UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE ALIMENTOS

BÁRBARA REGINA BAZZO

MACAUBEST: A COMPREHENSIVE TRANSCRIPTOME OF MACAÚBA PALM (*Acrocomia aculeata*)

MACAUBEST: ATLAS TRANSCRIPTÔMICO DA MACÁUBA (*Acrocomia aculeata*)

CAMPINAS
2018

BÁRBARA REGINA BAZZO

MACAUBEST: A COMPREHENSIVE TRANSCRIPTOME OF MACAÚBA PALM (*Acrocomia aculeata*)

MACAUBEST: ATLAS TRANSCRIPTÔMICO DA MACÁUBA (*Acrocomia aculeata*)

Thesis presented to the Faculty of Food Engineering of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Science.

Tese apresentada à Faculdade de Engenharia de Alimentos da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DE TESE DEFENDIDA PELA ALUNA BÁRBARA REGINA BAZZO E ORIENTADA PELO PROFESSOR DR. GONÇALO AMARANTE GUIMARÃES PEREIRA.

CAMPINAS
2018

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2014/07265-2; CNPq, 458045/2014-4
ORCID: https://orcid.org/0000-0002-0560-110X

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Engenharia de Alimentos
Márcia Regina Garbelini Sevillano - CRB 8/3647

B349m
Bazzo, Bárbara Regina, 1988-
Macaubest: a comprehensive transcriptome of macaúba palm (Acrocomia aculeata) / Bárbara Regina Bazzo. – Campinas, SP : [s.n.], 2018.

Orientador: Gonçalo Amarante Guimarães Pereira.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia de Alimentos.

1. RNA-seq. 2. Expressão gênica. 3. Marcadores moleculares. I. Pereira, Gonçalo Amarante Guimarães. II. Universidade Estadual de Campinas. Faculdade de Engenharia de Alimentos. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Macaubest: atlas transcriptômico da macaúba (Acrocomia aculeata).
Palavras-chave em inglês:
RNA-seq
Genic expression
Molecular markers
Área de concentração: Bioenergia
Titulação: Doutora em Ciências
Banca examinadora:
Gonçalo Amarante Guimarães Pereira [Orientador]
Maria Imaculada Zucchi
Jorge Mauricio Costa Mondego
Marco Aurelio Takita
Luiz Henrique Chorfi Berton
Data de defesa: 13-12-2018
Programa de Pós-Graduação: Bioenergia

EXAMINATION COMMITTEE

Dr. Gonçalo Amarante Guimarães Pereira – Presidente
Universidade de Campinas

Dra. Maria Imaculada Zucchi
Agência Paulista de Tecnologia dos Agronegócios - Polo

Dr. Jorge Mauricio Costa Mondego
Instituto Agronômico de Campinas

Dr. Marco Aurelio Takita
Instituto Agronômico de Campinas

Dr. Luiz Henrique Chorfi Berton
Instituto Agronômico de Campinas

DEDICATÓRIA

À minha vó Zelinda, que com sua sabedoria me ensinou muito mais do que os livros e artigos.

AGRADECIMENTOS

Agradeço ao prof. Gonçalo, que apostou na ideia da macaúba e deu todo apoio e orientação necessários para o trabalho saísse do papel e ganhasse o mundo.

Ao prof. Carlos Augusto Colombo, por compartilhar conhecimento, laboratório e alegrias por muitos anos da minha vida na ciência.

À FAPESP, pela concessão da bolsa de doutorado (processo 2014/07265-2), pois sem o auxílio financeiro a realização do trabalho não teria sido possível.

Ao CNPq, pelo financiamento do projeto de pesquisa (projeto CNPq Universal 458045/2014-4 – Coordenador: Carlos Augusto Colombo).

À CAPES, pelo apoio financeiro ao projeto e a pós-graduação.

Aos bioinformatas e amigos Marcelo e Lucas, pelo suporte e auxílio nas análises dos dados de RNA-Seq.

À minha família, que sempre me deu suporte e que me move para frente.

Ao meu namorado Renan, que me apoia nas ideias mirabolantes.

As minhas amigas Brenda, Manu, Fer, Paty, Daiane. Ao André, Denis, Renato, Carina, Bia, Cissa, Jonathan e Elivelton pela amizade durante o doutorado, que será para a vida.

Obrigada a todos.

ABSTRACT

The macaúba palm (*Acrocomia aculeata*) is a native and widespread palm tree, being present mainly in Brazilian Cerrado. It is considered an alternative oil-feedstock crop since it has high productivity, high oil content, high-quality oil, as well as it is possible to plant in drought areas. Consequently, molecular biology and genetic diversity studies may provide the basis for a successful breeding program. Therefore, this study aimed to build a macaúba palm gene bank and develop molecular markers for genetic diversity studies, marker-assisted selection, and population mapping. Moreover, the studies investigated possible differences between the macaúba pam and African oil palm, which are genetically close but with different edaphoclimatic requirements. For this, the RNA was extracted from the root, bulb, leaf, flower (female and male), and mesocarp of mature fruits. In the first experiment, De novo transcriptome assembly permitted to identify 34.293 transcripts of macaúba palm in all the libraries, gathered in a single reference. Gene expression analysis showed eight transcripts expressed only in fruit tissues, two in leaf tissue, eight in bulb tissue, five in female flowers, 14 in male flowers, 8053 in root and none in leaf sheath tissue. Furthermore, we observed 341 root-specific transcripts with high expression level. In addition, family turnover analysis between macaúba palm and African oil palm showed that macaúba has 24 expanded gene families, 11 families have at least one root-specific transcript and six families are exclusive of macaúba palm. Most transcripts were similarly to transduction proteins related to abiotic and biotic stresses responses, such as kinases, transcription factors, Ca^{2+}-binding proteins. The data suggest a possible regulation of signal transduction in response to many stresses. In the second experiment, the reference-based transcriptome assembly was performed. We identified 418 EST-SSRs and 232 EST-SSR were selected, with trinucleotide repeats being the most frequent motif (380-90.9%), followed by composited (4.5%), di- (3.6%), and hexanucleotides (3.6%). A total of 145 EST-SSR (62.5%) were validated on seventeen macaúba samples, and 100 were considered polymorphic with PIC values ranging from 0.25 to 0.77. Genetic diversity analysis performed with the 20 most informative EST-SSRs showed distinct separation of the different macaúba samples, according to the geographical location. Additionally, 145 markers were transferred in six other palm species resulting in 99% transferability rates (144) in *Acrocomia intumescens*, 98% (143) in *Acrocomia totai*, 80.7% (117 EST-EST) in samples of African oil palm (*Elaeis guineensis*) and peach palm (*Bactris gasipaes*), 70% (102) in juçara palm (*Euterpe edulis*) and 71.7% (104) in hat palm (*Sabal causiarum*). Analysis of genetic distance established distinct groups by genera. These are exploratory data but will support ongoing research with macaúba palm.

RESUMO

A macaúba (*Acrocomia aculeata*) é uma palmeira nativa de ampla ocorrência, estando presente principalmente em áreas de Cerrado. Sua elevada produtividade de frutos, altos teores e qualidade de óleo, além da possibilidade de plantio em áreas de déficit hídrico a credenciam como uma alternativa de matéria-prima para a produção de óleo para biodiesel e outras demandas. Assim, estudos na área de biologia molecular e diversidade genética formam a base para um programa de melhoramento de sucesso para uma espécie pouco estudada. Diante desse contexto, o trabalho visou construir um banco de genes da macaúba e o desenvolvimento de marcadores moleculares para estudos de caracterização da diversidade genética, seleção assistida e mapeamento das populações. Além disso, o trabalho investigou possíveis diferenças entre a macaúba e dendezeiro, espécies próximas geneticamente, porém com exigências edafoclimáticas distintas. Para isso, o RNA foi extraído de tecidos de raiz, bulbo, folha, flor (feminina e masculina) e mesocarpo de frutos maduros. No primeiro experimento, a montagem *De Novo* do transcriptoma foi utilizada, permitindo identificar 34.293 transcritos de macaúba em todas as bibliotecas, reunidas em uma única referência. O padrão de expressão gênica mostrou oito transcritos específicos de tecidos de fruto, dois de tecido foliar, oito de tecido de bulbo, cinco de flores femininas, 14 de flores masculinas, 8053 de tecido de raiz e nenhum em tecido de bainha foliar. Quando observados o nível de expressão dos transcritos tecido-específicos, o tecido de raiz apresentou 341 transcritos altamente expressos e específicos. Além disso, a comparação com o dendezeiro mostrou que a macaúba possui 24 famílias gênicas expandidas, 11 famílias possuem pelo menos um transcrito raiz- específico e seis famílias foram exclusivas de macaúba. A maioria dos transcritos expressos foi anotada como proteínas relacionadas a transdução de sinais em resposta a estresses abióticos e bióticos, como quinases, fatores de transcrição, proteínas ligantes de Ca^{2+}. Os dados sugerem uma possível regulação na transdução de sinal em resposta à diversos estresses. Na segunda abordagem foi utilizada a montagem com a referência do genoma do dendê. Neste estudo, um total de 418 EST-SSRs foram identificados e 232 EST-SSR foram selecionados, com repetições trinucleotídicas sendo o motivo mais frequente (380 -90,9%), seguido por compostas (4,5%), di- (3,6%) e hexanucleotídeos (3,6%). Um total de 145 EST-SSR (62,5%) foram validados em dezessete amostras de macaúba, e 100 foram considerados polimórficos com valores de PIC de 0.25 a 0.77. Análise de diversidade genética realizada com os 20 mais informativos EST-SSR mostraram uma separação distinta dos diferentes grupos de macaúba, de acordo com a localização geográfica das amostras. Além disso, 145 marcadores foram transferidos em outras seis espécies de palmeiras resultando em taxas de transferibilidade de 99% (144) em *Acrocomia intumescens*, 98% (143) em *Acrocomia totai*, 80,7% (117 EST-EST) em amostras de dendezeiro (*Elaeis guineensis*) e pupunha (*Bactris gasipaes*), 70% (102) na palma juçara (*Euterpe edulis*) e 71,7% (104) em sabal (*Sabal causiarum*). Análise da distância genética estabeleceu grupos distintos por gênero das palmeiras. Os dados obtidos são exploratórios, mas suportarão as pesquisas em andamento com a planta.

