

Transcriptome resources and genome-wide marker development for Japanese larch (*Larix kaempferi*)

Wanfeng LI¹, Suying HAN², Liwang QI¹, Shougong ZHANG (✉)¹

¹ State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China

² State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China

Abstract While the differential responses of trees to changes in climatic and environmental conditions have been demonstrated as they age, the underlying mechanisms and age control of tree growth and development are complex and poorly understood particularly at a molecular level. In this paper, we present a transcriptome analysis of *Larix kaempferi*, a deciduous conifer that is widely-grown in the northern hemisphere and of significant ecological and economic value. Using high-throughput RNA sequencing, we obtained about 26 million reads from the stems of 1-, 2-, 5-, 10-, 25- and 50-year-old *L. kaempferi* trees. Combining these with the published Roche 454 sequencing reads and the expressed sequence tags (both mainly from *Larix* embryogenic cell cultures), we assembled 26670549 reads into 146786 transcripts, of which we annotated 79182 to support investigations of the molecular basis of tree aging and adaption, somatic embryogenesis and wood formation. Using these sequences we also identified many single-nucleotide polymorphisms, simple sequence repeats, and insertion and deletion markers to assist breeding and genetic diversity studies of *Larix*.

Keywords *Larix*, transcriptome, age, wood formation, somatic embryogenesis, molecular marker

1 Introduction

Age is an important developmental cue for trees, and great changes occur as they age. For example, young trees reactivate earlier than old trees with increasing temperature in the spring, thus having a longer period of growth, which indicates that meristem activity is affected by age [1–4]. Notably, cambium cell division in locally-heated portions of the *Cryptomeria japonica* stem during the dormant

period restarts earlier in younger than in older trees [3], indicating that age affects the activity of the lateral meristem in response to temperature signals. The regular division of meristem cells in shoot and stem contributes to the continued growth in height and girth of a tree for many years, and this involves orderly cell-cycle progression and the expression of cell-cycle genes. The sensitivity of cambium to auxin, which is important in the control of cambium activity and wood formation [5–11], declines as trees age [12–14].

In addition to the differential responses to environmental and hormonal stimuli, the underlying mechanisms for transition from vegetative growth to flowering occurs as trees age are still poorly-defined. Given the age-control of meristem activity and identity, the expression of implicated genes appears to be regulated by age. However, limited information is available on the molecular basis of tree aging, especially the regulatory mechanisms of meristem activity.

Investigations of the molecular mechanisms of tree aging have identified many age-related genes [15–22]. Some of these genes participate in wood production [21,22] and the control of flowering time [23–26], while only a few have been associated with the regulation of meristem activity and identity. Due to the lack of genomic and transcriptome data, information about transcriptome reprogramming during tree aging is still limited. Thus the availability of transcriptome data for some trees, of which the genome has not been sequenced, would enable global analysis of transcriptome reprogramming during aging.

Larix kaempferi is a forest tree of important ecological and economic value widely-grown in the northern hemisphere, and its somatic embryogenesis has been used to study the regulatory mechanisms of plant development [27–35]. With the advantage of next-generation sequencing technologies, deep sequencing of the transcriptome has become rapid and economical, making it easier to develop molecular markers at the genome scale. Recently, transcriptome sequencing of *L. kaempferi* embryogenic

Received April 10, 2014; accepted May 5, 2014

Correspondence: shougong.zhang@caf.ac.cn

cell cultures and *L. gmelinii* needles using the Roche 454 and Illumina Solexa sequencing platforms has been reported [36,37], greatly adding to availability of transcriptome information for *Larix*. However, the *Larix* transcriptome is still one of the least explored among conifers [38] and remains insufficient for tree-aging studies. Here, we sequenced the transcriptome of the stems of 1-, 2-, 5-, 10-, 25- and 50-year-old *L. kaempferi*, and assembled it *de novo* with the transcriptome of the embryogenic cell cultures for global analysis of transcriptome resources and the development of genome-wide markers.

2 Materials and methods

2.1 Plant material

The materials were sampled in July 2011 at Dagujia seed orchard (42°22' N, 124°51' E), Liaoning Province, in North-east China. The uppermost main stems produced in the current year from three trees were harvested from 2-, 5-, 10-, 25- and 50-year-old *L. kaempferi* trees and ten stems from 1-year-old trees, which had been grown from seed. After removing the needles, the stems from trees of the same age were pooled, frozen in liquid nitrogen, and stored at -80°C for RNA extraction. We selected these ages because they constitute an entire rotation period from establishment to harvest, and include the vegetative and reproductive phases of *L. kaempferi*.

