Identification and expression analysis of *LATERAL ORGAN BOUNDARIES DOMAIN (LBD)* transcription factor genes in *Fragaria vesca*

Journal:	<i>Canadian Journal of Plant Science</i>
Manuscript ID	CJPS-2017-0098.R1
Manuscript Type:	Article
Date Submitted by the Author:	13-Jun-2017
Complete List of Authors:	Chen, Xiaodong; Institute of Pomology, Jiangsu Academy of Agricultural Sciences Wang, Jing; Institute of Pomology, Jiangsu Academy of Agricultural Sciences Zhao, Mizhen; Institute of Pomology, Jiangsu Academy of Agricultural Sciences Yuan, Huazhao; Institute of Pomology, Jiangsu Academy of Agricultural Sciences
Keywords:	Strawberry, Molecular biology, Horticulture

SCHOLARONE™
Manuscripts

Identification and expression analysis of *LATERAL ORGAN BOUNDARIES****DOMAIN (LBD)* transcription factor genes in *Fragaria vesca***

Xiaodong Chen, Jing Wang , Mizhen Zhao, Huazhao Yuan

Institute of Pomology, Jiangsu Academy of Agricultural Sciences/Jiangsu Key

Laboratory for Horticultural Crop Genetic Improvement.

Corresponding author: Mizhen Zhao (njzhaomz@163.com)

Short Title: Identification and expression analysis of *LBD* genes in strawberry

Abstract

The *LATERAL ORGAN BOUNDARIES DOMAIN (LBD)* gene family encodes plant-specific transcription factors that play crucial roles in the growth and development in many plant species. However, no systematic study of *LBD* genes has been conducted in strawberry. In this study, 35 *LBD (FvLBD)* genes were identified in the diploid woodland strawberry genome (*Fragaria vesca*). These LBD proteins could be classified into two groups based on the structure of their lateral organ boundaries (LOB) domain. The promoters of *FvLBD* genes contain different regulatory elements associated with potential response to different environmental stimuli and plant developmental signals. Furthermore, we analysed the expression patterns of the *LBD* genes during the callus formation in strawberry, and the results suggested that *FvLBD16* might play a prominent role in the regulation of callus formation. In addition, we investigated the expression profiles of *FvLBD* genes during early fruit development based on the transcriptome data. We found that some *FvLBD* genes show a specific expression pattern. These results indicate that the *FvLBD* genes may function in the early fruit development. Together, the present study provides insights into possible functions of the *FvLBD* genes and provides a basis for further functional research of *FvLBD* proteins.

Keywords: *LBD* transcription factors, *Fragaria vesca*, strawberry, callus formation, early fruit development.

Introduction

LATERAL ORGAN BOUNDARIES DOMAIN (LBD) transcription factors play versatile functions in the regulation of lateral organ development and metabolic processes in higher plants, such as anthocyanin and nitrogen metabolism (Shuai et al. 2002; Majer and Hochholdinger, 2011; Xu et al. 2016). The first *LATERAL ORGAN BOUNDARIES* (*LOB*) gene was identified in *Arabidopsis* and found to be expressed at the boundaries of lateral organs during plant development (Shuai et al. 2002; Luo et al. 2016). This class of transcription factors is characterized by a highly conserved LOB domain. The characteristic LOB domain comprises a C block containing four cysteine residues with spacing (CX₂CX₆CX₃C) required for DNA-binding activity, a Gly-Ala-Ser (GAS) block (beginning with an FX₂(V/A)H motif and ending with a DP(VI)YG motif) and a leucine zipper-like coiled-coil motif (LX₆LX₃LX₆L), responsible for protein dimerization (Shuai et al. 2002; Xu et al. 2016). According to the protein sequences, the LBD proteins could be further divided into two subclasses (Class I and Class II). All LBD proteins contain the C block, while the GAS block and the L-rich block are only observed in Class I LBD proteins (Shuai et al. 2002; Luo et al. 2016).

The LBD proteins play a versatile role in the regulation of plant growth and development. For example, LBD6/AS2 has a function in defining lateral organ boundaries and controlling leaf polarity and flower development. The mutation of *LBD6* causes a lotus-leaf structure and smaller sepals and petals in *Arabidopsis* (Xu, 2003; Xu et al. 2008). *Arabidopsis LBD18* was shown to have a role in regulating

lateral root initiation through regulating the cell cycle-controlling gene *E2Fa* and the cell wall-loosening factor *EXPANSIN4* (Lee et al. 2009; Berckmans et al. 2011; Lee et al. 2013). The *LBD* genes also function in regulating plant cell wall thickening and secondary growth. For example, *Arabidopsis LBD18* and *LBD30* were reported to be involved in xylem differentiation in the leaf and root (Soyano et al. 2008), and poplar *LBD1* had a positive role in secondary phloem growth (Yordanov et al. 2010).

In addition, the *LBD* genes were reported to be important in regulating callus formation. The *LBD16*, *LBD17*, *LBD18*, and *LBD29* genes were identified as key regulators of callus induction. The ectopic expression of each of these four genes was able to trigger callus formation without exogenous auxin application, while the suppression of these *LBD* genes blocked callus formation (Fan et al. 2012; Liu et al. 2014). *Arabidopsis LBD25* was reported to be involved in photomorphogenesis through regulating light/dark dependent hypocotyl elongation (Mangeon et al. 2011; Xu et al. 2016). LBD proteins are also involved in metabolic processes, as the Class II LBDs 37, 38, and 39 act as negative regulators of anthocyanin biosynthesis in *Arabidopsis* (Rubin et al. 2009). In addition, *LBD* genes also function in the plant defence response process. For example, *Arabidopsis LBD20* acted as a negative regulator of susceptibility to the root-infecting fungal pathogen *Fusarium oxysporum*. The expression of *LBD20* was induced after *F. oxysporum* inoculation. Furthermore, the *lbd20* mutant showed enhanced resistance to *F. oxysporum*, while plants overexpressing *LBD20* were sensitive to *F. oxysporum* infection (Thatcher et al. 2012; Xu et al. 2016).