SUMÁRIO

1. INTRODUCTION

Considering the increasing consumption of oil derivatives, reserves reduction and climate change resulting from the greenhouse effect, biofuels are promising alternatives as they allow the partial or total replacement of conventional fossil fuels.

According to the Ministry of Mines and Energy (MME), Brazil assumed to reduce greenhouse gas emissions by 37% below 2005 levels by 2025, with a subsequent contribution to reducing greenhouse gas emissions by 43% below 2005 levels by 2030. To this aim, the country committed itself to increase the share of sustainable bioenergy in its energy matrix to approximately 18% by 2030, to restoring and reforesting 12 million hectares of forests, as well as to achieve an estimated 45% share of renewable energy in the energy matrix by 2030. (MINISTRY OF MINES AND ENERGY, 2015). Moreover, increasing demand for energy based on biomass and industrial products translates into an expansion of agricultural production at the global level, which has economic and social advantages. This market generates impacts on commercial relations, generate employment and income throughout the production chain (SALLET; ALVIM, 2011).

Brazil is the second largest producer and consumer of ethanol (generated from sugar cane) and biodiesel, and it is projected that in 2020 an additional area of 0.5 million hectares will be dedicated to biofuels in Brazil. From 2000 to 2013, 26 million hectares of forest were lost to the production of biofuels, suggesting a role for the expansion of biofuels as an important driving force for deforestation in Brazil (LANGEVELD et al., 2013).

The use of clean energy is significant in Brazil. According to the Ministry of Mines and Energy, from 2005 to December 2012, 11 billion liters of biodiesel have already been added to fossil diesel. The mixture of 2%, initially on a voluntary basis and, as of 2008, has become mandatory and with increasing percentages. In March of 2018 had its percentage raised to 10%, which demand has grown in 1 billion liters and expectation of consumption in 5.3 billion liters this year (MINISTRY OF MINES AND ENERGY, 2018). According to IRENA (International Renewable Energy Agency), Brazil employed 202 000 people in biodiesel in 2017, up more than 30 000 from the previous year, and the production of green fuel added more than R$ 12 billion to the Gross Domestic Product (GDP) of the country between 2008 and 2012 (APROBIO, 2012; IRENA, 2018).

According to the survey, in the same four years, Brazilian trade balance saved R$ 11.5 billion in mineral diesel import (APROBIO, 2012).

Different vegetable feedstock for biodiesel production is appropriate for Brazil, since the diesel consumption rate is twice as high as gasoline, which has ethanol as a viable replacement. As for diesel, there is no oil resource that can meet the demand for biodiesel to the current situation in Brazil.

At the same time, biomass production is criticized for its negative environmental impact once it increases greenhouse gas emissions and biodiversity loss. This occurred by means of land use that changed from natural and semi-natural ecosystems to agricultural land. Currently, the main biodiesel production in Brazil is from soybean, accounting for 70.7% of all production, until June 2017. It is typically produced on a large-scale mechanized monoculture, which it does not make available social inclusion of small producers, and generates impact on forest systems, reducing biodiversity and soil quality.

Currently, the macaúba palm is considered a novel oil-feedstock crop with an environmentally and socially co-beneficial biofuel feedstock in South America. It presents high productivity and multiple potentialities since it can be used in the recovery, preservation, and restoration of permanent areas, restoring the soil and promoting sustainable extraction practices. It can be integrated with crop rotations, consortium or agrosilvopastoral system, atmospheric carbon sequestration and mitigating the negative impact of gases released by cattle (STEINFELD et al., 2006).

The macaúba palm (*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.) belongs to the Arecaceae family and is native to tropical America. It presents as a perennial, heliophilous, and arborescent palm with a single stem between 4- 15 m tall and 20-30 cm diameter (LORENZI H, NOBLICK L, 2010). Its fruit is an edible globose drupe, with mucilaginous fibrous mesocarp and strongly adheres endocarp that contains up to four oleaginous seeds. The bunch can weigh more than 25 kg (LORENZI H, NOBLICK L, 2010; SCARIOT; LLERAS; HAY, 1995) (Figure 1).

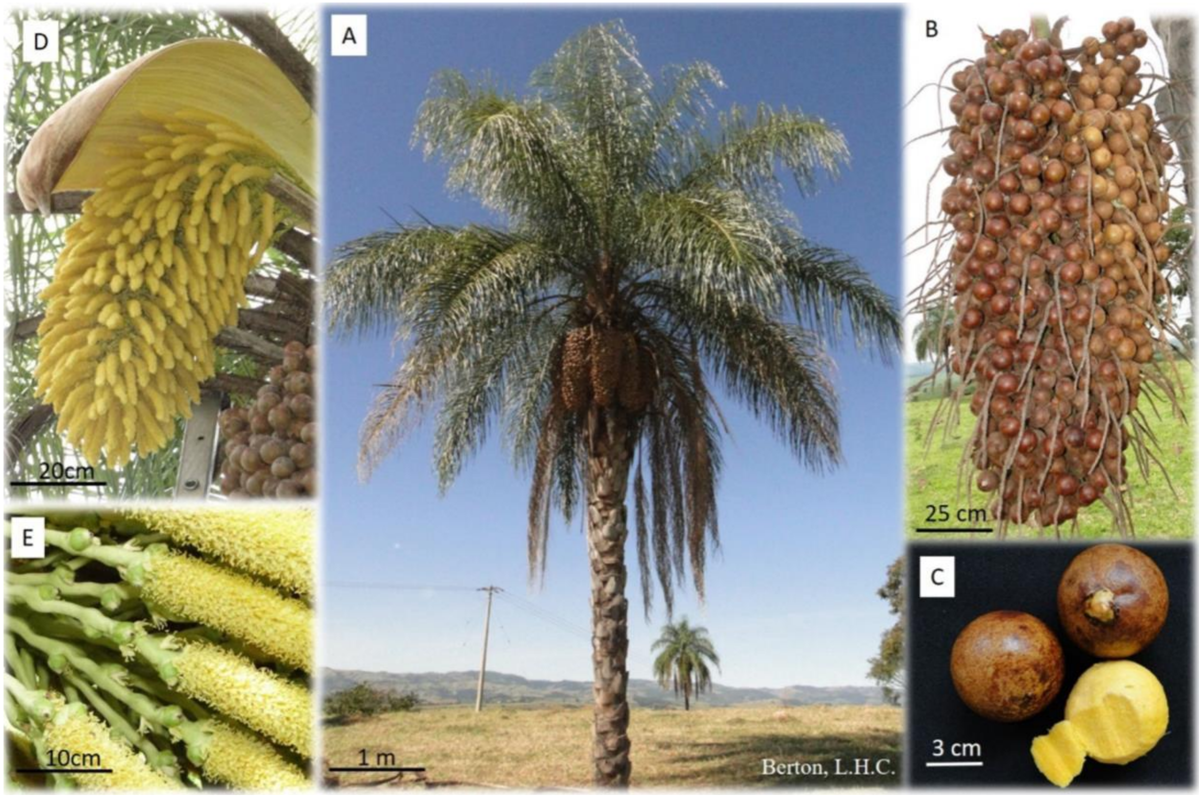

Figure 1. A. Adult plant of macaúba palm. B. Bunch of fruits; C. Ripe fruit in detail; fruit without epicarp (yellow mesocarp); D. Open inflorescence; E. Detail of inflorescence, with female flowers located in the basal region and male flowers in the apical region. (COLOMBO et al., 2018)

Macaúba is diploid (2n = 30) (ABREU et al., 2011), with a genome size 2C = 5.81 pg and base composition of AT = 58.3% and GC = 41.7%. It is monoecious and self-compatible, with a mixed reproductive system (ABREU et al., 2012; LANES et al., 2016; SCARIOT; LLERAS; HAY, 1995). The predominance is of open crosses, but there is a certain degree of self-crosses, most likely due to the limited flight range of pollinators and the restriction of the dispersion of the seeds. Although the seeds can be dispersed by large animals, most are dispersed by gravity, thus remaining under the mother tree (ABREU et al., 2012).

It is a native plant widely distributed in tropical America, occurring from Mexico, North Florida and West Indies to the South of Brazil, Paraguay and north Argentina (SCARIOT; LLERAS; HAY, 1995) and it is considered the most widespread palm tree, which indicates its adaptability to different soil, climatic conditions, and ecosystems (COLOMBO et al., 2018).

The fruit, the greatest product of macaúba palm, is rich in oil both in the pulp (mesocarp) and kernel (endosperm). Mesocarp oil is considered "high oleic" oil (CICONINI et al., 2013), which it is valuable for the food and biodiesel industry due to its greater oxidative stability and operability at low temperatures (BERTON, 2013). The endosperm oil is rich in lauric acid (short-chain fatty acid) and is a valuable source for pharmaceutical and cosmetic use (BELTRÃO; OLIVEIRA, 2007).

Macaúba palm has a great production potential of oil being similar with African oil palm (*Elaeis guineensis*), whose global production reached 69 million of tons in 2017, representing 34% of global oil production (USDA, 2018). The productivity of selected native plants of macaúba may reach 5000 kg of oil/ha/year (COLOMBO et al., 2018; EVARISTO et al., 2016), making it an inexpensive feedstock for oil production. Despite its incipient domestication and current use in extractive character, macaúba has several uses, with multiple products from its exploitation, like food, cosmetics, animal feed, and biofuels (COLOMBO et al., 2018). Nowadays, the renewed interest in this novel feedstock increased commercial interest and can lead to the propagation of plants without agronomic quality, which would make their competitiveness unfeasible.

