Journal of Proteomics & Bioinformatics - Open Access

Research Article JPB/Vol.1/October 2008

Identification and Analysis of the Arabidopsis Thaliana Atfas4 Gene Whose Overexpression Results in the Development of A Fasciated Stem

Gennady Pogorelko*, Oksana Fursova, Eugene Klimov

NI Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin Str., 3, 119991, Moscow, Russia

Corresponding author: Gennady Pogorelko, NI Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin Str., 3, 119991, Moscow, Russia, E-mail: gpogorelko@yandex.ru

Received September 25, 2008; Accepted October 10, 2008; Published October 10, 2008

Citation: Gennady P, Oksana F, Eugene K (2008) Identification and Analysis of the Arabidopsis Thaliana Atfas4 Gene Whose Overexpression Results in the Development of A Fasciated Stem. J Proteomics Bioinform 1: 329-335. doi:10.4172/jpb.1000041

Abstract

Using a new pEnLox vector employed to generate gain-of-function mutants in Arabidopsis thaliana, the AtFAS4 mutant has been obtained and analyzed. The mutant is characterized by super-expression of the At1g33390 gene, which leads to the occurrence of a mutant phenotype – stem fasciation. The level of expression of the AtFAS4 gene in normally developing A.thaliana plants is extremely low thus accounting for almost complete absence of information on EST's of this gene. The generated AtFAS4 mutant has permitted full-length cDNAof the At1g33390 gene to be obtained and analyzed for the first time.

Keywords: Activation tagging; *Arabidopsis thaliana*; Full length cDNA; Stem Fasciation

Abbreviations

CaMV35S: promoter of 35S ribosomal RNA gene from Cauliflower Mosaic Virus
Col-0: Columbia 0 arabidopsis thaliana ecotype
FSTs: Flanking Sequence Tags
T-DNA: Transferred DNA
TAIR: The Arabidopsis Information Resource

Introduction

Large collections of Arabidopsis thaliana transgenic plants are presently available with T-DNA or mobile elements used as insertions and can be found at the following locations; ABRC (http://www.arabidopsis.org/abrc/), NASC (http://nasc.nott.aac.uk/), SALK (http://signal.salk.edu/) and GABI (http://www.gabi-kat.de/). Most of these collections, however, are based on insertions causing so called loss of function mutations related to inactivation of genes carrying insertions (Feldmann, 1991). Other types of mutations, like mutations determined by super-expression of genes, are represented in the available collections rather poorly.

A system of two vectors, pEnLox (DQ645630) and pCre (DQ635631), has been previously constructed (Pogorelko et al., 2007). Vector pEnLox contains an enhancer (a tetramer of the cauliflower mosaic virus 35S RNA promoter) able to induce super-expression of genes adjacent to the site of insertion integration and directed to the start-codon, ith this direction corresponding to the direction of gene transcription.

Eight lines have been obtained with the use of the pEnLox vector carrying an insertion in such orientation that the enhancer in the T-DNA structure is directed to the start codon of translation of a nearby gene (Pogorelko et al., 2008). Thus this arrangement is potentially able to cause superexpression of the nearby gene. Three of the mutant lines differ by phenotype from wild-type plants. A molecular genetic analysis of the E78 line has been carried out and the mutant plants of this line have been designated as *AtFAS4*.

Materials And Methods

Plant Material And Growth

Arabidopsis thaliana plants of the *Col*-0 ecotype were grown at 21-23°C in normal day conditions (16 h of light and 8 h of dark) under Phillips BioLux fluorescent lights.

RNA Purification and cDNA Synthesis

RNA isolated from the plants using a RNA purification kit (Plant RNA Isolation Reagent, www.invitrogen.com) served as a template for synthesis of the first cDNA strand using an Invitrogen kit (5´ RACE System for Rapid Amplification of cDNA Ends; 3´ RACE System for Rapid Amplification of cDNA Ends; GeneRacer® Kit with AMV RT and TOPO TA Cloning® kit for Sequencing) (www.invitrogen.com).

