



RNASeq Analysis of the Shoot Apex of Flax (*Linum usitatissimum*) to Identify Phloem Fiber Specification Genes

Ningyu Zhang¹ and Michael K. Deyholos^{2*}

¹ Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada, ² Department of Biology, University of British Columbia, Kelowna, BC, Canada

Keywords: meristem, RNASeq, vascular differentiation, fiber, xylem, phloem

INTRODUCTION

All of the post-embryonic, above-ground structures of seed plants are generated from the shoot apical meristem (SAM), which acts as a reservoir of stem cells. Members of the flax genus (*Linum* spp.) have been used historically as models for the study of SAMs (Esau, 1942). Cultivated flax (*Linum usitatissimum*) is grown in more than 50 countries for its seeds or its stem phloem (bast) fibers (Rubilar et al., 2010). Due to prolonged intrusive growth, and a highly crystalline cellulosic secondary wall, flax phloem fibers are among the longest and strongest cells in plants (Mohanty et al., 2000). In flax, all phloem fibers are derived from primary growth in the shoot apex. Specification of phloem fibers occurs in the apical-most 0.5 mm of the stem, since young phloem fibers can be anatomically distinguished starting 0.4–0.5 mm from the shoot apex (Gorshkova et al., 2003). The molecular mechanisms that govern fiber identity are almost entirely unknown (Gorshkova et al., 2012). Also, in contrast to the significant progress obtained in the past decade toward understanding xylem differentiation, information about the phloem fiber differentiation is very scarce (De Rybel et al., 2016). In the past decade, shoot apex transcriptomes have been described in various plants, including maize, pea, soybean, rice, Arabidopsis and chickpea, but none of these produce significant primary phloem fibers (Ohtsu et al., 2007; Wong et al., 2008; Haerizadeh et al., 2009; Jiao et al., 2009; Yadav et al., 2009; Wang et al., 2014). Most molecular and cellular research on flax fiber has thus far focused on later stages of development (Day et al., 2005; Roach and Deyholos, 2007; Fenart et al., 2010). Differential transcript expression data from the region of the shoot apex in which fiber specification occurs would complement other approaches (e.g., mutant screening) aimed at understanding primary phloem fiber differentiation.

VALUE OF THE DATA

- The genetic basis of primary phloem fiber identity in any species is unknown. This limits basic research and crop improvement.
- Data were obtained from tissues at high spatial resolution, which allows the results to be correlated with specific developmental processes.
- The identification of transcripts enriched in the shoot apical region will help define mechanisms of phloem fiber specification, and contribute to improved understanding of the SAM in general.

OPEN ACCESS

Edited by:

Stewart Gillmor,
Centro de Investigación y de Estudios
Avanzados del Instituto Politécnico
Nacional, Mexico

Reviewed by:

Matthew R. Willmann,
Cornell University, USA
Berenice Garcia-Ponce,
Universidad Nacional Autónoma de
México, Mexico

*Correspondence:

Michael K. Deyholos
michael.deyholos@ubc.ca

Specialty section:

This article was submitted to
Plant Genetics and Genomics,
a section of the journal
Frontiers in Plant Science

Received: 28 April 2016

Accepted: 15 June 2016

Published: 28 June 2016

Citation:

Zhang N and Deyholos MK (2016)
RNASeq Analysis of the Shoot Apex
of Flax (*Linum usitatissimum*) to
Identify Phloem Fiber Specification
Genes. *Front. Plant Sci.* 7:950.
doi: 10.3389/fpls.2016.00950