Based on the fact the lack of breeded plants for commercial production and natural populations are very heterogeneous in production, agronomic, molecular, economic and social studies are essential to turn macaúba into a potential plant in the production of biodiesel and co-products. Precaution should be taken in the use of non-domesticated feedstock species because possible biological variations and genetic diversity are observed in the macaúba palm in different environments of occurrence (LANES et al., 2015).

Regardless of the increasing commercial interest, macaúba has limited genomic studies due to the small quantity of publicly available sequence data. In this context, High-throughput next-generation sequencing (NGS) technologies have revolutionized transcriptomics especially with the advent of RNA-sequencing (RNA-seq), giving us a deeper understanding of plant biology by the generation of biologically important data sets from different plant species. Most of the information we have about mechanisms underlying plant biological processes in palm trees come from investigations on model African oil palm, a palm tree extensively studied at the whole genome level to elucidate various complex biological phenomena.

RNA-Seq can be helpful for listing the transcripts and other RNAs from several tissues; studying post-transcriptional modification and mutations; quantifying gene expression; and identifying genic molecular markers (GMMs) (VARSHNEY et al., 2009). Molecular markers are considered an essential tool to identify and to select superior plants for adoption in large-scale commercial crops, the establishment of core collections, the creation of seed garden, and initiation of breeding studies (PALLIYARAKKAL; RAMASWAMY; VADIVEL, 2011; VARSHNEY; GRANER; SORRELLS, 2005). Microsatellites or Single Sequence Repeats (SSRs), mainly genic-SSR (or EST-SSRs), are widely employed in palm trees studies with commercial interest. In addition to being functional, EST-SSRs can lead to a gain or loss of gene function via frameshift mutation or other changes in the amino acid sequence (LI et al., 2004).

This study is an exploratory research and the genetic information obtained will allow identifying the molecular bases of macaúba palm, to investigate the genetic diversity through good molecular markers, forming the basis of a successful breeding program. These results may build a foundation for pieces of research with macaúba palm and the data would be helpful in selecting genes for macaúba palm breeding via genetic engineering. Therefore, this thesis is divided into two chapters. The first involves the comprehensive transcriptome sequencing of seven *A. aculeata* tissues, in which we identified many expressed transcripts of macaúba palm and further analyzed expanded gene families and stress-related transcripts in comparison with African oil palm. The tissue-specific genes and expression levels in different tissues were detected in many biological processes. The second chapter is a paper submitted to *BMC Plant Biology*, entitled **“***Development of novel EST-SSR markers in the macaúba palm (Acrocomia aculeata) using transcriptome sequencing and cross-species transferability in Arecaceae species*”, which provides a polymorphic set of genic microsatellite markers and will allow a deeper understanding of the genetic diversity, genotype characterization and genetic structure of *A. aculeata*, Acrocomia genera, and other environmental and commercial important palm species. Additionally, these markers will be useful in modern *A. aculeata* breeding programmes.

CHAPTER 1: “Novel Comprehensive multi-tissue transcriptome of *Acrocomia aculeata* – macaúba palm.”

Bárbara Regina Bazzo[1], Lucas Miguel de Carvalho[1], Carlos Augusto Colombo[2], Marcelo Falsarella Carazzolle[1], Gonçalo Amarante Guimarães Pereira[1].
Address:
[1]Institute of Biology, Laboratory of Genomic and Expression, State University of Campinas, Campinas, Brazil
[2]Agronomic Institute of São Paulo (IAC), Genetic Resources Center, Campinas, SP, Brazil

E-mail:
Bárbara Regina Bazzo – barbarareginabazzo@gmail.com
Lucas Miguel de Carvalho - lucasmigueel@gmail.com
Carlos Augusto Colombo – ccolombo@iac.sp.gov.br
Marcelo Falsarella Carazzolle - mcarazzo@unicamp.br
Gonçalo Amarante Guimarães Pereira - goncalo@unicamp.br

Abstract

Background The macaúba palm is a novel feedstock for oil production suitable for multiple uses, including as biodiesel and in the food and cosmetic industries. Despite the increasing commercial interest, macaúba is considered a non-model plant which genomic studies is limited by the small quantity of publicly available sequence data. We reported a comprehensive multi-tissue dataset of genes by using transcriptome sequencing and phylogeny analysis with principal palm trees genomes.

Results In this study, a total of 34.293 transcripts were identified in all libraries were joined in a single reference. 5062 transcripts (14.76%) had an annotation by PANNZER and 25.246 transcripts (73.61%) were annotated by CDD. Moreover, 48.545 terms were assigned to a molecular function category, 81.511 terms in the biological process category, and 20.421 terms in the cellular component category. Gene expression pattern showed eight transcripts expressed only in fruit tissues, two in leaf tissue, eight in bulb tissue, five in female flowers, fourteen in male flowers, 8053 in root tissues and none in leaf sheave tissues. In addition, family turnover analysis between macaúba palm and African oil palm showed that macaúba has 24 expanded gene families, and 11 families have at least one root-specific transcript.

Conclusions The data suggest a possible regulation of signal transduction in response to many stresses. The macaúba palm gene bank constructed in our study is a valuable resource and provide a genomic tool for further evolution, gene expression, and genomic studies.

1. BACKGROUND

The oleaginous macaúba palm (*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart., 2n = 30) is a native palm tree from America, belonging to the Arecaceae family, which is receiving attention as a novel oil-feedstock crop [1, 2]. It presents high productivity and multiple potentialities since it can be used in the recovery, preservation, and restoration of permanent areas, restoring the soil and promoting sustainable extraction practices. It can be integrated with crop rotations, consortium or agrosilvopastoral system, atmospheric carbon sequestration and mitigating the negative impact of gases released by cattle [3].

Comparing macaúba palm with oil palm, the most produced vegetable oil, the former is an environmentally and socially co-beneficial biofuel feedstock in South America, since oil palm is associated with sustainability problems such as deforestation of tropical areas (due to its high water requirement), biodiversity losses, and emerging social conflicts [4–6].

Macaúba palm grows naturally in large populations and its presence has been documented from north Florida, Mexico and West Indies to south Paraguay and north Argentina. It is considered the most widespread palm in Brazil, which indicates their adaptability to different soil and climatic conditions. Based on archaeological data, macaúba palm probably was dispersed from South America to Central America, once the oldest fossil sites are in Santarém, northeastern Brazil (11,200 B.P.) (as Acrocomia sp.), followed by Panama (8040 B.P.), and Mexico (6750 B.P.) [7].

The species is monoecious and the leading product of macaúba palm, the fruit, is rich in oil in two tissues: mesocarp and endosperm. Mesocarp oil is considered “high oleic” acid content and overtakes values up more than 70%. Ciconini *et al.* [8] observed oleic acid content in mesocarp oil above 80%, which is valuable for food industries and especially energy, due to higher oxidation stability and operability at low temperatures.

High-throughput next-generation sequencing (NGS) technologies have revolutionized transcriptomic studies especially with the advent of RNA-sequencing (RNA-seq), giving us a deeper understanding of plant biology by the generation of biologically important data sets from different non-model plant species [9]. Most of the information we have about mechanisms underlying plant biological processes in palm trees come from investigations on model plant African oil palm, a palm tree extensively studied at the whole genome level to elucidate various complex biological phenomena [10].

Transcriptome data is an efficient technique to discover genes or gene families involved in various metabolic pathways in plants without genome data. Transcriptome studies were performed in *Elaeis guineensis* for annotation of EST [11], SNPs detection [12], analyze differential expression of genes between normal and *mantled* genotypes (somaclonal variant that affects oil palm flower and fruit development) [13], and explain the mechanism of lipid biosynthesis in tissues that stores oil at developmental stages [14]. Furthermore, this technology was used to recognize genes at various stages of ripening fruit, as studied with date palm (*Phoenix dactylifera* L.). From this, it is possible to trace metabolic pathways of sugars accumulation, to study the ripening fruit or synthesis of enzymes [15]. Zhang *et al.* [16] produced a comprehensive transcriptome from expressed genes in different plant tissues (roots, branches, male and female flower, new and mature leaves and fruits), by showing possible metabolic pathways.

Despite the increasing commercial interest, the macaúba palm is considered a non-model plant which genomic studies are limited by the small quantity of available sequence data. Here, we identified many expressed genes in deeply sequencing pool based on RNA-Seq data from seven tissues (leaves, leaf sheaves, roots, bulbs, fruit mesocarp, male and female flowers) of macaúba palm and we analyzed expanded gene families in comparison with African oil palm. The tissue-specific genes, abundance level and the expression patterns of genes involved in mainly biological processes were analyzed. Moreover, we applied an orthology analysis with a drought gene database with principal palm trees genomes and related species.

These results may build a foundation for pieces of research with macaúba palm and the data would be helpful in selecting genes for future macaúba palm breeding via genetic engineering.

2. METHODS

2.1. Plant Material and RNA Isolation

The present study was performed using the tissues of leaves, leaf sheaves, roots, bulbs, pulp fruit and male and female flowers from the macaúba palm (*Acrocomia aculeata*) for RNA isolation and transcriptome sequencing.

Vegetative tissues were collected from eight-month-old seedlings from native plants of Dourado, Amparo, and Pedreira from São Paulo State/Brazil. The plants were acclimated for one month in a greenhouse. The male and female flowers were collected from adult plants at the Experimental Unit Santa Elisa – IAC (Agronomic Institute of São Paulo/ Campinas/ Sao Paulo State/Brazil). The flower bunches were removed from the base to avoid damage to these materials. The mesocarp tissue of ripe fruits was collected from Santo Antônio de Posse, Amparo from São Paulo State/Brazil, and Ibituruna from Minas Gerais State/Brazil.

