JOURNAL OF AGROBIOTECHNOLOGY 2022, VOL 13(1):61-76
e-ISSN: 2180-1983
http://dx.doi.org/10.37231/jab.2022.13.1.290
https://journal.unisza.edu.my/agrobiotechnology/index.php/agrobiotechnology/index

Comparative Transcriptome Analysis Identifies Potentially Relevant Genes in Rubber Clones with a High Latex Yield (*Hevea brasiliensis*)

Mohd Fahmi Abu Bakar [a]* and Ahmad Sofiman Othman[b]

[a] Faculty of Bioresources & Food Industry, Universiti Sultan Zainal Abidin, Besut Campus, 22200 Besut, Terengganu, Malaysia
[b] School of Biological Sciences, Universiti Sains Malaysia, 11700 Minden, Pulau Pinang, Malaysia

***Corresponding author: mohdfahmi@unisza.edu.my**

Received: 28/03/2022, Accepted: 02/06/2022, Available Online: 27/06/2022

ABSTRACT

Since the 1950s, the rubber tree (*Hevea brasiliensis*) has significantly contributed to Malaysia's agricultural economy due to its capability to produce high-value natural rubber. Due to the presence of cis-1,4-polyisoprene, natural rubber exhibits unique properties such as flexibility, high elasticity, and efficient heat dispersion. Cis-1,4-polyisoprene is synthesized via two distinct metabolic pathways: the isoprenoid and rubber biosynthesis. RNA sequencing was performed on 12 specimens from the bark, leaf, and latex tissues of two rubber clones, RRIM 3001 (high latex yield) and RRIM 712 (low latex yield). After quality assessment, these specimens generated a mean of 73,816,244 clean reads. The differential expression analysis revealed that the Acetyl-CoA-Acetyltransferase (*AACT*) and Diphosphomevalonate Decarboxylase (*MVD*) genes (via the MVA pathway), as well as the Small Rubber Particle Protein (SRPP) and Rubber Elongation Factor (REF) genes (via the rubber biosynthesis pathway), were overexpressed in the bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. The transcription factor Ethylene Response Factor (ERF) was also found to be more highly expressed in RRIM 3001 than in RRIM 712. Five rubber clones with random yield characteristics were used in quantitative amplification analysis to validate *AACT*, *MVD*, *SRPP*, *REF,* and *ERF* genes. The quantitative analysis shows that the five genes were consistently expressed with the yield characteristics of the five random rubber clones.

Keywords: *Hevea brasiliensis*, transcriptome, cis-1,4-polyisoprene, isoprenoid biosynthesis, rubber biosynthesis

INTRODUCTION

Hevea brasiliensis, more commonly referred to as the rubber tree, is a perennial, cross-pollinating, monoecious plant in the Euphorbiaceae family (Saha & Priyadarshan, 2012). It is found only in the Amazon rainforest in northern South America, specifically in Brazil, Venezuela, Colombia, Peru, and Bolivia (Mantello et al., 2014). Within the genus Hevea, only *H. brasiliensis* has been exploited as a high-value natural rubber producer, contributing significantly to Malaysia's agricultural economy (Rozhan, 2015).

Natural rubber is composed of cis-isoprene units, a byproduct of the synthesis of isopentenyl diphosphate (IPP) (Lau et al. 2016). Between 5,000 and 10,000 isoprene units are organized in an unbranched chain to form cis-

isopropene (Liu et al., 2016). IPP is synthesized in two distinct ways: through the cytoplasmic Mevalonate (MVA) pathway and the plastidic 2-C-Methyl-D-Erythritol 4-Phosphate (MEP) pathway (Wang et al. 2009). Additionally, Transcription Factors (TFs) and sucrose transport-related genes have been implicated in natural rubber biosynthesis.

Recent advances in high-throughput sequencing have increased the genetic resources available for *H. brasiliensis* (Rahman et al., 2013; Lau et al., 2016). Extracting genetic data from *H. brasiliensis* tissues using RNA-seq is critical for obtaining information at the chromosome and genome level and gaining insight into latex production traits. RNA-seq platforms enable studies of the transcriptome's or specific genes' expression levels, differential splicing, and allele-specific expression, as well as accurate transcriptome profiling that closely mimics a cell's biological processes. Numerous RNA-seq studies using rubber tissues as specimens have been conducted, for example, Li et al. (2015), Li et al. (2016), Montoro et al. (2018), Abdul Rahman et al. (2019), and Roy et al. (2019).

This study aimed to determine the related genes for natural rubber biosynthesis, the expression and differential expression of genes involved in isoprenoid biosynthesis, the rubber biosynthesis pathway, sucrose transport, and transcription factors in high and low yield rubber clones using the RNA-seq transcriptome database. The primary objective was to identify the genes that can contribute to the high production of natural rubber through their expression. This study compared gene expression between a low yield (RRIM 712) and high yield (RRIM 3001) rubber clones. The findings may be valuable for selecting offspring with desirable traits in breeding programs.

MATERIALS AND METHODS

Plant Materials

This study examines three tissues: bark, leaves, and latex. The specimens were collected from a ten-year-old rubber tree (*H. brasiliensis* clone; RRIM 712 and RRIM 3001) at RISDA Semaian in Bukit Perak, Pendang, Kedah, Malaysia. RRIM 712 a low-latex yielding rubber clone (1,516 kg/ha/yr), whereas RRIM 3001 is a high-latex yielding rubber clone (3,000 kg/ha/yr). RRIM 712 (43 cm) has a smaller circumference than RRIM 3001 (51 cm). The samples collected were immediately frozen in liquid nitrogen and transported to the laboratory on dry ice.

Total RNA Isolation

Total RNA was isolated from bark, leaves, and latex using the Qiagen RNAeasy Plant Minikit protocol (Qiagen Inc., Chatsworth, CA). We determined the quality and integrity of RNA using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

cDNA Library Preparation and Sequencing

According to the manufacturer's protocol, the Illumina NextSeq 500 v2 Kit was used to generate the paired-end Illumina mRNA libraries from 1 ug of total RNA. Each sample was sequenced in multiple HiSeq2000 lanes using the Illumina NextSeq 500 v2 Cycle Kit to generate two 75-bp reads.

Sequence Pre-processing

The raw reads generated from RNA sequencing in FASTQ format (obtained from Illumina platforms) were analyzed using the publicly available FastQC software (Andrew, 2015). The initial filtering steps included the elimination of adaptor reads and nucleotide ambiguity. Additionally, low-quality 3'-end reads were trimmed to eliminate low-quality reads. FastQC generates error-free reads, which were used for further analysis.