2.2 CDNA library preparation and transcriptome sequencing

CDNA (cDNA) library preparation and transcriptome sequencing were performed according to the manufacturer's instructions (Illumina, San Diego, CA). Briefly, total RNA was extracted from each sample using plant RNA purification reagent according to the manufacturer's instructions (Invitrogen, San Diego, CA). RNA was quantified on an ND-1000 Spectrophotometer (Thermo Fisher Scientific, Inc., Wilmington, DE). One microgram of total RNA from each sample was pooled and 5 µg of the mixed total RNA was used to generate an RNA library for Illumina paired-end sequencing. After isolation from 5 µg of the mixed total RNA, the mRNA was broken into short fragments. Using these short fragments as templates, a random hexamer primer was used to synthesize first-strand cDNA. Then second-strand cDNA was synthesized. The double-stranded cDNA was cleaned up and used for end repair and addition of poly A. Thereafter, these strands were connected to sequencing adapters and the desired cDNA was isolated through a 2% agarose gel. To enrich the desired cDNA, PCR amplification was performed. PCR products were isolated on 2% agarose gel and purified

using a PCR column purification kit. Finally, the library was sequenced using the Illumina HiSeq™ 2000 platform.

2.3 Transcriptome assembly

The raw Illumina RNA sequencing reads were first preprocessed by discarding reads with adaptors, with unknown nucleotides >5%, of low-quality (quality score < 20), or < 20 bp. Then, *de novo* assembly was performed using Trinity software [39] with default parameters. The Roche 454 raw reads (accession number: SRX110249) [36] were first preprocessed by trimming adaptors and Poly A, and then the preprocessed sequences were assembled using Newbler 2.6 (Roche; cDNA assembly mode) with default parameters. Another 3,874 expressed sequence tags (ESTs) [40,41] were used to increase the transcriptome coverage. Any vector contamination of the ESTs was removed using the seqclean program (<http://compbio.dfci.harvard.edu/tgi/software/>). The assembled Illumina transcripts, Roche 454 sequencing transcripts and ESTs were further assembled using CAP3 software [42] with default parameters.

2.4 Transcriptome annotation

To identify the *L. kaempferi* protein-coding genes, blastx (the basic local alignment search tool, search protein databases using a translated nucleotide query) [43] was used to search our assembled sequences against four sets of protein sequences from *Arabidopsis thaliana* (http://plants.ensembl.org/Arabidopsis_thaliana/Info/Index/), *Picea abies* (<http://congenie.org/>), *Populus trichocarpa* (<http://www.phytozome.net/poplar.php>) and *Vitis vinifera* (http://plants.ensembl.org/Vitis_vinifera/Info/Index/) with an *e*-value of 1e-5. The sequences with no blastx hits were then searched against the NR database (NCBI non-redundant protein database 'nr') with an *e*-value of 1e-5. To identify the biologic processes enriched in the *L. kaempferi* transcriptome, we assigned the Gene Ontology (GO) terms associated with the top hits in the four protein databases to the annotated transcripts.

2.5 Evolutionary analysis

To do evolutionary analysis, we extracted likely coding regions from the *L. kaempferi* annotated transcripts using TransDecoder, included with the Trinity software. First, protein sequences longer than 30 amino-acids from *A. thaliana*, *L. kaempferi*, *P. abies*, *P. trichocarpa* and *V. vinifera* were collected. Second, blastp (the basic local alignment search tool, simply compares a protein query to a protein database) [43] was used against a database containing a protein data set of the five species with the *e*-value of 1e-7. Then, the single-copy orthologous genes among the five species were used to construct a

phylogenetic tree using PhyML [44] with default parameters.

2.6 Molecular marker development

To aid *Larix* breeding and genetic diversity studies, we identified single nucleotide polymorphism (SNP), simple sequence repeat (SSR), and insertion and deletion (InDel) markers using the transcriptome data obtained in this work. To identify SNPs, the sequencing reads were aligned with the representative transcripts in the CLC genomic workbench (<http://www.clcbio.com/products/clc-genomics-workbench/>). The general alignment parameters were set to defaults except that non-specific matches were ignored to minimize read-alignment ambiguities. To capture reliable SNPs, we adjusted the minimum read coverage to 5.

SSRs were identified using Msatfinder (<http://www.genomics.ceh.ac.uk/msatfinder/>). The repeat thresholds for mononucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide motifs were set at 12, 8, 5, 5, 5, and 5, respectively. Only SSRs with flanking sequences longer than 50 bp on both sides were collected.

BWA and SAMtools software [45,46] were used to align reads to transcriptome reference and call InDels. The filtering threshold was set as follows: read depth no less than 10 and quality score no less than 20. The default parameter was used for quality control of flanking sequences in the SAMtools mpileup step.

3 Results

3.1 Transcriptome assembly

From the stem transcriptome sequencing of 1-, 2-, 5-, 10-, 25- and 50-year-old *L. kaempferi*, a total of 26074916 reads were generated and have been deposited in the NCBI SRA database (accession number: SRR1107838). Another 591759 Roche 454 sequencing reads and 3874 ESTs were also included in the *L. kaempferi* transcriptome sequence assembly. Assembly of 26670549 reads generated 146786 transcripts, ranging from 200 to 16701 bp, with a mean length of 849 bp and an N50 length of 1538 bp (Table 1).