RNA was isolated using the lithium chloride method [17, 18]; for fruit tissue, RNA was isolated using the perchlorate protocol [19]. Total RNA was treated with RNase-free DNase I (Takara, Kyoto, Japan) for 30 min at 37°C to remove residual DNA. The RNA quality was verified using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and RNase-free agarose gel electrophoresis. The concentration of the total RNA was further quantified using an RNA NanoDrop (Thermo Fisher Scientific Inc., Waltham, MA, USA).

2.2. Sequencing, filtering, and assembly

The mRNA libraries were synthesized using the TruSeq Stranded mRNA Library Preparation Kit and sequenced using the Illumina paired-end sequencing technology HiSeq^{TM} 2500 platform (Illumina, San Diego, CA, USA). Prior to assembly, the 100 bp paired-end reads were submitted to quality filtering and adapter trimming using the Trimmomatic software version 0.36 [20]. All reads with more than 10% of bases with a poor-quality score (Q<20) or non-coding RNA, as well as ambiguous sequences containing an excess of "N" nucleotide calls or adaptor contamination, were removed. The sequence data quality has been checked by software FasTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the reads have been then sorted to enriched 16S rRNA fragments and mRNA reads by SortMeRNA V 1.9 [21]. The high-quality filtered reads were assembled without a reference genome using the *de novo* transcriptome assembler Trinity 2.0.6 [22] with default parameters (k-mer size 25 and minimum transcript length > 200bp).

2.3. Expression Profile

Estimation of transcript abundance was done using program Kallisto [23]. This program uses pseudo-alignment for rapidly determining the compatibility of reads with targets. The Kallisto quantification algorithm was run with default parameters enabling sequence-based bias correction and plain text output of TPM and normalized count estimates.

The relative expression values (TPM) were transformed to allow the use of statistical tests that assume a normal data distribution. By adding one and taking the natural logarithm of the TPM values the range of the data is reduced and the residuals more frequently follow a normal distribution (determined by the distribution of P-values for the Shapiro–Wilk test for normality). All subsequent calculations used the transformed values.

2.4. Filtering transcripts

All ORFS (open read frame) were identified within the assembled transcripts using the TransDecoder program v4.1.0 (https://github.com/TransDecoder) with ORFs at least 100 codons. For further analysis, the final assembly contains only one transcript as a representative at each locus, with at least TPM more or equal 1 in at least one tissue. After the filtering, 34.293 transcripts were used to further analysis.

2.5. Annotation and Analysis of orthologous protein groups

The final assembled transcripts were annotated with descriptions and gene ontology (GO) classes by using the PANNZER software v.1.0 [24] against UniProtKB/Swiss to map alternative descriptions for the UniProt entries. We also compared all the transcripts against the NCBI CDD database [25] using an expect value threshold of 1e-10 and extracting only the top hit for each sequence.

Protein sequences from four other plant species were obtained to identify ortholog groups. *Elaeis guineensis, Musa acuminata, Phoenix dactylifera and Oryza sativa* (output group) protein sequences were downloaded from NCBI. OrthoFinder v1.1.8 [26] was used to identify groups of orthologs between the species. We estimated the rate of gene duplication with the software BadiRate v.1.35 [27] using the five species described (including *Acrocomia aculeata* – macaúba palm).

2.6. Phylogenetic Analysis

Therefore, the obtained sequences were aligned using alignment algorithms as implemented in MAFFT v.7.310 [28]. MAFFT was run in default mode with the maximum number of iterations set to 1000. Maximum likelihood phylogenetic analysis was performed using the RAxML software v.8.2.11 [29]. Bootstrap values were generated with 1.000 replicates.

2.7. Gene Specificity Analysis

In Tissue-specific Genes Database (TiSGeD), a statistic specificity measure (SPM), was proposed to measure the specificity of genes [30]. Following this method, we transformed each expression profile of gene *x* into a vector Xp, then the tissue specificity of gene *x* in tissue *i* was determined by the ratio of vector Xi's scalar projection in the direction of vector Xp against the length of Xp.

2.7.1. Specificity Measure (SPM) = ||X||/||Xp||

The SPM ranges from 0 to 1.0, and 1.0 means the highest tissue specificity. In every experiment, the SPM value for each transcript was calculated among all tissues, and those transcripts detected as top 5 % with the average SPM higher than 0.95 were selected as TS candidates.

3. RESULTS

3.1. Sequencing of macaúba palm transcriptome and De novo assembly

About 567 million raw reads with an average read length of 100 bp were generated, with an average of 27 million per tissue from sequencing the transcriptomes from the leaves, leaf sheath, bulb, root, female flowers, male flowers samples (Table 1). Subsequently, the adaptor sequences and low-quality reads were trimmed during sequence preprocessing. Following the pre-processing of reads, 522.490.624 high-quality reads (92.08 %) were retained for assembly.

Table 1. Number of reads and rRNA contamination percentage of samples of leaves, leaf sheaves, bulb, root, male and female flowers (plant 4 – Dourado; plant 12 – Amparo; plant 20 - Pedreira), and ripening fruits (plant 7 - S. Antônio de Posse, plant 5 - Amparo, plant 8 – Ibituruna).

Tissue	Sample	RNAr %	Reads
Leaf	Sample 4	0.89	20938255
	Sample 12	0.93	24284522
	Sample 20	0.79	23053873
Leaf Sheath	Sample 4	0.87	25450315
	Sample 12	1.77	29394232
	Sample 20	1.22	27501990
Bulb	Sample 4	1.55	23313242
	Sample 12	2.44	20773350
	Sample 20	2.37	25893211
Root	Sample 4	1.13	23561512
	Sample 12	1.26	24696967
	Sample 20	1.25	21595079
Female Flower	♀1	13.18	27824979
	♀2	46.13	28635910
	♀3	13.30	78987778
Male Flower	♂1	6.43	22487082
	♂2	19.52	22564627
	♂3	22.9	24850587
Ripe fruit	Sample 7	3.18	27617509
	Sample 5	1.34	21286807
	Sample 8	3.59	22681101
Total			**567392928**

Using the Trinity assembler and TransDecoder on the cleaned RNA-seq of the individual tissue, a sum of 34.293 transcripts were generated. All libraries were joined in a single reference and the transcripts were assembled according to predicted ORFs longer than 100 amino acids, representing a locus in macaúba genome, and FPKM > 1 in at least one tissue. Since they included various transcript isoforms, we removed redundant contigs. The removal of these contigs generated a non-redundant set of transcripts.

The final set of representative transcripts resulted in transcriptome with an N50 length of 1483 bp, the longest transcript length of 8698 bp and the shortest transcript length of 297 bp.

PANNZER software and CDD database were applied to the derived protein sequences. 5062 transcripts (14.76%) had an annotation in UniProtKB database by PANNZER and 25.246 transcripts (73.61%) were annotated by CDD.

3.2. Gene expression

The abundance of transcripts was quantified using the high-speed transcript quantification tool Kallisto as the median TPM (transcripts per million) value calculated from counts of reads mapped against the transcripts of the 7 tissues and 21 samples. The highest expressed transcript was TRINITY_DN35918_c5_g2, which was mainly expressed in female flower tissue with TPM values of 43707.01, 159627.97, 47684.97 in the three samples; the same transcripts had TPM values of 21968.53, 67238.47, 63189.58 in male flower tissues. According to the TPM values, the ten more expressed transcripts are listed in Table 2.

Table 2. List of ten highly expressed transcript of all macaúba palm tissues with PANNZER annotation, TPM value, CCD Annotation and Identity percentage by CDD.

	Transcript ID	PANNZER	TPM (Transcripts per million)			CDD Description
			S 4	S 12	S 20	
Leaf	TRINITY_DN11684_c0_g1	--	60623.45	56434.69	55816.77	RuBisCO_small, Ribulose bisphosphate carboxylase, small chain.
	TRINITY_DN28304_c0_g1	--	25267.01	26505.83	20911.79	ribulose bisphosphate carboxylase / (RCA)
	TRINITY_DN5419_c0_g1	--	24155.56	10646.39	8801.6	Chloroa_b-bind, Chlorophyll A-B binding protein.
	TRINITY_DN6089_c0_g1	--	9493.15	2881.74	4130.04	Chloroa_b-bind, Chlorophyll A-B binding protein.
	TRINITY_DN57842_c0_g1	--	9325	3049.13	3523.75	PSI_PSAK, Photosystem I psaG/psaK.
	TRINITY_DN15380_c0_g1	Chloroplast chlorophyll a/b-binding protein	8055.05	1159.92	2033.15	Chloroa_b-bind, Chlorophyll A-B binding protein.
	TRINITY_DN54193_c0_g1	Chloroplast chlorophyll a/b binding protein	7420.21	2716.88	3733.43	Chloroa_b-bind, Chlorophyll A-B binding protein.
	TRINITY_DN27835_c0_g1	--	7132.8	8861.94	8566.97	Glycolytic, Fructose-bisphosphate aldolase class-I.
	TRINITY_DN75268_c0_g1	--	7086.51	3207.4	5283.61	--
	TRINITY_DN25375_c0_g1	--	6923.87	2987.69	3157.69	Chloroa_b-bind, Chlorophyll A-B binding protein.
			S 4	S 12	S 20	
Leaf sheath	TRINITY_DN32518_c0_g1	--	16867.43	8440.59	35.66	B_lectin, Bulb-type mannose-specific lectin.
	TRINITY_DN11684_c0_g1	--	14705.26	4006.02	13264.63	RuBisCO_small, Ribulose bisphosphate carboxylase
	TRINITY_DN5419_c0_g1	--	14352.69	6660.69	2155.09	Chloroa_b-bind, Chlorophyll A-B binding protein.
	TRINITY_DN67292_c0_g1	--	8783.24	12964.53	1544.2	nsLTP1: Non-specific lipid-

