Fasciclin-like arabinogalactan gene family in *Nicotiana benthamiana*: genome-wide identification, classification and expression in response to pathogens

Xinyang Wu¹,², Yuchao Lai², Lanqing Lv², Mengfei Ji¹,², Kelei Han², Dankan Yan², Yuwen Lu², Jiejun Peng², Shaofei Rao², Fei Yan², Hongying Zheng²* and Jianping Chen¹,²*

Abstract

**Background:** *Nicotiana benthamiana* is widely used as a model plant to study plant-pathogen interactions. Fasciclin-like arabinogalactan proteins (FLAs), a subclass of arabinogalactan proteins (AGPs), participate in mediating plant growth, development and response to abiotic stress. However, the members of FLAs in *N. benthamiana* and their response to plant pathogens are unknown.

**Results:** 38 *NbFLAs* were identified from a genome-wide study. *NbFLAs* could be divided into four subclasses, and their gene structure and motif composition were conserved in each subclass. *NbFLAs* may be regulated by cis-acting elements such as STRE and MBS, and may be the targets of transcription factors like C2H2. Quantitative real time polymerase chain reaction (RT-qPCR) results showed that selected *NbFLAs* were differentially expressed in different tissues. All of the selected *NbFLAs* were significantly downregulated following infection by turnip mosaic virus (TuMV) and most of them also by *Pseudomonas syringae pv tomato* strain DC3000 (*Pst* DC3000), suggesting possible roles in response to pathogenic infection.

**Conclusions:** This study systematically identified FLAs in *N. benthamiana*, and indicates their potential roles in response to biotic stress. The identification of *NbFLAs* will facilitate further studies of their role in plant immunity in *N. benthamiana*.

**Keywords:** *N. benthamiana*, Fasciclin-like arabinogalactan proteins, Gene expression, Abiotic stress, Turnip mosaic virus
Background

The plant cell wall is a dynamic and complex organelle, which is mainly composed of cellulose, hemicellulose, pectins, glycans and proteins. It is not only involved in mechanical protection and structural support, but also in signal transduction, intercellular communication and immunity [1–3].

Hydroxyproline-rich glycoproteins (HRGPs) are typical cell-wall proteins that participate in plant growth, development and immunity [4, 5]. HRGPs have a few repetitive glycosylation motifs containing hydroxyproline (Hyp) residues that are glycosylation sites. Based on the different levels of O-glycosylation, the HRGP superfamily can be classified into three subfamilies: the hyperglycosylated arabinogalactan proteins (AGPs), the minimally glycosylated Pro-rich proteins (PRPs) and the moderately glycosylated extensins (EXTs) [5]. AGPs are abundant in plants, and can themselves be subdivided into six main subclasses: the classical AGPs, AG peptides, Lys-rich AGPs, FLAs, non-classical AGPs and chimeric AGPs [6]. FLAs generally have one or two fasciclin domains, and have been discovered in fruit flies, mammals, sea urchins, plants, yeast and bacteria. Besides fasciclin domains, FLAs often contain an N-terminal signal peptide as well as a C-terminal glycosylphosphatidylinositol (GPI) anchor signal peptide. The GPI and fasciclin domains are functionally important and are believed to mediate cell adhesion [7, 8].

So far, the FLA family members have been identified in several plant species. 21 FLAs have been identified in Arabidopsis thaliana [8], 27 in rice (Oryza sativa) [9, 10], 34 in wheat (Triticum aestivum) [10], 35 in poplar (Populus trichocarpa) [11], 19 in cotton (Gossypium hirsutum) [12], 33 in Chinese cabbage (Brassica rapa) [13], 18 in Eucalyptus grandis [14] and 23 in textile hemp (Cannabis sativa) [15]. FLAs are cell wall structural glycoproteins that mediate cellulose deposition and cell wall development. They are believed to participate in fiber development, elongation and stem dynamics, affecting the quality of fiber and wood in cotton and woody plants like poplar and eucalyptus [16] and are abundant in the xylem [17]. Knock down of PtFLA6 resulted in a decrease of stem hardness and xylem cellulose lignin, and down-regulation of genes involved in cell wall synthesis [18]. Overexpression of GhGalT1 promoted cotton fiber development by controlling the glycosylation of FLAs [19] and in plants where GhAGP4 was knocked down, fiber initiation and elongation were strongly inhibited and there was suppression of the cytoskeleton network and of cellulose deposition in fiber cells [20]. During cell wall regeneration from cotton protoplasts, there is up regulation of proline-rich protein (PRPL), glycine-rich protein (GRP), and extensin (EPR1) but also of FLA2, which may mediate the construction and modification of the cell wall [21]. In addition, AtFLA11, AtFLA12, EgrFLA2 and EgrFLA3 have similar functions [14, 22]. FLAs can also regulate pollen development. In Arabidopsis and maize, AtFLA9 and ZmFLA7 showed negative correlation with abortion, and reductions in the expression of FLAs increased the abortion of fertilized ovaries [23]. AtFLA3-silenced Arabidopsis had abnormal pollen grains, also suggesting a function in pollen formation [24]. FLAs have also been implicated in cell-to-cell communication [13], shoot development [25, 26], seed mucilage adherence [27], glycan stabilization [28] and in response to stresses from salt [29–31], cold [32] and hydrogen peroxide [33].

Although FLAs have multiple roles in plant growth and development, very little is known about any involvement they may have in response to pathogens. N. benthamiana is a model plant for studying plant immunity, but the structure, function and expression of its FLA gene family members is unknown. In this study, we have identified and characterized the members of the FLA gene family in N. benthamiana and also reported their subcellular localization, expression patterns, and their response to viral and bacterial pathogens.

Results

Identification of members of the NbFLA family

Based on previous studies [8], FLAs have an AGP-like glycosylated region, a fasciclin domain and an N-terminal signal peptide. We followed these criteria to identify putative FLAs in N. benthamiana. The sequences of the 21 identified AtFLAs were downloaded [8] and the N. benthamiana genome was downloaded from the Sol Genomics Network (https://solgenomics.net/) [34]. A total of 38 NbFLAs were identified by two round BLASTP and signal peptide prediction (Table 1 and Additional file 1: Table S1). Most of these (66%) have lengths of 200-300aa, while the largest (NbFLA10) has 495aa and the smallest (NbFLA26) has only 182aa. The predicted isoelectric points range from 4.29 to 9.77, and the molecular weights (MWs) derived only from the amino acid sequences (not including glycans) are in the range 19.68–52.32 kDa. The protein properties of the NbFLAs are similar to those of other plant species [8, 11].