Transcriptome Mapping and Assembly

Bowtie2 was then used to map the clean reads to the draught *H. brasiliensis* genome (Lau et al. 2016) (accession: PRJDB4387) using the DDBJ/EMBL/GenBank BioProject database and the TopHat software version 2.2.1 (Trapnell et al. 2010). The reads were then mapped to the genome using default mapping parameters. The

mapped reads were then assembled using default parameters in Cufflink v2.2.1 (Trapnell et al. 2010). To ensure the identification of isoforms with a high degree of confidence, only assembled transcripts with a fragment per kilobase of exon per million reads (FPKM) value greater than 0 were retained for further analysis. The assembled transcript sequence generated by Cufflink was then extracted from the rubber genome sequence. The unmapped reads were extracted from the unmapped BAM file. De novo assembly of the reference reads was performed using Trinity v2.6.5 (Grabherr et al. 2011) with default parameters to generate the reference transcriptome. Using the TopHat software version 2.2.1, unmapped reads were mapped to the reference transcriptome (Trapnell et al. 2010). The assembled sequences from Cufflink's genome mapping and TopHat2's reference-transcriptome mapping were combined. Finally, CD-HIT software combined assembled sequences from both sources and reduced sequence redundancy (Huang et al. 2010). The output transcripts were used as reference sequences for *H. brasiliensis*.

Identification of Genes Related to Natural Rubber Biosynthesis

The reference sequence for *H. brasiliensis* was annotated using multiple databases, including the NCBI NR (Non-Redundant) database and the UniProtKB/Swiss-Prot (Universal Protein Resources) database, the GO (Gene Ontology) database, and the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Blast2GO software (Conesa et al. 2005) was used to conduct the searches, configured with an e-value cut-off of 1e-10. Each transcript was assigned the description and gene names of the top BLAST hit.

Digital Gene-Expression Analysis

Each tissue sample's expression profile was calculated using the relative abundance of transcripts. Bowtie2 (Langmead & Salzberg, 2012) was used to align clean reads from each sample to the assembled reference, followed by quantification of gene expression in fragments per kilobase of exon per million reads mapped (FPKM) (Abdul Rahman et al., 2019). The EdgeR software (Robinson et al. 2010) was used to normalize each gene's expression level for each tissue sample and identify differentially expressed genes using pairwise comparisons. The significance level for DEG values was set to log2FC 1 and p 0.05.

Quantification of Expressed Genes

Six genes were subjected to relative quantification or gene expression analysis using a Real-Time PCR System, including the housekeeping gene (Biorad, USA). The Real-Time PCR reaction mixtures contained 5 l of SENSIFAST SYBR® Green Master Mix, 0.5 l of each 10 M primer, and 1 l of 20 ng cDNA in a ten-liter volume. The PCR cycling conditions included a 15-second denaturation cycle at 95 °C, 35 15-second amplification cycles between 53.3 and 56.5 °C, and a 20-second extension cycle at 72 °C. The standard curve was constructed using a tenfold dilution series of eight points from cDNA samples in three technical replicates. The housekeeping gene (β-actin) was used to normalize each sample's relative quantification as an internal reference gene. Livak and Schmittgen compared the relative transcript abundance data for each gene and sample using the 2-Ct method (2001). The qPCR data were then analyzed to determine the expression level of each targeted gene in different individuals using CT or CQ. The CQ data generated by the BIORAD system's qPCR were derived from the emission of a fluorescent reporter dye (SYBR Green). The software automatically calculated the CQ value, and the results were then analyzed for statistically significant differences.

Statistical Data Analysis

The correlation between biological replicates and the transcriptome dataset was determined using the correlation coefficient (R^2). Due to the small sample size and the presence of a single variable, a one-way ANOVA was chosen to determine the probability level (p 0.05) using Microsoft Excel 2013.

RESULTS AND DISCUSSION

Transcriptome Assembly and Analysis

The high-throughput sequencing generated raw 75,742,501 paired-end reads from the library, while the pre-processing step recovered a total of 73,816,244 (97%) paired-end high-quality reads. The clean reads were then mapped to the *H. brasiliensis* genome (accession: PRJDB4387). A total of 101,269 transcripts have been

generated, corresponding to 48,281 unique genes between 424 and 10,503 bp in length. The mean length of the transcript was 1,046 bp, Of the 101,269 transcripts, 50,375 (49.7%) ranged from 1 to 2 kb, and 25,940 (25.6%) were longer than 2 kb. A summary of the transcriptome assembly is shown in Table 1.

Table 1. Summary of the transcriptome assembly of *H. brasiliensis*

Summary	Number of Transcripts
Number of transcripts	101,269
Total raw read count	15,862,548
Total clean read count	14,931,274
Mean transcript length	1,046
Maximum transcript length	10,503
Minimum transcript length	424
N50 length	1,313
GC% content	49.8%

Functional Annotation

The annotations analysis showed that of the 101,269 transcripts, 83,771 (83%) showed significant matches in the NCBI NR database, 61,917 (61%) hits in the UniprotKB/Swiss-Prot database, 71,902 (71%) hits in the GO annotations, and 18,158 (18%) hits in KEGG annotations. Moreover, 38,411 transcripts (38%) matched one of the protein domains in the Pfam database. On the other hand, the top species with the highest number of hits in BLASTx analysis against the NR database is *Manihot esculenta* with 58,881 hits, *Jatropha curcas* with 8,229, *Ricinus communis* with 4,767 hits, *Hevea brasiliensis* had 1,836 hits and *Populus trichocarpa* had 752 hits. Table 2 shows the summary of functional annotations of the transcripts.

Table 2. Summary of the functional annotations of *H. brasiliensis* transcript

Databases	Hits	Hits percentage
NCBI non-redundant protein (NR)	83,771	83%
UniprotKB/Swiss-Prot	61,917	61%
GO	71,902	71%
KEGG	18,158	18%
Pfam	38,411	38%

Analysis of DEGs Among Different Tissues of RRIM 3001 and RRIM 712

The differential expression analysis was determined by comparing the gene expression level of bark, leaf, and latex tissues between RRIM 3001 and RRIM 712; at thresholds of $|\log2\ (FC)| \geq 1$ and $p \leq 0.05$. The analysis showed that both clones produced a total number of 55,594 (bark), 48,446 (leaf), and 55,561 (latex) expressed transcripts. Based on the analysis, the number of upregulated transcripts within RRIM 3001 was higher than that of downregulated transcripts within bark, leaf, and latex tissues compared to RRIM 712. Most upregulated transcripts are associated with defense responses, photosynthesis, and rubber biosynthesis, while most downregulated transcripts are involved in ATP formation processes such as glycolysis and transcription. Figure 1 shows the number of differential expressions of genes in bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712.