Recently, a number of LBD proteins have been identified in many plant species, such as rice (*O. sativa* (Yang et al. 2006)), maize (*Zea mays* (Zhang et al. 2014)), grape (*Vitis vinifera* (Cao et al. 2016)), apple (*Malus domestica* (Wang et al. 2013)), and *Brachypodium distachyon* (Gombos et al. 2017). However, the *LBD* gene family has not been identified in the strawberry genome.

The cultivated strawberry (*Fragaria* × *ananassa* Duch.) is an important horticultural crop for the production of fresh fruit and juice and is also an excellent model plant for fleshy fruit development (Kang et al. 2013; Wei et al. 2016). *Fragaria* × *ananassa* has an extremely complex genome, harbouring eight sets of chromosomes ($2n = 8x = 56$) derived from as many as four different diploid ancestors, which makes genetic studies very complicated (Shulaev et al. 2011). Thus, the sequenced diploid woodland strawberry (*Fragaria vesca*) is emerging as a model for fruit crop plant species (Shulaev et al. 2011; Gu et al. 2016). In the present study, we performed a broad analysis of the *LBD* genes in the diploid woodland strawberry by investigating their chromosome position, evolutionary relationships, protein motifs, gene structure, cis-acting elements, and expression profiles during callus formation and early fruit development. Our research will provide a better understanding of the function of the *FvLBD* gene family in strawberry.

Materials and Methods

Plant materials and culture conditions

Seeds of *F. vesca* ssp. *vesca* accession Hawaii 4 (Kindly provided by Janet Slovin) were treated with 75% alcohol for 20 min and then washed four times with sterilized

water. The seeds were kept at 4°C for 4 days and then placed on 1/2 MS medium (half-strength MS basal medium with 2% sucrose, 1% agar, and 0.5 g L⁻¹ of MES, pH 5.7 (Murashige and Skoog, 1962)). Strawberries were grown at 23°C with a 16-h light and 8-h dark photoperiod in a phytotron. Fully expanded leaves of strawberry plants were placed on callus-inducing medium (1/2 MS basal medium with 2% sucrose, 0.8% agar, 0.5 g L⁻¹ of MES, 3 mg L⁻¹ of 6-benzyladenine (6-BA), and 0.2 mg L⁻¹ of indole-3-butyric acid (IBA)) followed by incubation at 23°C in the dark.

Identification of *LBD* genes from strawberry genome

We downloaded the Hidden Markov Model (HMM) profile of the LBD family (DUF260, Pfam number: PF03195) from the Pfam database (Finn et al. 2016). We then searched the *Fragaria vesca* genome (fvesca_v1.0_genemark_hybrid) database (downloaded from Genome Database for Rosaceae (GDR, <http://www.rosaceae.org>) (Jung et al. 2014)) using HMMER3.0 to identify genes with typical DUF260 motifs using the default parameters. All of the putative genes were further verified to have the conserved *LOB* domain using the Conserved Domain Database at NCBI (CDD, <https://www.ncbi.nlm.nih.gov/cdd> (Marchler-Bauer et al. 2017)) and the Pfam database (<http://pfam.xfam.org/search>) (Finn et al. 2016). The information regarding the gene name, gene ID, gene length, and number of amino acids was retrieved from the diploid woodland strawberry genome database at GDR (Jung et al. 2014).

Chromosomal location, phylogenetic analysis, protein domains, and gene

structure analysis of strawberry *LBD* Genes

The detailed positions of all of the *FvLBD* genes were retrieved from NCBI. MapInspect software was used to clarify the locations of the *LBD* genes in the strawberry genome (<http://www.files32.com/MapInspect-i366074.asp>). The *LBD* protein sequences of *F. vesca* and *Arabidopsis* were downloaded from GDR (Jung et al. 2014) and TAIR (<http://www.arabidopsis.org/>) (Lamesch et al. 2012), respectively. The phylogenetic analysis of the *LBD* protein sequences of *F. vesca* and *Arabidopsis* was performed using MEGA 5.0 software (Tamura et al. 2011). A phylogenetic tree was generated using the neighbour-joining method with the p-distance model, pairwise deletion of gaps, and bootstraps (1000 replicates; random seed). The protein domains of the *LBD* genes were identified using MEME online tools (<http://meme-suite.org/tools/meme>) (Bailey et al. 2009) with the following parameters: minimum length of the motif, 6; maximum length of the motif, 100; and maximum number of motifs, 7 (*F. vesca*) and 10 (*Arabidopsis*). The exon-intron gene structures of the *LBD* genes were visualized using the online programme GSD2.0 (<http://gsds.cbi.pku.edu.cn/>) (Hu et al. 2015).

Promoter analysis

The *FvLBD* nucleotide sequences were used to retrieve whole-genomic sequences from NCBI (Table S1) using the Nucleotide BLAST programme. Then the upstream 1.5-kb region of the translation start site of the *FvLBD* genes was obtained from NCBI. PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)

(Lescot et al. 2002) was used to analyse the cis-element(s) in the promoter regions of *FvLBD* genes.

Expression analysis

The transcriptome data for the early stages of fruit (*Fragaria vesca*, Yellow Wonder 5AF7) development were downloaded from (<http://bioinformatics.towson.edu/strawberry/>; (Kang et al. 2013)). The transcriptome data (RPKM) were calculated using a log2 scale, and the heatmap was plotted with HemI software (Deng et al. 2014).