Table 1 Summary of transcriptome assembly

Types of data	Total sequences	Total bases/bp	Mean length/bp
Illumina	26074916	2607491600	100
Roche 454	591759	211291630	357
Sanger	3874	2956198	763
Total	26670549		
Assembled transcripts	146786	124640235	849

3.2 Transcriptome annotation

Transcripts were first annotated by blastx of four protein databases from *A. thaliana*, *P. abies*, *P. trichocarpa*, and *V. vinifera*, and then the NR database (e -value < 1e-5). Altogether, 79182 (53.9%) of the 146786 transcripts had significant matches, at least one hit in these databases. Among the 79182 annotated transcripts, 74744 (94.4%) were matched to *P. abies*, 56573 (71.4%) to *A. thaliana*, 55807 (70.5%) to *V. vinifera*, 54687 (69.1%) to *P. trichocarpa*, and 3022 (3.8%) to the NR database.

GO annotation was performed for the 79182 annotated transcripts in terms of ‘biologic process’, ‘molecular function’ and ‘cellular component’. In total, 55349 (69.9%) transcripts were assigned to a total of 2024491 GO terms: 53241 transcripts were assigned to the biologic process category, 53015 to molecular function, and 51800 to cellular component. In the biologic process category, ‘response to stress’ was the most abundant GO term (21280, 40.0%), followed by ‘biosynthetic process’ (20855, 39.2%), and ‘development of an anatomical structure’ (17558, 33.0%) (Fig. 1).

3.3 Evolutionary analysis

Seventy-two single-copy gene families across *A. thaliana*, *L. kaempferi*, *P. abies*, *P. trichocarpa* and *V. vinifera* were obtained. The gene families were concatenated to five super-peptides (100084 peptide sites) to construct a phylogenetic tree. The phylogenetic tree analysis showed that *L. kaempferi* was closely related to *P. abies* (Fig. 2), consistent with their taxonomic classification and evolutionary relationship.

3.4 Identification of gene-associated markers

In total, 463482 high-quality SNPs from 48578 transcripts were identified, including 272417 transitions and 191065 transversions (Table 2). The minor allele frequencies of SNPs were estimated from the sequence data (Fig. 3a). The distribution of SNPs per transcript was estimated (Fig. 3b) and the overall frequency of all types of SNPs in the transcriptome was one per 108 bp. Among these SNPs, 364227 (78.6%) were identified from transcripts with annotation information, and they were distributed in 32453 known genes (Table 2).

The 4756 SSRs detected in 4438 transcripts (Table 2) included 2165 (45.5%) mononucleotide, 569 (12.0%) dinucleotide, 1941 (40.8%) trinucleotide, 57 (1.2%) tetranucleotide, 17 (0.4%) pentanucleotide and 7 (0.1%) hexanucleotide motifs. Apart from the mononucleotide motifs, the most abundant was ACG\CGT (9.7%), followed by AT\AT (8.5%), and AAG\CTT (8.5%) (Fig. 3c). Suitable PCR primers were designed for 3595 SSRs using primer 3 [47].