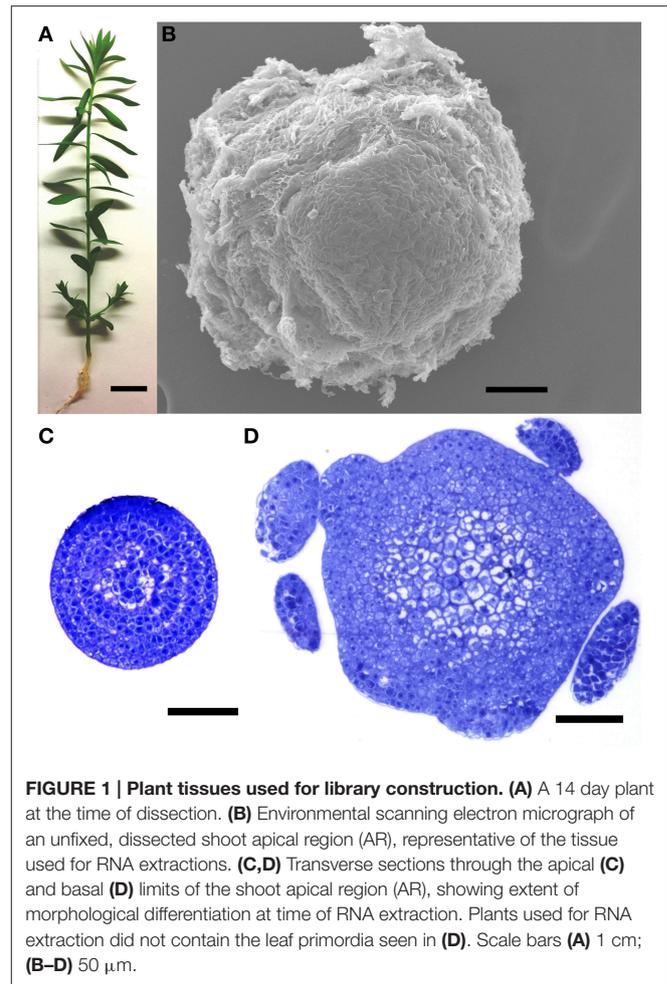
DATA

We used RNAseq to compare transcript expression patterns in two segments of the vegetative stem of 14d flax plants, from which all visible leaves had been removed. The segments were: (i) the apical region (AR) of the shoot apex, which contained the apical-most 0.5 mm of the stem, including the SAM and its immediate derivatives; and (ii) the basal region (BR), which contained the entire stem except for the apical-most 1 cm, and therefore represented all stem and vascular tissues at later stages of differentiation as compared to the AR. Four biologically independent replicates of AR (AR1, AR2, AR3, AR4), and two biologically independent replicates of BR (BR1, BR2) were sequenced on the Illumina HiSeq platform in a total of nine runs. Data were deposited in the Sequence Read Archive (SRA) as the following accessions: AR1: SRR1056618; AR2: SRR1056620, SRR1056621; AR3: SRR1056622, SRR1056623; AR4: SRR1056624, SRR1056625; BR1: SRR1038482; BR2: SRR1421513 (<http://www.ncbi.nlm.nih.gov/sra?term=SRP033325>). In total, 117.5 million clean reads (21 Gbp) were obtained and mapped to the reference flax genome to detect transcripts enriched in the AR as compared to the BR. These differential transcript expression data (measured as normalized FPKM) are available at NCBI GEO (GSE80718; <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80718>), and an annotated version of the file is available as Supplemental Table 1.

EXPERIMENTAL DESIGN, MATERIALS, AND METHODS

Plant Materials

Flax (i.e., linseed) plants (*L. usitatissimum* L. cv. CDC Bethune; Rowland et al., 2002) were grown in potting mix in an environmental chamber at 22°C, with a cycle of 16 h light and 8 h dark, as previously described (Wang et al., 2012). Fourteen days after germination (Figure 1A), approximately 0.5 mm of the apical-most part of each stem (the apical region, AR) was dissected under a Leica S6D stereo microscope, all visible leaf primordia were removed, and the tissue was frozen in liquid nitrogen. A representative dissection, visualized under an environmental scanning electron microscope, is shown in Figure 1B, and transverse sections of a shoot apex, corresponding to the apical and basal-most tissues sampled, are shown in Figures 1C,D. Shoot apices were similarly dissected from approximately 200 plants and pooled prior to each RNA extraction. After collecting the shoot apex, the remainder of the stem (i.e., the basal region, BR) from 1 cm below the shoot apex to the stem base was also dissected, stripped of leaves, visible lateral branches and axillary meristems, and frozen in liquid nitrogen. In this way, maturing stems from at least six plants were pooled for each RNA extraction. For RNASeq of the AR, samples were harvested from four biological replicates (i.e., four sets of plants that were grown spatially and temporally independently from each other), and tissues were obtained from two biologically independent replicates were used for the BR. For qRT-PCR, three additional, independent biological replicates (i.e., different plants



than those used for RNASeq) were obtained from each of the AR and BR.