Phylogenetic analysis and multiple sequence alignment of NbFLAs

To better reveal their evolutionary relationships and to help the classification of NbFLAs, the sequences of all 21 AtFLAs and 38 NbFLAs were used to construct a phylogenetic tree (Fig. 1). Because of the low sequence similarity between some FLAs, phylogenetic analysis alone could be misleading and therefore pair-wise sequence similarity, presence and number of fasciclin domains and GPI were also used to create a classification, as previously described [8]. Most NbFLAs were
### Table 1: Putative FLAs in *N. benthamiana*

| gene symbol | gene locus | gene position | chromosome location | strand | CDS(bp) | protein length(aa) | Pi | amino acids | MW(kDa) | predicted localization | FAS | GPI |
|-------------|------------|---------------|---------------------|--------|---------|-------------------|----|-------------|---------|------------------------|------|------|
| NbFLA1      | Niben101Scf00321g04012.1 | Niben101Scf00321 | Niben101Scf00321:422864,423997 | +      | 1131    | 377               | 6.59 | 4064 | 2 N                  |
| NbFLA2      | Niben101Scf00509g01008.1 | Niben101Scf00509 | Niben101Scf00509:90123 | -      | 741     | 247               | 6.39 | 2608 | 1 Y                  |
| NbFLA3      | Niben101Scf00550g00014.1 | Niben101Scf00550 | Niben101Scf00550:55345 | +      | 1290    | 430               | 6.25 | 4448 | 1 N                  |
| NbFLA4      | Niben101Scf00788g03209.1 | Niben101Scf00788 | Niben101Scf00788:461639 | +      | 825     | 275               | 4.62 | 2895 | 1 Y                  |
| NbFLA5      | Niben101Scf00788g03035.1 | Niben101Scf00788 | Niben101Scf00788:243441 | -      | 819     | 273               | 4.74 | 2855 | 1 Y                  |
| NbFLA6      | Niben101Scf00861g07009.1 | Niben101Scf00861 | Niben101Scf00861:10031,815156 | +      | 1218    | 406               | 6.40 | 4496 | 2 N                  |
| NbFLA7      | Niben101Scf00905g07001.1 | Niben101Scf00905 | Niben101Scf00905:720957 | -      | 774     | 258               | 4.80 | 2697 | 1 Y                  |
| NbFLA8      | Niben101Scf01037g09001.1 | Niben101Scf01037 | Niben101Scf01037:933101 | +      | 801     | 267               | 5.60 | 2817 | 1 Y                  |
| NbFLA9      | Niben101Scf01484g0009.1 | Niben101Scf01484 | Niben101Scf01484:48168 | +      | 1377    | 459               | 5.75 | 5065 | 2 N                  |
| NbFLA10     | Niben101Scf01535g01003.1 | Niben101Scf01535 | Niben101Scf01535:175164 | -      | 1485    | 495               | 5.74 | 5232 | 2 Y                  |
| NbFLA11     | Niben101Scf01685g06002.1 | Niben101Scf01685 | Niben101Scf01685:175164 | -      | 1053    | 351               | 7.76 | 3887 | 2 N                  |
| NbFLA12     | Niben101Scf01911g00001.1 | Niben101Scf01911 | Niben101Scf01911:73669 | -      | 744     | 248               | 5.81 | 2608 | 1 Y                  |
| NbFLA13     | Niben101Scf02934g01012.1 | Niben101Scf02934 | Niben101Scf02934:129519 | -      | 1074    | 358               | 4.40 | 3973 | 1 N                  |
| NbFLA14     | Niben101Scf03271g02005.1 | Niben101Scf03271 | Niben101Scf03271:312206 | +      | 810     | 270               | 6.42 | 2836 | 1 Y                  |
| NbFLA15     | Niben101Scf03563g00015.1 | Niben101Scf03563 | Niben101Scf03563:165737 | +      | 801     | 267               | 5.60 | 2820 | 1 Y                  |
| NbFLA16     | Niben101Scf03634g11004.1 | Niben101Scf03634 | Niben101Scf03634:1157457 | -      | 1344    | 448               | 6.14 | 4940 | 2 N                  |
| NbFLA17     | Niben101Scf03757g02001.1 | Niben101Scf03757 | Niben101Scf03757:19902,202071 | +      | 1377    | 459               | 5.76 | 5079 | 2 N                  |
| NbFLA18     | Niben101Scf04444g06007.1 | Niben101Scf04444 | Niben101Scf04444:679736 | -      | 1248    | 416               | 5.93 | 4384 | 2 Y                  |
| NbFLA19     | Niben101Scf04650g07041.1 | Niben101Scf04650 | Niben101Scf04650:712304 | +      | 1221    | 407               | 5.90 | 4386 | 2 Y                  |
| NbFLA20     | Niben101Scf04731g11005.1 | Niben101Scf04731 | Niben101Scf04731:1263224 | +      | 1038    | 346               | 9.41 | 3788 | 1 N                  |
| NbFLA21     | Niben101Scf04765g02001.1 | Niben101Scf04765 | Niben101Scf04765:1263224 | +      | 1080    | 360               | 5.44 | 3875 | 1 N                  |
| NbFLA22     | Niben101Scf04792g02004.1 | Niben101Scf04792 | Niben101Scf04792:335740 | -      | 735     | 251               | 6.83 | 2701 | 1 Y                  |
| NbFLA23     | Niben101Scf04813g00014.1 | Niben101Scf04813 | Niben101Scf04813:69262 | +      | 1386    | 462               | 5.50 | 4834 | 2 N                  |
| NbFLA24     | Niben101Scf04847g00091.1 | Niben101Scf04847 | Niben101Scf04847:67150 | -      | 1248    | 416               | 5.81 | 4405 | 2 Y                  |
| NbFLA25     | Niben101Scf05372g02008.1 | Niben101Scf05372 | Niben101Scf05372:251817 | +      | 750     | 250               | 6.03 | 2610 | 1 Y                  |
| NbFLA26     | Niben101Scf05486g01006.1 | Niben101Scf05486 | Niben101Scf05486:76890 | +      | 546     | 182               | 9.23 | 1967 | 1 Y                  |
| NbFLA27     | Niben101Scf05486g01015.1 | Niben101Scf05486 | Niben101Scf05486:168845 | +      | 774     | 258               | 6.06 | 2736 | 1 Y                  |
| NbFLA28     | Niben101Scf06087g12013.1 | Niben101Scf06087 | Niben101Scf06087:1272274 | +      | 1146    | 382               | 9.77 | 4113 | 1 N                  |
| NbFLA29     | Niben101Scf06123g01025.1 | Niben101Scf06123 | Niben101Scf06123:104049 | +      | 708     | 236               | 8.79 | 2561 | 1 Y                  |
| Flavin-like A (FLA) | mRNA Sequence | AmiTag2 Primers | Sequence Length | GC % | Motif | Location | CDS Length |
|---------------------|--------------|----------------|----------------|------|-------|----------|-------------|
| NbFLA30             | Niben101Scf06193g02005.1 Niben101Scf06193 Niben101Scf06193268833,269871 | + | 747 249 | 9.36 | 27.50 | Cell membrane | 1 N |
| NbFLA31             | Niben101Scf06203g03009.1 Niben101Scf06203 Niben101Scf06203290365,293147 | - | 1326 442 | 6.26 | 47.82 | Cell membrane | 2 N |
| NbFLA32             | Niben101Scf08300g00008.1 Niben101Scf08300 Niben101Scf0830060937,61692 | + | 753 251 | 8.59 | 26.95 | Cell membrane | 1 Y |
| NbFLA33             | Niben101Scf11969g00007.1 Niben101Scf11969 Niben101Scf1196932722,34415 | + | 741 247 | 6.39 | 26.13 | Cell membrane | 1 Y |
| NbFLA34             | Niben101Scf12431g00010.1 Niben101Scf12431 Niben101Scf124311378,2604 | + | 1224 408 | 5.41 | 43.98 | Cell membrane | 2 Y |
| NbFLA35             | Niben101Scf13776g02002.1 Niben101Scf13776 Niben101Scf1377633230,335138 | + | 1068 356 | 4.29 | 39.38 | Cell membrane | 1 N |
| NbFLA36             | Niben101Scf19479g00006.1 Niben101Scf19479 Niben101Scf1947991773,104152 | - | 819 273 | 9.63 | 29.11 | Cell membrane | 1 N |
| NbFLA37             | Niben101Scf22195g00016.1 Niben101Scf22195 Niben101Scf2219562492,63565 | + | 1071 357 | 8.56 | 39.46 | Cell membrane | 1 N |
| NbFLA38             | Niben101Scf32632g00003.1 Niben101Scf32632 Niben101Scf326327941,10147 | + | 1128 376 | 5.30 | 41.56 | Cell membrane | 1 N |