Identification of full-length cDNA of the *At1g33390* gene

Full-length cDNA was synthesized using a special kit from Evrogen (www.evrogen.ru) strictly following the instructions. Since *AtFAS4* cDNA is more than 4 kb in size, its amplification was carried out in 5 sequentially overlapping PCR fragments with the following primers:

Fragment 1 (5'end)
M1 (Evrogen)
At1g33390_Rev1 ACGACAAGTTGTCAGTGGGA
Fragment 2
At1g33390_Forw2 GCAACTGGTTCATGCTGATC
At1g33390_Rev2 GACTCTTTTGTTGTTCCTCG
Fragment 3
At1g33390_Forw3 CTTATTGAGGCGTTATCTTG
At1g33390_Rev3 ACAGTGACCAGGTCCAGTTC
Fragment 4
At1g33390_Forw4 ACTTCTCTGACTATTCCTGG
At1g33390_Rev4 GGAGCTGAGTTTATGAGAGA
Fragment 5 (3'-end)
At1g33390_Forw5 GTGTAGCCAGAAAGACCAGAG
M1 (Evrogen)

Bioinformatic tools

Promoter regions were identified using the PromoterInspector (http://www.genomatix.de/) and Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) programs which allow to predict promoter regions and initial transcription sites. The MatInspector program with plant filter (http://www.genomatix.de/) was used for search of binding sites of transcription factors.

Results

Phenotype And Genotype Of The Mutant

The morphological distinction of the E78 line is a clearly visible fasciated stem consisting of 7-12 stems (Fig.1). Segregation analysis showed the mutant trait to be inheritable in a monogenic fashion (3:1 segregation, 0.56<3.84 for p=0.05) (Table 1). Southern blot-hybridization demonstrated the presence of a single T-DNA insertion copy in the genome of E78 plants (Suppl. fig. 1). The modified inverse PCR method was used to identify the site of T-DNA insertion in the genome of E78 line plants (Pogorelko and Fursova, 2008). Sequencing and subsequent analysis *in silico* of the obtained DNA fragment allowed us to determine the place of T-DNA integration in the genome of E78 plants (GenBank Acc.: EI183464). The insertion is localized in the precentromeric region of chromosome 1. A more detailed analysis showed the insertion to precede the start-codon of the *AtFAS4* gene (approximately at a distance of 0.5 kb) and the enhancer to be oriented to the start-codon. This enhancer acts at a distance of 380 bp-3.6 kb from the target gene (Weigel et al., 2000).

Analysis of *AtFAS4* Expression

Since the mutant phenotype of E78 plants is determined by the activity of the enhancer contained in T-DNA, we analyzed *AtFAS4* gene expression in 4 independent E78 mutant sub lines (T1 generation plants that had been selected by "Basta") and in wild-type plants (Fig.2).

The AtFAS4 gene was found to be expressed in the mutant plants and not expressed in the wild-type plants. Analysis of the databases containing information on ESTs demonstrated the availability of only two experimentally obtained partial cDNAs from the 3'-region of At1g33390 mRNA (GenBank Acc.: AU235278 and AU225941) (http://rarge.gsc.riken.jp/). mRNA of the At1g33390 gene represented in the databases was obtained by the methods of computer analysis of the genome (NM_103064).

Figure 1: Comparison of the phenotypes of plants E78 and A.thaliana Col0. Age – 7 weeks . (A) AtFAS4 mutant, (B) Wild-type A.thaliana Col-0 mutant

Line Name	Gene/Length	Protein Description	Localization/ Orientation of T-DNA Insertion	GeneBank accession no.	Number of plants, resistant to herbicide in T2 generation	Number of plants, NOT resistant to herbicide in T2 generation	χ^2	Amount of insertions
E78	*AtFAS4 (AT1G33390) 4371 bp*	DEAH-box ATP-dependent helicase	Localized in 530bp before Start-Codon. Enhancer directed on Start-Codon	EI183464	139	51	0,56	1

Table 1: Description of E78 mutant: predicted in silico gene length and putative protein product, localization and amount of insertions.

