Identification of laticifer-specific genes and their promoter regions from a natural rubber producing plant *Hevea brasiliensis*®

Yuichi Aoki,a Seiji Takahashi,a−, Daisuke Takayama,b Yoshiyuki Ogata,c,1 Nozomu Sakurai,c Hideyuki Suzuki,c Kasem Asawatreratanakul,b,d Dhirayos Wititsuwannakul,e Rapeutn Wititsuwannakul,f Daisuke Shibata,c Tanetoshi Koyama,b Toru Nakayamaa

a Graduate School of Engineering, Tohoku University, Sendai, Miyagi 980-8579, Japan
b Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai 980-8577, Japan
c Research Institute of Innovative Technology for the Earth, Tohoku University, Sendai 980-8577, Japan
d Research and Development Institute, Thaksin University, Songkhla, Thailand
e Department of Biochemistry, Mahidol University, Bangkok, Thailand
f Department of Biochemistry, Prince of Songkla University, Hat-Yai, Thailand

**A R T I C L E   I N F O**

**Article history:**
Received 20 November 2013
Received in revised form 2 May 2014
Accepted 5 May 2014
Available online 12 May 2014

**Keywords:**
cis-prenyltransferase
Expressed sequence tag
*Hevea brasiliensis*
Natural rubber
Latex
Laticifer

**A B S T R A C T**

Latex, the milky cytoplasm of highly differentiated cells called laticifers, from *Hevea brasiliensis* is a key source of commercial natural rubber production. One way to enhance natural rubber production would be to express genes involved in natural rubber biosynthesis by a laticifer-specific overexpression system. As a first step to identify promoters which could regulate the laticifer-specific expression, we identified random clones from a cDNA library of *H. brasiliensis* latex, resulting in 4325 expressed sequence tags (ESTs) assembled into 1308 unigenes (692 contigs and 617 singletons). Quantitative analyses of the transcription levels of high redundancy clones in the ESTs revealed genes highly and predominantly expressed in laticifers, such as Rubber Elongation Factor (REF), Small Rubber Particle Proteins and putative protease inhibitor proteins. HRT1 and HRT2, cis-prenyltransferases involved in rubber biosynthesis, was also expressed predominantly in laticifers, although these transcript levels were 80-fold lower than that of REF. The 5′-upstream regions of these laticifer-specific genes were cloned and analyzed in *sito*, revealing seven common motifs consisting of eight bases. Furthermore, transcription factors specifically expressed in laticifers were also identified. The common motifs in the laticifer-specific genes and the laticifer-specific transcription factors are potentially involved in the regulation of gene expression in laticifers.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Natural rubber consists mainly of cis-1,4-polyisoprene and is the most important raw material produced by plants, because of its unique physical properties, including resilience, elasticity, abrasion and impact resistance, efficient heat dispersion and malleability at cold temperatures [1]. Although more than 2500 higher plants produce natural rubber, only the Para rubber tree (*Hevea brasiliensis*) has been established as a key commercial source of natural rubber because of its high yield and excellent physical properties. In *H. brasiliensis*, natural rubber is obtained from latex, the milky cytoplasm of laticifers, which are highly specialized cells in the vascular tissues [2].

Although the precise structure and biosynthetic pathway of natural rubber has not been fully elucidated, many genes have
been suggested to participate in the natural rubber biosynthesis. The basic backbone structure of natural rubber, comprising of cis-1,4-polyisoprene with two or three trans-isoprene units at the \( \omega \)-terminus [3], suggests that its biosynthesis involves sequential condensation of the C\(_5\) isoprene unit, isopentenyl diphosphate (IPP), with cis-configuration onto an all-trans short-chain prenyl diphosphate, such as E,E-farnesyl diphosphate (FPP, C\(_{15}\)) and E,E,E-geranylgeranyl diphosphate (GGPP, C\(_{20}\)). Therefore, the key enzyme responsible for the biosynthesis of natural rubber is considered to be a member of the cis-prenyltransferase (cPT) enzyme family, which catalyze the sequential cis-1,4-condensation of IPP and exist ubiquitously in all organisms for the formation of ZE-mixed isoprenoids, such as polyisoprene, and dolichols. Two cPT homologs, HRT1 and HRT2, cloned from \( H. \) brasiliensis latex, are proposed to be key enzymes responsible for the formation of natural rubber. Recombinant HRT2 protein expressed in \( E. \) coli is significantly activated by the addition of a centrifuged fraction of latex, resulting in the formation of polyisoprenes corresponding to natural rubber [4]. The involvement of cPT in natural rubber biosynthesis is also indicated by the observation that RNA interference-mediated depletion of cPTs in the laticifers of transgenic \( T. \) breviconicum plants resulted in significant reduction in rubber biosynthesis [5].