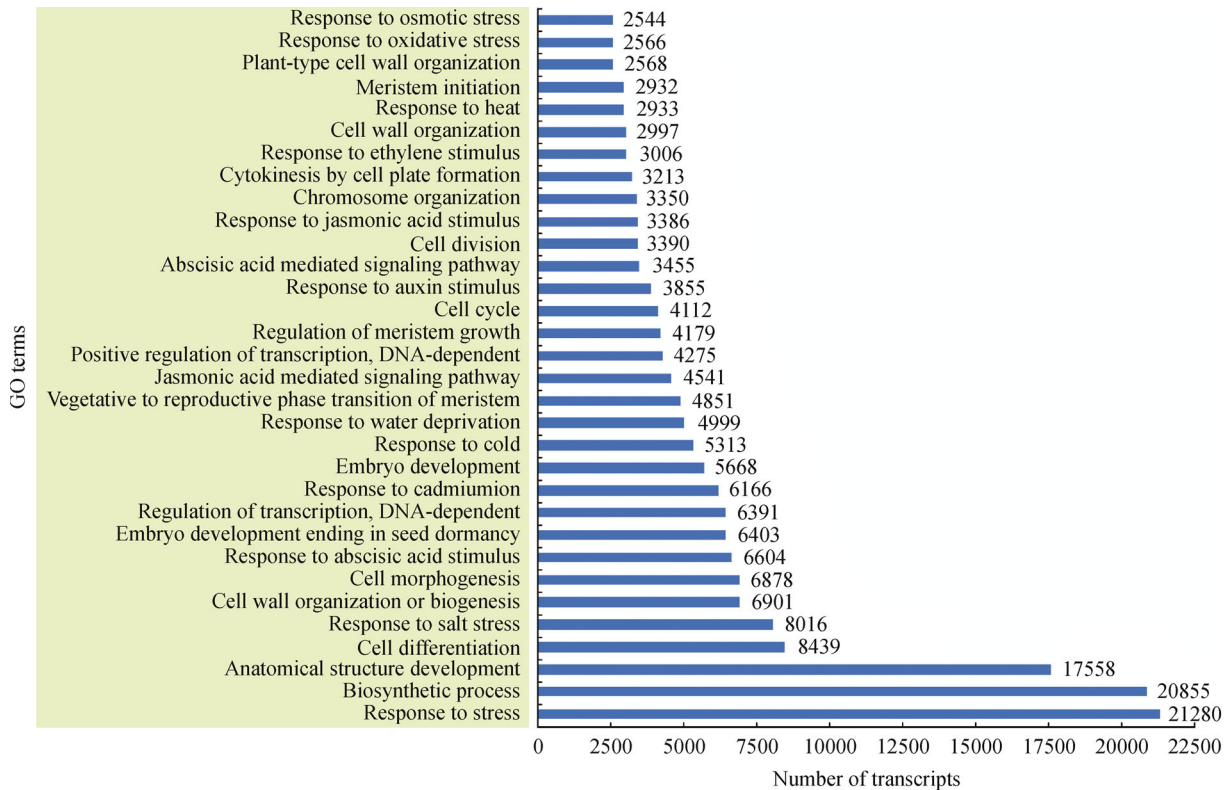


Fig. 1 Representation of Gene Ontology (GO) classification in terms of ‘biologic processes’

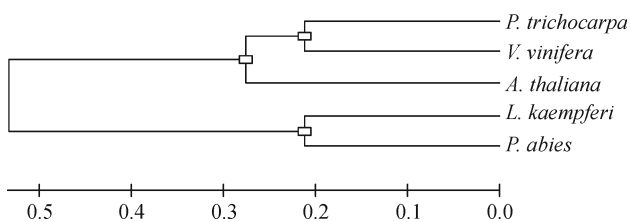


Fig. 2 Phylogenetic tree analysis from the orthologous data set across *A. thaliana*, *L. kaempferi*, *P. abies*, *P. trichocarpa* and *V. vinifera*

Table 2 Molecular markers identified from *L. kaempferi*

Marker information	SNP	SSR	InDel
Total number of identified markers	463482	4756	12434
Number of markers within annotated transcripts	364227	3520	8723
Number of transcripts with markers	48578	4438	10357
Number of annotated transcripts with markers	32453	3264	7035

A total of 12434 InDels were identified in 10357 transcripts (Table 2), including 4080 deletions and 8354 insertions. Only 1640 (15.8%) of the sequences had more than one InDel. Among the 4080 deletions, >68% were mononucleotide deletions, followed by dinucleotide (16.7%), trinucleotide (11.0%), tetranucleotide (3.1%), and pentanucleotide (1.0%) deletions. Among the 8354

insertions, mononucleotide insertions were also prevalent (59.8%), while insertions of dinucleotide (19.7%), trinucleotide (13.6%), tetranucleotide (4.3%), and pentanucleotide (2.6%) accounted for small proportions.

4 Discussion

Environmental signals control tree growth and development. Notably, age affects their responses to these environmental signals [3,48]. Studying the underlying mechanisms is important not only for understanding the adaptation of trees in the context of global climate change, but also for forest breeding and management. For example, in one rotation period of *Larix*, the rate of increase in height peaks at the age of about 15 [49]; so determining the optimal time to harvest forest trees is important. Here, large-scale identification of the transcripts associated with ‘response to stress’ (21280), ‘response to salt stress’ (8016), ‘response to cadmium ion’ (6166), ‘response to cold’ (5313), ‘response to water deprivation’ (4999), ‘response to heat’ (2933), ‘response to oxidative stress’ (2566), and ‘response to osmotic stress’ (2544) will facilitate studies of the environmental regulation of tree growth and development and the mechanisms of age control of the responses to these environmental signals.

As a tree ages, floral meristems develop. Based on the functional annotation by GO, 4851 transcripts were