RNA Extraction and Sequencing

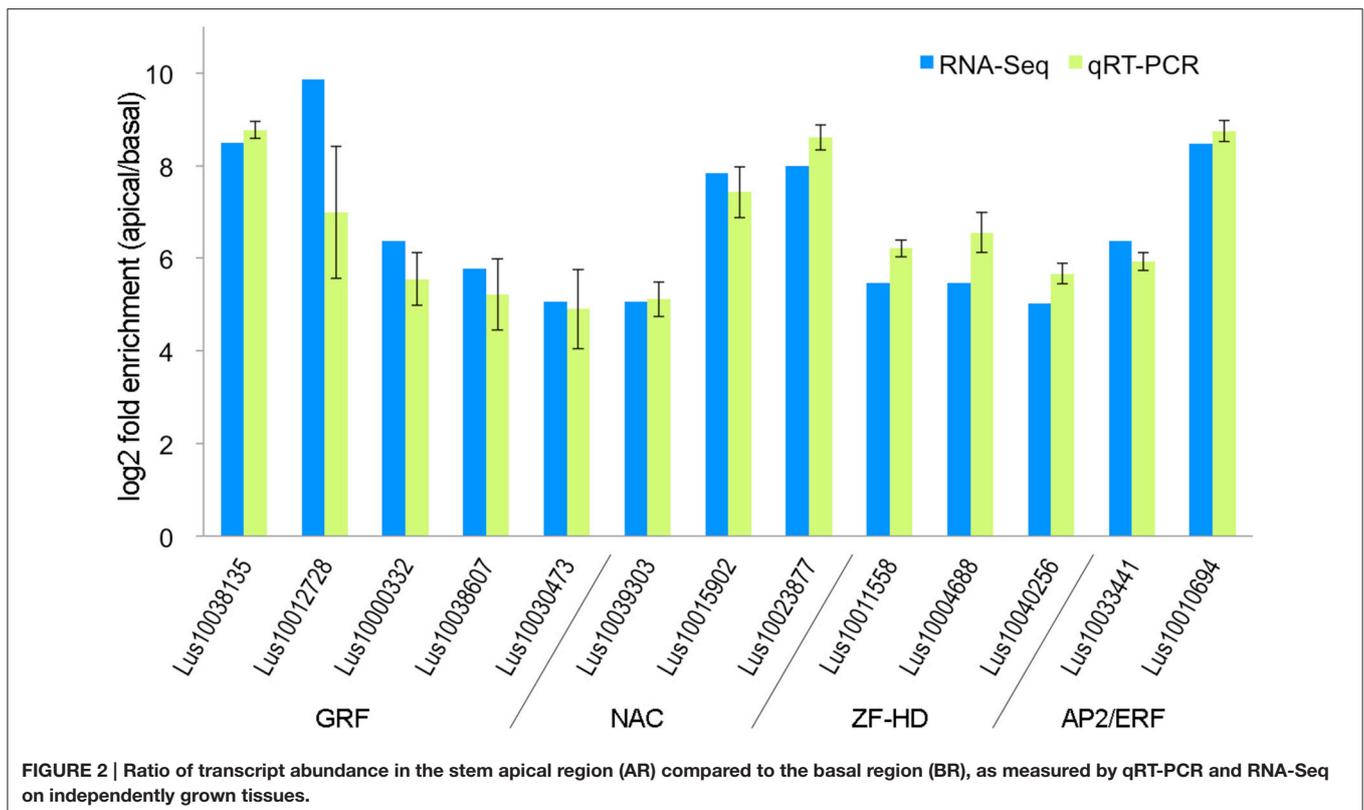
RNA from each biological replicate (Section Plant Materials) was extracted separately. RNeasy Micro Kit (Qiagen) and RNeasy Plant Mini Kit (Qiagen) were used to isolate RNA from the AR and BR samples, respectively. Extracted RNA was then digested with TURBO DNA-free™ Kit (Life Technologies) to remove DNA contamination and their quality was evaluated using a RNA 6000 Nano chip (Agilent Technologies) on an Agilent 2100 Bioanalyzer. Total RNA was delivered to the service provider, BGI (Shenzhen, China), where each biological replicate was sequenced separately. oligodT coupled magnetic beads were used to isolate poly-A+ mRNA, which was used as a template for cDNA synthesis (Superscript II, Invitrogen) primed by random hexamers, followed by second strand synthesis using *E. coli* DNA PolII (Invitrogen). Double-stranded cDNA (Qiaquick PCR Purification Kit, Qiagen), was sheared with a nebulizer, end repaired, and ligated to Illumina PE adapter oligos, and the products size-selected by gel purification to produce 200 bp fragments. These were PCR amplified through 15 cycles to prior