Notes: +: The mRNA sequence is positive strand; -: The mRNA sequence is negative strand; Y: The sequence has GPI signal; N: The sequence has no GPI signal.
sufficiently classified by phylogenetic analysis, but for a few (NbFLA8/15 and NbFLA10/14) their protein properties including the presence and number of fasciclin domains and GPI had also to be taken into account.

The 38 NbFLAs we identified could be divided into the same four subclasses previously reported for the AtFLAs [8], named I to IV (Fig. 1). NbFLA2/8/12/15/22/25/26/27/29/32/33/36 belong to subclass I, and have a single fasciclin domain and GPI anchored signal (except NbFLA36), as do the related AtFLAs and PtrFLAs [8, 11]. NbFLA6/9/16/17 belong to subclass II. Subclass II is the smallest group and members contain two fasciclin domains but have no C-terminal GPI anchor site. Members of subclass III (NbFLA3/4/5/7/10/14/18/19/23/24/34/38) have either one or two fasciclin domains, and most (77%) have a C-terminal GPI anchor site. The remaining NbFLAs (NbFLA1/11/13/20/21/28/30/31/35/37) constitute subclass IV, which contains NbFLAs that are quite distantly related to the other NbFLAs and which have no consistent pattern in the number of fasciclin domains or the presence of a GPI signal.

We also constructed separate phylogenetic trees for each subclass of NbFLAs, including the sequences from the other 8 plant species in which FLAs have been identified (Arabidopsis, rice, wheat, poplar, cotton, Chinese cabbage, Eucalyptus grandis and textile hemp) (Additional file 2: Fig. S1). In general, FLAs have a relatively high homology among closely related species, like AtFLAs/BrFLAs and OsFLAs/TaFLAs. FLAs from the same species often exist in pairs, like NbFLA26/29 and TaFLA19/27, suggesting that they may be paralogous genes. Subclasses I and III are the two largest groups and the clustering patterns are complicated. FLAs from the same species do not generally group together, and there are some closely-related pairs from different species suggesting that they are orthologous genes (e.g. NbFLA12/BrFLA22 and TaFLA2/OsFLA2). In subclasses II and IV, most FLAs from the same species group together (e.g. NbFLA6/9/16/17 and TaFLA6/7/8/29). Subclass II has fewest members and most of them are not GPI anchored, but the OsFLAs are a significant exception.
Previously reported fasciclin domains contain about 110–150 amino acid residues and have two highly conserved regions (H1 and H2) and a [Phe/Tyr]-His ([Y/F]H) motif [12]. An alignment of the amino acid sequences of the fasciclin domains of the NbFLAs constructed using MUSCLE and some manual analysis showed a similar pattern (Fig. 2). The Thr residue in the H1 region is highly conserved and is followed by other conserved residues such as Val/ Ile (one position after Thr) and Asn/Asp (six positions after Thr). These residues may play a role in maintaining the structure of the fasciclin domain and/or cell adhesion [12]. As reported for other fasciclin domains [11, 31, 35], small hydrophobic amino acids such as Leu, Val and Ile are abundant in the H2 region. In the [Y/F]H motif, His and Pro residues are also relatively conserved.