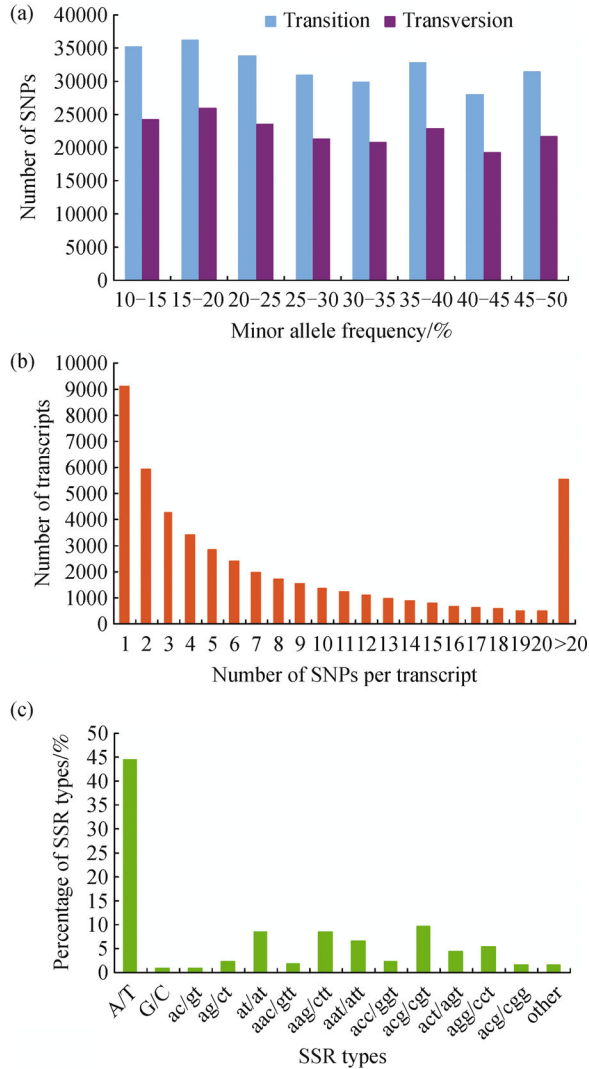


Fig. 3 Summary analysis of single-nucleotide polymorphisms (SNPs) and single-nucleotide polymorphisms (SSRs). (a) Distribution of minor allele frequencies of SNPs identified in *L. kaempferi*. The x-axis represents the SNP sequence-derived minor allele frequency, while the y-axis represents the number of SNPs with a given minor allele frequency; (b) distribution of SNPs per transcript. The x-axis represents the number of SNPs per transcript, while the y-axis represents the number of transcripts with a given number of SNPs; (c) frequency of SSRs. The x-axis represents the SSR types and the y-axis represents the percentage of a given SSR type.

assigned to the GO term ‘vegetative to reproductive phase transition of meristem’, suggesting that the transcriptome in meristem is reprogrammed during aging. The transcripts assigned to ‘regulation of transcription, DNA-dependent’ (6391), ‘positive regulation of transcription, DNA-dependent’ (4275), and ‘chromosome organization’ (3350) might participate in these transcriptome reprogramming processes. These results provide important information for studies of the phase-transition of meristem from vegetative to reproductive during tree aging.

When we collected the samples for sequencing, the *L. kaempferi* trees were actively growing. Active cambium cell division and wood formation were reflected by the assignment of 8439 transcripts to ‘cell differentiation’, 6901 to ‘cell wall organization or biogenesis’, 4112 to ‘cell cycle’, 3390 to ‘cell division’, 3213 to ‘cytokinesis by cell plate formation’, 2997 to ‘cell wall organization’ and 2568 to ‘plant-type cell wall organization’. Hormones control cambium activity and wood formation [50,51], and this was supported by the assignment of 4541 transcripts to ‘jasmonic acid mediated signaling pathway’, 3855 to ‘response to auxin stimulus’, 3386 to ‘response to jasmonic acid stimulus’, and 3006 to ‘response to ethylene stimulus’. These results will help to understand the molecular basis of wood formation and the regulation of cambium activity by hormones.

Somatic embryogenesis is not only a valuable technique in clonal propagation, but also provides a useful experimental system to study the regulatory mechanisms of plant development [52–54]. Hormonal signals also play regulatory roles in somatic embryogenesis. For example, abscisic acid promotes the maturation of the somatic embryo, regulates the synchronization of its development, maintains its dormant state, and suppresses its germination [29,31,35,55,56]. Based on the functional annotation by GO, 17558 transcripts were assigned to the GO term ‘anatomical structure development’, 6878 to ‘cell morphogenesis’, 6604 to ‘response to abscisic acid stimulus’, 6403 to ‘embryo development ending in seed dormancy’, 5668 to ‘embryo development’, and 3455 to ‘abscisic acid mediated signaling pathway’. Consistent with the reorganization of meristem during somatic embryogenesis, 4179 transcripts were assigned to ‘regulation of meristem growth’ and 2932 to ‘meristem initiation’. These data depict the features of somatic embryogenesis and give prominence to the regulation of somatic embryo development by abscisic acid.

Molecular markers are important for the assessment of genetic diversity [57–61], pedigree and mating system analyses [62], hybrid identification [63], the development of genetic maps, and marker-assisted breeding. In *Larix*, EST-SSR markers have been developed [64,65] and two highly-informative SSRs have also been identified recently [66], but other types of markers, such as SNPs and InDels, have not been well developed. Here, using *Larix* transcriptome sequences, we identified 463482 SNPs, 4756 SSRs, and 12434 InDels, which will have considerable utility for the assessment of genetic diversity and marker-assisted breeding in *Larix*.

5 Conclusions

Here, we present an analysis of *Larix* transcriptomes. All of these data enrich the transcriptome resources of *Larix*, are publicly available online, and will serve as useful tools

for understanding the molecular basis of tree growth and development and for breeding and genetic diversity studies of *Larix*.

Acknowledgements This work was supported by the National Natural Science Foundation of China (31200464 and 31330017), and the National High Technology Research and Development Program of China (2011AA100203 and 2013AA102704). The authors thank Dr. IC Bruce (Zhejiang University) and Dr. Yong Guo (Institute of Crop Science, Chinese Academy of Agricultural Sciences) for critical reading of the manuscript, and Dr. Tao Wu for sample collection.