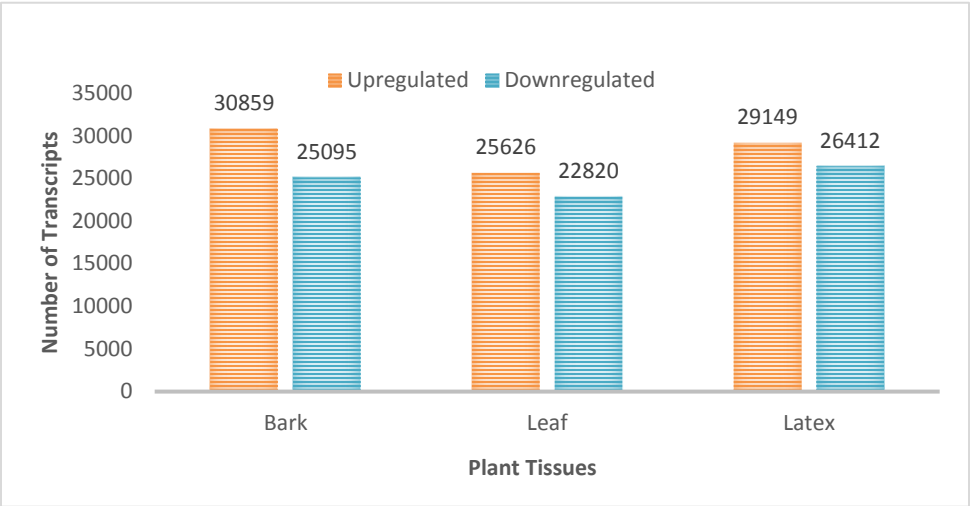

Fig. 1. The number of upregulated and downregulated transcripts in bark, leaf, and latex tissues of RRIM 3001 compared with RRIM 712.

Identification of Genes Related to Latex Biosynthesis

Latex biosynthesis is related to the production of isopentenyl pyrophosphate (IPP). IPP biosynthesis proceeds via two distinct routes: the MVA pathway and the MEP pathway. There is a total of six genes encoding six enzymes responsible for IPP production through the MVA pathway, including Acetyl-CoA-Acetyltransferase (*AACT*) Hydroxymethylglutaryl-CoA Synthase (*HMGS*), Hydroxymethylglutaryl-CoA Reductase (*HMGR*), Mevalonate Kinase (*MVK*), Phosphomevalonate Kinase (*PMK*) and Diphosphomevalonate Decarboxylase (*MVD*). On the other hand, seven genes are responsible for the IPP production through the MEP pathway, including 1-Deoxy-D-Xylulose-5-Phosphate Synthase (*DXS*), 1-Deoxy-D-Xylulose-5-Phosphate Reductoisomerase (*DXR*), 2-C-Methyl-D-Erythritol 4-Phosphate Cytidylyltransferase (*MCT*), 4-(Cytidine 5'-Diphospho)-2-C-Methyl-D-Erythritol Kinase (*CMK*), 2-C-Methyl-D-Erythritol 2,4-Cyclodiphosphate Synthase (*MDS*), 4-Hydroxy-3-Methylbut-2-en-1-yl Diphosphate Synthase (*HDS*) and 4-Hydroxy-3-Methylbut-2-enyl Diphosphate Reductase (*HDR*).

The IPP produced earlier requires polymerization through the Rubber Biosynthesis (RB) pathway to produce *cis*-1,4-polyisoprene. A total of three genes are responsible for producing *cis*-1,4-polyisoprene through the RB pathway, including *cis*-Prenyltransferase (*CPT*), Small Rubber Particle Protein (*SRPP*), and Rubber Elongation Factor (*REF*). Moreover, sucrose products are a precursor for IPP production via MVA (acetyl-CoA) and MEP (glyceraldehyde-3-phosphate) pathways (Chow et al. 2007). Since laticifer cells are a strong sink for sucrose, sucrose needs to be supplied continuously for latex production and influences latex biosynthesis. Here we identify two sucrose-related genes in *Hevea* tissues: Sucrose Transport (*SUT*) and Sucrose Synthase (*SuSy*). Table 3 shows the information on genes related to latex biosynthesis.

Table 3. The information on genes related to latex biosynthesis

	Enzyme Symbol	Enzyme Commission Number
MVA Pathway		
Acetyl-CoA-Acetyltransferase	*AACT*	[EC:2.3.1.9]
Hydroxymethylglutaryl-CoA Synthase	HMGS	[EC: 2.3.1.10]
Hydroxymethylglutaryl-CoA Reductase	HMGR	[EC:1.1.1.88]
Mevalonate Kinase	MVK	[EC:2.7.1.36]
Phosphomevalonate Kinase	PMK	[EC:2.7.4.2]
Diphosphomevalonate Decarboxylase	*MVD*	[EC:4.1.1.33]
MEP Pathway		
1-Deoxy-D-Xylulose-5-Phosphate Synthase	DXS	[EC:2.2.1.7]
1-Deoxy-D-Xylulose-5-Phosphate Reductoisomerase	DXR	[EC:1.1.1.267]
2-C-Methyl-D-Erythritol 4-Phosphate Cytidylyltransferase	MCT	[EC:2.7.7.60]
4-(Cytidine 5'-Diphospho)-2-C-Methyl-D-Erythritol Kinase	CMK	[EC:2.7.1.148]
2-C-Methyl-D-Erythritol 2,4-Cyclodiphosphate Synthase	MDS	[EC:4.6.1.12]
4-Hydroxy-3-Methylbut-2-en-1-yl Diphosphate Synthase	HDS	[EC:1.17.7.3]
4-Hydroxy-3-Methylbut-2-enyl Diphosphate Reductase	HDR	[EC:1.17.7.4]
RB Pathway		
cis-prenyltransferase	CPT	[EC:2.5.1.20]
Small Rubber Particle Protein	SRPP	N/A
Rubber Elongation Factor	REF	N/A
Sucrose Related Genes		
Sucrose Transport Protein	SUT	N/A
Sucrose synthase	SuSy	[EC:2.4.1.13]

Differential Expression Analysis of Genes Related to the MVA Pathway

The differential expression analysis shows that within RRIM 3001, the *AACT* and *MVD* genes have a high number of upregulated transcripts compared with RRIM 712. The differential expression of transcripts derived from the *AACT* gene showed that the highest fold-change values were 7.81 (bark tissue), 7.61 (leaf tissue), and 8.04 (latex tissue) in RRIM 3001 compared with RRIM 712. In addition, the differential expression of transcripts represented by the *MVD* gene showed fold-change values of 13.82 (bark tissue), 8.72 (leaf tissue), and 14.62 (latex tissue) in RRIM 3001 compared with RRIM 712 (Fig. 2). In searching for important genes responsible for higher latex production, the fold change must be at least five times to be accepted (Makita et al., 2017). A fold change higher than five showed that transcripts represented by the *AACT* and *MVD* genes are highly expressed within all tissues of RRIM 3001 compared with RRIM 712. Figure 2 shows the relative mRNA levels of genes involved in the MVA pathway within bark, leaf, and latex of RRIM 3001.