Analysis of the structural and conserved motifs of NbFLAs

Further analysis of gene structure and motifs of the NbFLAs is shown in Fig. 3. The phylogenetic tree confirmed that NbFLAs could be grouped into four subclasses (Fig. 3a). Analysis of the genomic DNA sequences showed that NbFLAs usually had 0, 1 or 2 introns (Fig. 3b). All of the members in subclass II have one or two introns while most members of subclasses I and III have none (Fig. 3b). The most closely related members of each subclass, usually have a similar exon/intron structure, with little difference in the length of introns and exons. However, a few NbFLA gene pairs showed different intron/exon arrangements. For example, NbFLA1 and NbFLA31 have high sequence similarity, but NbFLA1 has no introns while NbFLA31 has one.

An online MEME analysis was done to identify additional motifs among the 38 NbFLAs. Twenty conserved motifs were predicted (Fig. 3c and Additional file 3: Table S2) and each NbFLA contained between five and ten of these. Some motifs were common to most members, while the others were unique to one or few subclasses. For example, NbFLA1 and NbFLA31 have high sequence similarity, but NbFLA1 has no introns while NbFLA31 has one.

An online MEME analysis was done to identify additional motifs among the 38 NbFLAs. Twenty conserved motifs were predicted (Fig. 3c and Additional file 3: Table S2) and each NbFLA contained between five and ten of these. Some motifs were common to most members, while the others were unique to one or few subclasses. For example, NbFLA1 and NbFLA31 have high sequence similarity, but NbFLA1 has no introns while NbFLA31 has one.
Fig. 3 Phylogenetic relationship, gene structure and architecture of the conserved protein motifs in NbFLAs. a The phylogenetic tree was constructed based on the full-length sequences of NbFLA proteins. b Exon-intron structure of NbFLAs. Pink boxes indicate untranslated 5' and 3' regions; green boxes indicate exons; and black lines indicate introns. The fasciclin domains are shown by yellow boxes. c The motif composition. The motifs, numbered 1–20, are displayed in different colored boxes. The sequence information for each motif is provided in Additional file 1: Table S2.

Fig. 4 Prediction of cis-acting elements in NbFLAs. a Numbers of cis-acting elements detected in the promoter region of each NbFLA gene. All cis-acting elements were divided into seven types. b Kind, quantity and position of environmental stress-related elements in NbFLAs. c Kind, quantity and position of hormone responsive elements in NbFLAs.
most members of subclasses I and III contained both motifs 3 and 8 except NbFLA4/5/7/26/38. Subclass IV was clearly less closely related to the other subclasses, and motifs 12, 13 and 15 were unique to this subclass.

**Prediction of cis-acting elements and transcription factors among the NbFLAs**

The cis-acting elements in the promoter regions of the NbFLAs were analyzed and a totally 105 cis-acting elements were predicted (Fig. 4 and Additional file 4: Table S3). Those cis-acting elements were related to environmental stress, hormone response, development, light response, promoter, site binding and other functions (Fig. 4a). The most abundant elements were light-responsive elements, including G-box, GT1-motif and GATA-motif. 15 hormone responsive elements were identified and these are mainly involved in response to abscisic acid (ABA) or methyl jasmonate (MeJA) (Fig. 4b). Among the predicted environmental stress-related elements, STRE, MBS and ARE were the most abundant (Fig. 4c). Several abundant predicted cis-acting elements are known to mediate plant immunity. For example, VdMYB1 binds to the MBS in the VdSTS2 gene promoter, thus activating VdSTS2 transcription and positively regulating defense responses [36]. Machi3–1 and TaRIM1 also bind MBS cis-acting elements to increase host resistance [37, 38].

By binding to transcription factors (TFs), cis-acting elements regulate the precise initiation and efficiency of gene transcription. We then therefore predicted potential TFs which may regulate the transcription of NbFLAs (Fig. 5 and Additional file 5: Table S4). The NbFLAs had

---

**Fig. 5 Regulation network between NbFLAs and potential TFs. Green hexagons represent transcription factors, blue rectangles represent NbFLAs, and black lines represent potential regulatory relationships**
an average of five TFs, but it appears that \( \text{NbFLA4} \) and \( \text{NbFLA27} \) may be regulated by more TFs, including specific TFs like RAV and CPP, while \( \text{NbFLA8/15/38} \) may each be regulated by only two TFs. In total, 25 TFs were predicted of which C2H2, BBR-BPC, Dof, Myb and MIKC were the most abundant. Previous studies have demonstrated the role of TFs in regulating plant immunity. \( \text{NbCZF1} \), a novel C2H2-Type zinc finger protein, is a regulator of plant defense [39] and \( \text{VvDOF3} \) enhances powdery mildew resistance in \( \text{Vitis vinifera} \) [40]. In addition, \( \text{AtMyb15} \) and \( \text{MdMyb30} \) also participate in enhancing disease resistance [41, 42].

**Subcellular localization analysis of NbFLAs**

Bioinformatics analysis based on the NbFLA amino acid sequences suggested that all of them could locate to membranes, and only NbFLA4 was predicted to locate in both the nucleus and membranes (Table 1). To validate these predictions, we selected one NbFLA in each subclass (NbFLA4/6/31/32) to analyze their localization by laser confocal microscopy. \( \text{AtP12A-GFP} \) was used as membrane marker [43]. The results showed that while NbFLA6 and NbFLA32 were only located in membranes, NbFLA4 was present both in the nucleus and the nucleus, consistent with the predictions (Fig. 6).

A GPI anchored signal is vital for membrane localization and is predicted in about two thirds of AtFLAs and PtrFLAs and in 20 of 38 (53%) of NbFLAs (Table 1). Among the four selected NbFLAs, only NbFLA31 was not GPI anchored. Correspondingly, although a plasmolysis experiment confirmed the membrane localization of NbFLA31, a diffused red fluorescence could also be observed in the cytoplasm (Fig. 6 and Additional file 6: Fig. S2).

