

Note

Establishment of expressed sequence tags from *Taiwania* (*Taiwania cryptomerioides* Hayata) seedling cDNA

Yu-Rong Chen^a, Yi-Ru Lee^b, Sheng-Yang Wang^c, Shang-Tzen Chang^a,
Jei-Fu Shaw^b, Fang-Hua Chu^{a,*}

^a School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan

^b Institute of Botany, Academia Sinica, Nankang, Taipei, Taiwan

^c Institute of BioAgricultural Sciences, Academia Sinica, Nankang, Taipei, Taiwan

Received 22 March 2004; received in revised form 23 May 2004; accepted 24 May 2004

Available online 15 June 2004

Abstract

Taiwania (*Taiwania cryptomerioides* Hayata), native to Taiwan, is one of the “living fossils” from the Tertiary period of the Cenozoic era. To isolate genes involved in wood formation and biochemical synthesis from this tree, 436 randomly selected clones from a cDNA library derived from seedling were sequenced and analyzed firstly. Contig analysis of these expressed sequence tags (ESTs) identified a total of 246 unigene sets. Based on the results obtained by BLASTX analysis, 10.1% of these ESTs were involved in protein synthesis and processing, including highly expressed small heat shock proteins (smHSPs); 3.9% were related to the cell rescue and defense, including highly expressed cysteine proteinase inhibitors; 0.7% were involved in secondary metabolism. However, 39.0% of ESTs showed no significant similarity to any other protein sequences in public databases. These sequences indicate the uniqueness of *Taiwania*, and it is revealed that the remarkable value of *Taiwania*.

© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: *Taiwania* (*Taiwania cryptomerioides* Hayata); Seedling; Expressed sequence tags (ESTs)

1. Introduction

Taiwania (*Taiwania cryptomerioides* Hayata) is an indigenous Taxodiaceae conifer tree growing at elevations from 1800 to 2600 m in Taiwan's central mountains. *Taiwania* as well as *Ginkgo biloba*, *Sequoiadendron giganteum* and *Metasequoia glyptostroboides* are the relict plants from the Tertiary period of the Cenozoic era. It is one of the most economically important tree species in Taiwan because of its highly wood utilization properties, such as excellent durability. According to the reviews by Wang et al. [1], Kuo et al. [2] and Chang et al. [3], more than 100 secondary metabolite, including terpenoids, lignans, isoflavones, and other compounds, have been isolated from this species during the past 70 years [1–3]. There have also been many studies dealing with the bioactivities of chemical ingredients from *Taiwania*,

including antibacterial, antifungal, antitermite, antioxidant, antimite and antitumor activities [3]. In order to obtain more information about the regulation of wood and metabolite formation at a molecular level in *Taiwania*, the pilot expressed sequence tags (ESTs) study of *Taiwania* seedling cDNA library was established. By comparing with identified genes from available public databases, obtained putative identification of cDNAs will help researchers to understand the biochemical systems, physiological mechanisms, and wood formation at molecular level in general, and of *Taiwania* in particular.

2. Materials and methods

Two-year-old *Taiwania* seedlings were obtained from the Experimental Forest of National Taiwan University. Total RNA extraction was performed by the method used by Chang et al. [4] (with slight modification). Poly (A)⁺ RNA was then purified with Oligotex mRNA Mini Kit (Qiagen).

* Corresponding author. Tel.: +886 2 33665261; fax: +886 2 23654520.

E-mail address: fhchu@ntu.edu.tw (F.-H. Chu).

First-strand cDNA was synthesized by SuperScript™ II Reverse Transcriptase (Invitrogen) and the double-strand cDNA was amplified using PCR Plus Master Mix Kit (GeneMark). Sized-fractionation (>100 bp) and purification of the cDNA were performed using DNA Clean/Extraction Kit (GeneMark). The resulting cDNA was cloned into pGEM-T Easy vector system (Promega). Plasmids DNA were isolated from overnight cultures by the Plasmid Miniprep Purification Kit (GeneMark). Sequencing was performed with an ABI 377 automatic sequencer (Perkin-Elmer), using T7 sequencing primer to obtain the 5'-end of the insert. Individual ESTs were assembled into contigs using the Contig Express program of Vector NTI Suite 8, with parameters optimized for ESTs rather than for genomic clones. DNA sequence analysis, according to the procedure described by Shelton et al. [5], was performed using the WU-BLAST 2.0 search of the NCBI's NR Protein Database (BLASTX 2.2.8) at the Baylor College of Medicine Search Launcher (http://www.searchlauncher.bcm.tmc.edu/seq-search/nucleic_acid-search.html). *P*-value results obtained from the BLASTX: $<10^{-10}$ were characterized to have significant homology and categorized according to putative function; $>10^{-10}$ were deemed to no hit.