						transfer protein type 1 subfamily.
	TRINITY_DN22886_c0_g1	--	8584.03	8021.54	10520.34	--
	TRINITY_DN32159_c0_g1	--	5353.76	5318.99	2549.89	TT_ORF1, TT viral orf 1.
	TRINITY_DN28304_c0_g1	--	4259.91	1043.93	3394.72	ribulose bisphosphate carboxylase/ (RCA)
	TRINITY_DN32851_c0_g1	--	4213.37	4749.34	4234.07	TCTP, Translationally controlled tumour protein.
	TRINITY_DN32165_c0_g1	--	3396.96	2289.71	6961.88	Peptidase_C1, Papain family cysteine protease.
	TRINITY_DN31485_c0_g1	--	2790.72	3574.21	1989.66	ChaC, ChaC-like protein.
			S 4	S 12	S 20	
Bulb	TRINITY_DN35302_c5_g1	--	10164.7	8234.2	9967.61	alpha_tubulin, The alpha-tubulin family.
	TRINITY_DN34463_c0_g1	--	9865.71	1305.79	2038.22	BURP domain.
	TRINITY_DN22886_c0_g1	--	9398.93	5658.15	4363.61	--
	TRINITY_DN32851_c0_g1	--	8144.09	6461.18	6886.83	TCTP, Translationally controlled tumour protein.
	TRINITY_DN27205_c0_g1	--	6484.29	7863.73	3180.74	--
	TRINITY_DN35610_c0_g3	--	4604.02	2273.83	256.65	BURP domain.
	TRINITY_DN19170_c0_g1	--	4568.36	613.62	1108.18	Metallothio_2, Metallothionein
	TRINITY_DN33510_c0_g1	--	4431.64	6095.2	3879.91	RRM_HP0827_like, RNA recognition motif in Helicobacter pylori
	TRINITY_DN35918_c5_g2	--	3670.86	7394.08	8914.78	--
	TRINITY_DN23914_c0_g1	--	3476.81	2466.47	2367.01	UBQ, Ubiquitin
			S 4	S 12	S 20	
Root	TRINITY_DN22886_c0_g1	--	6742.68	5560.18	6052.43	--
	TRINITY_DN23914_c0_g1	--	4691.48	4600.04	5412.8	UBQ, Ubiquitin
	TRINITY_DN32851_c0_g1	--	4654.19	4650.69	6228.46	TCTP, Translationally controlled tumour protein.

	TRINITY_DN33485_c0_g1	--	3460.89	2837.31	1938.43	Peptidase_C1, Papain family cysteine protease.
	TRINITY_DN34068_c4_g1	--	3337.56	1426.79	1588.01	Dimerization domain.
	TRINITY_DN34775_c0_g2	--	3334.94	553.54	2018.38	GST_C_Tau, C-terminal, alpha helical domain of Class Tau Glutathione S-transferases.
	TRINITY_DN28574_c0_g1	--	2691.5	4024.34	4873.73	TT_ORF1, TT viral orf 1.
	TRINITY_DN35302_c5_g1	--	2374.81	1890.9	1606.66	alpha_tubulin family.
	TRINITY_DN33510_c0_g1	--	2300.15	3297.07	2878.16	RRM_HP0827_like, RNA recognition motif in Helicobacter pylori
	TRINITY_DN45547_c0_g1	Polyubiquitin	2195.83	1677.99	1447.17	UBQ, Ubiquitin
Female flower			**1 ♀**	**2 ♀**	**3 ♀**	
	TRINITY_DN35918_c5_g2	--	43707.01	159627.97	47684.97	--
	TRINITY_DN26358_c0_g1	Transcript antisense to ribosomal RNA protein	32613.16	130600.43	34496.28	--
	TRINITY_DN35918_c6_g1	--	20527.75	80360.71	21747.01	Transcriptional regulator ICP4
	TRINITY_DN31899_c0_g1	--	16254.47	62531.76	14557.84	DNA polymerase III subunits gamma and tau
	TRINITY_DN27350_c0_g1	--	8881.3	2834.87	8276	nsLTP1: Non-specific lipid-transfer protein type 1 subfamily.
	TRINITY_DN32851_c0_g1	--	4257.61	2459.87	5377.2	TCTP, Translationally controlled tumour protein.
	TRINITY_DN45547_c0_g1	Polyubiquitin	4253.25	1886.38	3453.88	UBQ, Ubiquitin
	TRINITY_DN16559_c0_g1	--	4226.11	825.58	2050.3	HSP20/alpha crystallin family.
	TRINITY_DN35579_c0_g3	--	3645.45	1898.98	4057	NAD_binding_5, Myo-inositol-1-phosphate synthase.

	TRINITY_DN35302_c5_g1	--	3447.73	2464.77	1732.32	alpha_tubulin family.
Male flower			1 ♂	2 ♂	3 ♂	
	TRINITY_DN76919_c0_g1	--	56094.67	10.78	11.4	--
	TRINITY_DN34227_c1_g1	--	33301.19	416.23	2476.57	chitinase_glyco_hydro_19, Glycoside hydrolase family 19 chitinase domain.
	TRINITY_DN35918_c5_g2	--	21968.53	67238.47	63189.58	--
	TRINITY_DN27350_c0_g1	--	16771.95	5542.12	7925.39	nsLTP1: Non-specific lipid-transfer protein type 1 subfamily.
	TRINITY_DN26358_c0_g1	Transcript antisense to ribosomal RNA protein	15679.84	40813.54	42151.39	--
	TRINITY_DN35918_c6_g1	--	10528.91	37996.77	47077.82	transcriptional regulator ICP4
	TRINITY_DN24722_c0_g1	--	7551.31	1622.31	1686.02	ACBP, Acyl CoA binding protein (ACBP).
	TRINITY_DN88158_c0_g1	--	7170.38	3532.29	3920.9	--
	TRINITY_DN35751_c1_g1	--	6763.7	3323.63	2451.09	--
	TRINITY_DN31899_c0_g1	--	6600.77	16304.15	17436.55	DNA polymerase III subunits gamma and tau.
Fruit			S 7	S 5	S 8	
	TRINITY_DN32263_c0_g1	--	20049.09	900.04	6941.41	THI4, Ribulose 1,5-bisphosphate synthetase.
	TRINITY_DN34838_c1_g2	--	9830.79	1967.95	3125.25	acyl- [acyl-carrier protein] desaturase.
	TRINITY_DN35918_c5_g2	--	8120.33	2358.12	8611.7	--
	TRINITY_DN57470_c0_g1	--	6506.86	866.54	6981.03	--
	TRINITY_DN32051_c0_g1	--	6418.36	4090.9	2544.56	Wali7.
	TRINITY_DN26358_c0_g1	Transcript antisense to ribosomal RNA protein	5372.7	1565.49	7168.95	--
	TRINITY_DN32851_c0_g1	--	5145.57	1153.95	4444.43	TCTP, Translationally controlled tumour protein.

TRINITY_DN35918_c6_g1	--	4977.51	1543.55	4371.98	transcriptional regulator ICP4.
TRINITY_DN35580_c0_g1	--	4509.37	1956.9	1311.51	--
TRINITY_DN31243_c1_g1	--	4252.69	983.59	2593.17	rne, ribonuclease E.

Almost all transcripts highly expressed in leaf and leaf sheaves tissues are related to photosynthesis and carbon metabolism. Also, the transcript TRINITY_DN32851_c0_g1 annotated as TCTP (Translationally controlled tumor protein) was highly expressed in leaf sheaves, bulb, root, female flower and fruit tissues. TCTP is regulated at both translational and posttranslational levels and it is associated with several biochemical and cellular functions [31]. In *Arabidopsis thaliana,* knock-out of TCTP leads to impaired pollen tube growth; TCTP silencing by RNA interference showed slows vegetative growth, reduced leaf expansion, and lateral root formation, and root hair development is impaired [32].

The SPM (specificity measure) value for each transcript was calculated and those transcripts detected as top 5 % in at least two different experiments with the average SPM higher than 0.95 were considered tissue-specific (TS) transcripts.

Considering SPM values higher than 0.95, eight transcripts are specific for fruit tissue, two in leaf tissue, eight in bulb tissue, five in female flowers, 14 in male flowers, 8053 in root tissues and none in leaf sheath tissues. Also, 1651 transcripts had no specificity in anyone tissue. To observe target transcripts related to desirable traits, a TPM cutoff ≥ 20 in at least two samples was applied to consider a TS transcript and highly expressed. Since the applied threshold, it was no observed TS transcripts with high expression in leaf, bulb and female flower tissues.

In male flower tissue, eight TS transcripts were highly expressed (Table 3) and two transcripts, TRINITY_DN14033_c0_g1 and TRINITY_DN22118_c0_g1, were annotated as Non-specific lipid-transfer protein (nsLTP) with 33,33% and 35,61% of identity, respectively. Also, other transcripts annotated as nsLTP were highly expressed in female and male flowers (TRINITY_DN27350_c0_g1 and TRINITY_DN27350_c0_g1, respectively), but not specific in these tissues (Table 1). It is a well-known multigene family, and *nsLTP* expression has been detected in the flowers of many plant species and may be associated with reproduction [33, 34]. The protein plays a role in pollen and/or anther development, such as in pollen formation and germination [35] and in the protection of reproductive tissues from environmental stress [36].