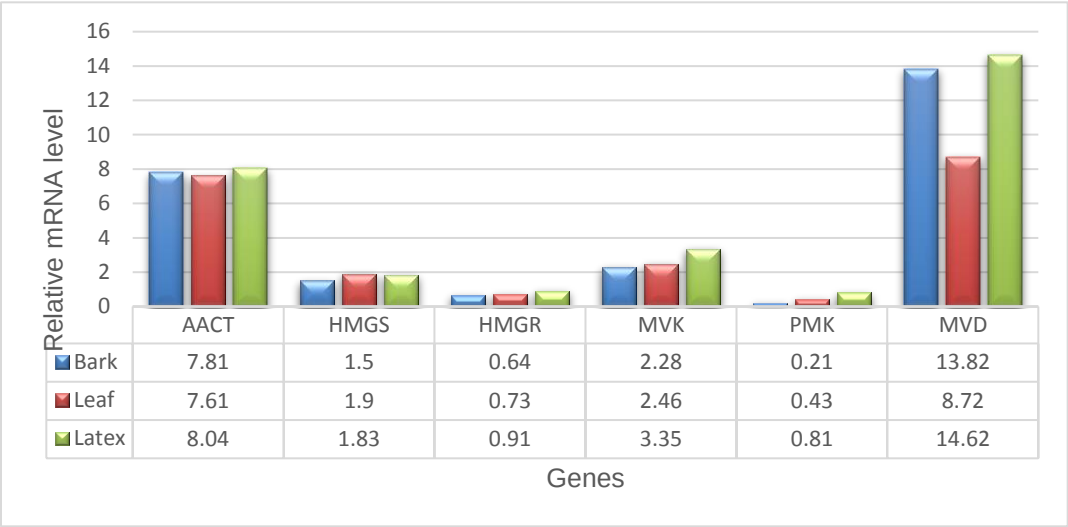

Fig. 2. The mRNA level of six genes expression related to the MVA pathway within bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. RRIM 3001 has high expression of *AACT*, and *MVD* genes within bark, leaf and latex tissues compared to RRIM 712.

Differential Expression Analysis of Genes Related to the MEP Pathway

The differential expression analysis of transcripts represented by genes related to the MEP pathway of RRIM 3001 compared with RRIM 712 shows different results compared with the MVA pathway. According to the analysis, latex tissues show higher fold-change values than bark and leaf tissues. However, the fold-change value of all transcripts represented by genes related to the MEP pathway within RRIM 3001 tissues is not as high as in the MVA pathway. The highest fold-change value was the transcript represented by the *DXS* gene within latex tissue, which expressed 3.92 times compared to RRIM 712, followed by the *HDR* gene with 2.33 times (Fig. 3). The low fold-change value shows that the expression genes of latex biosynthesis within RRIM 3001 tissue from genes related to the MEP pathway has almost similar expression to RRIM 712 tissues. Figure 3 shows relative mRNA levels of genes involved in the MEP pathway within bark, leaf, and latex of RRIM 3001.

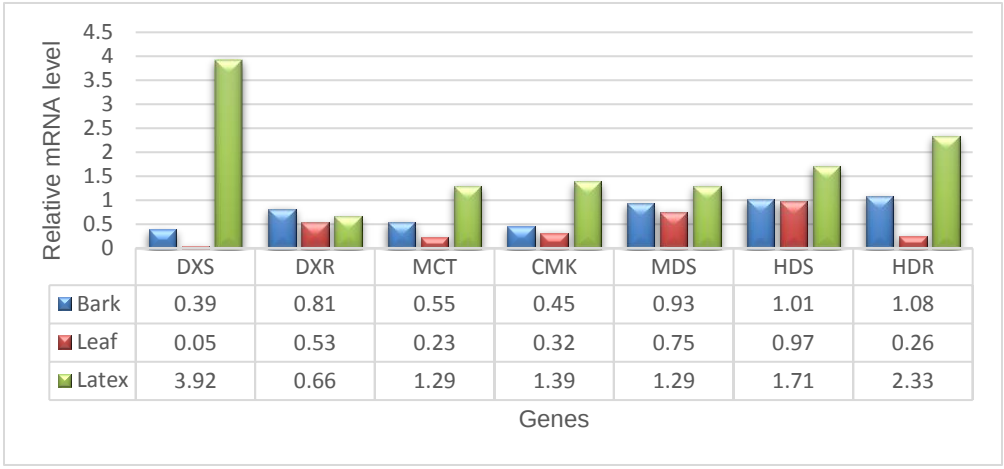

Fig. 3. The mRNA level of seven genes expression related to MEP pathway within bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. RRIM 3001 has high expression of *DXS* gene in latex tissue compared to RRIM 712.

Differential Expression Analysis of Genes Related to Rubber Biosynthesis Pathway

The differential expression analysis shows that genes related to RB biosynthesis within latex tissue have a higher expression compared with bark and leaf tissues of RRIM 3001. Within latex tissue of RRIM 3001, transcripts represented by *CPT*, *SRPP,* and *REF* genes have expression values 1.99, 7.77, and 10.81 times higher than latex tissue of RRIM 712 (Fig. 4). Both *SRPP* and *REF* genes were highly expressed where the fold-change value was higher than five times the expression value in RRIM 3001 compared with RRIM 712. Figure 4 shows relative mRNA levels of genes involved in the RB pathway within bark, leaf, and latex in RRIM 3001.

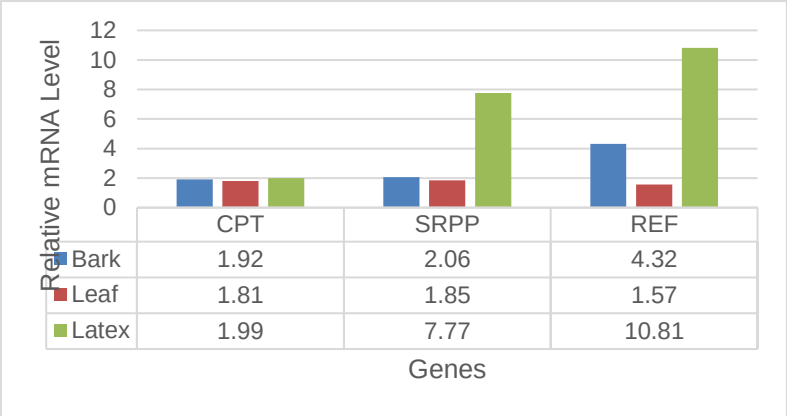

Fig. 4. The mRNA level of three genes expression related to rubber biosynthesis pathway within bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. RRIM 3001 has high expression of *SRPP* and *REF* genes within latex tissue compared to RRIM 712.

Differential Expression Analysis of Sucrose Related Genes

The differential expression analysis showed that both *SUT* and *SuSy* genes were expressed moderately within all tissues of RRIM 3001 compared with RRIM 712 (Fig. 5). *SuSy* gene expressed higher than *SUT* gene within bark tissue but lowered within leaf and latex tissues. The differential expression analysis of the *SUT* gene showed that the upregulated transcripts with the highest fold-change value were 2.64 (bark tissue), 2.99 (leaf tissue), and 1.68 (latex tissue). The differential expression analysis of the *SuSy* gene showed that the upregulated transcripts with the highest fold-change value were 4.52 (bark tissue), 0.15 (leaf tissue), and 0.23 (bark tissue) compared with RRIM 712. Figure 5 shows the relative mRNA level of genes involved in sucrose transportation within the bark, leaf, and latex tissues of RRIM 3001.