To meet the ever-increasing demand for natural rubber, especially for the manufacture of tires, metabolic engineering of \( H. \) brasiliensis to improve its natural rubber production is required. One method of enhancing the productivity of natural rubber would be the simple overexpression of genes involved in natural rubber biosynthesis, such as HRT2. However, constitutive overexpression of genes for a specific isoprenoid biosynthetic pathways in whole plant tissues is predicted to affect plant growth and development because of the metabolic disorder of other physiologically important isoprenoids, such as quinones, sterols, and phytohormones [6]. To avoid this problem, the development of a laticifer-specific high-level gene expression system is required. However, no cis- and trans-acting regulatory elements that function in the laticifer-specific expression of genes in \( H. \) brasiliensis have been elucidated.

Expressed sequence tags (ESTs), which are single-pass sequence reads from reverse-transcribed mRNAs, provide a substantial representation of the transcriptome. ESTs have played significant roles in accelerating gene discovery; several reports conducted EST analysis on latex of \( H. \) brasiliensis to provide a comprehensive view of the transcriptional regulation of natural rubber biosynthesis [7–9]. These reports revealed unique transcriptional profiles in latex, showing high-redundancy of ESTs for two major genes, Rubber Elongation Factor (REF) [10] and Small Rubber Particle Protein (SRPP) [11], corresponding to about 30% of total ESTs. The high redundancy of these genes among the EST is considered to correlate with the high expression levels of the corresponding genes in latex, which have also been reported at the transcriptional [9,11,12] and the protein levels [10,13,14]. In addition, the expression levels of REF and SRPP in laticifers (latex) are much higher than those in leaves [7,8,11,12]. However, neither the precise expression levels nor the transcriptional regulation mechanisms of these laticifer-abundant genes have been elucidated.

To develop a laticifer-specific overexpression system, the regulation mechanism for gene expressions of laticifer-specific highly expressed genes must be determined. Here, we report the construction and analyses of ESTs comprising 57,600 sequences obtained by sequencing a cDNA library from latex of \( H. \) brasiliensis to discover genes specifically expressed in laticifers. In addition, we successfully cloned the promoter regions of those genes, which are expected to be useful for laticifer-specific high level gene expression.

2. Materials and methods

### 2.1. Plant materials and total RNA extraction

Latex and various tissue samples were obtained from ten-year-old rubber plants (\( H. \) brasiliensis clone RRIM 600) grown at the Rubber Research Center of Songkla, Thailand. Latex collection by tapping of the bark of the tree trunks was performed basically as described by Kush et al. [15]. Young leaves with size from 15 to 20 cm, green stems which were 5 cm below the shoot apex and less than 1.5 cm in diameter, and lateral roots less than 0.5 mm in diameter were collected, immediately frozen in liquid nitrogen, and stored at \(-80^\circ\)C until the total RNA extraction. Total RNA from latex for the cDNA library construction was extracted using RNA-gents Total RNA Isolation System (Promega, Fitchburg, WI, USA), and Poly(A) + RNA was isolated from the total RNA with Oligodex-dT30 mRNA Purification kit (TakaRa, Ohtsu, Japan). Total RNAs from other tissues were extracted using the TRIzol Reagent (Life Technologies, Carlsbad, CA, USA). Contaminating DNA in the total RNA sample was eliminated by treatment with DNase I (RNase-free; Takara Bio, Ohtsu, Japan) at 37 °C for 30 min.

### 2.2. cDNA library construction and DNA sequencing

A latex cDNA library was constructed using a ZAP-cDNA Synthesis Kit (Stratagene, La Jolla, CA, USA), according to the manufacturer’s instructions. The cDNA library was converted into a phagemid form by mass \textit{in vivo} excision and transfected into \( E. \) coli SOLR strain. The resultant SOLR cells were plated onto Luria and Bertani (LB) media containing 50 \( \mu \)g/ml ampicillin and incubated for 16 h at 37 °C. 5760 colonies were picked by sterile toothpicks and inoculated into 384-well plastic plates containing 60 \( \mu \)L LB medium supplemented with 36 mM K\(_2\)HPO\(_4\), 13.2 mM KH\(_2\)PO\(_4\), 1.7 mM sodium citrate, 0.4 mM MgSO\(_4\), 6.8 mM ammonium sulfate, 4.4% (v/v) glycerol [16] and 50 \( \mu \)g/ml ampicillin, and incubated for 16 h at 37 °C. The insert of each plasmid was amplified by direct PCR from the 384-well plates using primers T3 (5’-ATTAAACCTC ACTAA AGGC-3’) and T7 (5’-GTAATACGACTCACTATA GGCC-3’). The PCR program was as follows: 94 °C for 2 min; 35 cycles of 94 °C for 45 s, 52 °C for 45 s, and 72 °C for 3 min; and 72 °C for 10 min. The PCR products were purified by ethanol precipitation and sequenced using primer T3 and a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) with an automated DNA sequencer (ABI PRISM 3730xl DNA sequencer, Applied Biosystems).