**Tissue-specific expression of NbFLAs**

To comprehensively understand the functions of \( \text{NbFLAs} \), two or three \( \text{NbFLAs} \) from each subclass were randomly selected to analyze their expression in five different tissues (root, stem, young leaf, mature leaf and flower) by RT-\( \text{qPCR} \) (Fig. 7 and Additional file 7: Fig. S3). The expression level of all selected \( \text{NbFLAs} \) (except \( \text{NbFLA4} \)) was higher in young leaves than in mature ones. \( \text{NbFLA11/18/31/32/34} \) were highly expressed in young leaves, and \( \text{NbFLA4} \) were expressed highly in flowers. It was earlier reported that \( \text{PtFLA6} \) is specifically expressed in tension wood (TW) and that decreased transcripts of \( \text{PtFLA6} \) influenced stem dynamics [18]. In this study, \( \text{NbFLA2/6/15/17} \), belonging to subclasses I and II, were highly expressed in stems, suggesting that they may play a role in stem dynamics.

**Expression of NbFLAs under biotic stress**

To investigate whether \( \text{NbFLAs} \) participate in the response to pathogens, leaves of \( \text{N. benthamiana} \) were inoculated with turnip mosaic virus (TuMV), potato virus X (PVX), pepper mosaic virus (PMoV) and the bacterial pathogen \( \text{Pseudomonas syringae pv tomato} \) strain DC3000 \( \text{(Pst DC3000)} \). At 5 days post virus inoculation (dpi), or 2 days post \( \text{Pst DC3000} \) infection, leaves were collected to study the expression pattern of 11 \( \text{NbFLA} \) genes by RT-qPCR (Fig. 8).

TuMV infection led to a huge reduction in expression of all the \( \text{NbFLAs} \) tested, especially \( \text{NbFLA15/18/32/34} \), which all decreased by more than 99%. PVX or PMoV infection usually induced a modest reduction in expression, although \( \text{NbFLA6} \) was slightly upregulated by PVX. The bacterial pathogen \( \text{Pst DC3000} \) decreased expression of most \( \text{NbFLAs} \) by 73–99% but, in contrast, \( \text{NbFLA4} \) and \( \text{NbFLA7} \) were substantially upregulated. These results show that most \( \text{NbFLAs} \) are substantially affected by TuMV and \( \text{Pst DC3000} \) and may therefore play roles in post-infection responses.

**Discussion**

\( \text{FLA} \) families have been identified and characterized in several plants including Arabidopsis [8], rice [9, 10], wheat [10], poplar [11], cotton [12], Chinese cabbage [13], \( \text{Eucalyptus grandis} \) [14] and textile hemp [15]. In this study, we identified 38 \( \text{FLA} \)s in \( \text{N. benthamiana} \) and found that their structural domains were conserved by studying phylogenetic trees, gene structure and conserved motifs (Fig. 3). In general, \( \text{NbFLAs} \) could be divided into four subclasses and \( \text{NbFLAs} \) in each subclass had similar gene structure, motifs and conserved domains. Consistent with the \( \text{FLAs} \) in Arabidopsis [8], subclass II contained fewest \( \text{NbFLAs} \) and \( \text{NbFLAs} \) in subclass IV were the most variable. The \( \text{FLAs} \) of other dicotyledonous plant species had similar properties in each subclass, while dicot members of subclass II have no GPI, most OsFLAs and TaFLAs in the subclass are GPI anchored [10]. In addition, OsFLAs in subclass II have only one fasciclin domain, unlike the \( \text{FLAs} \) of the dicotyledonous species [10]. Thus a different classification of \( \text{FLAs} \) in monocotyledonous plants may be required.

Twenty-five of the 38 \( \text{NbFLAs} \) had a single fasciclin domain, 13 of them had two domains and 20 of the 38 were GPI anchored. A GPI-anchored signal together with a fasciclin domain are known to be important for cell adhesion, for membrane localization and for enabling more stable interactions between adhesion complexes. It has been suggested that plants may have \( \text{FLAs} \) with GPI-anchoring for maintaining the integrity of the plasma membrane and FLAs that are not GPI-anchored for mediating cell expansion [8].

Previous studies have shown different expression patterns of \( \text{FLAs} \) in the tissues of other plants. For example, \( \text{AtFLA11/12} \) were highly expressed in stems [22], as were
BrFLA6/9/22 (homologous to AtFLA11). Some EgrFLAs were also highly expressed in stems [14, 22] and 10 PopFLAs were highly expressed in poplar tension wood [35]. PtFLA6 and ZeFLA11 were exclusively expressed in xylem tissues [18, 44]. These studies suggest that some FLAs play important roles in stem dynamics and cell...
wall elongation. In our study, \( NbFLA2/6/15 \) were also expressed highly in stems whereas \( NbFLA7/34 \) were highly expressed in roots, as were \( PtrFLA12/21/22/24/27/28/30 \) [11], indicating that they may participate in root apical meristem development. Many \( NbFLA\)s were expressed highly in young leaves [11], as reported for \( GhFLA5/8/9/12 \) and \( Br4/5/10/21/27/33 \) [8, 12, 13], but no \( PtrFLA\)s tested had high expression in young leaves [11]. This may be because \( N.\ benthamiana \) more closely resembles cotton and Chinese cabbage in being a herbaceous annual.

Some biotic and abiotic stresses lead to significant changes in the transcription of \( FLA\)s. For example, Under \( H_2O_2 \) stress, the expression levels of wheat \( FLA\) proteins were increased, which may contribute to \( H_2O_2 \) tolerance [33]. Similarly, \( AtFLA3 \) was expressed more highly under cold stress [32]. Under salt stress, \( OsFLA10/18 \) expression was reduced [9] while \( PtrFLA2/12/20/21/24/30 \) were upregulated [11]. In addition, \( TaFLA3/4/9 \) were downregulated after heat, ABA or \( NaCl \) treatment [10]. \( OsFLA24 \) and \( AtFLA1/2/8 \) were also significantly reduced following ABA treatment [8, 9]. Many of the frequently predicted TEs in the \( NbFLA\)s, including \( C2H2, Dof\) and \( Myb \), have been reported to play a role in the ABA pathway [45–48] and therefore, as in other species, \( NbFLA\)s may be regulated by the ABA pathway. While the function of \( FLA\)s in the signaling pathway during abiotic stresses has been investigated, little is known about their potential role in response to pathogens. \( AtFLA1/2/8 \) were decreased by pathogen challenge, oxidative stress and in ascorbate-deficient vtc mutants [49]. The fungus \( Ophiostoma novo-ulmi \) reduced the expression of \( FLA\)s in English elm ramets [50]. Our results show that almost all \( NbFLA\)s were specifically downregulated by TuMV and \( Pst\) DC3000 infection and this suggests that \( NbFLA\)s may have specific roles in pathogen infection.