RNA extraction and qRT-PCR

Total RNA samples were extracted from leaf explants using the BioTeke Plant Total RNA Extraction Kit (RP3301, Beijing, China) and then reverse-transcribed into cDNA using the BioTeke supermoIII RT Kit. The quantitative reverse transcription-polymerase chain reaction (qRT-PCR) was carried out on a 7500 Real Time PCR system (Applied Biosystems, New York, USA), using a SYBR Premix Ex Taq (TaKaRa) reaction SYBR kit. The GADPH gene was used as a reference (Gu et al. 2016). The primers used for qRT-PCR are listed in Supplemental Table S3. The qRT-PCR results were confirmed in three biological repeats. Each biological repetition was performed with three technical repetitions.

RESULTS

Identification of *LATERAL ORGAN BOUNDARIES DOMAIN (LBD)* genes in *F. vesca*

First, we searched the Pfam database and obtained the Hidden Markov Model (HMM) profile of the LBD family (DUF260, Pfam number: PF03195) (Finn et al. 2016). We next used HMMER3.0 to search the *Fragaria vesca* genome database using LBD HMM profiles (Eddy, 1998). We subsequently identified a total of 37 candidate genes. The sequences of the 37 candidate genes were submitted to the CDD (Conserved Domain Database at NCBI) and Pfam database of the NCBI to confirm their DUF260 domains. Based on the presence of complete DUF260 domains, 1 candidate gene (gene02445) was discarded due to the absence of a complete predicted DUF260 domain, and 1 gene (gene29787) was found to belong to the ERF/AP2 transcription factor family. Ultimately, a total of 35 *LBD* genes were identified in the diploid woodland strawberry. Detailed information about the *FvLBD* genes, including the gene IDs, gene locations, length of the gene sequences, and characteristics of the *FvLBD* proteins, is shown in Table 1. The deduced length of the *FvLBD* proteins ranged from 145 (*FvLBD24B*) to 744 amino acids (*FvLBD2*), while the pI (isoelectric point) values ranged from 5.14 (*FvLBD24A*) to 9.62 (*FvLBD22C*), which suggested that different LBD proteins might operate in different cellular microenvironments. Thirty-four *FvLBD* genes in total were distributed throughout all seven *F. vesca* pseudochromosomes (Fig. 1). Of these pseudochromosomes, pseudochromosome 7 had the most *FvLBD* genes (9), followed by pseudochromosomes 2 and 3, which had

seven and six *FvLBD* genes, respectively. Pseudochromosome 6 had five *FvLBD* genes. Three genes mapped to pseudochromosomes 1 and 5, while pseudochromosome 4 had just one *FvLBD* gene. According to a previous study, a chromosome region containing two or more genes within 200 kb can be defined as a gene cluster (Holub, 2001). In strawberry, ten *FvLBD* genes were present in four clusters, of which two were on pseudochromosome 2 and two on pseudochromosome 7 (Fig. 1).

Phylogenetic analysis and distribution of conserved motifs of FvLBD proteins and gene structure analysis of *FvLBD* genes

To gain insight into the potential function of FvLBD proteins from well-studied LBDs in the model plant *Arabidopsis*, an unrooted phylogenetic tree of LBDs from *F. vesca* and *Arabidopsis* was constructed (Fig. 2). In addition, to better understand the structure of FvLBD proteins, multiple sequence alignment was performed that indicated a sequence with approximately 100 amino acids was conserved among FvLBD proteins (Supplemental Fig. S1). The phylogenetic analysis and multiple sequence alignment showed that the strawberry FvLBD proteins could be classified into two classes that are characterized by the presence (Class I) or absence (Class II) of functional leucine zipper-like domains (Shuai et al. 2002; Majer and Hochholding, 2011). Six of the FvLBD proteins were classified into Class II, and 29 genes were assigned to Class I. We used the online MEME tool to predict the conserved FvLBD protein domains, identifying 7 conserved motifs (Fig. 3). The logos

of these motifs were obtained from the MEME Suite web server (Supplemental Fig. S2). We found that most FvLBD proteins shared similar motifs within subfamilies. Generally, all 35 of the FvLBDs showed the presence of the C block (Shuai et al. 2002). Motif 3 was identified in all Class I members, while motif 5 and motif 6 were present in all Class II members. Motif 4 was found in all Class I members and one Class II subfamily member (*FvLBD2*). In addition, 11 Class I members showed the presence of motif 7 (Fig. 3).

To investigate the structural evolution of the *F. vesca* LBD gene family, the exon-intron patterns were analysed. As shown in Fig. 3, the Class I *FvLBD* genes exhibited a highly divergent exon-intron organization of the 29 Class I *FvLBD* genes: 15 genes contained one intron, 4 genes contained two introns, and 3 genes contained three introns. *FvLBD12A* contained six introns, and 6 genes had no introns. On the other hand, the Class II *FvLBD* genes showed a highly conserved exon-intron organization: most of the Class II *FvLBD* genes had one intron, and all of the Class II genes were phase-one introns (Intron phases, the positions of introns between or within codons)(Sharp, 1981; Long and Deutsch, 1999), with the exception of *FvLBD2*, which had nine introns (Fig. 3).