### 2.3. Bioinformatics

The sequence trace files were submitted to the PHRED program [17] using the parameter –trim_alt 0.01 (quality value = 20). Sequences longer than 100 bases were selected and submitted to the CROSS_MATCH program using the parameters, –minmatch: 12 and –minscore: 20. Sequence assembly was performed using the PHRAP program with the default parameters, which resulted in clustered sequences (contigs) and unclustered sequences (singletons). The unigene set comprising the contigs and singletons was subjected to blastn and blastx searches. The blastn searches were performed against the GenBank nt nucleotide database at NCBI using default parameters, and the blastx searches were conducted against the NCBI nr protein database and the TAIR9 peptide database (The Arabidopsis Information Resource version 9; http://www.arabidopsis.org). For gene ontology (GO) analysis, GO slim term annotation of the best-matched Arabidopsis proteins, according to the blastx search results, was assigned to each unigene.
2.4. Quantitative reverse transcription PCR

Quantitative reverse transcription (RT)-PCR analyses were carried out using a One-step SYBR RT-PCR kit (Takara Bio), according to the manufacturer’s instructions, with a LightCycler 1.2 apparatus (Roche Diagnostics, Tokyo, Japan), using gene-specific primers and thermal cycling conditions (Supplemental Table S1).

2.5. Isolation of 5’-upstream flanking regions

Genomic DNA from *H. brasiliensis* leaves was prepared by a standard genomic DNA isolation method for plant tissue using cetyltrimethylammonium bromide [18]. For *REF, SRPP*, and *HRT1*, the genomic DNA segments of the 5’ upstream flanking regions of each gene were isolated by an inverse polymerase chain reaction (IPCR) method [19]. About 1 μg of genomic DNA was separately digested with *XbaI*, *SpeI* or *SacI* restriction enzymes (Takara Bio). After phenol/chloroform (1:1, v/v) extractions and ethanol precipitations, DNA fragments were self-circularized using a Ligation High kit (TOYOBO, Osaka, Japan). The primer sets used for the primary and secondary IPCRs were designed according to the coding sequences for *REF*, *SRPP*, and *HRT1* (GenBank accession numbers: X56535, AF051317 and AB061234) (Supplemental Table S2). The primary IPCR with KOD FX (TOYOBO), using the each DNA fragment as a template, was performed with the following thermal program: 94°C for 2 min; and 30 cycles of 98°C for 10 s, 55°C for 30 s, and 68°C for 3 min. The secondary IPCR with nested primer sets, using the PCR product from the first IPCR as a template, was performed with the following thermal program: 94°C for 2 min; and 30 cycles of 98°C for 10 s, 57°C for 30 s, and 68°C for 3 min. For *Pl*, a 5’ upstream flanking region was amplified by a PCR-based genome walking method using DNA Walking SpeedUp Premix Kit (Seegene, Seoul, Korea), according to the manufacturer’s instructions, with primers TSP1, TSP2 and TSP3 (Table S2), designed according to the *Pl* coding sequences (GenBank accession number: AY221985). Amplified products were analyzed by agarose gel electrophoresis and purified from the agarose gel using UltraClean15 DNA Purification Kit (MO BIO Laboratories, Carlsbad, CA, USA). After the addition of 3’ A-overhangs using a Mighty TA-cloning Reagent Set for PrimeSTAR (Takara Bio), purified products were subcloned into pGEM-T Easy Vector (Promega, Fitchburg, WI, USA) and sequenced using a CEQ 2000XL DNA Analysis System (BECKMAN COULTER, Brea, CA, USA).