Because of their role in cell adhesion and their membrane localization, AGPs (including \( FLA\)s) may interact with receptor-like kinases as wall-associated kinases and
Thus be involved in signal transduction [51]. For example, AtFLA4 (SOS5) mediated root growth and seed adhesion through cell wall receptor-like kinase (FEI1/2) [27], and modulated ABA signaling to regulate cell wall biosynthesis and root growth [25, 27]. The known functions of GPI and the fasciclin domain suggest that NbFLAs might be involved in host-pathogen interactions. Thus, a further role of NbFLAs in plant resistance is worth exploring.

**Conclusion**

In this study, 38 NbFLAs were identified and could be divided into four subclasses. In general, the closest members of NbFLAs from the same subclass have similar structure and conserved motifs. The expression patterns of selected NbFLAs in different tissues were diverse and selected NbFLAs were downregulated following infection by TuMV or Pst DC3000. Our results will help to lay the foundation for understanding of the structure and characteristics of the FLA family and for exploring the relationship between FLAs and immunity in *N. benthamiana*.

**Methods**

**Identification of the NbFLAs family**

The sequences of the 21 identified AtFLAs were downloaded and the *N. benthamiana* genome was downloaded from the Sol Genomics Network (https://solgenomics.net/) [34]. NbFLAs were identified by two rounds of BLASTP. Firstly, all AtFLAs were used to search possible NbFLAs using TBtools [52]. Then NCBI Batch CD-Search [53, 54] was used to confirm whether candidate NbFLAs contained a fasciclin domain including FAS1 (smart00554), Fasciclin superfamily (cl02663) or Fasciclin (pfam02469). Next, we predicted the N-terminal signal peptide by SignalP5.0 [55], the C-terminal GPI anchor addition signal by big-PI Plant Predictor [56], and the glycosylation site by NetGlycate 1.0 [57]. Finally, using criteria previously established, sequences that contained an AGP-like glycosylated region, fasciclin domains and an N-terminal signal peptide were considered as NbFLAs [11]. The CDS length, pI and molecular weights (MW) of all predicted NbFLAs were then determined by ExPASy [58] and their subcellular localization predicted by Plant–mPLoc [59].

**Phylogenetic analysis and multiple sequence alignment**

Sequences of AtFLA proteins were obtained from the NCBI protein database (http://www.ncbi. nlm.nih.gov/protein/). A neighbor-joining (NJ) phylogenetic tree of full-length sequences of AtFLAs and NbFLAs was constructed with 1000 bootstrap replicates using MEGA7.0. A multiple sequence alignment of all NbFLAs was also created by Clustal X 2.0 [60].
Gene structure and conserved domain analysis

Gene structure and conserved domains were analyzed and visualized using NCBI Batch CD-Search [53, 54] and TBtools [52]. Conserved motifs of the genes were analyzed by the MEME program [61] with the following parameters: optimum motif width was set to 30–70, the number of repetitions was set to zero or one, the maximum number of motifs was set to identify 15 motifs.

Promoter cis-acting elements and TFs prediction

The promoter cis-acting elements were predicted by PlantCARE [62] and transcription factors were predicted by PlantRegMap [63], with N. sylvestris as the target species.

Plasmid construction and Agroinfection assays in N. benthamiana

Based on the sequences above, we cloned the CDS sequences of NbFLA4/6/31/32 and constructed them into a transient expression vector with red fluorescent label. All primers used for plasmid construction are listed in Additional file 8: Table S5. Agroinfection assays were conducted as previously described [64]. Briefly, the constructs were transformed into A. tumefaciens (strain GV3101) by electroporation. The transformants were cultured and re-suspended in the inoculation buffer [10 mM MgCl₂, 2 mM acetosyringone, 100 mM MES (pH 5.7)] for 3–5 h at room temperature. The suspensions were then adjusted to OD₆₀₀ = 0.1 and were infiltrated into leaves of 4- to 6-week old N. benthamiana plants with needleless syringes.

Plant growth and pathogen inoculation

N. benthamiana seeds were donated by Dr. Yule Liu (Tsinghua University, China) and grown in mixed soil matrix (peat: vermiculite = 1:1) under a 16-h light (2000 lx)/8-h dark photoperiod at 26 ± 2 °C with relative humidity 60 ± 5%. A TuMV infectious clone was kindly provided by Dr. Fernando Ponz (INIA, Laboratorio de Virología, Spain), a PVX infectious clone was kindly provided by Dr. Stuart MacFarlane (James Hutton Institute, UK) and a PMMoV infectious clone was constructed by Neighbor-joining using MEGA7 software and with 1000 bootstrap replicates.

Expression analysis by RT-qPCR

RT-qPCR analysis was performed to confirm the expression of representative NbFLA genes. We used at least three independent biological replicates and technical replicates. First-strand cDNA was synthesized from 0.5 mg of RNA with PrimeScript RT reagent (TaKaRa). RT-qPCR was carried out by SYBR-green fluorescence using the Roche LightCycler®480 Real-Time PCR System. Relative gene expression levels were calculated according to the ΔΔCT method [66] and visualized in a heat map by Tbtools [52]. All primers used for RT-qPCR are listed in Additional file 8: Table S5.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12870-020-02501-5.