to sequencing using an Illumina HiSeq 2000 with 90 bp, paired-end reads. The quality of the sample during processing prior to sequencing was monitored using the Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System. Because the sequencing output for samples AR2, AR3, and AR4 was slightly lower than expected (9.6 million reads output per sample), additional aliquots of each of these three samples were sequenced in three additional runs. Raw reads from all runs were filtered to remove adapter sequences, contamination, and low-quality reads, and the filtered raw reads were deposited in the SRA archive. Each of the nine paired read files were uploaded to SRA in fastq format.

RNASeq Analysis of Differential Transcript Abundance

To quantify the relative abundance of transcripts in the shoot apex (AR) as compared to the remainder of the stem (BR), the clean sequencing reads described in Section RNA Extraction and Sequencing were mapped to the flax reference genome (Wang et al., 2012; downloaded from Phytozome 9 as *Lusitatissimum_200.fa*) using Tophat2 (Trapnell et al., 2012), and the accepted hits were used as input for cufflinks, with default parameters. The resulting assemblies were merged and with the reference genome annotation (downloaded from Phytozome 9 as *Lusitatissimum_200_gene.gff3*) with cuffmerge, and finally Cuffdiff was used to calculate normalized differential transcript abundance between the samples. The output of cuffdiff

(*gene_exp.diff*) is available at NCBI GEO as accession GSE80718, and an annotated version of this file is available as Supplemental Table 1. The merged.gtf file is available as Supplemental Table 2, and defines positions of a locus identified by cufflinks, in reference to the scaffolds in the *Lusitatissimum_200.fa* genome assembly. Within these results, transcripts for 6207 genes were significantly ($q < 0.05$) more abundant in AR compared to BR, and 4405 of these were enriched at least 2-fold in the AR. Conversely, transcripts for 8388 genes were significantly ($q < 0.05$) more abundant in BR compared to AR and 7901 of these were enriched at least 2-fold in the BR. Inspection of the data showed that several markers of shoot apex tissues were highly enriched in the AR sample. For example, *PROTODERMAL FACTOR 1 (PDF1)* transcripts have been reported to be expressed exclusively in the L1 layer of meristems and the protoderm of organ primordia (Abe et al., 1999). In our results, transcripts of putative *PDF1* genes (Lus10007351, Lus10031390, Lus10010941) were at least 19.5-fold more abundant in AR than BR (Supplemental Table 1). Similarly, *CUP-SHAPED COTYLEDON (CUC)* genes are required for SAM function and organ separation (Hasson et al., 2011). Transcripts of three putative *CUC* genes (Lus10041924, Lus10005537, Lus10013205) were at least 45-fold more abundant in AR than BR; two other putative *CUC* genes (Lus10037106, Lus10003458) were not detected in either sample. As a third example, the SHOOT MERISTEMLESS (*STM*) transcription factor is essential for SAM formation and maintenance (Endrizzi et al., 1996); a putative *STM* gene (Lus10030003) was 4.8-fold enriched in the AR



sample compared to BR. Conversely, several markers of late differentiation were more enriched in the BR compared to the AR. For example, *CELLULOSE SYNTHASE A (CESA)* genes *CESA4*, *CESA7*, and *CESA8* are associated with secondary wall synthesis (Chantreau et al., 2015); we observed transcripts of flax genes annotated as *CESA4* (Lus10008225, Lus10008226), and *CESA8* (Lus10007296, Lus10029245) to be at least 125-fold enriched in the BR compared to the AR (no *CESA7* genes were identified in the original flax genome annotation used in this study). Another well-established marker of xylem differentiation, *XYLEM CYSTEINE PROTEINASE-2 (XCP2)*; (Avci et al., 2008). The two putative flax *XCP2* genes (Lus10030722, Lus10013204) were enriched 106-fold in the BR compared to the AR. Thus, expression of at least some well-known markers of early and late stem development were observed in patterns that matched expectations.