2.6. Consensus motif prediction and database-assisted motif annotation

POCO (http://ekhidna.biocenter.helsinki.fi/poxo/poco/) [20] was used to predict eight base motifs that were common among the 5’ upstream regions of *REF, SRPP, PI* and *HRT1*. The predicted motifs were used in homology searches of the Plant Cis-acting Regulatory DNA Elements (PLACE) database (http://www.dna.affrc.go.jp/PLACE/) [21] for similarity to previously described cis-regulatory elements.

3. Results and discussion

3.1. Transcriptome analyses of *H. brasiliensis* latex

In total, 5760 independent colonies of the latex cDNA library were isolated and subjected to 5’-end single-pass sequencing. After eliminating low quality sequences (quality value <20 or sequenced length <100 base), we obtained 4325 high quality sequences (average size: 613 base), which represented the EST dataset (Table 1). Assembly of the ESTs using PHRAP algorithm resulted in 1309 unigenes, comprising 692 contigs and 617 singletons (Table 1, Supplemental Table S3 and S2), which were analyzed with the blastn and blastx algorithm against the non-redundant (nr) database of NCBI. Although 1045 (80%) unigenes were assigned as identical or homologous to identified genes, exhibiting E-values <10⁻⁵, many of the matched genes were annotated as putative proteins with unknown function. These result reflected, at least in part, the increased number of genes of unknown function deposited in the nr database from the Euphorbiaceae family, to which *H. brasiliensis* belongs, such as *Ricinus communis* [22], and *Jatropha curcas* [23]. To predict the biological processes in which the ESTs are involved, unigenes were also matched using the blastx algorithm against the TAIR-9 Arabidopsis database. Gene Ontology (GO) analysis was then performed, assigning GO slim terms to Arabidopsis genes that were the best matches for each EST. Among 12,929 GO slim terms belonging to the Biological Process, the top three GO slim terms assigned to the ESTs were “unknown biological processes” (51.0%), “other cellular processes” (21.2%), and “response to stress” (5.6%) (Fig. 1). The high frequency of gene of “unknown biological processes” may reflect the higher expression levels of Euphorbiaceae family-specific genes and/or laticifer-specific genes involved in specialized physiological roles during laticifer development. Comparative genomic analyses among *A. thaliana*, *Oryza sativa*, *Populus trichocarpa*, *Vitis vinifera*, *Lotus japonicus*, *Glycine max* and Euphorbiaceae families, *R. communis*, *J. curcas* and *Manihot esculenta*, revealed that 1529 genes in *J. curcas* (4% of the putative protein-encoding genes) are specific to the Euphorbiaceae family [23], suggesting that such Euphorbiaceae-specific genes, which cannot be annotated accurately based on the Arabidopsis datasets, are expressed in laticifers of *H. brasiliensis*. Latex, estimated to occur in more than 2000 plant species in 40 families [24], is suggested to have defensive roles against herbivorous insects; however, the precise roles and functions of various latex proteins and chemicals in the plant defense systems have not been well documented [25]. Therefore, the latex ESTs categorized in “unknown biological

Table 1

Summary of EST analysis of *H. brasiliensis* latex.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of clones sequenced</td>
<td>5760</td>
</tr>
<tr>
<td>Number of high-quality sequences</td>
<td>4325</td>
</tr>
<tr>
<td>Average length of high-quality ESTs (bp)</td>
<td>613 ± 179</td>
</tr>
<tr>
<td>Number of contigs (A)</td>
<td>386</td>
</tr>
<tr>
<td>Number of ESTs in contigs</td>
<td>3402 (79%)</td>
</tr>
<tr>
<td>Number of singletons (B)</td>
<td>923 (21%)</td>
</tr>
<tr>
<td>Number of UniESTs (A + B)</td>
<td>1309</td>
</tr>
</tbody>
</table>

Fig. 1. Distribution of gene ontology slim term assignments.
Table 2

List of unigenes with high redundancy. Unigenes are listed according to the EST redundancy as the number of ESTs included in each contig. In this table, the top 20 unigenes are listed. The sequences were searched against the nr (NCBI-nr) and Arabidopsis TAIR-9 (TAIR9-pep) databases using the blastx program, and the annotations and corresponding E-values for top hit genes are listed.