Analysis of the Genevestigator database (https://www.genevestigator.eyhz.ch/) containing information on the expression of genes in different organisms with the use of the Microarray technology provided information on the expression of the *AtFAS4* gene (Fig.3) using mRNA extracted from the *A.thaliana* plants grown under normal conditions.

According to these data, the *AtFAS4* gene has a very low expression as compared to the constitutively expressed "house-keeping" *At3g18780* gene whose level of transcription is characteristic of genes controlled by the CaMV35S promoter.

Thus, the analysis of information from different databases confirmed our experimental results.

Analysis of *AtFAS4* Gene Structure

We obtained and analyzed full-length cDNA of the *AtFAS4* gene from a mutant plant (EF630362). As a result, the gene was found to have 8 exons, which is in agreement with the predicted gene model (NM_103064). However the synthesis of *AtFAS4* mRNA in a mutant plant of the E78 line starts from the promoter of the vector and as a consequence the point of initiation of *AtFAS4* gene transcription remains unknown. We used "PromoterInspector" and "Promoter predictions" programs to determine *in silico* the site of transcription initiation and 5'-UTR.

Therefore our results permit elaboration of the primary structure of the AtFAS4 gene that was previously predicted in silico (NM_103064). The promoter predicted with the Promoter predictions program corresponds to the promoter predicted on a Genomatix server. However the transcription start point predicted with the Promoter predictions program is at a distance of 15 bp after the start of translation in the gene model (NM_103064). The start of translation determined by us is in position 26 from the start of transcription predicted by the programs. Further analysis of the amino acid sequence permits our variant of the 5'-region structure to be considered as more correct.

Journal of Proteomics & Bioinformatics - Open Access
Research Article JPB/Vol.1/October 2008

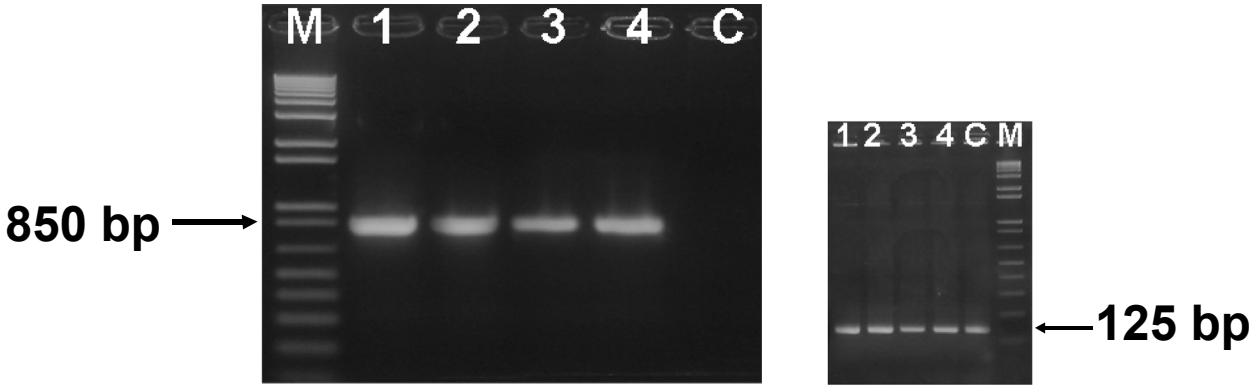

Figure 2: An electrophoregram of RT-PCR products of an AtFAS4 fragment (A) and control Actin-2 fragment (B). The overall pools of RNA isolated from 4 independent mutant plants and Col-0 wild type Arabidopsis plant were used as a template for synthesis of the first cDNA strand using an Invitrogen kit (www.invitrogen.com). The RT products were used as a template for PCR with primers At1g33390_ Forw2 (GCAACTGGTTCATGCTGATC), At1g33390_Rev2 (GACTCTTTTGTTGTTCCTCG). The actin-2 house-keeping gene was used to normalize the amount of RNA and to control RNA quality. For the Actin-2 amplifying two primers, Actin-2For (CTCTCCCGCTATGTATGTCGC) and Actin-2Rev (GAAACCCTCGTAGATTGGCA) were used.