Regulatory elements in the *FvLBD* promoter

To understand the specific expression of *FvLBD* genes and their roles in abiotic stress and hormone responses, a search for the existence of cis-acting regulatory elements in the 1500-bp upstream region of the translation initiation codon sites of

FvLBD genes was performed by using PlantCare online tool (Lescot et al. 2002). As a result, many cis-acting regulatory elements involved in environmental or hormone responses and in tissue/organ-specific expression were identified (Table 2, Table S2). For example, all of the *FvLBD* genes showed the presence of the light-responsive element. In addition, elements associated with heat stress and low temperature stress responses were found in 14 and 10 *FvLBD* genes, respectively. Twenty-three *FvLBD* genes had regulatory elements involved in the defence and stress responses and circadian control. The fungal elicitor response element was found in 10 *FvLBD* genes, the MYB binding site was observed in 28 *FvLBD* genes, and a cis-acting regulatory element in response to wounding was found in 5 *FvLBD* genes. In addition, many hormone-response elements were identified in *FvLBD* genes: MeJA and SA response elements were observed in 18 and 16 *FvLBD* genes, respectively; auxin response elements were observed in 8 *FvLBD* genes; GA response elements were identified in 25 *FvLBD* genes; and ethylene and ABA response elements were found in 6 and 13 *FvLBD* genes, respectively. In addition, the palisade mesophyll cell differentiation element, leaf morphology development element, endosperm expression element, meristem expression element, cell cycle-control element and root-specific expression elements were also identified in the promoters of *FvLBD* genes. Notably, many cis-acting regulatory elements involved in plant development were identified in the *FvLBD* promoter sequences, suggesting that the *FvLBD* genes may have important roles in controlling strawberry development.

Expression analysis of the *FvLBD* genes during callus formation

Unlike animals, many plants have the remarkable ability to regenerate and form an entire plant body from various tissues or organs. Callus induction is often the initial step in a typical *in vitro* plant regeneration system (Fan et al. 2012; Liu et al. 2014). The *LBD* transcription factors have been reported to be important in regulating callus formation in *Arabidopsis* (Fan et al. 2012). To gain insight into the function of *FvLBD* genes during the callus formation in strawberry, we performed quantitative reverse transcription-polymerase chain reaction (qRT-PCR) to evaluate the expression profiles of the *FvLBD* genes during callus formation (Fig. 4). As shown in Fig. 4, most of the *FvLBD* genes were downregulated with the exception of *FvLBD16*, which showed a significant increased expression profile during the process of callus formation.

Expression analysis of the *FvLBD* genes during early fruit development

The *LBD* genes have been shown to have an important role in controlling plant development in *Arabidopsis* and rice. To investigate whether and how *FvLBD* genes might be involved in early strawberry fruit development, the transcriptome data during the early stages of fruit development (Kang et al. 2013) were analysed. The genomic transcriptome data showed that the *LBD* genes are expressed differently in different tissues during early fruit development (Fig. 5). For example, *FvLBD38A* showed a continuous high level of expression during the development of both the cortex and the pith. *FvLBD25A* showed an increased expression profile during the

early cortex and pith development. *FvLBD12A* and *FvLBD12B* showed increased expression profiles from stage 1 to stage 2, although their expression levels were downregulated from stage 3 to stage 5 in the cortex and pith, respectively. Four *FvLBD* genes (*FvLBD15A*, *FvLBD18*, *FvLBD38A* and *FvLBD19*) exhibited high transcript accumulation in the wall. *FvLBD42* showed high accumulation patterns in the embryo. The *FvLBD9* and *FvLBD18* genes showed increased expression profiles during early ovule development. Three *FvLBD* genes (*FvLBD4B*, *FvLBD16*, and *FvLBD36A*) exhibited high transcript accumulation in the ghost (entire seed with its embryo removed). In summary, *FvLBD* genes show clear changes in expression in specific tissues and stages, which might be consistent with their specific functions during early fruit development.

DISCUSSION

In this study, a broad analysis of the *LBD* gene family was conducted. The chromosome position, protein motifs, gene structure and cis-acting elements of the genes were analysed in silico. qRT-PCR was subsequently performed to reveal the expression patterns during the callus formation. In addition, we further investigated the expression profiles of the *FvLBD* genes during early fruit development in silico. Genome-wide analysis of the *LBD* genes in *F. vesca* will greatly contribute to understanding their role during strawberry growth and development.

Evolution and structure of the *LBD* gene family in *F. vesca*

A total of 35 *FvLBD* genes were identified in the diploid woodland strawberry genome (accession Hawaii-4 (Shulaev et al. 2011)), whose number is fewer than that in *Arabidopsis* (43). These *FvLBD* genes were localized to 7 chromosomes, although the location of *FvLBD20* remains unknown. Unlike in *Arabidopsis*, many *FvLBD* genes (11) have more than one copy in the genome. This redundancy might be due to gene duplication in the strawberry genome. Based on the sequence and phylogenetic analyses, the *FvLBD* proteins were classified into two subclasses (Class I and Class II). Both subclasses contain the C Block, while the GAS block and L-rich block only appear in Class I. In our research, *FvLBD2* was separated into the Class II subfamily due to the absence of the L-rich motif. The gene structure and protein domains might reflect the phylogenetic relationships and duplication events within the gene family. In strawberry, the exon numbers of the *LBD* genes ranged from 1 to 10, and most of these genes contained 2 exons. Additionally, the exon numbers vary much more compared with *Arabidopsis*. For example, three *FvLBD* genes (*FvLBD25A*, *FvLBD29B*, and *FvLBD20*) have four exons, *FvLBD12A* has 7 exons and *FvLBD2* even has 10 exons. In contrast, most of the *AtLBD* genes (27) contain 2 exons, and only 2 genes contain 3 exons (Fig. 3, Supplemental Fig. S3).

FvLBD16* might have a prominent role in callus formation in *F. vesca

Many plants have the remarkable ability to regenerate and form an entire plant body from various tissues or organs. Callus induction is often the initial step in a typical in vitro plant regeneration system (Fan et al. 2012; Liu et al. 2014). The callus

formation process is mainly regulated by auxin and its downstream regulators. It was previously reported that four *LBD* genes downstream of auxin participate in regulating callus formation in *Arabidopsis* (Fan et al. 2012). However, unlike *Arabidopsis*, only *FvLBD16* was upregulated during the callus formation in strawberry. This result suggests that *FvLBD16* might play a prominent role in regulating callus formation in *F. vesca*.