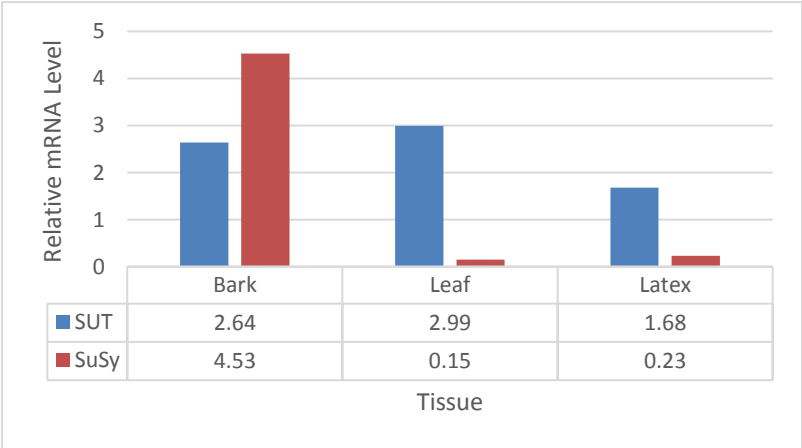

Fig. 5. The mRNA level of two genes expression related to sucrose synthesis and transportation genes within bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. RRIM 3001 have almost similar expressions in bark, leaf and latex tissues with RRIM 712.

Investigation of Transcription Factors Involved in Natural Rubber Production

Many studies have suggested that Transcription Factors (TFs) are essential in biological pathways, including natural rubber production. Since TFs are related to the gene expression process, differential expression analysis was conducted to compare the expression value of TFs between RRIM 3001 and RRIM 712 tissues. The analysis of the transcriptome of tissues within RRIM 3001 showed that most of the transcripts expressed represented basic Helix-Loop-Helix (bHLH) followed by WRKY and Ethylene Responsive Factor (ERF). On the other hand, the analysis of the transcriptome of tissues within RRIM 712 showed that most of the transcripts expressed represented bHLH followed by WRKY, Nuclear Transcription Factor Subunit Y (NF-Y) (bark and latex tissues), and ERF (leaf tissue). However, only the ERF was highly expressed with relative mRNA values of more than five times compared with others within bark, leaf, and latex tissues of RRIM 3001. Within RRIM 3001 tissues, ERF is expressed at 5.79 (bark), 5.51 (leaf), and 5.94 (latex) times higher than other TFs. Figure 6 shows the relative mRNA levels of genes involved in transcription factors within bark, leaf, and latex of RRIM 3001.

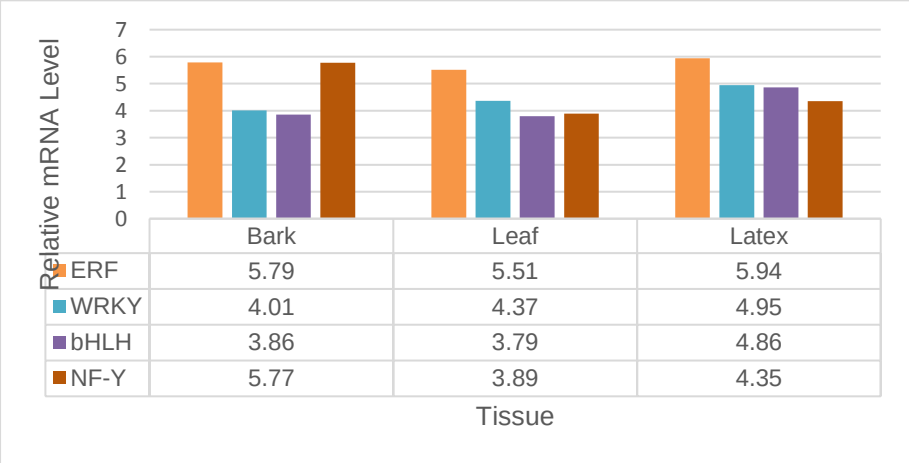

Fig. 6. The mRNA level of genes expression related to transcription factors within bark, leaf, and latex tissues of RRIM 3001 compared to RRIM 712. RRIM 3001 has high expression on ERF within bark, leaf and latex tissue compared to RRIM 712.

Relative Quantification Analysis

The relative quantification for target gene expression was calculated based on the ΔΔCt method described by Livak and Schmittgen (2001). Five rubber clones with random natural rubber yield are used to quantify the expression value of selected genes, including RRIM 3001, RRIM 2025, RRIM 712, PB 350, and PB 260. Using RRIM 712 as a control, the rubber clone had upregulated CT on selected genes when the fold-change value ratio was > 1.0 but downregulated CT on selected genes when the fold-change value ratio was < 1.0 CT. The significance of the upregulated or downregulated dataset was adopted by the ratio of the relative transcript abundance with a p-value ≤ 0.05. The analysis shows that RRIM 3001 has the highest expression within all selected genes compared with other rubber clones, followed by PB 350. On the other hand, the highest expression value can be found within REF, which is 12.98 times higher in RRIM 3001 compared with RRIM 712, followed by *MVD* (12.01 times) and *AACT* (8.78 times). Figure 7 shows the relative gene expression of five genes from four different clones that were differentially expressed compared with RRIM 712.