(A) M. 1kb+ Invitrogen DNA marker
1. E78/1-mutant plant
2. E78/2-mutant plant
3. E78/3-mutant plant
4. E78/4-mutant plant
C. Wild-type *A.thaliana* Col0

(B) 1. E78/1-mutant plant
2. E78/2-mutant plant
3. E78/3-mutant plant
4. E78/4-mutant plant
C. Wild-type *A.thaliana* Col0
M. 1kb+ Invitrogen DNA marker

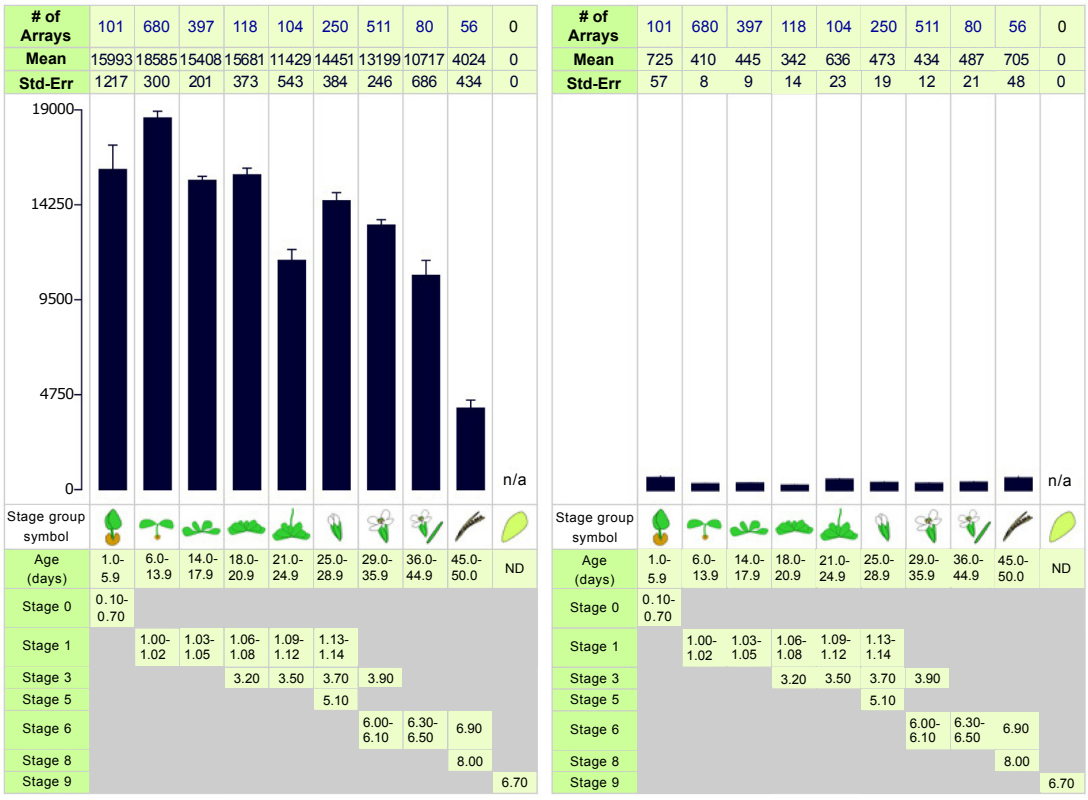

Figure 3: Comparison of expression of the genes At3g18780 (Housekeeping gene, transcriptional factor with constitutive expression) (left) and AtFAS4 (right) in wild-type plants.

ISSN:0974-276X JPB, an open access journal

Analysis of the promoter, using MatInspector program, revealed the presence of 12 sites of binding of transcriptional factors (TF) on the positive chain. These transcriptional factors belong to 5 TF families characteristic of plants (Table 2, Fig. 4). The optional parameters used for TF searching are presented in Table 2.