***FvLBD* genes are likely to play a role during early fruit development**

Many studies have revealed that *LBD* genes play important roles in the regulation of plant development. The cis-element analysis of the *FvLBD* genes identified the endosperm, meristem, and cell cycle-control related elements. These data revealed the potential role of *FvLBD* genes in regulating strawberry development. Based on the transcriptome data, we found that some *FvLBD* genes are induced in specific tissues during the early stages of fruit development. For example, the *FvLBD18* gene was induced during the development of the cell wall and ovule, and the *FvLBD25A* gene was upregulated during the early stages of the cortex and pith development. These results suggest that the *FvLBD* genes might take part in early-stage fruit development. It has been reported that the *AtLBD18* gene could regulate the cell cycle-controlling gene *E2Fa* and the cell wall-loosening factor gene *EXPANSIN14* (Berckmans et al. 2011; Lee et al. 2013). These findings suggest that the *LBD* genes might function as a cell cycle regulator during the early stages of fruit development in *F. vesca*.

In conclusion, a total number of 35 *FvLBD* genes were identified in the diploid

woodland strawberry genome. The evolutionary relationships and comparisons with their counterparts in *Arabidopsis* aided in identifying the potential functions of the *FvLBD* genes. A number of cis-acting elements were found in the promoter regions of *FvLBD* genes, increasing our knowledge of the regulatory mechanisms of *LBD* genes in strawberry. By analysing the expression profiles of the *FvLBD* genes during callus formation, we found that *FvLBD16* might play an important role in regulating this process. In addition, since fruit development is very special and important for strawberry, we further investigated the expression profiles of the *FvLBD* genes during the early stages of fruit development by analysing the transcriptome data. We found that some *FvLBD* genes were specifically induced during early fruit development, suggesting that the *FvLBD* genes might play a role in strawberry fruit development. Our research has laid the foundation for future studies to reveal the functions of the *LBD* genes during the growth and development of strawberry.

Acknowledgements

This work was supported by Jiangsu Planned Projects for Postdoctoral Research Funds (NO. 1601103B); Jiangsu Provincial Natural Science Foundation of China (NO. BK20130722) and National Natural Science Foundation of China (NO. 31301764).

References

- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**: W202-W208. doi: 10.1093/nar/gkp335. PMID: 19458158.
- Berckmans, B., Vassileva, V., Schmid, S. P., Maes, S., Parizot, B., Naramoto, S., Magyar, Z., Alvim Kamei, C. L., Koncz, C., Bogre, L. and others. 2011. Auxin-dependent cell cycle reactivation through transcriptional regulation of *Arabidopsis E2Fa* by lateral organ boundary proteins. *Plant Cell.* **23**: 3671-3683. doi: 10.1105/tpc.111.088377. PMID: 22003076.
- Cao, H., Liu, C., Liu, C., Zhao, Y., and Xu, R. 2016. Genomewide analysis of the lateral organ boundaries domain gene family in *Vitis vinifera*. *J Genet.* **95**: 515-526. PMID: 27659322.
- Deng, W. K., Wang, Y. B., Liu, Z. X., Cheng, H., and Xue, Y. 2014. HemI: A Toolkit for Illustrating Heatmaps. *PLoS One.* **9**: e111988. doi: 10.1371/journal.pone.0111988. PMID: 25372567.
- Eddy, S. R. 1998. Profile hidden Markov models. *Bioinformatics.* **14**: 755-763. doi: 10.1093/bioinformatics/14.9.755.
- Fan, M., Xu, C., Xu, K., and Hu, Y. 2012. LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in *Arabidopsis* regeneration. *Cell Res.* **22**: 1169-1180. doi: 10.1038/cr.2012.63. PMID: 22508267.
- Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L.,

- Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A. and others. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**: D279-D285. doi: 10.1093/nar/gkv1344. PMID: 26673716.
- Gombos, M., Zombori, Z., Szecsenyi, M., Sandor, G., Kovacs, H., and Gyorgyey, J. 2017. Characterization of the LBD gene family in *Brachypodium*: a phylogenetic and transcriptional study. *Plant Cell Rep.* **36**: 61-79. doi: 10.1007/s00299-016-2057-0. PMID: 27686461.
- Gu, T., Ren, S., Wang, Y., Han, Y., and Li, Y. 2016. Characterization of DNA methyltransferase and demethylase genes in *Fragaria vesca*. *Mol Genet Genomics.* **291**: 1333-1345. doi: 10.1007/s00438-016-1187-y. PMID: 26956009.
- Holub, E. B. 2001. The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat Rev Genet.* **2**: 516-527. doi: 10.1038/35080508. PMID: 11433358.
- Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics.* **31**: 1296-1297. doi: 10.1093/bioinformatics/btu817. PMID: 25504850.
- Jung, S., Ficklin, S. P., Lee, T., Cheng, C. H., Blenda, A., Zheng, P., Yu, J., Bombarely, A., Cho, I., Ru, S. and others. 2014. The Genome Database for Rosaceae (GDR): year 10 update. *Nucleic Acids Res.* **42**: D1237-D1244. doi: 10.1093/nar/gkt1012. PMID: 24225320.
- Kang, C., Darwish, O., Geretz, A., Shahan, R., Alkharouf, N., and Liu, Z. 2013. Genome-scale transcriptomic insights into early-stage fruit development in woodland strawberry *Fragaria vesca*. *Plant Cell.* **25**: 1960-1978. doi:

- 10.1105/tpc.113.111732. PMID: 23898027.
- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M. and others. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**: D1202-10. doi: 10.1093/nar/gkr1090. PMID: 22140109.
- Lee, H. W., Kim, M. J., Kim, N. Y., Lee, S. H., and Kim, J. 2013. *LBD18* acts as a transcriptional activator that directly binds to the *EXPANSIN14* promoter in promoting lateral root emergence of *Arabidopsis*. *Plant J.* **73**: 212-224. doi: 10.1111/tpj.12013. PMID: 22974309.
- Lee, H. W., Kim, N. Y., Lee, D. J., and Kim, J. 2009. *LBD18/ASL20* regulates lateral root formation in combination with *LBD16/ASL18* downstream of *ARF7* and *ARF19* in *Arabidopsis*. *Plant Physiol.* **151**: 1377-1389. doi: 10.1104/pp.109.143685. PMID: 19717544.
- Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P., and Rombauts, S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**: 325-327. PMID: 11752327.
- Liu, J., Sheng, L., Xu, Y., Li, J., Yang, Z., Huang, H., and Xu, L. 2014. *WOX11* and *12* are involved in the first-step cell fate transition during de novo root organogenesis in *Arabidopsis*. *Plant Cell.* **26**: 1081-1093. doi: 10.1105/tpc.114.122887. PMID: 24642937.

- Long, M. Y., and Deutsch, M. 1999. Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. *Molecular Biology and Evolution*. **16**: 1528-1534. PMID: 10555284.
- Luo, Y., Ma, B., Zeng, Q., Xiang, Z., and He, N. 2016. Identification and characterization of *Lateral Organ Boundaries Domain* genes in mulberry, *Morus notabilis*. *Meta Gene*. **8**: 44-50. doi: 10.1016/j.mgene.2014.04.004. PMID: 27014591.
- Majer, C., and Hochholdinger, F. 2011. Defining the boundaries: structure and function of LOB domain proteins. *Trends in Plant Science*. **16**: 47-52. doi: 10.1016/j.tplants.2010.09.009. PMID: 20961800.
- Mangeon, A., Bell, E. M., Lin, W. C., Jablonska, B., and Springer, P. S. 2011. Misregulation of the LOB domain gene DDA1 suggests possible functions in auxin signalling and photomorphogenesis. *Journal of Experimental Botany*. **62**: 221-233. doi: 10.1093/jxb/erq259. PMID: 20797997.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R. and others. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*. **45**: D200-D203. doi: 10.1093/nar/gkw1129. PMID: 27899674.
- Murashige, T., and Skoog, F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol Plant*. **80**: 662-668. doi: 10.1111/j.1399-3054.1962.tb08052.x.
- Rubin, G., Tohge, T., Matsuda, F., Saito, K., and Scheible, W. R. 2009. Members of

- the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in *Arabidopsis*. *Plant Cell*. **21**: 3567-3584. doi: 10.1105/tpc.109.067041. PMID: 19933203.
- Sharp, P. A. 1981. Speculations on RNA Splicing. *Cell*. **23**: 643-646. PMID: 7226224.
- Shuai, B., Reynaga-Pena, C. G., and Springer, P. S. 2002. The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant Physiol*. **129**: 747-761. doi: 10.1104/pp.010926. PMID: 12068116.
- Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkerts, O., Delcher, A. L., Jaiswal, P., Mockaitis, K., Liston, A., Mane, S. P., and others. 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nat Genet*. **43**: 109-116. doi: 10.1038/ng.740. PMID: 21186353.
- Soyano, T., Thitamadee, S., Machida, Y., and Chua, N. H. 2008. *ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30* and *ASL20/LBD18* Regulate Tracheary Element Differentiation in *Arabidopsis*. *Plant Cell*. **20**: 3359-3373. doi: 10.1105/tpc.108.061796. PMID: 19088331.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*. **28**: 2731-2739. doi: 10.1093/molbev/msr121. PMID: 21546353.
- Thatcher, L. F., Powell, J. J., Aitken, E. A., Kazan, K., and Manners, J. M. 2012. The lateral organ boundaries domain transcription factor LBD20 functions in Fusarium wilt Susceptibility and jasmonate signaling in *Arabidopsis*. *Plant Physiol*. **160**: 407-

418. doi: 10.1104/pp.112.199067. PMID: 22786889.
- Wang, X., Zhang, S., Su, L., Liu, X., and Hao, Y. 2013. A genome-wide analysis of the *LBD* (*LATERAL ORGAN BOUNDARIES domain*) gene family in *Malus domestica* with a functional characterization of *MdLBD11*. PLoS One. **8**: e57044. doi: 10.1371/journal.pone.0057044. PMID: 23468909.
- Wei, W., Hu, Y., Cui, M. Y., Han, Y. T., Gao, K., and Feng, J. Y. 2016. Identification and Transcript Analysis of the TCP Transcription Factors in the Diploid Woodland Strawberry *Fragaria vesca*. Front Plant Sci. **7**: 1937. doi: 10.3389/fpls.2016.01937. PMID: 28066489.
- Xu, B., Li, Z., Zhu, Y., Wang, H., Ma, H., Dong, A., and Huang, H. 2008. *Arabidopsis* genes *AS1*, *AS2*, and *JAG* negatively regulate boundary-specifying genes to promote sepal and petal development. Plant Physiol. **146**: 566-575. doi:10.1104/pp.107.113787. PMID: 18156293.
- Xu, C., Luo, F., and Hochholdinger, F. 2016. LOB Domain Proteins: Beyond Lateral Organ Boundaries. Trends Plant Sci. **21**: 159-167. doi: 10.1016/j.tplants.2015.10.010. PMID: 26616195.
- Xu, L. 2003. Novel *as1* and *as2* defects in leaf adaxial-abaxial polarity reveal the requirement for *ASYMMETRIC LEAVES1* and *2* and *ERECTA* functions in specifying leaf adaxial identity. Development. **130**: 4097-4107. PMID: 12874130.
- Yang, Y., Yu, X., and Wu, P. 2006. Comparison and evolution analysis of two rice subspecies *LATERAL ORGAN BOUNDARIES* domain gene family and their evolutionary characterization from *Arabidopsis*. Mol Phylogenet Evol. **39**: 248-262.

doi: 10.1016/j.ympev.2005.09.016. PMID: 16290186.