Four transcripts (TRINITY_DN64786_c0_g1, TRINITY_DN25411_c0_g1, TRINITY_DN83931_c0_g1, TRINITY_DN68783_c0_g1) presented high TPM values In fruit tissue in fruit tissue. TRINITY_DN64786_c0_g1 and TRINITY_DN25411_c0_g1 were the main expressed transcripts, which they had an identity of 34,74% with Enoyl-[acyl-carrier-protein (ACP)] reductase – FabL and 27,47% of identity with Trans-2-enoyl-CoA reductase (TER), an enoyl-ACP reductase isoform - FabI. Enoyl-ACP reductases are known to have isoforms of this enzyme, which FabL was described in *Brassica napus* [37] (Table 3); Fawcett et al. [38] in *Brassica napus* reported a correlation between the increased level of FabL transcript and lipid deposition in the seed. Also, Budiani et al. [39] observed the same in African oil palm, which expression increased at the beginning and just before the period of active oil biosynthesis, respectively. Trans-2-enoyl-CoA reductase (TER) was observed by Wong et al. [40] in high-oil yielding oil palm in comparison with low-oil yielding oil palm. Ruzlan et al. [41] provide a novel method for predicting oil palm yield with a set of SNPs. Notable, one SNP into exon 3 of Mitochondrial Trans-2-enoyl-CoA reductase (TER) is associated with the high-oil-production trait in the population studied. Interestingly, these transcripts are involved in the fatty acid biosynthesis, and this stage displays major differences between oil palm and date palm during fruit ripening [42].

Table 3. TPM values of tissue-specific transcripts (SPM ≥ 0.95) male flowers and fruit tissues. TPM values ≥ 0.95 in pink-colored box.

TISSUE	Transcript ID	SPM	Male Flower			Ripe fruits		
			1 ♂	2 ♂	3 ♂	S 7	S 5	S 8
MALE FLOWER	TRINITY_DN14033_c0_g1	1.0	6.8	**31103.8**	**34146.8**	0.0	0.0	0.0
	TRINITY_DN29914_c0_g1	1.0	**264.9**	**30177.9**	**49469.3**	0.0	0.0	0.0
	TRINITY_DN22118_c0_g1	1.0	5.5	**28276.2**	**25560.8**	0.0	0.0	0.3
	TRINITY_DN59694_c0_g1	1.0	5.0	**17070.1**	**37987.3**	0.0	0.1	0.4
	TRINITY_DN8432_c0_g1	1.0	0.8	**1632.0**	**1576.7**	0.5	0.7	0.0
	TRINITY_DN542_c0_g1	1.0	2.7	**139.4**	**221.7**	0.0	0.0	0.3
	TRINITY_DN38356_c0_g1	1.0	0.0	**80.6**	**20.4**	0.0	0.0	0.0
	TRINITY_DN81374_c0_g1	1.0	**97.7**	**22.6**	13.3	0.0	0.0	0.0
	TRINITY_DN16557_c0_g1	1.0	0.1	8.4	0.6	3.6	0.0	0.0
	TRINITY_DN85322_c0_g1	1.0	**25.4**	4.6	2.5	0.0	0.0	0.0
	TRINITY_DN56097_c0_g1	1.0	0.1	2.1	0.3	3.7	0.0	0.0
	TRINITY_DN38933_c0_g1	1.0	8.9	0.9	0.5	0.0	0.0	0.0
	TRINITY_DN85764_c0_g1	1.0	2.5	0.2	1.8	0.0	0.0	0.6
	TRINITY_DN68484_c0_g1	1.0	0.1	0.0	0.1	0.0	0.0	0.0
FRUIT	TRINITY_DN64786_c0_g1	1.0	3.9	0.0	0.1	**2132.5**	**12252.1**	**146.4**
	TRINITY_DN25411_c0_g1	1.0	1.3	0.2	0.4	**1500.2**	**9657.1**	**78.6**
	TRINITY_DN83931_c0_g1	1.0	0.3	0.5	0.0	**486.5**	**944.5**	**133.9**
	TRINITY_DN68783_c0_g1	1.0	0.2	0.0	0.0	**164.2**	**165.5**	5.6
	TRINITY_DN31531_c0_g1	1.0	0.0	8.8	0.0	2.1	1.2	0.0
	TRINITY_DN85732_c0_g1	1.0	0.0	0.0	0.0	1.0	0.0	2.1
	TRINITY_DN78977_c0_g1	1.0	0.0	0.0	0.0	0.1	0.0	3.3
	TRINITY_DN4080_c0_g1	1.0	0.0	0.3	0.0	0.0	4.0	0.4

From root tissues, twelve transcripts were annotated as serine/threonine kinases Interleukin-1 Receptor Associated Kinases, twelve were similarly to CAMK (calcium and calmodulin (CaM) stimulated STKs), and forty transcripts were annotated transcript as a mitogen-activated protein kinase (MAPK). Also, other 72 transcripts related to STKs were observed as highly expressed and root specific. STKs play an important role in signal transduction in plants, being involved in many cellular processes, including developmental, hormonal, biotic, and abiotic stress signaling [43, 44]. Patankar *et al.* [44] observed a high salt-tolerance of recombinant yeasts with date palm genes of serine/threonine protein kinases (*STKs*). Moreover, we found 341 TS transcripts highly expressed in root tissue (additional file 1). Including all expressed TS transcripts, important classes of proteins were annotated, suggesting that macaúba palm has a rusticity and high tolerance to abiotic and biotic stress. Transcripts similarly to Serine/threonine kinases (nine transcripts) and Cytochrome P450 (14 transcripts) presented high TPM values in root tissues. Cytochrome P450 is involved in a wide range of secondary metabolic pathways, since biosynthesis of important molecules, biosynthesis of plant growth regulators, and involved in stress tolerance [45].

Additionally, family proteins related to pathogen defense were found in root tissue. Hevamine, a chitinase pathogenesis-related proteins (PR) (TRINITY_DN34687_c0_g1), C2 domain-SCR2 (TRINITY_DN24495_c0_g1), pathogenesis-related (PR-1) protein (TRINITY_DN29805_c0_g1, TRINITY_DN4120_c0_g1, TRINITY_DN41521_c0_g1), PR10, first identified as a major pollen allergen (Bet-v1) (TRINITY_DN17745_c0_g1_i1, TRINITY_DN21727_c0_g1, TRINITY_DN69779_c0_g1), PR-5 protein family, as known thaumatin-like proteins (TLPs) (TRINITY_DN24378_c0_g1), conserved fungal-specific extra-cellular membrane-spanning (CFEM) protein domain (TRINITY_DN34234_c0_g1, TRINITY_DN72190_c0_g1, TRINITY_DN29454_c0_g2, TRINITY_DN49946_c0_g1, TRINITY_DN1278_c0_g1). Three transcripts annotated as pathogenesis-related proteins presented high TPM values (TRINITY_DN29805_c0_g1 - PR-1; TRINITY_DN17745_c0_g1 – Bet-v1; TRINITY_DN24378_c0_g1 – thaumatin-like protein) (additional file 1). All these protein families play a major role in plant defense against various biotic and abiotic stresses, mainly in fungal and bacterial defense [46, 47].

3.3. Expanded gene families in the macaúba palm

Orthogroups (i.e., gene families) in OrthoFinder are defined as homologous genes that have evolved by speciation event in their last common ancestor.

Phylogenetic analysis was undertaken with the transcriptome data of *Elaeis guineensis, Musa acuminata, Phoenix dactylifera, and Oryza sativa* genomes. Based on the concatenated alignment of orthologs from the five species with transcriptome data (macaúba palm) and genome data, we used both Maximum Likelihood (ML) approach to reconstruct phylogenetic trees. Maximum likelihood bootstrap values (shown as percentages) were 100% at all nodes of the phylogenetic tree (Fig. 1).