<table>
<thead>
<tr>
<th>Contig ID</th>
<th>Number of ESTs</th>
<th>NCBI-nr</th>
<th>Accession number</th>
<th>Definition</th>
<th>E-value</th>
<th>TAIR9-pep</th>
<th>AGI number</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>692</td>
<td>635</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.2E−70</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>691</td>
<td>378</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.1E−70</td>
<td>AT3G05500</td>
<td>1.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>690</td>
<td>148</td>
<td>A0CZ2355</td>
<td>Small Rubber Particle Protein (SRPP)</td>
<td>3.9E−97</td>
<td>AT3G05500</td>
<td>2.0E−48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>689</td>
<td>127</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.1E−70</td>
<td>AT3G05500</td>
<td>1.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>688</td>
<td>60</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.8E−70</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>687</td>
<td>52</td>
<td>AB834280</td>
<td>Unknown Protein 1 [Populus trichocarpa] (UP1)</td>
<td>1.2E−31</td>
<td>AT1G31812</td>
<td>1.0E−25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>686</td>
<td>46</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.0E−68</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>685</td>
<td>40</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>2.7E−70</td>
<td>AT3G05500</td>
<td>3.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>684</td>
<td>35</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.8E−70</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>683</td>
<td>34</td>
<td>AAP46156</td>
<td>Trypsin Inhibitor-like Protease Inhibitor Protein (PI)</td>
<td>8.9E−25</td>
<td>AT2G38870</td>
<td>8.0E−16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>682</td>
<td>33</td>
<td>CA793933</td>
<td>Unknown Protein 2 [Vitis vinifera] (UP2)</td>
<td>3.7E−30</td>
<td>AT3G46310</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>681</td>
<td>33</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>9.5E−69</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>680</td>
<td>31</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.0E−70</td>
<td>AT3G05500</td>
<td>1.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>679</td>
<td>31</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.1E−68</td>
<td>AT3G05500</td>
<td>3.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>678</td>
<td>29</td>
<td>AAP46156</td>
<td>Trypsin Inhibitor-like Protease Inhibitor Protein (PI)</td>
<td>1.2E−23</td>
<td>AT2G38870</td>
<td>1.0E−14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>677</td>
<td>28</td>
<td>AAP46156</td>
<td>Trypsin Inhibitor-like Protease Inhibitor Protein (PI)</td>
<td>9.8E−33</td>
<td>AT3G21070</td>
<td>7.0E−18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>676</td>
<td>28</td>
<td>AAP46155</td>
<td>Abscisic Acid-, Stress-, Ripening-induced (ASR)-like Protein</td>
<td>4.0E−08</td>
<td>AT4G15070</td>
<td>9.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>675</td>
<td>26</td>
<td>CA39880</td>
<td>Rubber Elongation Factor (REF)</td>
<td>1.3E−70</td>
<td>AT3G05500</td>
<td>2.0E−24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>674</td>
<td>25</td>
<td>AAP46156</td>
<td>Trypsin Inhibitor-like Protease Inhibitor Protein (PI)</td>
<td>5.8E−33</td>
<td>AT2G38870</td>
<td>5.0E−18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>673</td>
<td>24</td>
<td>AAP46156</td>
<td>Trypsin Inhibitor-like Protease Inhibitor Protein (PI)</td>
<td>5.5E−33</td>
<td>AT2G38870</td>
<td>6.0E−18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* EMBL Accession No.

processes” might correlate with the, as yet undiscovered, plant defense systems.

Based on the EST redundancy of each unigene, the top 20 high redundancy unigenes, having 25–635 ESTs, corresponded to ~40% of the total number of ESTs, indicating that a limited number of genes are expressed at exceptionally high levels in laticifers (Table 2), which correlates with pervious reports of latex ESTs of *H. brasiliensis* [7–9]. Among the highly expressed genes, three gene families were notable, i.e., *REF*, *SRPP*, and genes for protease inhibitor-like protein (PI). Some isoforms of *REF* and *SRPP* have been identified from latex [9]. In the top 20 high redundancy unigenes, 11 unigenes showing extremely high homology to *REF*, corresponding to 33% of the total ESTs, and one unigene of *SRPP*, corresponding to 3.4% of total ESTs, were included. *REF* is a well-studied latex abundant protein that is suggested to be related to the natural rubber biosynthesis in laticifers. *REF* is localized on the surface of rubber particles (RPs) [10,26,27], which are mainly composed of rubber molecules surrounded by a lipid monolayer and membrane proteins [28], and on which rubber biosynthesis takes place. Comparative analyses of different *H. brasiliensis* clones with contrasting latex yields demonstrated a positive correlation between the transcript levels of *REF* and latex yield [12]. In addition, inhibition of the rubber biosynthesis activity of latex in vitro by the addition of *REF*-specific IgG suggested the involvement of *REF* in natural rubber biosynthesis [10]. SRPP, which has a domain with high similarity to *REF* [11] and is localized on the RPs [26,27], is also a latex-abundant protein that is proposed to play a role in natural rubber biosynthesis, having a positive effect on the rubber biosynthesis activity of *Hevea* latex in vitro [11]. As reported in other latex-producing plant species, such as *Ficus carica* [29] and *Carica papaya* [30], putative serine protease inhibitors (trypsin inhibitors) were identified among the *Hevea* latex ESTs, showing relatively high redundancy (five unigenes in the top 20 high frequency unigenes, Table 2). Protease inhibitor proteins in laticiferous cells are thought to play defensive roles against insect herbivores, causing inhibition of their digestive processes [25]. The defensive roles of protease inhibitor proteins in latex of *H. brasiliensis* was suggested by a study that showed that the inhibitory effect of C-serum from routinely wounded trees on trypsin was higher than that from single wounded trees, indicating that the trypsin inhibitory activity in latex was induced by wounding [31]. In the top 20 high redundancy unigenes, two unigenes coding an unknown protein were included. Unknown protein 1 (UP1) had high similarities to an unknown function protein found in *P. trichocarpa* and other plant species, but unknown protein 2 (UP2) was a novel gene because there was no gene showing significant similarity to it in public databases.