Additional file 1: Table S1. List of NbFLA CDS and protein sequences.
Additional file 2: Figure S1. Unrooted phylogenetic trees showing the relationships among FLA proteins of 9 plant species in each subclass. a, b, c, d represent subclasses I, II, III, IV, respectively. The phylogenetic trees were constructed by Neighbor-joining using MEGA7 software and with 1000 bootstrap replicates.
Additional file 3: Table S2. The MEME motif sequences and length of NbFLAs.
Additional file 4: Table S3. Cis-acting elements in NbFLAs.
Additional file 5: Table S4. Potential transcription factors of NbFLAs.
Additional file 6: Figure S2. Plasmolysis experiment of NbFLA31.
Additional file 7: Figure S3. The differential expressions of representative NbFLA genes in different tissues by RT-qPCR (raw data). YL: young leaf; MF: mature leaf; ST: stem; RO: root; FL: flower. The mean expression values were calculated from three independent biological replicates and are relative to that in young leaves.
Additional file 8: Table S5. Primers used in this study.

Abbreviations

FLAs: Fasciclin-like arabinogalactan proteins; AGPs: Arabinogalactan proteins; GPI: Glycosylphosphatidylinositol; TuMV: Tumip mosaic virus; Pst DC3000: Pseudomonas syringae pv tomato (Pst) strain DC3000; HRGPs: Hydroxyproline-rich glycoproteins; TFs: Transcription factors; ABA: Abscisic acid; MeJA: Methyl jasmonate; TW: Tension wood; PVX: Tomato virus X; PMMoV: Pepper mottle mosaic virus; FAS1: Fasciclin 1

Acknowledgements

We thank Prof. M. J. Adams for correcting the English of the manuscript. We thank Dr. Fernando Ponz for providing the TuMV infectious clone, Dr. Stuart MacFarlane for providing the PVX infectious clone and Dr. Yule Liu for providing Nicotiana benthamiana seeds and the Pst DC3000 strain.

Authors’ contributions

WXY, ZHY initiated and designed the experiments. WXY, LYC, LLQ, JMF, HKL, YDK, LYW, PJJ and RSF performed the experiments and collected the data. WXY analyzed the data and wrote the manuscript; ZHY, YF and CJP revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was financially supported by the National key research and development program (2018YFD0201200), China Agriculture Research System (CARS-24-C-04) and sponsored by K. C. Wong Magna Fund in Ningbo.
University. The funders had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials
All data generated or analyzed during this study are included in this published article and its Additional files. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 29 March 2020 Accepted: 16 June 2020