3. Results and discussion

Of the 436 high quality clones analyzed, the average length of sequence was 497 nucleotides, which have been submitted to dbEST (accession number: CK900683–CK901118). A total of 56 contigs were formed after assembly of the 436 ESTs. The contigs contained a total

of 246 sequences, whereas 190 sequences remained as singleton ESTs. This corresponds to a redundancy of 56%. As a result of the contig analysis, we obtained a total of 246 (56 + 190) unigene sets. Among the 436 submitted sequences, 266 (61.0%) of the ESTs were similar to sequences in public databases and 228 (52.3%) of the ESTs analyzed were homologous to genes of known function after the BLASTX analysis. The remaining 170 clones were not similar to any sequences in the available databases. The functional classification of the 228 putatively identified cDNA clones and highly expressed transcripts were listed in Table 1. It is notable that small heat shock proteins (smHSPs) were expressed abundantly in the *Taiwania* seedling cDNA. The similar results were also observed in cDNA libraries of *Citrus sinensis* young seedlings and *Cryptomeria japonica* (Taxodiaceae) [6,7]. It was also found that genes related to cell rescue and defense, such as cysteine proteinase inhibitors were expressed abundantly. Although there were few secondary metabolite related genes (0.7%) obtained, many studies of phytochemistry in *Taiwania*, which have been reported still revealed the important of researching the secondary metabolites of *Taiwania*. Therefore, we will still focus on secondary metabolites synthesis within *Taiwania* and their related enzymes in further investigations. A high percentage of no hit (39.0%) was also revealed in this ESTs analysis. The absence of these sequences from public databases may indicate specific roles in *Taiwania*, and stress the necessary to future studying *Taiwania*. Results of the ESTs analysis described here provide a preliminary data for understanding the genes expressed in the seedlings of *Taiwania*. Further investigation into the functional genomics of *Taiwania*, especially mature woody samples, is on going.

Table 1
Functional classification of the *Taiwania cryptomerioides* seedling cDNA library

Functional classification	NC. ^a	Per. ^b (%)	Example of genes expressed abundantly	<i>P</i> -value
Cell wall	1	0.2	Proline-rich protein (<i>Arabidopsis thaliana</i>)	1.00E-14
Chromatin and DNA metabolism	1	0.2	H2A histone protein (<i>Petroselinum crispum</i>)	7.20E-35
Energy	2	0.5	Phosphoenolpyruvate synthase (<i>Arabidopsis thaliana</i>)	2.90E-33
Secondary metabolism	3	0.7	Flavonoid 3'-hydroxylase (<i>Pelargonium x hortorum</i>)	2.10E-14
Membrane transport	4	0.9	ABC transporter (<i>Arabidopsis thaliana</i>)	1.90E-84
Cell cycle and cell division	7	1.6	14-3-3 protein (<i>Picea glauca</i>)	2.30E-77
Transcription, gene expression and RNA metabolism	8	1.8	Glycine-rich RNA-binding protein (<i>Picea glauca</i>)	2.00E-39
Cell rescue and defense	17	3.9	Cysteine proteinase inhibitor (<i>Glycine max</i>)	3.80E-53
Signal transduction and cellular communication	25	5.7	Mago Nashi-like protein (<i>Euphorbia lagascae</i>)	2.10E-66
Protein synthesis and processing	44	10.1	Small heat shock protein (<i>Arabidopsis thaliana</i>)	5.30E-16
Primary metabolism	50	11.5	Glyceraldehyde-phosphate dehydrogenase (<i>Pinus sylvestris</i>)	1.80E-72
Photosynthesis	58	13.3	Type II chlorophyll a/b-binding protein (<i>Pinus sylvestris</i>)	3.40E-92
Other	8	1.9	Translationally controlled tumor protein (<i>Pseudotsuga menziesii</i>)	3.30E-71
Unknown	38	8.7		
No hit	170	39.0		
Total	436			

^a Number of clones.

^b Percentage.

Acknowledgements

We thank Ms. May-Chih Fen, Institute of Botany, Academia Sinica, for DNA sequencing. This research was supported by a research grant from the National Science Council, Republic of China (NSC-92-2313-B002-128).

References

- [1] S.Y. Wang, S.T. Chang, Y.C. Su, Y.H. Kuo, Studies on the extractives of Taiwanian (*Taiwania cryptomerioides* Hayata): a review, Q. J. Exp. Formos. Nat. Taiwan Univ. 11 (1997) 67–81.
- [2] Y.H. Kuo, M.T. Yu, L.L. Shiu, Natural products research in Taiwan V, Formos. Sci. 52 (1999) 1–145.
- [3] S.T. Chang, S.Y. Wang, Y.H. Kuo, Resources and bioactive substances from Taiwanian, J. Wood Sci. 49 (2003) 1–4.
- [4] S. Chang, J. Puryear, J. Cairney, A simple and efficient method for isolation RNA from pine trees, Plant Mol. Biol. Rep. 11 (1993) 113–116.
- [5] D. Shelton, D. Leach, P. Baverstock, R. Henry, Isolation of genes involved in secondary metabolism from *Melaleuca alternifolia* (Cheel) using expressed sequence tags (ESTs), Plant Sci. 162 (2002) 9–15.
- [6] M. Bausher, R. Shatters, J. Chaparro, P. Dang, W. Hunter, An expressed sequence tag (EST) set from *Citrus sinensis* L. Osbeck whole seedlings and the implications of further perennial source investigations, Plant Sci. 165 (2003) 415–422.
- [7] T. Ujjino-Ihara, K. Yoshimura, Y. Ugawa, H. Yoshimaru, K. Nagasaka, Y. Tsumura, Expression analysis of ESTs derived from the inner bark of *Cryptomeria japonica*, Plant Mol. Biol. 43 (2000) 451–457.