By contrast, among the ESTs in this study, a limited number of ESTs were identified as representing genes for enzymes in the isoprenoid biosynthetic pathway, which is considered to be involved in natural rubber biosynthesis (Table 3), suggesting relatively low expression levels of these genes, despite of the exceptionally high accumulation levels of polyisoprenoids, i.e., natural rubber, in laticifers.

3.2. Identification of laticifer-specific highly expressed genes

Genes represented by unigenes that showed high EST redundancy were expected to be highly expressed in laticiferous cells. To identify laticifer-specific genes with high expression levels among the ESTs, the transcript levels of the unigenes with high redundancy, represented by *REF* [EST Clone No.: Hebr-OJ23], *SRPP* [EST Clone No.: Hebr-HL13], *PI* [EST Clone No.: Hebr-EL15], Unknown Protein 1 (UP1) [EST Clone No.: Hebr-ID22] and Unknown Protein 2 (UP2) [EST Clone No.: Hebr-BB02], in various plant tissues, such as leaves, green stems, roots and latex (laticifers), were analyzed using quantitative RT-PCR. All five genes analyzed were expressed predominantly in latex: the expression levels of each gene in latex were much higher than those in other tissues analyzed (Fig. 2). Since all tissues of *H. brasiliensis* analyzed in this study contained latex, which was impossible to be eliminated thoroughly, the transcripts in each tissue are considered to include those originated from laticifers in each tissue. The higher accumulation levels of latex in stems, compared with those in leaves and roots, might reflect relatively higher transcript levels of each gene in stems. Consequently, these five genes were identified as laticifer-specific highly expressed genes. To compare the expression levels of enzymes corresponding to the natural rubber biosynthesis with these laticifer-specific highly expressed genes, the transcript levels of two cPTs, *HRT1* and *HRT2* [4,32], were also analyzed (Fig. 2). Surprisingly, transcripts of *HRT1* were detected only in latex among the tissues analyzed, while transcripts of *HRT2* were detected in all tissues analyzed, the highest level in latex. Extraordinary
laticifer-specific expression of HRT1, despite inclusion of latex in all tissues analyzed, indicated that HRT1 could be considered as a trunk laticifer-specific gene, at least among the tissues analyzed, and that the regulation mechanism of gene expression of HRT1 is different from that of HRT2, despite the 92% nucleotide sequence identity of their coding regions. This suggested different physiological roles of these cPTs in laticifers of *H. brasiliensis*, concordant with the fact that HRT1 shows a lower functional complementation ability of a temperature-sensitive phenotype of rer2-2, the yeast mutant of cPT, than HRT2 although in vitro enzymatic function of recombinant proteins of HRT1 and HRT2 expressed in yeast are almost identical [32]. In terms of the transcriptional levels in latex, these cPTs showed 80-fold lower expression than REF, consistent with their low representations among the ESTs.