Table 2.

Family/matrix	Further Information	Opt.	Matrix sim.	Sequence (capitals: core sequence)
P$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	0.957	gaattgtcAAAGgtttt
P$TALE/HVH21.01	Homeodomain protein of the Knotted class1	1.00	0.969	taTGACataactt
P$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	0.975	tgaatttgAAAGctata
P$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	0.985	gggattgtAAAGtgttt
P$CAAT/CAAT.02	CCAAT-box in plant promoters	1.00	0.993	aatCAATta
P$SEF4/SEF4.01	Soybean embryo factor 4	0.98	0.958	tgTTTTttttt
P$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	0.974	gtagaaacAAAGtggtt
P$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	0.963	gttgtTAATca
P$SEF4/SEF4.01	Soybean embryo factor 4	0.98	0.980	agTTTTtgttt
P$CAAT/CAAT.02	CCAAT-box in plant promoters	1.00	0.991	ttgCAATtt
P$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	0.976	agaagaatAAAGgatcc
P$CAAT/CAAT.02	CCAAT-box in plant promoters	1.00	0.998	atcCAATaa

Core similarity for all TF is 1.00.

Table 2: Plant specific transcription factors predictably controlling the At1g33390 gene expression.

ISSN:0974-276X JPB, an open access journal

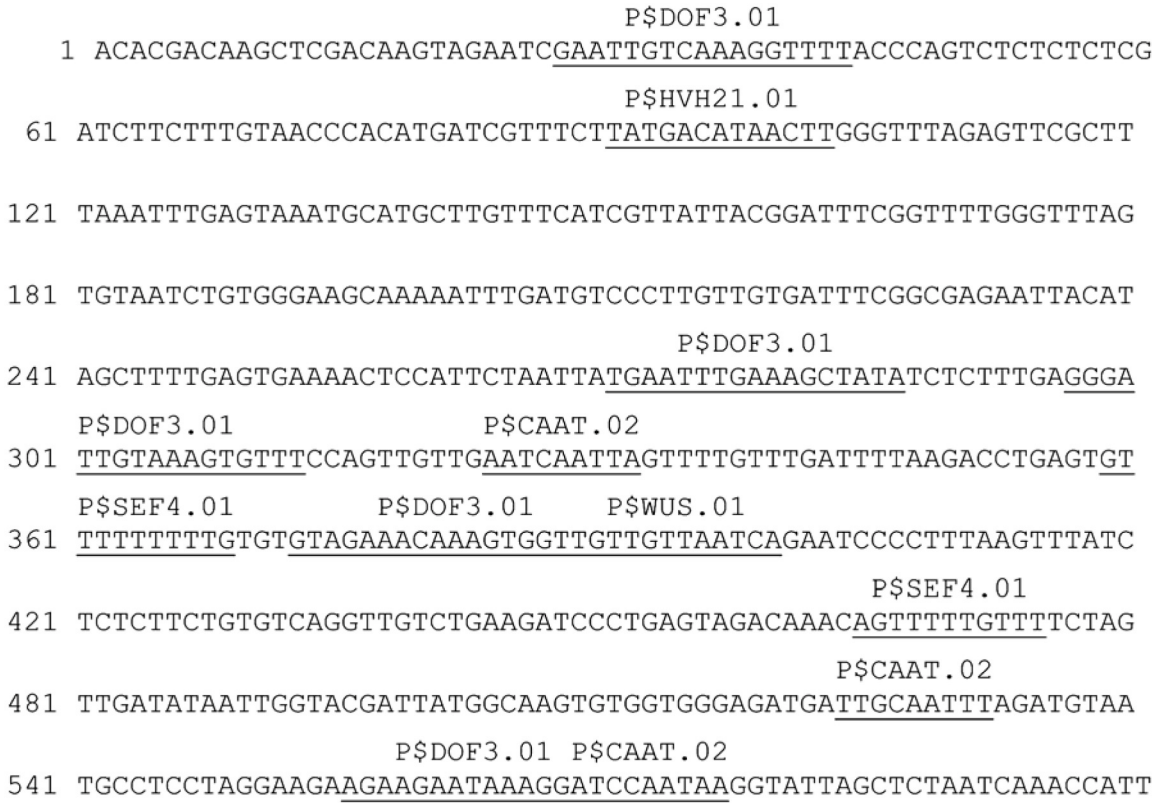

Figure 4: Promoter region of the AtFAS4 gene with recognition sites of the transcription factors.

In silico Analysis of the *At1g33390* Amino Acid Sequence

Comparison of the *AtFAS4* amino acid sequence with the known or predicted proteins showed the presence of two domains and fragments of two more domains. The omain of DEXH-box or DEAD-box helicases is located at the N-end of the protein. It participates in ATP-dependent denaturation of DNA or RNA. The C-end of the protein contains the DUF1605 domain. Its function is unknown but it is always found at the C-end of DEAD-box helicases. A fragment of the HrpA domain (the central part of the protein) is found in HrpAlike helicases involved in DNA replication, recombination and repair. A fragment in proximity to the C-end of the protein represents the HA2 domain (associated with helicases). Its function is unknown but it is supposed to be involved in the process of binding of the protein with nucleic acids.

Discussion