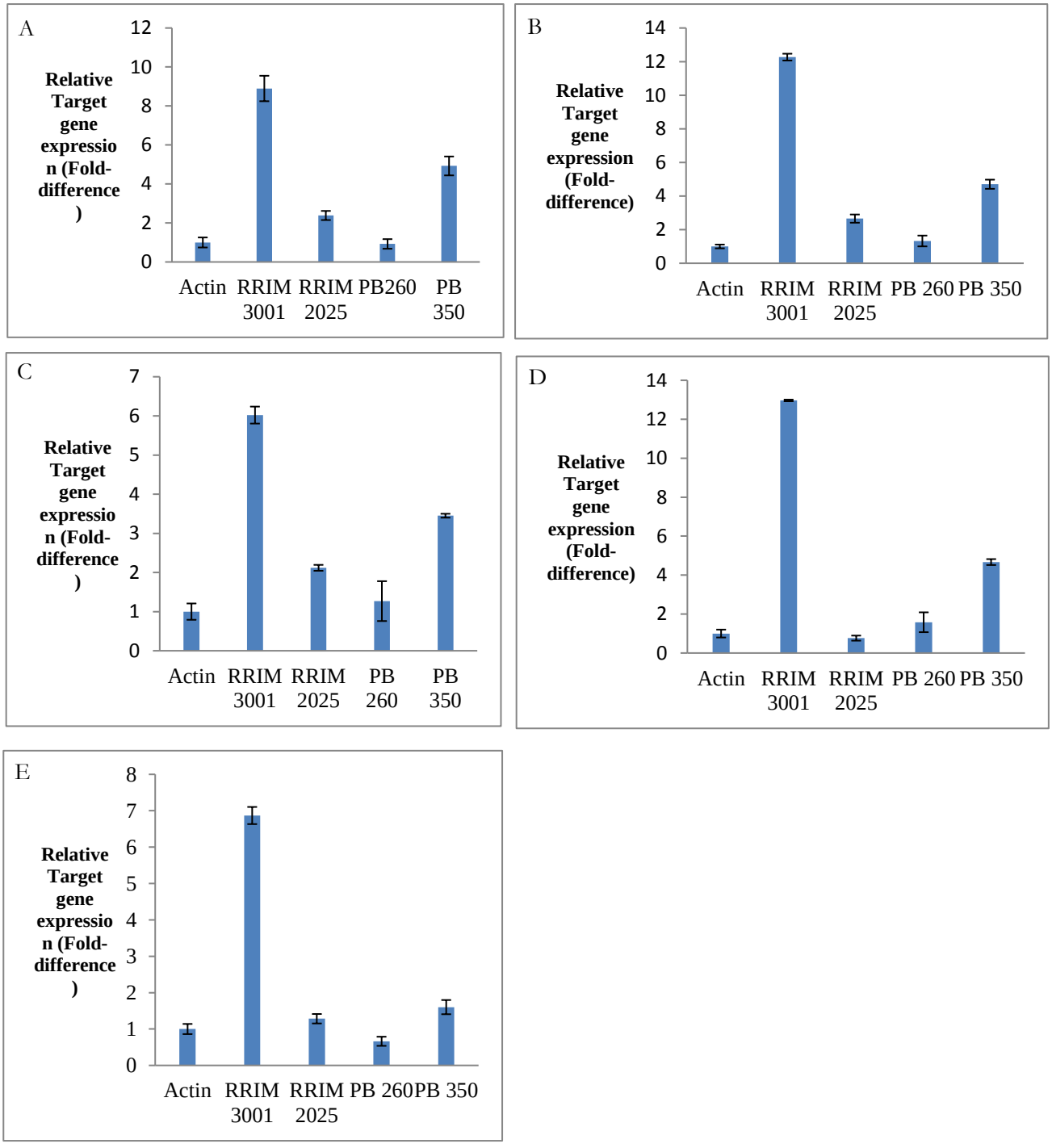

Fig. 7. The relative gene expression of five genes from four different clones was differentially expressed compared with RRIM 712. A: *AACT* (Acetyl-CoA-Acetyltransferase); B: *MVD* (Diphosphomevalonate Decarboxylase); C: *SRPP* (Small Rubber Particle Protein); D: *REF* (Rubber Elongation Factor); E: *ERF* (Ethylene Response Factor). RRIM 3001 consistently have highest expression in all five selected genes compared to other rubber clones.

Discussion

The high expression of MVA and MEP pathway genes in the bark, leaf, and latex tissues of the high yield rubber clone (RRIM 3001) suggested that increased activity could contribute to increased latex output. Gene expression is frequently higher in high latex-yielding rubber clones than in low latex-yielding rubber clones. Due to the increased expression of genes linked to the MVA pathway compared to the MEP pathway, expression analysis of genes related to the MVA and MEP routes revealed that the MVA pathway is the main route for IPP generation in Hevea. Previous research showed similarities with this finding. Genes involved in the MVA pathway were expressed more frequently than genes involved in the MEP pathway (Tang et al., 2013). IPP production in Hevea also has been found mainly through the MVA pathway (Makita et al., 2017). In addition, most MVA pathway genes were more strongly expressed than MEP pathway genes, with the *AACT* gene being the most highly expressed (Tan et al., 2017).

Two genes related to the MVA pathway were highly expressed: *AACT* and *MVD* genes. In RRIM 3001 compared to RRIM 712, transcripts encoded by the *AACT* gene had an expression value between 7.61 to 8.04 times greater. Tan et al. (2017) found that all genes participating in the MVA route have higher expression than those engaged in the MEP pathway, with the *AACT* gene being the most highly expressed. Chao et al. (2015) found that the MVA pathway's *AACT* gene was substantially expressed within the CATAS8-78 clone. The results shows that transcripts encoded by the *MVD* gene had expression levels 8.72 to 14.62 times higher in RRIM 3001 bark, leaf, and latex tissues than in RRIM 712. Sando et al. (2008) discovered that Hb*MVD* was highly expressed in latex tissue and performed a crucial role in rubber biosynthesis, and their findings are similar. Wu et al. (2017) found that the *MVD* gene was strongly expressed in the rubber manufacturing process in the rubber tree's bark, latex, leaf tissues, and female and male flowers. The comparison between high latex-yielding rubber clones and low latex-yielding rubber clones showed that the MVD gene was highly expressed (Ambily et al., 2018).

Rubber particles must be polymerized after IPP manufacture via isoprenoid biosynthesis. Rubber particles play a vital role in rubber production and are responsible for the rubber's high molecular weight, which defines its quality (Cherian et al., 2019). Rubber polymerization process is included in the rubber biosynthetic pathway. Rubber polymerization is catalyzed by many enzymes, including those encoded by CPT, SRPP, and REF, according to Akhtar et al. (2017). SRPP and REF were found to be highly expressed inside RRIM 3001. When RRIM 3001 latex tissue was compared to RRIM 712 latex tissue, transcripts representing the SRPP and REF genes were expressed 7.77 and 10.81 times greater, respectively. Several earlier investigations also have similar findings. REF and SRPP were shown to be highly expressed in the latex transcriptome by Chow et al. (2007). Chotigeat et al. (2010) also discovered that the REF gene was the most abundant transcript in latex tissue, followed by the SRPP gene. Tan et al. (2014) also discovered that, compared to other genes, REF and SRPP genes were substantially expressed within laticifer cells. REF genes were the most highly expressed compared to SRPP and CPT, where REF gene expression in latex tissue was nearly 250 times higher than in leaf tissue (Makita et al., 2017).

TFs are usually master regulators for expressing many genes in plant tissues (Gong et al., 2018). TFs control gene expression by binding to cis-regulatory specific regions in their target genes' promoters (Franco-Zorilla et al., 2014). According to our comparative analysis, ERF is significantly expressed in high latex-yielding rubber clones across all tissues. Furthermore, NF-Y and WRKY expression is slightly higher than other genes but not consistently across all tissues. Although there is little information on how TFs act in latex synthesis, various discoveries have added to our knowledge of TFs. ERF has also been identified as a highly expressed TF in previous investigations. Lindemose et al. (2013) state that AP2/ERF regulates gene expression and is implicated in the abiotic stress response. Because of its association with the stress response, AP2/ability ERFs to act as a gene expression regulator may impact high latex production (Makita et al. 2017). RRIM 600 treated with jasmonic acid for laticifer tissue growth, AP2/ERF was shown to be strongly expressed (Loh et al., 2019).

Sucrose is delivered to the laticifer cells for latex manufacture within Hevea tissue via a specialized network of cells known as the sieve elements. According to Lemoine et al. (2013), the idea of sucrose loading entails transporting sucrose from source to sink to complete the biological process. Sucrose transport is required to continue natural rubber production due to its relevance as a precursor for IPP manufacture, as natural rubber leaking from tapping procedures might be considered a powerful sink for the plant. According to the differential expression analysis, both genes were moderately expressed inside the high latex-yielding rubber clone. Transcripts encoded by the SUT gene were expressed 1.68 to 2.99 times greater in RRIM 3001 than in RRIM 712. SuSy gene transcripts were expressed 0.15 to 4.52 times greater in the bark and leaf tissues of RRIM 3001 than in the same tissues of RRIM 712. Both genes' decreased expression is linked to their functions in latex biosynthesis. Laticifers require a significant amount of carbon to renew their cytoplasm and replace latex lost during the tapping process (Duangngam et al., 2020). The cytoplasm can take anywhere from 48 to 72 hours to regenerate. The moderate expression of the SuSy gene in the bark tissue proves that sucrose synthesis commences promptly after the latex is released. HbSUT gene expression, on the other hand, takes hours to rise after hormone stimulation (Duangngam et al., 2020). The study explains why the SUT gene is expressed lower in RRIM 3001 tissues than in the SuSy gene.