3.3. Cloning of the promoter regions of laticifer-specific expressed genes

To obtain laticifer-specific promoters, the 5' upstream regions of the coding sequences of REF, SRPP, PI and HRT1 were cloned using thermal asymmetric interlaced (TAIL) PCR or inverse PCR. As a results, 3199, 2129, 924, and 1087 bp of sequences for REF, SRPP, PI, and HRT1, respectively, were identified as putative promoter regions (Supplemental Figs. S1–S4). To identify cis-regulatory elements involved in laticifer-specific promoter activity, common motifs among the four promoter regions were explored using the POCO algorithm (http://ekhidna.biocenter.helsinki.fi/poco/poco/) [20]. As a result, seven common motifs consisting of eight bases were identified (Fig. 3, Supplemental Figs. S1–S4). Similarity search against a database of Plant Cis-acting Regulatory DNA Elements (PLACE) [21] revealed these motifs were similar to cis-regulatory elements involved in several biological phenomena in other plant species (Fig. 3).

Motif 4, “CTCCGTG”, commonly found in the 5’ upstream region of REF, SRPP and PI, resembles VSF1PVGRP18 (PLACE ID: S000249), a cis-acting regulatory element identified as a binding sequence of VFS-1, a bZIP transcription factor participate in plant vascular development by regulating vascular-specific gene expression [33]. Latex is the cytoplasm of highly specialized cells, laticifers, which are extremely long cells ramified at maturity throughout various tissues of the plant, such as the cortex, pith, foliar mesophyll, and especially the vascular tissues, phloem, and xylem rays [2,34]. Different expression of articulated laticifers in *H. brasiliensis* is coordinated with the development of other phloem tissues [2]; therefore, motif 4 is potentially involved in a laticifer-specific gene expression.

Common motifs 1, 2 and 3 were found in the promoter regions of four laticifer-specific genes. Motif 1, “AAAATTA”, bears a close resemblance to LECPLAC52 (PLACE ID: S000465), which is a cis-acting regulatory element of an elicitor-induced expression of a gene encoding 1-aminoacyclopropane-1-caboxyl acid synthase *Acs2*, a key enzyme in ethylene biosynthesis in *Solanum lycopersicum* [35]. Motif 3, “AAAAATGA”, was similar to the pathogen-responsive element GT1GSMC4 (PLACE ID: S000453), identified as a target sequence of a GT-1-like transcription factor involved in the pathogen- and salinity-induced expression of *ScAM-F* [36]. Some latex genes are potentially involved in pathogen responses, such as the response to herbivores [25], Indeed, the antifungal activity of *Hevein*, one of the laticifer-specific genes [37], against several fungi has been reported [38]. Therefore, these motifs commonly found in all four laticifer-specific genes may participate in the control of anti-pathogenic gene expressions in laticifers, concomitant with the upregulation of natural rubber biosynthesis involving in HRT1, REF and SRPP.

Motif 2, “ATTATTA”, has significant sequence similarity to ABA-responsive element ATHB6COREAT (PLACE ID: S000399) [39], identified as a target sequence of a homeodomain protein in *A. thaliana*, ATHB6, a regulator of the abscisic acid (ABA)-responsive gene expression. Many studies have shown the involvement of certain phytohormones in laticifer development or latex yields in *H. brasiliensis*, making them important factors in aspects of the natural rubber production. In *H. brasiliensis*, ethylene stimulates latex production. Bark treatment with ethephon (an ethylene releaser) increases the latex yield by 1.5–2-fold [40,41]. In addition, treatment of *H. brasiliensis* bark with ABA and auxin also affected latex yield positively, resulting from the sugar transporter-mediated turgor pressure change in laticifers [42]. In addition, in plant immunosystems, ABA acts agonistically on jasmonic acid (JA)-mediated signaling in response to herbivores, whereas it supresses salicylic acid (SA)- and ethylene-mediated immuno-signaling pathways [43]. Motif 2 may contribute to the ABA-mediated upregulation of natural rubber biosynthesis as one of the plant immunosystems specialized in laticifers. Similar to ABA, other common motifs, which have similarity with cis-acting regulatory elements identified to be involved in phytohormone responses, were found. Motif 6, “TGTTGCT”, has significant sequence similarity with a cis-acting element D3GMAUX28 (PLACE ID: S000330), identified in the promoter region of an auxin responsive gene, *GmAUX28* [44]. Motif 5, “CCAAATTA”, has high sequence similarity with JASE1ATOPR1 (PLACE ID: S000388), a cis-acting regulatory element identified as a positive regulator of a senescence associate gene in response to JA and senescence in *A. thaliana* [45]. In addition to the role as a signaling molecule in the plant defense system against necrotrophic pathogens and herbivorous insects, JA also plays an important role as an inducer of facultative cell differentiation in *H. brasiliensis* [46], implying a principal
role of latex as plant defensive agent and an involvement of phytohormone-mediated expression of laticifer-specific genes in laticifer differentiation.

