022, Research Progress on Cloning and Molecular Mechanism of Cytoliasmic Male Sterility Genes in Plants, ZHIwu Shengli Xuebao (Plant Physiology Journal), 58(01): 61-76 （燕厚兴，林春晶，范亚军，张春宝， 2022， 生进 植种霜男死基因衣衣及分子机制研究前进， 植植孔学报， 58(01): 61-76 ）

Yang W., Lou

Yao XF, Dong JJ, Liu SF, Zhang Y., Tang RM, Wang WB, Xie HE, Wu YH, Wu ZX, Li RZ, and Jia XY, 2021, Genome-wide Identification and Tissue-specific Expression Analysis of MADS-box Gene Family in Ipomoea trifida, Jiguang Shengwu Xuebao (Acta Laser Biology Sinica), 30(03): 223-235 (姚旭峰,董静静, 刘世芳, 张毅, 唐锐敏, 王文斌, 解红娥, 吴宇浩, 吴宗信, 贺立恒, 李润植, 贾小云, 2021, Genome-wide Identification and Tissue-specific Expression Analysis of MADS-box Gene Family in Ipomoea trifida, Journal of Laser Biology, 30(03): 223-235)

Zhang YM, 2021, Analysis of Anther Proteomics and Ubiquitinome and Functional Analysis of E3 Ligase TaHUB1 */2* Genes in P-type Cytoplasmic Male Sterility in Wheat, Dissertation for Ph.D., Northwest A&F University .