Quantitative Real-Time PCR Analysis of Differential Transcript Abundance

To evaluate the accuracy of the differential transcript expression measurements that we obtained (Section RNASeq Analysis of Differential Transcript Abundance) we used qRT-PCR to measure transcript abundance in independently grown replicates of the same tissues that were used for RNA-Seq. In order to select an appropriate reference gene for the qRT-PCR, GeNorm was used to determine the expression stability of nine commonly used reference genes among tissues assayed in our study (Huis et al., 2010). *GADPH* and *ETIF5A* were found to be the most stable, and *ETIF5A* gene chosen arbitrarily from this pair as the internal control. Thirteen genes were selected for qRT-PCR, as an independent validation of the accuracy of the RNA-Seq results (Figure 2). These genes were selected in part because they were all transcription factors from gene families that could be potentially associated with early differentiation events in the shoot apex including specification of vascular/phloem identity

REFERENCES

- Abe, M., Takahashi, T., and Komeda, Y. (1999). Cloning and characterization of an L1 layer-specific gene in *Arabidopsis thaliana*. *Plant Cell Physiol.* 40, 571–580. doi: 10.1093/oxfordjournals.pcp.a029579
- Avci, U., Petzold, H. E., Ismail, I. O., Beers, E. P., and Haigler, C. H. (2008). Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in *Arabidopsis* roots. *Plant J.* 56, 303–315. doi: 10.1111/j.1365-313X.2008.03592.x
- Chantreau, M., Chabbert, B., Billiard, S., Hawkins, S., and Neutelings, G. (2015). Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing. *Plant Biotechnol. J.* 13, 1312–1324. doi: 10.1111/pbi.12350
- Day, A., Addi, M., Kim, W., David, H., Bert, F., Mesnage, P., et al. (2005). ESTs from the fibre-bearing stem tissues of flax (*Linum usitatissimum* L.): expression analyses of sequences related to cell wall development. *Plant Biol.* 7, 23–32. doi: 10.1055/s-2004-830462
- De Rybel, B., Mahonen, A. P., Helariutta, Y., and Weijers, D. (2016). Plant vascular development: from early specification to differentiation. *Nat. Rev. Mol. Cell Biol.* 17, 30–40. doi: 10.1038/nrm.2015.6
- Endrizzi, K., Moussian, B., Haecker, A., Levin, J. Z., and Laux, T. (1996). The shoot meristemless gene is required for maintenance of undifferentiated cells

(Zhao et al., 2005; Kalve et al., 2014; De Rybel et al., 2016). Real-time PCR was performed in Applied Biosystems 7500 Fast Real-time PCR System following the manufacturer's protocol. Each amplification reaction was 10 μ l and it consisted of 0.4 μ M of each primer, 5 μ l SYBR Green Master Mix and 2.5 μ l 16-fold diluted cDNA. Threshold cycles (CT) were determined through 7500 Fast Software. The PCR program used was as follows: 95°C for 2 min, 40 cycles of 95°C for 10 s and 60°C for 30 s, then 72°C for 30 s and 72°C for 3 min; fluorescence data was collected at 60°C. Data were analyzed using the $2^{-\Delta\Delta C_T}$ method. Primer sequences used are listed in the Supplemental Table 3. As shown in Figure 2, the RNA-seq and qRT-PCR analysis showed highly consistent expression patterns for the 13 genes tested. We therefore conclude that that RNA-Seq data presented here accurately represents differences in transcript expression between the shoot apical region (AR) and the bulk of the stem (BR).

AUTHOR CONTRIBUTIONS

NZ conducted all experiments, and assisted in analysis and writing of the manuscript. MD designed experiments and assisted in analysis and writing of the manuscript.

FUNDING

Genome Canada ABC Program grant TUFGEN; Natural Sciences and Engineering Council (Canada) Discovery Grant 2014-03596 to MD; China Scholarship Council fellowship to NZ.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.00950>