Published online: 01 July 2020

References
1. De Lorenzo G, et al. Cell wall traits that influence plant development, immunity, and bioconversion. Plant J. 2019;97(1):134–47.
2. Bacete L, et al. Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. Plant J. 2018;93(4):614–36.
3. Rui Y, Dinneny JR. A wall with integrity: surveillance and maintenance of the plant cell wall under stress. New Phytol. 2019;225(4):1428–39.
4. Showalter AM, et al. A bioinformatics approach to the identification, identification and bioinformatic analyses. Mol Gen Genomics. 2006;276(5):478–94.
5. Hijioli M, et al. An update on post-translational modifications of hydroxyproline-rich glycoproteins. Plant Physiol. 2010;153(2):485–513.
6. Clout NJ, Tisi D, Hohenester E. Novel fold revealed by the structure of a fasciclin-like arabinogalactan proteins. Cell Mol Life Sci. 2001;58(10):1399–417.
7. Clout NJ, Tisi D, Hohenester E. Novel fold revealed by the structure of a FAS1 domain pair from the insect cell adhesion molecule fasciclin I. Structure. 2003;11(2):197–203.
8. Johnson KL, et al. The fasciclin-like arabinogalactan proteins of Arabidopsis: A multigene family of putative cell adhesion molecules. Plant Physiol. 2003;133(4):1911–25.
9. Ma H, Zhao J. Genome-wide identification, classification, and expression analysis of the arabinogalactan protein gene family in rice (Oryza sativa L.). J Exp Bot. 2010;61(10):2647–68.
10. Falk A, Abouzouhair J, Sarhan F. Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (Triticum aestivum) and rice (Oryza sativa): identification and bioinformatic analyses. Mol Gen Genomics. 2006;276(5):478–94.
11. Zang L, et al. Genome-wide analysis of the Fasciclin-like Arabinogalactan protein gene family reveals differential expression patterns, localization, and salt stress response in Populus. Front Plant Sci. 2015;6:1140.
12. Huang GQ, et al. Characterization of 19 novel cotton FLA genes and their expression profiling in fiber development and in response to phytohormones and salt stress. Physiol Plant. 2008;134(2):348–59.
13. Jun L, Xiaoming W. Genome-wide identification, classification and expression analysis of genes encoding putative fasciclin-like arabinogalactan proteins in Chinese cabbage (Brassica rapa L.). Mol Bio Rep. 2012;391(12):10541–55.
14. MacMillan CP, et al. The fasciclin-like arabinogalactan protein family of Eucalyptus grandis contains members that impact wood biology and biomechanics. New Phytol. 2015;206(4):1314–27.
15. Guerriero G, et al. Identification of fasciclin-like arabinogalactan proteins in textile hemp (Cannabis sativa L.) in silico analyses and gene expression patterns in different tissues. BMC Genomics. 2017;18(1741).
16. Wang H, et al. Fasciclin-like arabinogalactan proteins, PfLAs, play important roles in GA-mediated tension wood formation in Populus. Sci Rep. 2017;7(1):6182–13.
17. Zhang Z, et al. Xylem sap in cotton contains proteins that contribute to environmental stress response and cell wall development. Funct Integr Genomics. 2015;15(1):17–26.
18. Wang H, et al. Antisense expression of the fasciclin-like arabinogalactan protein FLA6 gene in Populus inhibits expression of its homologous genes and alters stem biomechanics and cell wall composition in transgenic trees. J Exp Bot. 2015;66(5):1291–302.
19. Qin LX, et al. The cotton β-galactosyltransferase 1 (GaT1) that galactosylates arabinogalactan proteins participates in controlling fiber development. Plant J. 2017;89(5):957–71.
20. Li Y, et al. Suppression of GaAGP4 gene expression repressed the initiation and elongation of cotton fiber. Plant Cell Rep. 2012;29(2):193–202.
21. Yang X, et al. Expression profile analysis of genes involved in cell wall regeneration during protoplast culture in cotton by suppression subtractive hybridization and macroarray. J Exp Bot. 2008;59(13):3661–74.
22. MacMillan CP, et al. Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus. Plant J. 2010;62(4):689–703.
23. Cagnola JI, et al. Reduced expression of selected FASCI LIN-LIKE ARAB INOGALACTAN PROTEIN genes associates with the abortion of kernels in field crops of Zea mays (maize) and of Arabidopsis seeds. Plant Cell Environ. 2018(41)(3):661–74.
24. Li J, et al. The fasciclin-like arabinogalactan protein gene, FLA3, is involved in microspore development of Arabidopsis. Plant J. 2010;64(3):482–97.
25. Seifert GJ, Xue H, Acet T. The Arabidopsis thaliana Fasciclin LIKE ARABINOGALACTAN PROTEIN 4 gene acts synergistically with abscisic acid signalling to control root growth. Ann Bot. 2014;114(6):1125–33.
26. Johnson KL, et al. A fasciclin-like arabinogalactan-protein (FLA) mutant of Arabidopsis thaliana, fla1, shows defects in shoot regeneration. PLoS One. 2011;6(9):e25154.
27. Basu D, et al. Glycosylation of a fasciclin-like arabinogalactan-protein (SOS5) mediates root growth and seed mucilage adherence via a cell wall receptor-like kinase (FEI1/FEI2) pathway in Arabidopsis. PLoS One. 2016;11(1):e0154092.
28. Xue H, et al. Arabidopsis thaliana FLA4 functions as a glycan-stabilized soluble factor via its carboxy-proximal Fasciclin 1 domain. Plant J. 2017;91(4):613–30.
29. Li W, et al. Identification of early salt stress responsive proteins in seedling roots of upland cotton (Gossypium hirsutum L.) employing iTRAQ-based proteomic technique. Front Plant Sci. 2015;6:732.
30. Guerriero G, et al. Textile Hemp vs. Salinity: Insights from a targeted gene expression analysis. Genes. 2017;8(10):242.
31. Shi H, et al. The Arabidopsis SOS5 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. Plant Cell. 2003;15(1):19–32.
32. Takahashi D, Kawamura Y, Uemura M. Cold acclimation is accompanied by complex responses of glycosylphosphatidylinositol (GPI)-anchored proteins in Arabidopsis. J Exp Bot. 2016;67(17):5203–15.
33. Ge P, et al. ITRAQ-based quantitative proteomic analysis reveals new metabolic pathways of wheat seedling growth under hydrogen peroxide stress. Proteomics. 2013;13(20):3046–58.
34. Bombaraie A, et al. A draft genome sequence of Nicotiana benthamiana to enhance molecular plant-microbe biology research. Mol Plant Microbe Interact. 2012;25(12):1523.
35. Lafarguette F, et al. Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phytol. 2004;164(1):102–21.
36. Yu Y, et al. The grapevine R2R3-type MYB transcription factor VdMYB1 positively regulates defense responses by activating the stilbene synthase gene 2 (VdST52). BMC Plant Biol. 2019;19(1):478.
37. Kurtilla A, et al. Nectar- and stigma exudate-specific expression of an acidic chitinase could partially protect certain apple cultivars against fire blight disease. Planta. 2019;251(1):20.
38. Shan T, et al. The wheat R2R3-MYB transcription factor TaRIM1 participates in resistance response against the pathogen R. cerealis infection through regulating defense genes. Sci Rep. 2016;6:28777.
39. Zhang H, et al. NkCZF1, a novel C2H2-type zinc finger protein, as a new regulator of SsCut-induced plant immunity in Nicotiana benthamiana. Plant Cell Physiol. 2016;6:5149.
40. Yu YH, et al. Grape (Vitis vinifera) VvDOF3 functions as a transcription activator and enhances powdery mildew resistance. Plant Physiol Biochem. 2019;143:183–9.
41. Chezem WR, et al. SG2-type R2R3-MYB transcription factor MYB15 controls defense-induced lignification and basal immunity in Arabidopsis. Plant Cell. 2017;29(8):1907–26.
42. Zhang Y, et al. The R3R MYB transcription factor MdMYB30 modulates plant resistance against pathogens by regulating cuticular wax biosynthesis. BMC Plant Biol. 2019;19(1):362.

43. Cutler SR, et al. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. Proc Natl Acad Sci U S A. 2000;97(7):3718–23.

44. Dahia P, et al. A fasciclin-domain containing gene, ZeFAN11, is expressed exclusively in xylem elements that have reticulate wall thickenings in the stem vascular system of Zinnia elegans cv envy. Planta. 2006;223(6):1281–91.

45. Fang Q, et al. A salt-stress-regulator from the poplar R3R MYB family integrates the regulation of lateral root emergence and ABA signaling to mediate salt stress tolerance in Arabidopsis. Plant Physiol Biochem. 2017;114:100–10.

46. Fang Q, et al. AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. Plant Cell Rep. 2018;37(11):1499–511.

47. Sun B, et al. TaZFP1, a C2H2 type-ZFP gene of T. aestivum, mediates salt stress tolerance of plants by modulating diverse stress-defensive physiological processes. Plant Physiol Biochem. 2019;136:127–42.

48. Lorrai R, et al. Genome-wide RNA-seq analysis indicates that the DAG1 transcription factor promotes hypocotyl elongation acting on ABA, ethylene and auxin signaling. Sci Rep. 2018;8(1):15895–13.

49. Sultana N, et al. Ascorbate deficiency influences the leaf cell wall glycoproteome in Arabidopsis rht(1s). Plant Cell Environ. 2015;38(2):375–84.

50. Pérignon R, et al. Gene expression trade-offs between defence and growth in English elm induced by Ophiostoma novo-ulmi. Plant Cell Environ. 2018;41(1):198–214.

51. Scott Gens J, Fujiki M, Pickard BG. A fasciclin-domain containing gene, ZeFLA11, is expressed exclusively in xylem elements that have reticulate wall thickenings in the stem vascular system of Zinnia elegans cv envy. Planta. 2006;223(6):1281–91.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.