Compliance with ethics guidelines Wanfeng Li, Suying Han, Liwang Qi and Shougong Zhang declare that they have no conflict of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

- Rossi S, Deslauriers A, Anfodillo T, Carrer M. Age-dependent xylogenesis in timberline conifers. *New Phytologist*, 2008, 177(1): 199–208
- Rossi S, Deslauriers A, Gričar J, Seo J W, Rathgeber C B K, Anfodillo T, Morin H, Levanic T, Oven P, Jalkanen R. Critical temperatures for xylogenesis in conifers of cold climates. *Global Ecology and Biogeography*, 2008, 17(6): 696–707
- Begum S, Nakaba S, Oribe Y, Kubo T, Funada R. Cambial sensitivity to rising temperatures by natural condition and artificial heating from late winter to early spring in the evergreen conifer *Cryptomeria japonica*. *Trees-Structure and Function*, 2010, 24(1): 43–52
- Li X, Liang E, Gričar J, Prislan P, Rossi S, Čufar K. Age dependence of xylogenesis and its climatic sensitivity in Smith fir on the south-eastern Tibetan Plateau. *Tree Physiology*, 2013, 33(1): 48–56
- Li W F, Ding Q, Chen J J, Cui K M, He X Q. Induction of *PtoCDKB* and *PtoCYCB* transcription by temperature during cambium reactivation in *Populus tomentosa* Carr. *Journal of Experimental Botany*, 2009, 60(9): 2621–2630
- Little C H A, Bonga J M. Rest in cambium of *Abies balsamea*. *Canadian Journal of Botany*, 1974, 52(7): 1723–1730
- Mwange K N, Wang X W, Cui K M. Mechanism of dormancy in the buds and cambium of *Eucommia ulmoides*. *Acta Botanica Sinica*, 2003, 45(6): 698–704
- Mwange K N, Hou H W, Wang Y Q, He X Q, Cui K M. Opposite patterns in the annual distribution and time-course of endogenous abscisic acid and indole-3-acetic acid in relation to the periodicity of cambial activity in *Eucommia ulmoides* Oliv. *Journal of Experimental Botany*, 2005, 56(413): 1017–1028
- Sundberg B, Little C H A. Tracheid production in response to changes in the internal level of indole-3-acetic Acid in 1-year-old shoots of scots pine. *Plant Physiology*, 1990, 94(4): 1721–1727
- Baba K, Karlberg A, Schmidt J, Schrader J, Hvidsten T R, Bako L, Bhalerao R P. Activity-dormancy transition in the cambial meristem involves stage-specific modulation of auxin response in hybrid aspen. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(8): 3418–3423
- Nilsson J, Karlberg A, Antti H, Lopez-Vernaza M, Mellerowicz E, Perrot-Rechenmann C, Sandberg G, Bhalerao R P. Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell*, 2008, 20(4): 843–855
- Savidge R A, Wareing P F. A tracheid-differentiation factor from pine needles. *Planta*, 1981, 153(5): 395–404
- Little C H A, Sundberg B. Tracheid production in response to indole-3-acetic-acid varies with internode age in *Pinus sylvestris* stems. *Trees-Structure and Function*, 1991, 5(2): 101–106
- Savidge R A. The role of plant hormones in higher plant cellular differentiation. II. Experiments with the vascular cambium, and sclereid and tracheid differentiation in the pine, *Pinus contorta*. *Histochemical Journal*, 1983, 15(5): 447–466
- Alvarez C, Valledor L. R. H, Sanchez-Olate M, Ríos D. Variation in gene expression profile with aging of *Pinus radiata* D. Don. *BMC Proceedings*, 2011, 5 (Suppl 7): 62
- Busov V B, Johannes E, Whetten R W, Sederoff R R, Spiker S L, Lanz-Garcia C, Goldfarb B. An auxin-inducible gene from loblolly pine (*Pinus taeda* L.) is differentially expressed in mature and juvenile-phase shoots and encodes a putative transmembrane protein. *Planta*, 2004, 218(6): 916–927
- Carlsbecker A, Tandere K, Johanson U, Englund M, Engström P. The MADS-box gene *DALI* is a potential mediator of the juvenile-to-adult transition in Norway spruce (*Picea abies*). *Plant Journal*, 2004, 40(4): 546–557
- Diego L B, Berdasco M, Fraga M F, Cañal M J, Rodríguez R, Castresana C. A *Pinus radiata* AAA-ATPase, the expression of which increases with tree ageing. *Journal of Experimental Botany*, 2004, 55(402): 1597–1599
- Fernández-Ocaña A, Carmen García-López M, Jiménez-Ruiz J, Saniger L, Macías D, Navarro F, Oya R, Belaj A, de la Rosa R, Corpas F J, Bautista Barroso J, Luque F. Identification of a gene involved in the juvenile-to-adult transition (JAT) in cultivated olive trees. *Tree Genetics & Genomes*, 2010, 6(6): 891–903
- Hutchison K W, Sherman C D, Weber J, Smith S S, Singer P B, Greenwood M S. Maturation in larch: II. effects of age on photosynthesis and gene expression in developing foliage. *Plant Physiology*, 1990, 94(3): 1308–1315
- Li X, Wu H X, Southerton S G. Seasonal reorganization of the xylem transcriptome at different tree ages reveals novel insights into wood formation in *Pinus radiata*. *New Phytologist*, 2010, 187(3): 764–776
- Li X, Wu H X, Southerton S G. Transcriptome profiling of wood maturation in *Pinus radiata* identifies differentially expressed genes with implications in juvenile and mature wood variation. *Gene*, 2011, 487(1): 62–71
- Wang J W, Park M Y, Wang L J, Koo Y, Chen X Y, Weigel D, Poethig R S. miRNA control of vegetative phase change in trees. *PLoS Genetics*, 2011, 7(2): e1002012
- Hsu C Y, Liu Y, Luthe D S, Yuceer C. Poplar *FT2* shortens the juvenile phase and promotes seasonal flowering. *Plant Cell*, 2006, 18(8): 1846–1861
- Hsu C Y, Adams J P, Kim H, No K, Ma C, Strauss S H, Drnevich J, Vandervelde L, Ellis J D, Rice B M, Wickett N, Gunter L E, Tuskan G A, Brunner A M, Page G P, Barakat A, Carlson J E, DePamphilis