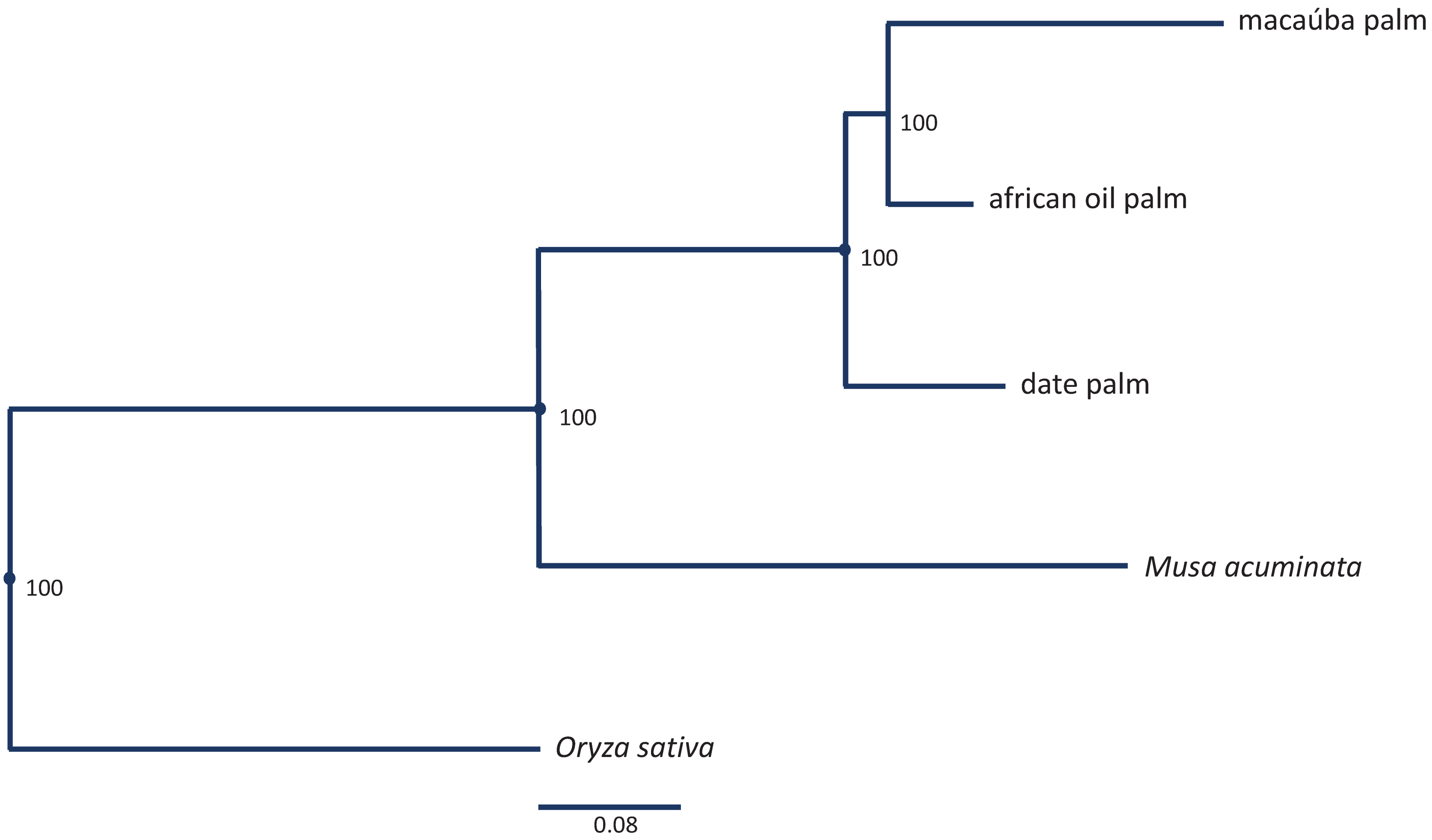

Fig. 1. Phylogenetic relationships between macaúba palm, *Elaeis guineensis* (African oil palm), *Musa acuminata* (Banana), *Phoenix dactylifera* (date palm) *and Oryza sativa* (rice) by transcriptome data. Numbers at each node are the bootstrap values percentages.

Macaúba palm and African oil palm were first clustered into the same clade, which date palm as a sister clade, establishing a palm tree cluster. That data is consistent since African Oil Palm and Macaúba palm are from the same Tribe (Tribe Cocoseae; Family Arecoideae) and Date palm belongs to another subfamily (Coryphoideae). All the palm trees (Arecaceae) belong to a monophyletic group including five subfamilies, 28 tribes and 27 subtribes [48]. Despite a reduced number of palm tree samples in our data, monophyly of Arecaceae family is supported by the 100% of bootstraps.

A total of 15.283 orthogroups (at least one species representing the group) were identified using *Elaeis guineensis, Musa acuminata, Phoenix dactylifera,* and *Oryza sativa* protein dataset. Among all groups, 12.664 family groups have one representant of macaúba palm.

BadiRate analysis was applied to examine to characterize the expansion of family genes. The analysis of turnover rates showed 24 gene families in macaúba palm species expanded through gene gain compared with African oil palm (Table 4), indicating that the gene gain confers an ecological advantage for survival and reproduction in different ecological ambient between macaúba palm and oil palm.

Table 4. Gene families expanded through gene gain in macaúba palm compared with African Oil Palm (*Elaeis guineensis*) and a common ancestor. In the table: Gene family name; the number of African Oil Palm genes; the number of macaúba palm genes; the number of common ancestor genes; P value; Gene function according to BlastX.

Gene family	*Elaeis guineensis*	Macaúba palm	Ancestor: Macaúba palm/ Elaeis sp.	P value (FDR)	BlastX
OG0000015	10	18	10	0.0066990	--
OG0000024	8	47	5	0.0000000	Probably Retrovirus-related Pol polyprotein from transposon TNT 1-94
OG0000038	7	25	4	0.0000000	Polyubiquitin
OG0000057	3	35	3	0.0000000	Pentatricopeptide repeat-containing protein
OG0000077	33	10	4	0.0202369	--
OG0000193	5	11	5	0.0290704	--

OG0000487	4	9	3	0.0118412	SCAR-like protein 1
OG0000659	2	7	2	0.0318502	Type IV inositol polyphosphate 5-phosphatase 9
OG0000697	6	7	2	0.0318502	STRUBBELIG-RECEPTOR FAMILY 5-like
OG0000748	1	19	1	0.0000000	Pentatricopeptide repeat-containing protein
OG0001176	0	19	0	0.0000000	--
OG0001613	2	6	1	0.0118412	Serine/threonine-protein kinase D6PKL
OG0001853	2	6	1	0.0118412	--
OG0002240	1	10	1	0.0000024	Pentatricopeptide repeat-containing protein
OG0002627	4	8	2	0.0065115	Pentatricopeptide repeat-containing protein
OG0002773	1	7	1	0.0018332	IKU2 protein (Ser/Thr protein kinase Family)
OG0003836	0	4	0	0.0090916	Putative disease resistance protein RGA
OG0004289	2	7	1	0.0018332	Heat shock 70 kda protein-like
OG0005442	1	6	1	0.0118412	--
OG0005452	0	7	1	0.0018332	Elongation factor-1 gene
OG0010763	0	4	0	0.0090916	Heat shock cognate 70 kda protein
OG0010767	0	4	0	0.0090916	--
OG0012394	0	4	0	0.0090916	Glyceraldehyde-3-phosphate dehydrogenase 2
OG0012409	0	4	0	0.0090916	Elongation factor 2-like protein

Four families were like to Pentatricopeptide repeat-containing protein family (PPR) (OG0000057, OG0000748, OG0002240, OG0002627). They are one of the major mediators of organelle post-transcriptional control and all PPR proteins contain tandemly repeated sequence motifs (the PPR motifs) which can vary in number. A typical PPR protein is targeted to mitochondria or chloroplasts, binds one or several organellar transcripts, and influences their expression by altering RNA sequence, turnover, processing, or translation. Their combined action has profound effects on organelle biogenesis and function and, consequently, on photosynthesis, respiration, plant development, and environmental responses [49, 50].

The second major group of protein families is leucine-rich repeat receptor-like protein kinase (LRR-RLK) (family OG0000697) and serine/threonine-protein kinase (families OG0001613, OG0002773). The first family, OG0000697, has one transcript highly expressed - TRINITY_DN6019_c0_g1 – and similar with *STRUBBELIG-RECEPTOR FAMILY* (*SRF*), which it is a monophyletic family represented by *STRUBBELIG* (*SUB*) and eight additional members. Kinases from this family *SUB* plays an important role in plant development as it affects cellular morphogenesis, root hair specification (named as *SCRAMBLED* - *SCM*) and is required for the orientation of the cell division plane and the control of cell number, cell size, and cell shape [51–53]. Moreover, the other families (OG0002773 and OG0001613) were similar to proteins IRAK and D6PKL, respectively. The family OG0002773, which contain four transcripts highly expressed, was annotated as IRAK serine/threonine-protein kinase, interleukin receptor-associated kinases. In plants, this RLK (receptor-like kinase) family is related to two categories of biological function. The categories include growth/development control and plant-microbe interactions/defense responses, as *Arabidopsis thaliana* BAK1, involved in plant resistance to pathogen infection and herbivore attack, and BR (brassinosteroid)-regulated plant development; and CLAVATA1 (CLV1) in meristem maintenance [43, 54, 55].

Four transcripts with high TPM values in family OG0001613 (TRINITY_DN11200_c0_g1, TRINITY_DN25915_c0_g1, TRINITY_DN34203_c0_g1, TRINITY_DN34203_c1_g3) were like D6PKL (D6 PROTEIN KINASE-LIKE), a kinase that regulates the auxin transport activity of PIN auxin efflux facilitators by direct phosphorylation. D6PK-mediated PIN phosphorylation promotes auxin transport in the hypocotyl and this is a prerequisite for PHOT1-dependent hypocotyl bending (PHOT1 – phototropin) which regulates light – responses and promote stomatal opening to regulate CO_2 uptake for photosynthesis and water loss by transpiration [56–58].

Moreover, it was shown six exclusive families of macaúba palm and seven had no similarity with proteins in sequence databases (Table 4). Seventeen gene families had orthologous genes with high similarity (> 90%) in the NCBI database. Of all gene families observed, 11 families contain at least one TS transcript, ranging from 1 transcript for families OG0000077, OG0001853, OG0010763, OG0012394, OG0012409 up to 11 TS transcripts for the

OG0001176 family (Table 5). All expanded family with respective transcripts and SPM values are listed in additional file 2.

Table 5. Expanded families of macaúba palm with respective transcripts and SPM values for each macaúba palm tissue.

Expanded Family	Fruit pulp	Leaf	Leaf sheath	Bulb	Root	Female flower	Male flower
OG0000024							
TRINITY_DN17383_c0_g1	0.01	0.00	0.13	0.21	0.97	0.06	0.03
TRINITY_DN5271_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN55745_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN64334_c0_g1	0.00	0.00	0.00	0.04	1.00	0.00	0.00
TRINITY_DN87291_c0_g1	0.00	0.07	0.00	0.00	1.00	0.00	0.00
OG0000038							
TRINITY_DN11382_c0_g2	0.00	0.00	0.00	0.19	0.98	0.00	0.00
TRINITY_DN14629_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN16579_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN17682_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN19913_c0_g1	0.04	0.05	0.04	0.20	0.96	0.16	0.09
TRINITY_DN21380_c0_g1	0.00	0.04	0.00	0.00	1.00	0.00	0.00
TRINITY_DN25769_c1_g1	0.05	0.09	0.09	0.05	0.99	0.00	0.00
TRINITY_DN26294_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN34075_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g2	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g3	0.00	0.00	0.00	0.00	1.00	0.01	0.03
OG0000077							
TRINITY_DN24002_c0_g1	0.03	0.02	0.16	0.08	0.98	0.06	0.05
OG0000659							
TRINITY_DN16790_c0_g1	0.00	0.00	0.00	0.03	0.99	0.00	0.16
TRINITY_DN25697_c0_g1	0.00	0.00	0.00	0.01	0.99	0.00	0.15
TRINITY_DN58518_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN67071_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0001176							
TRINITY_DN12752_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN15658_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN16748_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN19613_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00