Yordanov, Y. S., Regan, S., and Busov, V. 2010. Members of the LATERAL ORGAN BOUNDARIES DOMAIN Transcription Factor Family Are Involved in the Regulation of Secondary Growth in *Populus*. *Plant Cell*. **22**: 3662-3677. doi: 10.1105/tpc.110.078634. PMID: 21097711.

Zhang, Y., Zhang, S., and Zheng, C. 2014. Genomewide analysis of LATERAL ORGAN BOUNDARIES Domain gene family in *Zea mays*. *J Genet*. **93**: 79-91. PMID: 24840825.

Figure legends

Fig. 1. Chromosomal distribution of *FvLBD* genes (fvesca_v1.0_genemark_hybrid). Pseudochromosome numbers are provided at the top of each pseudochromosome. The names on the left side of each pseudochromosome correspond to the approximate location of each *LBD* gene. Class I members are indicated in black font; Class II members are indicated in red. The unmapped *LBD* gene (*FvLBD20*) is not shown. The scale is in megabases (Mb).

Fig. 2. Phylogenetic analysis of LBD proteins. The full-length amino acid sequences of the LBD proteins from *F. vesca* (red) and *Arabidopsis* (green) were used to construct the phylogenetic tree using the MEGA 5.0 software programme. Bootstrap values from 1000 replicates are indicated at each node.

Fig. 3. Structural analysis of strawberry LBD transcription factors. The unrooted phylogenetic tree was constructed using full-length protein sequences of 35

strawberry LBD proteins by the neighbour-joining method using 1000 bootstrap replicates. The two classes are indicated as different greyscale boxes. The protein domains are shown on the left and denoted by rectangles with different colours. The exon-intron organization is shown on the right, with exons and introns represented by green wedges and black lines, respectively. Untranslated regions (UTRs) are indicated by the blue boxes. Intron phases 0 (introns are located between codons), 1 (introns are located between the first and second nucleotides of a codon) and 2 (introns are located between the second and third nucleotides) (Long and Deutsch, 1999) are indicated by numbers 0, 1 and 2, respectively.

Fig. 4. Expression patterns of *FvLBD* genes during callus formation. Transcript levels were analysed from time 0 to 16 DAC. The expression levels relative to GADPH were measured by qRT-PCR. Thirty-four *FvLBD* genes' expression profiles were shown, while *FvLBD6* was not detectable during the callus formation. Bars show SD from three PCR experiments. Results were confirmed in three biological repeats. DAC (Days After Culture). Time 0 (i.e. explants prior to culturing).

Fig. 5. Expression profiles of *FvLBD* genes during early fruit development. The expression levels (RPKM) were directly downloaded from <http://bioinformatics.tows-on.edu/strawberry/> and plotted as a log₂ scale. Red and blue boxes indicate high and low expression levels, respectively. Wall: is the carpel wall that encloses individual seed inside. Ghost: refers to the entire seed with its embryo removed. Embryo: the tiny embryo dissected out of seed. Cortex: the upper layer of the receptacle or the fleshy tissue immediately underneath achenes. Pith: the interior tissue of the

receptacle.

For Review Only

Table 1. Characteristics of strawberry *LBD* genes.

Name	Gene Id ^a	Gene location ^b	Length (bp)	Number of aa	Mw(kDa) ^c	pI ^c
<i>FvLBD1A</i>	gene25294	LG3:19319878-19321366	1489	236	25796.09	6.34
<i>FvLBD1B</i>	gene23226	LG7:15560904-15562076	1173	224	24187.56	5.35
<i>FvLBD2</i>	gene17794	LG6:15395267-15403403	8137	744	83765.23	7.85
<i>FvLBD4A</i>	gene20083	LG3:5259862-5261170	1309	176	19063.6	8.2
<i>FvLBD4B</i>	gene30046	LG3:9695830-9698193	2364	177	19520.1	7.67
<i>FvLBD6</i>	gene21541	LG6:2081548-2082294	747	248	27153.35	8.23
<i>FvLBD9</i>	gene12646	LG1:6210726-6211877	947	213	24512.68	5.9
<i>FvLBD12A</i>	gene28202	LG3:21111780-21118566	6787	387	43322.46	6.43
<i>FvLBD12B</i>	gene21039	LG7:18560030-18560687	658	173	19016.64	5.65
<i>FvLBD15A</i>	gene21311	LG7:18832799-18833716	918	237	25944.26	8.6
<i>FvLBD15B</i>	gene28137	LG3:20743221-20749193	1169	381	41483.54	9.56
<i>FvLBD16</i>	gene08250	LG2:5876546-5877666	1121	215	23571.61	7.57
<i>FvLBD18</i>	gene19355	LG7:6731369-6734081	2713	244	25145.55	8.15
<i>FvLBD19</i>	gene19356	LG7:6743759-6746127	2369	218	23875.36	7.67
<i>FvLBD20</i>	gene24461	Unknown	3926	362	40205.32	5.91
<i>FvLBD21</i>	gene21008	LG7:18422127-18422666	540	179	19316.81	7.66
<i>FvLBD22A</i>	gene23043	LG4:20902590-20904097	1508	327	36637.93	5.65
<i>FvLBD22B</i>	gene02697	LG2:20532826-20533423	598	168	18559.44	8.95
<i>FvLBD22C</i>	gene02696	LG2:20530558-20531142	585	194	21908.64	9.62
<i>FvLBD24A</i>	gene26806	LG7:1687372-1688189	818	146	16328.3	5.14
<i>FvLBD24B</i>	gene13324	LG7:22829838-22830417	580	145	16298.54	6.19
<i>FvLBD25A</i>	gene28478	LG6:37376976-37380152	3177	238	25838.87	8.06
<i>FvLBD25B</i>	gene31797	LG5:3437944-3438525	582	193	21297.98	7.65
<i>FvLBD27A</i>	gene09561	LG5:11054580-11056034	1455	409	46780.67	5.66
<i>FvLBD27B</i>	gene01286	LG6:4414727-4415747	1021	286	32043.69	5.75
<i>FvLBD29A</i>	gene08253	LG2:5889625-5890424	800	227	24982.48	5.64
<i>FvLBD29B</i>	gene08254	LG2:5896933-5898130	1198	212	23318.46	8.38
<i>FvLBD33</i>	gene26330	LG7:16245172-16245964	793	236	26519.74	5.19
<i>FvLBD36A</i>	gene24088	LG1:11255419-11256339	921	306	33753.9	7.14
<i>FvLBD36B</i>	gene30401	LG3:2877654-2878457	804	267	28801.36	6.29
<i>FvLBD38A</i>	gene05722	LG1:14168889-14169737	849	250	27016.68	8.56
<i>FvLBD38B</i>	gene25541	LG2:3876673-3877412	740	213	23577.61	8.8
<i>FvLBD39</i>	gene02557	LG2:19820897-19821653	757	214	22723.86	8.1
<i>FvLBD41</i>	gene10300	LG6:30073221-30074211	991	301	32338.44	8.81
<i>FvLBD42</i>	gene03741	LG5:4408528-4409322	795	231	24915.3	8.37