Quantification analysis with qPCR is a typical method for experimentally verifying RNA-seq data. Because of its stability, β-actin was chosen as a reference gene in this investigation (Ma et al., 2020). Compared to RRIM 2025, RRIM 712, PB 350, and PB 260, the relative transcript abundance study revealed that RRIM 3001 had the highest expression among all genes. RRIM 3001 has a higher expression in jasmonic acid and percentage of dry rubber content following induction than RRIM 2025 and RRIM 600, according to a study by Loh et al. (2019). RRIM 3001 has higher expression than RRIM 712 in the *AACT*, *MVD*, *SRPP, REF, ERF, DXS*, and *HMGS* genes, according to Abdul Rahman et al. (2019). There is a link between relative transcript abundance and the RNA-seq dataset in this work, with expression in RRIM 3001 tissues remaining higher than in RRIM 712. Previous research showed that the relative transcript abundance and RNA-seq datasets are 91.67 percent identical regarding the fold change expression of selected genes (Loh et al., 2019), Abdul Rahman et al. (2019) claimed that most of the fold change expression between relative transcript abundance and RNA-seq data was the same in both datasets. The interaction of both methodologies and the accuracy of the datasets may be seen in the correlation of gene expression between RNA-seq and relative transcript abundance datasets. This suggestion is based on the data's similarity, in which the expression of the selected genes is expressed at a certain level, as proven by cross-validation of both datasets. The greater expression of some genes in RRIM 3001 could bolster evidence that this rubber clone is one of the most latex yieldings. This data could help determine which genes are responsible for high latex production.

CONCLUSION

Based on differential gene expression analysis, we compared the transcriptome datasets from RRIM 3001 (high-yield) and RRIM 712 (low-yield) clones to find potential genes that influence high natural rubber production. Over twenty genes associated with natural rubber production have been discovered, with five putative genes (*AACT*, *MVD*, *SRPP*, *REF,* and *ERF*) having higher expression than others. Quantitative amplification with five rubber clones was used to validate gene expression from these five putative genes. According to the findings, the five genes were highly expressed in high-yielding clones and lowly expressed in low-yielding clones. As a result, the five genes might be used as a biomarker in molecular-assisted breeding to create new rubber clones more quickly and at a lesser cost.

ACKNOWLEDGMENTS

We thank Yue Keong Choon (Universiti Sains Malaysia) for collecting the samples used in this work and Mohd Khairul Luqman Mohd Sakaf and Khairul Nasirudin Abu Mangsor (Universiti Sains Malaysia) for assisting in the analysis.

REFERENCES

Akhtar, T. A., Surowiecki, P., Siekierska, H., Kania, M., Gelder K. V., Rea, K. A., Virta, L. K. A., Vatta, M., Gawarecka, K. W., Jacek, D., Witold, B., Swiezewska, D. & Surmacz, L. (2017). Polyprenols are synthesized by a plastidial cis-prenyltransferase and influence photosynthetic performance. *The plant cell, 29*(7), 1709-1725.

Ambily, P.K., Thomas, M., Sreelatha, S., Krishnakumar, R., Annamalainathan, K. & Jacob, J. (2018). Expression analysis of rubber biosynthetic pathway genes in *Hevea brasiliensis*. *Journal of plantation crops, 46*(2), 102-111,

Chao, J., Chen, Y., Wu, S. & Tian, W. (2015). Comparative transcriptome analysis of latex from rubber tree clone CATAS8-79 and PR107 reveals new cues for the regulation of latex regeneration and duration of latex flow. *BMC plant biology, 15*(104), 1-12.

Cherian, S., Ryu, S.B. & Cornish, K. (2019). Natural rubber biosynthesis in plants, the rubber transferase complex, and metabolic engineering progress and prospects. *Plant Biotechnology Journal, 17*, 2041-2061.

Chotigeat, W., Duangchu, S. & Phongdara. A. (2010). cDNA library from the latex of *Hevea brasiliensis*. *Songklanakarin J. Sci. Technol., 32*, 555-559.

Chow, K., Wan, K., Noor, M., Isa, M., Bahari, A. & Tan, S. (2007). Insights into rubber biosynthesis from transcriptome analysis of *Hevea brasiliensis* latex. *Journal of experimental botany, 58*(10), 2429-2440.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X. & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome biology, 17*(13), 1-19.

Duangngam, O., Desalme, D., Thaler, P., Kasemsap, P., Sathornkich, J., Satakhun, D., Chayawat, C., Angeli, N., Chantuma, P. & Epron, D. (2020). In situ CO_2 labelling of rubber trees reveals a seasonal shift in the contribution of the carbon sources involved in latex regeneration. *J. Exp. Bot., 71*, 2028–2039.

Dusotoit-Coucaud, A., Brunel, N., Kongsawadworakul, P., Viboonjun, U., Lacointe, A. & Julien, J. L. (2009). Sucrose importation into laticifers of *Hevea brasiliensis*, in relation to ethylene stimulation of latex production. *Annals of botany, 104*(4), 635-647.

Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P. & Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc Natl Acad Sci USA, 111*(6), 2367-2372.

Gong, X. X., Yan, B. Y., Hu, J., Yang, C. P., Li, Y. J., Liu, J. P. & Liao, W. B. (2018). Transcriptome profiling of rubber tree (*Hevea brasiliensis*) discovers candidate regulators of the cold stress response. *Genes & Genomics, 40*, 1181–1197.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I.,… & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology 29*(7), 644–652.

Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatic, 26*(5), 680-682.

Kim, I. J., Ryu, S. B., Kwak, Y. J. & Kang, H. (2004). A novel cDNA from *Parthenium argentatum* Gray enhances the rubber biosynthetic activity in vitro. *Journal of experimental botany, 55*(396), 377-385.

Kush, A. (1994). Isoprenoid biosynthesis: the Hevea factory. *Plant physiology and biochemistry, 32*(6), 761-767.

Langmead, B. & Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods, 9*, 357–359.

Lau, N. S., Makita, Y., Kawashima, M., Taylor, T., Kondo, K., Othman, A. S., Shu-Chien, A. C. & Matsui, M. (2016). The rubber tree genome shows expansion of gene family associated with rubber biosynthesis. *Scientific reports*, 6(2859), 1-14.

Lemoine, R. (2000). Sucrose transporter in plants: update on function and structure. *Biochimica et Biophysica Acta, 1465*, 246-262.

Li, D., Deng, Z., Qin, B., Liu, X. & Men, Z. (2012). De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). *BMC Genomics, 13*(192), 1-14.