- in *Arabidopsis* shoot and floral meristems and acts at a different regulatory level than the meristem genes *Wuschel* and *Zwille*. *Plant J.* 10, 967–979. doi: 10.1046/j.1365-313X.1996.10060967.x
- Esau, K. (1942). Vascular differentiation in the vegetative shoot of *Linum*. I. The procambium. *Am. J. Bot.* 29, 738–747. doi: 10.2307/2437727
- Fenart, S., Ndong, Y.-P. A., Duarte, J., Riviere, N., Wilmer, J., van Wuytswinkel, O., et al. (2010). Development and validation of a flax (*Linum usitatissimum* L.) gene expression oligo microarray. *BMC Genomics* 11:592. doi: 10.1186/1471-2164-11-592
- Gorshkova, T. A., Sal'nikov, V. V., Chemiksova, S. B., Ageeva, M. V., Pavlencheva, N. V., and van Dam, J. E. G. (2003). The snap point: a transition point in *Linum usitatissimum* bast fiber development. *Ind. Crops Prod.* 18, 213–221. doi: 10.1016/S0926-6690(03)00043-8
- Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., et al. (2012). Plant fiber formation: state of the art, recent and expected progress, and open questions. *Crit. Rev. Plant Sci.* 31, 201–228. doi: 10.1080/07352689.2011.616096
- Haerizadeh, F., Wong, C. E., Singh, M. B., and Balla, P. L. (2009). Genome-wide analysis of gene expression in soybean shoot apical meristem. *Plant Mol. Biol.* 69, 711–727. doi: 10.1007/s11103-008-9450-1
- Hasson, A., Plessis, A., Blein, T., Adroher, B., Grigg, S., Tsiantis, M., et al. (2011). Evolution and diverse roles of the CUP-SHAPED COTYLEDON

- genes in Arabidopsis leaf development. *Plant Cell* 23, 54–68. doi: 10.1105/tpc.110.081448
- Huis, R., Hawkins, S., and Neutelings, G. (2010). Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). *BMC Plant Biol.* 10:71. doi: 10.1186/1471-2229-10-71
- Jiao, Y., Tausta, S. L., Gandotra, N., Sun, N., Liu, T., Clay, N. K., et al. (2009). A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nat. Genetics* 41, 258–263. doi: 10.1038/ng.282
- Kalve, S., De Vos, D., and Beemster, G. T. S. (2014). Leaf development: a cellular perspective. *Front. Plant Sci.* 5:362. doi: 10.3389/fpls.2014.00362
- Mohanty, A. K., Misra, M., and Hinrichsen, G. (2000). Biofibres, biodegradable polymers and biocomposites: an overview. *Macromol. Mater. Eng.* 276–277, 1–24. doi: 10.1002/(SICI)1439-2054(20000301)276:1<1::AID-MAME1>3.0.CO;2-W
- Ohtsu, K., Smith, M. B., Emrich, S. J., Borsuk, L. A., Zhou, R., Chen, T., et al. (2007). Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.). *Plant J.* 52, 391–404. doi: 10.1111/j.1365-313X.2007.03244.x
- Roach, M. J., and Deyholos, M. K. (2007). Microarray analysis of flax (*Linum usitatissimum* L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. *Mol. Genet. Genomics* 278, 149–165. doi: 10.1007/s00438-007-0241-1
- Rowland, G. G., Hormis, Y. A., and Rashid, K. Y. (2002). CDC Bethune flax. *Can. J. Plant Sci.* 82, 101–102. doi: 10.4141/P01-066
- Rubilar, M., Gutierrez, C., Verdugo, M., Shene, C., and Sineiro, J. (2010). Flaxseed as a source of functional ingredients. *J. Soil Sci. Plant Nutr.* 10, 373–377. doi: 10.4067/s0718-95162010000100010
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016
- Wang, L., Cao, C., Ma, Q., Zeng, Q., Wang, H., Cheng, Z., et al. (2014). RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. *BMC Plant Biol.* 14:169. doi: 10.1186/1471-2229-14-169
- Wang, Z. W., Hobson, N., Galindo, L., Zhu, S. L., Shi, D. H., McDill, J., et al. (2012). The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. *Plant J.* 72, 461–473. doi: 10.1111/j.1365-313X.2012.05093.x
- Wong, C. E., Bhalla, P. L., Ottenhof, H., and Singh, M. B. (2008). Transcriptional profiling of the pea shoot apical meristem reveals processes underlying its function and maintenance. *BMC Plant Biol.* 8:73. doi: 10.1186/1471-2229-8-73
- Yadav, R. K., Girke, T., Pasala, S., Xie, M., and Reddy, V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4941–4946. doi: 10.1073/pnas.0900843106
- Zhao, C. S., Craig, J. C., Petzold, H. E., Dickerman, A. W., and Beers, E. P. (2005). The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. *Plant Physiol.* 138, 803–818. doi: 10.1104/pp.105.060202

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhang and Deyholos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.