^aIDs are available in the strawberry Genome Database (<http://www.rosaceae.org>).

^b Genome assembly version (fvesca_v1.0_genemark_hybrid).

^cDetailed FvLBD protein characteristics were predicted by ExPASy online

(http://web.expasy.org/compute_pi/).

For Review Only

Table 2. The cis-elements in the promoter regions of *FvLBD* genes.

	1 ^a	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
<i>FvLBD1A</i>	+ ^b	+		+	+	+	+	+		+			+			+			+	+		
<i>FvLBD1B</i>	+	+	+	+	+		+		+	+			+		+				+	+		
<i>FvLBD2</i>	+	+	+	+	+		+			+	+	+	+		+				+			
<i>FvLBD4A</i>	+			+	+		+			+	+								+			
<i>FvLBD6</i>	+			+	+		+				+		+			+			+			
<i>FvLBD9</i>	+				+		+			+			+							+		
<i>FvLBD12A</i>	+			+			+			+	+	+				+			+		+	
<i>FvLBD12B</i>	+	+		+	+		+					+			+	+	+		+			
<i>FvLBD15A</i>	+			+	+		+						+						+			
<i>FvLBD15B</i>	+			+			+			+									+			
<i>FvLBD18</i>	+	+		+	+	+	+				+		+		+	+			+			
<i>FvLBD19</i>	+	+	+	+	+	+	+			+	+		+		+	+			+	+		
<i>FvLBD20</i>	+			+		+				+	+		+		+	+	+	+	+	+		
<i>FvLBD21</i>	+				+	+	+				+	+	+		+				+			
<i>FvLBD22A</i>	+		+	+		+	+			+	+		+									
<i>FvLBD22B</i>	+		+		+					+					+	+			+			
<i>FvLBD22C</i>	+			+	+		+					+	+	+					+	+		
<i>FvLBD24A</i>	+	+	+	+		+							+		+				+			+
<i>FvLBD24B</i>	+	+		+	+		+							+					+			+
<i>FvLBD25A</i>	+			+	+		+			+			+	+					+	+		
<i>FvLBD25B</i>	+		+						+				+						+			
<i>FvLBD27A</i>	+		+		+		+			+			+						+			
<i>FvLBD27B</i>	+	+					+			+	+	+							+			
<i>FvLBD29A</i>	+			+	+						+		+			+			+			
<i>FvLBD29B</i>	+			+	+	+	+		+				+		+				+			
<i>FvLBD33</i>	+	+			+	+	+			+			+						+	+		
<i>FvLBD36A</i>	+	+			+		+			+		+	+						+			
<i>FvLBD36B</i>	+			+		+	+	+	+	+	+	+	+									
<i>FvLBD38A</i>	+	+	+	+	+		+			+				+		+			+	+		
<i>FvLBD38B</i>	+	+		+	+				+				+	+		+			+	+		
<i>FvLBD39</i>	+	+			+		+			+	+	+	+		+		+	+	+	+		+
<i>FvLBD41</i>	+		+	+			+			+			+		+				+	+		
<i>FvLBD42</i>	+			+		+	+				+			+	+							

^aThe different cis-element(s) are represented by numbers:1 (Light); 2 (Heat stress); 3 (Low temperature); 4(Defense and stress); 5(Circadian); 6 (Fungal elicitor); 7 (MYB binding site); 8 (Dehydration,Low-temp,Salt stress(DRE)); 9 (wound); 10 (MeJA); 11 (SA); 12 (Auxin); 13 (GA); 14 (Ethylene); 15 (ABA); 16 (Zein metabolism); 17 (Palisade mesophyll cells differentiation); 18 (Leaf morphology development); 19 (Endosperm); 20 (Meristem); 21 (Root); 22 (Cell cycle).

^bindicates the presence of the cis-element(s).

For Review Only

74x31mm (300 x 300 DPI)

249x245mm (96 x 96 DPI)

223x286mm (300 x 300 DPI)

205x242mm (300 x 300 DPI)