Li, D., Hao, L., Liu, H., Zhao, M., Deng, Z., Li, Y., Zeng, R. & Tian, W. (2015). Next-generation sequencing, assembly, and comparative analyses of the latex transcriptomes from two elite *Hevea brasiliensis* varieties. *Tree genetics and genomes, 11*(5), 1-9.

Li, D., Wang, X., Deng, Z., Liu, H., Yang, H. & He, G. (2016). Transcriptome analyses reveal molecular mechanism underlying tapping panel dryness of rubber tree (*Hevea brasiliensis*). *Scientific reports, 6*(1), 1-10.

Lindemose, S., Shea, C. O., Jensen, M. K. & Skriver, K. (2013). Structure, function and networks of transcription factors involved in abiotic stress responses. *International journal of molecular sciences, 14*(3), 5842-5878.

Liu, J., Zhuang, Y., Guo, X. & Li, Y. (2016). Molecular mechanism of ethylene stimulation of latex yield in rubber tree (*Hevea brasiliensis*) revealed by *de novo* sequencing and transcriptome analysis. *BMC Genomics, 17*(257), 1–13.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C (T)) Method. *Methods, 25*(4). 402-408.

Loh, S. C., Othman, A.S. & Veera, S. G. (2019). Identification and characterization of jasmonic acid- and linolenic acid-mediated transcriptional regulation of secondary laticifer differentiation in *Hevea brasiliensis*. *Scientific reports, 9*(14296), 1-16.

Ma, L., Wu, J., Qi, W., Coulter, J.A., Fang Y,… & Wanchang, S. (2020). Screening and verification of reference genes for analysis of gene expression in winter rapeseed (*Brassica rapa* L.) under abiotic stress. *PLoS One, 15*(9), 1-16.

Makita, Y., Ng, K. K., Singham, G. V., Kawashima, M., Hirakawa, H., Sato, S., Othman, A. S. & Matsui, M. (2017). Large-scale collection of full-length cDNA and transcriptome analysis in *Hevea brasiliensis*. *DNA research, 24*(2), 159-167.

Mantello, C. C., Cardoso-Silva, C. B., da Silva, C. C., de Souza, L. M., Junior, E. J. S., Goncalves, P. D. S., Vicentini, R. & de Souza, A. P. (2014). De novo assembly and transcriptome analysis of the rubber tree (*Hevea brasiliensis*) and SNP markers development for rubber biosynthesis pathways. *PLoS One, 9*(7), 1-14.

Montoro, P., Wu, S., Favreau, B., Herlinawati, E., Labrune, C., Martin-Magniette, M. L., Pointet, S., Rio, M., Lecclerrcq, J., Ismawanto, S. & Kuswanhadi (2018). Transcriptome analysis in *Hevea brasiliensis* latex revealed

changes in hormone signalling pathways during ethephon stimulation and consequent tapping panel dryness. *Scientific reports, 8*(8483), 1-12.

Priyadarshan, P. M. (2003). Breeding *Hevea brasiliensis* for environmental constraints. *Advances in agronomy, 79*, 351-400.

Rahman, A. Y. A., Usharraj, A. O., Misra, B. B.,… & Maqsudul, A. (2013). Draft genome sequence of the rubber tree *Hevea brasiliensis*. *BMC Genomics, 14*(75), 1-15.

Rahman, S. N. A., Bakar, M. F. B., Singham, G. V. & Othman, A. S. (2019). Single-nucleotide polymorphism markers within MVA and MEP pathways among *Hevea brasiliensis* clones through transcriptomic analysis. *3 Biotech, 9*(388), 1-10.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. (2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics, 26*(1), 139-140.

Roy, C. B., Liu, H., Rajamani, A. & Saha, T. (2019). Transcriptome profiling reveals genetic basis of disease resistance against Corynespora cassiicola in rubber tree (*Hevea brasiliensis*). *Current plant biology, 17*, 2-16.

Rozhan, A. D. (2015). Transformation of agricultural sector in Malaysia through agricultural policy. Agriculture policy platform. Retrieved from FFTC Agricultural Policy Platform: http://ap.fftc.agnet.org.

Saha, T. & Priyadarshan, P. M. (2012). *Genomics of Hevea Rubber*. New York, NY. Springer.

Sando, T., Takaoka, C., Mukai, Y., Yamashita, A., Hattori, M., Ogasawara, N., Fukusaki, E. & Kobayashi, A. (2008). Cloning and Characterization of Mevalonate Pathway Genes in a Natural Rubber Producing Plant, *Hevea brasiliensis*. *Bioscience, Biotechnology, and Biochemistry, 72*(8), 2049-2060.

Sando, T., Hayashi, T., Takeda, T., Akiyama, Y., Nakazawa, Y., Fukusaki, E. & Kobayashi, A. (2009). Histochemical study of detailed laticifer structure and rubber biosynthesis-related protein localization in *Hevea brasiliensis* using spectral confocal laser scanning microscopy. *Planta, 230*, 215-225.

Tan, D., Sun, X. & Zhang, J. (2014). Age-dependent and jasmonic acid-induced laticifer-cell differentiation in anther callus cultures of rubber tree. *Planta, 240*(2), 337-344.

Tan, D., Hu, X., Fu, L., Kumpeangkeaw, A., Ding, Z. & Sun, X. (2017). Comparative morphology and transcriptome analysis reveals distinct functions of the primary and secondary laticifer cells in the rubber tree. *Scientific reports, 7*(3126), 1-17.

Tang, C., Xiao, X., Li, H., Fan, Y., Yang, J., Qi, J. & Li, H. (2013). Comparative Analysis of Latex Transcriptome Reveals Putative Molecular Mechanisms Underlying Super Productivity of *Hevea brasiliensis*. *PLoS One, 8*(9), 1-18.

Trapnell, C., Pachter, L. & Salzberg, S. L. (2010). TopHat: discovering splice junctions with RNA-seq. *Bioinformatics, 25*(9), 1105-1111.

Tritsch, D., Hemmerlin, A., Bach, T. J. & Rohmer, M. (2010). Plant isoprenoid biosynthesis via the MEP pathway: In vivo IPP / DMAPP ratio produced by (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase in tobacco BY-2 cell cultures. *FEBS letters, 584*(1), 129-134.

Wang, Z., Gerstein, M. & Snyder, M. (2009). RNA-seq: a revolutionary tool for transcriptomics. *Nature reviews, 10*(1), 57-63.

Wilhelm, B. T. & Landry, J. R. (2009). RNA-seq: quantitative measurement of expression through massively parallel RNA-sequencing. *Methods, 48*(3), 249-257.

Wu, C., Li, Y., Nie, Z., Dai, L., Kang, G. & Zeng, R. (2017). Molecular cloning and expression analysis of the mevalonate diphosphate decarboxylase gene from the latex of *Hevea brasiliensis*. *Tree Genetics & Genomes, 13*(22), 1-16.

How to cite this paper:
Bakar, M.F.A. & Othman, A.S. (2022). Comparative Transcriptome Analysis Identifies Potentially Relevant Genes in Rubber Clones with a High Latex Yield (*Hevea brasiliensis*). Journal of Agrobiotechnology, *13*(1), 61-76.