Our analysis of the function of the *AtFAS4* gene permits the following conclusions to be made. The gene is likely to be expressed at certain stages of plant development or under the action of some factors that depend on the stage of development and on environmental conditions. Considering the presence in the *AtFAS4* protein of domains characteristic of helicases (for which the substrate can be both DNA and RNA), it can be assumed that the synthesis of *AtFAS4* mRNA can occur in short periods of plant ontogenesis when it is necessary to activate certain RNA or DNA molecules. Overexpression of the helicase coding gene causes development of a fascinated stem that was also shown in previously published work of Soichi Inagaki with colleagues (Inagaki et al., 2006). This assumption is supported by a short lifetime of the protein and its susceptibility to a large number of peptidases. Thus, cDNA of the *AtFAS4* gene has been experimentally obtained for the first time and the mutant phenotype of A.thaliana characterized by the development of fasciated stems under the action of AtFAS4 gene super-expression has been described.

So we would suggest that overexpressing mutants like AtFAS4 can be used as a stand-alone convenient technique for EST and cDNA studying in case of genes that have no or low transcription level in wild type plants.

Acknowledgments

We thank to Jean-Denis Faure and Francois Roudier from the Cell Biology Laboratory of Versailles INRA for assistance at the initial stages of this work.

References

1. Feldmann KA (1991) T-DNA insertion mutagenesis in Arabidopsis. Mutational spectrum. The Plant Journal 1: 71-82. » CrossRef » Google Scholar

2. Pogorelko GV, Fursova OV, Ogarkova OA, Tarasov VA (2007) New vector system for induction of gene expression in dicotyledonous plants. Genetika (RUS) 43:194-201. » CrossRef » Pubmed » Google Scholar

3. Pogorelko GV, Fursova OV, Ogarkova OA, Tarasov VA (2008) A new technique for activation tagging in Arabidopsis. Gene 414: 67-75. » CrossRef » Pubmed » Google Scholar

4. Pogorelko GV, Fursova OV (2008)A highly efficient miPCR method for isolating FSTs from transgenic *Arabidopsis thaliana* plants. J Gent 87: 133-140. 133-140. » CrossRef » Pubmed » Google Scholar

5. Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, et al. (2000) Activation tagging in Arabidopsis. Plant Physiol 122: 1003-13. » CrossRef » Pubmed » Google Scholar

6. Inagaki S, Suzuki T, Ohto M, Urawa H, Horiuchi T, et al. (2006) *Arabidopsis* TEBICHI, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems. Plant Cell 18: 879-892. » CrossRef » Pubmed » Google Scholar