Meanwhile, one putative cis-regulatory element involved in metabolic regulation was discovered. Motif 7, “GCTATCAA”, has sequence similarity with a cis-acting regulatory element TATCCAOSAMY (PLACE ID: S000403, TATCCA element) and functions in sugar starvation-induced gene expression of α-amylases in rice, which is regulated by MYB transcription factors with one DNA-binding repeat [47,48]. It has been suggested [49] that the TATCCA element could be a common cis-acting regulatory element functioning during sugar starvation, because some variants of the cis-acting elements are found not only in the 5′-upstream regions of α-amylases in monocots, but also in those of other sugar starvation-induced genes, such as the cucumber malate synthase and isocitrate lyase genes, and the maize sucrose synthase gene. Although a correlation between the laticifer-specific gene expression and sugar starvation-induced responses has not been uncovered, motif 7 “GCTATCAA” might participate in metabolic regulation in laticifers, involving MYB transcription factor-mediated gene expression.

3.4. Identification of laticifer-specific transcription factors

In the EST analysis, genes that have similarities to transcription factors were identified (Table 4, Supplemental Table S3 and S2). To identify potential transcription factors that could regulate the expressions of laticifer-specific genes in H. brasiliensis, transcript levels of the putative transcription factors with high EST redundancy were analyzed using total RNAs from various tissues as templates, by means of quantitative RT-PCR (Fig. 4). Notably, the redundancies of the putative transcription factors were lower than those of the laticifer-specific highly expressed genes. All of the putative transcription factors analyzed were revealed to be expressed in latex predominantly. Among them, a unigene represented by

Fig. 4. Tissue specific transcript levels of putative transcription factors. Detailed annotations of putative transcription factors, indicated as representative EST IDs, were summarized in Table 4. The transcript levels of genes of interest (GOI) were determined as the copy number in an amount of total RNA extracted from latex, leaves, and stems, and then normalized by the level of the 18S rRNA in corresponding tissues. The results are presented as the means of three independent determinations ±SDs.

Table 4

Putative transcription factors analyzed in this study. Putative transcription factors were annotated by searching against the Arabidopsis TAIR-9 databases with the blastx program. The family types of the TFs were determined using A. thaliana transcription factor database (AtTFDB) provided by the Arabidopsis Gene Regulatory Information Server (AGRIS). In this table, only putative transcription factor genes subjected to quantitative RT-PCR analysis are listed.

<table>
<thead>
<tr>
<th>Representative EST ID (contig ID)</th>
<th>Number of ESTs</th>
<th>Transcription Factor Family</th>
<th>Arabidopsis protein name</th>
<th>AGI number</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ14 (Contig554)</td>
<td>6</td>
<td>MYB Family</td>
<td>AtMYB17</td>
<td>AT3G61250</td>
<td>8.0E-12</td>
</tr>
<tr>
<td>IN24 (Contig430)</td>
<td>2</td>
<td>Zinc Finger, LSD1-type</td>
<td>LOL2</td>
<td>AT4G21610</td>
<td>1.0E-30</td>
</tr>
<tr>
<td>MB14 (Contig402)</td>
<td>2</td>
<td>bHLH Family</td>
<td>AtbHLH6</td>
<td>AT1G32640</td>
<td>2.0E-29</td>
</tr>
<tr>
<td>AP07 (Contig325)</td>
<td>2</td>
<td>bZIP Family</td>
<td>AtbZIP60</td>
<td>AT1G2990</td>
<td>7.0E-39</td>
</tr>
<tr>
<td>NG02</td>
<td>1</td>
<td>Alfin-like Family</td>
<td>PHD Finger Family Protein</td>
<td>AT5G05610</td>
<td>3.0E-34</td>
</tr>
<tr>
<td>IA23</td>
<td>1</td>
<td>NAC Family</td>
<td>NAC Domain-containing Protein</td>
<td>AT1G73230</td>
<td>5.0E-63</td>
</tr>
<tr>
<td>GF14</td>
<td>1</td>
<td>NAC Family</td>
<td>NAC Domain-containing Protein</td>
<td>AT1G73230</td>
<td>9.0E-56</td>
</tr>
</tbody>
</table>
Acknowledgement

This work was supported in part by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 22780298 to S.T.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plantsci.2014.05.003.

References

[29] J.S. Kim, Y.O. Kim, H.J. Ryu, Y.S. Kwak, J.Y. Lee, H. Kang, Isolation of stress-related genes of rubber particles and latex in fig tree (Ficus carica) and their expressions