- C W, Luthe D S, Yuceer C. *FLOWERING LOCUS T* duplication coordinates reproductive and vegetative growth in perennial poplar. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(26): 10756–10761
26. Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner A M, Jansson S, Strauss S H, Nilsson O. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science*, 2006, 312(5776): 1040–1043
27. Li S G, Li W F, Han S Y, Yang W H, Qi L W. Stage-specific regulation of four HD-ZIP III transcription factors during polar pattern formation in *Larix leptolepis* somatic embryos. *Gene*, 2013, 522(2): 177–183
28. Li W F, Zhang S G, Han S Y, Wu T, Zhang J H, Qi L W. Regulation of *LaMYB33* by miR159 during maintenance of embryogenic potential and somatic embryo maturation in *Larix kaempferi* (Lamb.) Carr. *Plant Cell, Tissue and Organ Culture*, 2013, 113(1): 131–136
29. Zhang J, Zhang S, Han S, Li X, Tong Z, Qi L. Deciphering small noncoding RNAs during the transition from dormant embryo to germinated embryo in Larches (*Larix leptolepis*). *PLoS ONE*, 2013, 8(12): e81452
30. Zhang J, Zhang S, Han S, Wu T, Li X, Li W, Qi L. Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. *Planta*, 2012, 236(2): 647–657
31. Zhang L F, Li W F, Han S Y, Yang W H, Qi L W. cDNA cloning, genomic organization and expression analysis during somatic embryogenesis of the translationally controlled tumor protein (TCTP) gene from Japanese larch (*Larix leptolepis*). *Gene*, 2013, 529(1): 150–158
32. Zhang S, Zhou J, Han S, Yang W, Li W, Wei H, Li X, Qi L. Four abiotic stress-induced miRNA families differentially regulated in the embryogenic and non-embryogenic callus tissues of *Larix leptolepis*. *Biochemical and Biophysical Research Communications*, 2010, 398(3): 355–360
33. Zhang S G, Han S Y, Yang W H, Wei H L, Zhang M, Qi L W. Changes in H₂O₂ content and antioxidant enzyme gene expression during the somatic embryogenesis of *Larix leptolepis*. *Plant Cell, Tissue and Organ Culture*, 2010, 100(1): 21–29
34. Li W F, Zhang S G, Han S Y, Wu T, Zhang J H, Qi L W. The post-transcriptional regulation of *LaSCL6* by miR171 during maintenance of embryogenic potential in *Larix kaempferi* (Lamb.). *Carr. Tree Genetics & Genomes*, 2014, 10(1): 223–229
35. Zhang J H, Zhang S G, Li S G, Han S Y, Li W F, Li X M, Qi L W. Regulation of synchronism by abscisic-acid-responsive small noncoding RNAs during somatic embryogenesis in larch (*Larix leptolepis*). *Plant Cell, Tissue and Organ Culture*, 2014, 116(3): 361–370
36. Zhang Y, Zhang S, Han S, Li X, Qi L. Transcriptome profiling and in silico analysis of somatic embryos in Japanese larch (*Larix leptolepis*). *Plant Cell Reports*, 2012, 31(9): 1637–1657
37. Men L, Yan S, Liu G. *De novo* characterization of *Larix gmelinii* (Rupr.) Rupr. transcriptome and analysis of its gene expression induced by jasmonates. *BMC Genomics*, 2013, 14(1): 548
38. Mackay J, Dean J F, Plomion C, Peterson D G, Cánovas F M, Pavy N, Ingvarsson P K, Savolainen O, Guevara M Á, Fluch S, Vinceti B, Abarca D, Díaz-Sala C, Cervera M T. Towards decoding the conifer giga-genome. *Plant Molecular Biology*, 2012, 80(6): 555–569
39. Grabherr M G, Haas B J, Yassour M, Levin J Z, Thompson D A, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren B W, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 2011, 29(7): 644–652
40. Zhang L, Qi L W, Han S Y. [Differentially expressed genes during *Larix* somatic embryomaturation and the expression profile of partial genes. *Hereditas*, 2009, 31(5): 540–545 (in Chinese)
41. Zhang L, Qi L, Han S. Construction and analysis of differentially expressed cDNA library of larch somatic embryo at the stage of proembryogenic mass. *Molecular Plant Breeding*, 2008, 6(4): 675–682 (in Chinese)
42. Huang X, Madan A. CAP3: A DNA sequence assembly program. *Genome Research*, 1999, 9(9): 868–877
43. Altschul S F, Gish W, Miller W, Myers E W, Lipman D J. Basic local alignment search tool. *Journal of Molecular Biology*, 1990, 215(3): 403–410
44. Guindon S, Dufayard J F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 2010, 59(3): 307–321
45. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009, 25(14): 1754–1760
46. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 2009, 25(16): 2078–2079
47. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth B C, Remm M, Rozen S G. Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 2012, 40(15): e115
48. Carrer M, Urbinati C. Age-dependent tree-ring growth responses to climate in *Larix decidua* and *Pinus cembra*. *Ecology*, 2004, 85(3): 730–740
49. Tian Z H, Dong J, Wang X W, Huang G R. *Silviculture of Larix kaempferi*. 1995, Beijing: Beijing Agriculture University Press 94–105 (in Chinese)
50. Sorce C, Giovannelli A, Sebastiani L, Anfodillo T. Hormonal signals involved in the regulation of cambial activity, xylogenesis and vessel patterning in trees. *Plant Cell Reports*, 2013, 32(6): 885–898
51. Sehr E M, Agusti J, Lehner R, Farmer E E, Schwarz M, Greb T. Analysis of secondary growth in the Arabidopsis shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant Journal*, 2010, 63(5): 811–822
52. Cairney J, Pullman G S. The cellular and molecular biology of conifer embryogenesis. *New Phytologist*, 2007, 176(3): 511–536
53. Quiroz-Figueroa F R, Rojas-Herrera R, Galaz-Avalos R M, Loyola-Vargas V M. Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell, Tissue and Organ Culture*, 2006, 86(3): 285–301
54. Zimmerman J L. Somatic embryogenesis: a model for early development in higher plants. *Plant Cell*, 1993, 5(10): 1411–1423

55. Gutmann M, vonAderkas P, Label P, Lelu M A. Effects of abscisic acid on somatic embryo maturation of hybrid larch. *Journal of Experimental Botany*, 1996, 47(12): 1905–1917
56. Rai M K, Shekhawat N S, Harish, Gupta A K, Phulwaria M, Ram K, Jaiswal U. Harish, Gupta A K, Phulwaria M, Ram K, Jaiswal U. The role of abscisic acid in plant tissue culture: a review of recent progress. *Plant Cell, Tissue and Organ Culture*, 2011, 106(2): 179–190
57. Khasa D P, Jaramillo-Correa J P, Jaquish B, Bousquet J. Contrasting microsatellite variation between subalpine and western larch, two closely related species with different distribution patterns. *Molecular Ecology*, 2006, 15(13): 3907–3918
58. Pluess A R. Pursuing glacier retreat: genetic structure of a rapidly expanding *Larix decidua* population. *Molecular Ecology*, 2011, 20(3): 473–485
59. Oreshkova N V, Belokon M M, Jamiyansuren S. Genetic diversity, population structure, and differentiation of Siberian larch, Gmelin larch and Cajander larch on SSR-markers data. *Russian Journal of Genetics*, 2013, 49(2): 178–186
60. Kozyrenko M M, Artyukova E V, Reunova G D, Levina E A, Zhuravlev Y N. Genetic diversity and relationships among Siberian and Far Eastern larches inferred from RAPD analysis. *Russian Journal of Genetics*, 2004, 40(4): 401–409
61. Yu X M, Zhou Q, Qian Z Q, Li S, Zhao G F. Analysis of genetic diversity and population differentiation of *Larix potaninii* var. *chinensis* using microsatellite DNA. *Biochemical Genetics*, 2006, 44(11–12): 483–493
62. Funda T, Chen C, Liewlaksaneeyanawin C, Kenawy A M A, El-Kassaby Y A.C.. Liewlaksaneeyanawin C, Kenawy A M A, El-Kassaby Y A.. Pedigree and mating system analyses in a western larch (*Larix occidentalis* Nutt.) experimental population. *Annals of Forest Science*, 2008, 65(7): 705
63. Acheré V, Faivre Rampant P, Pâques L E, Prat D. Chloroplast and mitochondrial molecular tests identify European × Japanese larch hybrids. *Theoretical and Applied Genetics*, 2004, 108(8): 1643–1649
64. Yang X, Sun X, Zhang S, Xie Y, Han H. Development of EST-SSR markers and genetic diversity analysis of the second cycle elite population in *Larix kaempferi*. *Scientia Silvae Sinicae*, 2011, 47(11): 52–58 (in Chinese)
65. Liu C, Zhang L P, Wang C G, Song W Q, Chen C B. Development and Characterization of EST-SSR Molecular Markers in *Larix kaempferi*. *Forest research*, 2013, 26(S1): 60–68 (in Chinese)
66. Wagner S, Gerber S, Petit R J. Two highly informative dinucleotide SSR multiplexes for the conifer *Larix decidua* (European larch). *Molecular ecology resources*, 2012, 12(4): 717–25