TRINITY_DN21803_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN21989_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN22432_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN23091_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN25199_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN27832_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN28824_c0_g1	0.01	0.18	0.14	0.10	0.97	0.00	0.00
TRINITY_DN54299_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN57180_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN57390_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN62765_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN65272_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN68960_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0001853							
TRINITY_DN23105_c0_g1	0.00	0.00	0.00	0.00	1.00	0.04	0.06
OG0005452							
TRINITY_DN14350_c0_g1	0.00	0.00	0.00	0.05	1.00	0.00	0.00
TRINITY_DN30150_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN33407_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN73525_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0010763							
TRINITY_DN31034_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0010767							
TRINITY_DN16764_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN9921_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN9921_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0012394							
TRINITY_DN73669_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0012409							
TRINITY_DN25641_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00

Applying a TPM cutoff of 20 to estimate the expression of gene families, a small number of transcripts had significant values, when considering the analyzed tissues (additional file 3). Including all data, family OG0000038 presented 12 TS transcripts, being six expressed at a higher level in root: TRINITY_DN21380_c0_g1, TRINITY_DN25769_c1_g1, TRINITY_DN26294_c0_g1, TRINITY_DN34075_c1_g1, TRINITY_DN35187_c4_g1, TRINITY_DN35187_c4_g2, TRINITY_DN35187_c4_g3. (additional file 4). OG0000038 family was annotated as

Polyubiquitin/Ubiquitin, a highly conserved family with many regulatory functions. The covalent attachment of ubiquitin to proteins (ubiquitination) can influence activity, abundance, or localization. The most prominent function of the ubiquitin is its role in selective proteolysis. Interestingly, all TS transcripts were specific for root tissue, suggesting differences between oil palm and macaúba palm.

3.4. Transcripts related to Drought stress

We analyzed the gene expression related to drought stress processes. The transcripts of macaúba palm were mapped to set of orthologs in *Arabidopsis thaliana, Zea mays*, *Oryza sativa*, *Elaeis guineensis* from Drought DB – Drought Stress Gene Database [59]. The putative orthologs were grouped by family or orthogroups.

An amount of 227 macaúba palm transcripts were found in 114 families being each family containing at least one macaúba palm transcript as a member (additional file 5 and additional file 6). Two TS transcripts derived from seven tissue-specific (all being root specific) were highly expressed: TRINITY_DN28917_c0_g1 (calcium binding motif; calmodulin) and TRINITY_DN34169_c0_g1 (14-3-3 family protein). The latter transcript belongs to root-specific family OG0000024. Considering all families, 41 transcripts had no considerable expression (applying TPM cutoff = 20) (additional file 5 and additional file 6).

Families OG0001144, OG0001847, OG0002138, OG0000985, OG0000371, OG0000580, OG0000595, OG0000919, OG0000262, OG0000119 had no considerable expression in at least one sample analyzed.

3.4.1 Tissue-specific transcripts

Among families found by Orthofinder, two transcripts were highly expressed and specific to the root tissue: TRINITY_DN28917_c0_g1, similar in orthology with calmodulin and TRINITY_DN34169_c0_g1, like 14-3-3 family proteins (Fig. 2).

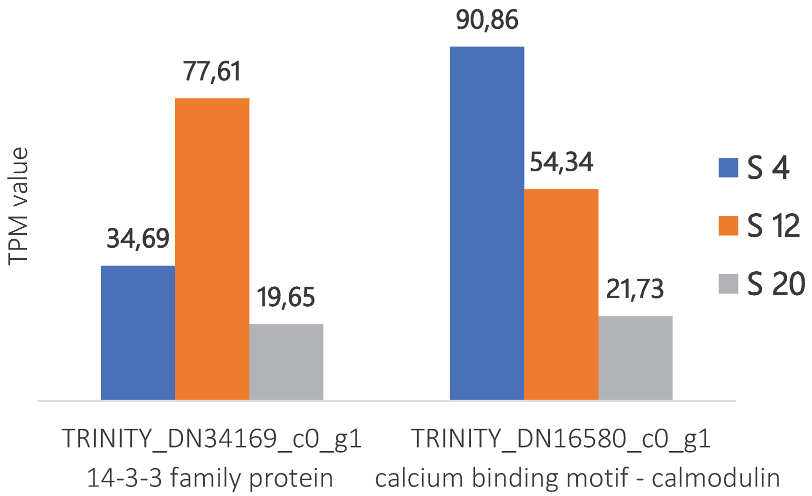

Fig. 2. TPM values of tissue-specific transcripts related to drought stress: TRINITY_DN28917_c0_g1 and TRINITY_DN34169_c0_g1 in samples 4, 12 and 20.

TRINITY_DN34169_c0_g1 transcript of the OG0000024 family was highly expressed in root tissue (TPM values of 34.69 – plant 4; 77.61 – plant 12; 19.65 – plant 20) and orthology data showed similarity to 14-3-3 family protein. This family protein represents a phosphoserine-binding protein that regulates a wide range of targets by protein-protein bonds. They are widely known to play a vital role in responses to abiotic and biotic stresses [60]. Yan et al. [61] provide one of the earliest evidence of the role in drought stress response. They observed that the overexpression of *Arabidopsis thaliana* 14-3-3λ in cotton results a 'stay-green' phenotype in transgenic plants, and increased drought tolerance, which wilted less and maintained higher photosynthesis. He et al. [62] observed that *Arabidopsis* plants overexpressing GRF9, a 14-3-3 protein, had higher root and plant growth, and carbon allocation in the root, essential mechanisms for an effective response to drought stress. Although we found the expression was high in two of the three root samples, it was expressed only in root tissue, showing a possible drought tolerance function.

The second-high expressed transcript is TRINITY_DN16580_c0_g1 (TPM values of 90.86 – plant 4; 54.34 – plant 12; 21.73 plant 20), which is orthologous to calmodulin (CaM), a Ca^{2+} binding protein. It is closely related to drought stress since osmotic stress leads to a transient increase in Ca^{2+} cellular, phytotoxic at high levels, and consequently activation of sensors and proteins involved in the cellular signaling process [63]. Xu et al. [64] observed an up-regulation calmodulin (OsMSR2) by different abiotic stresses, including cold, dry and heat in different tissues at different stages of development in rice; moreover, the ectopic expression in *Arabidopsis* conferred higher salt and drought tolerance, as well as altered expression of stress-responsive/ABA genes. Tomatoes overexpressing ShCML44 (like-calmodulin) presented higher tolerance to cold, drought and salinity and a higher germination rate and better seedling growth. Transgenic tomato showed more activity of antioxidant enzymes, gas exchange and water retention capacity with a lower accumulation of malondialdehyde and membrane damage under water stress and cold compared to wild-type [65].

3.4.2 Transcription factors

Transcription factors (TF) are essential elements for gene expression regulation, being relatively conserved from bacteria to humans and plants. From all the families obtained and considering all genes related to stress to drought, transcription factors were highly expressed in all tissues, as transcription factor MYC (family OG0000052), bZIP (families OG0000057 and OG0000069), MYB (OG0000315 and OG0000932), NF-Y (OG0000005 and OG0000012), ERF-ethylene-responsive transcription factor (OG0000796, OG0000112, OG0000209), AP2 (OG0001372), WRKY (OG0002791) and HSFA1 (family OG0000123) (Figure 3).

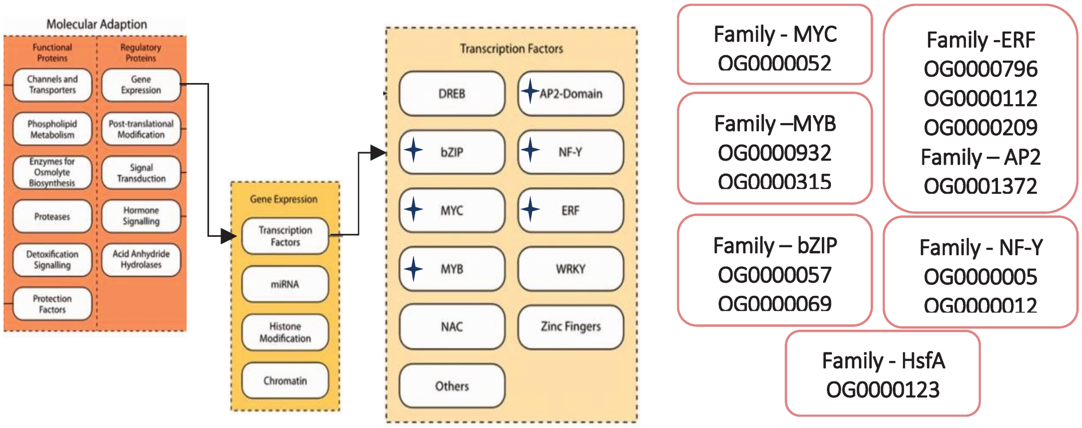

Fig. 3. Schematic overview of transcription factors families found in macaúba palm data. Layer 1 consists of the block molecular adaptation (orange). Layer 2 (dark yellow) shows a categorization

of gene expression pathways. Layer 3 (light yellow) represents specific transcription factors families involved in drought stress response.

The transcription factors MYC, bZIP, and MYB are regulated by ABA (abscisic acid), which plays an important role in stresses response activating signaling cascade by dehydration signal. In our data, the transcripts TRINITY_DN2149_c0_g1, TRINITY_DN27638_c0_g2, and TRINITY_DN31421_c1_g1 of the MYC family (OG0000052) showed high expression in all tissues. MYB families OG0000932 and OG0000315 showed that only the transcript TRINITY_DN33427_c1_g2 was highly expressed, with TPM ≥ 20 in the vegetative tissues and fruit. The MYC transcription factor has been reported providing cold tolerance ectopic expression of maize MYC-type ICE-like transcription factor in *Arabidopsis thaliana*, and both MYC and MYB function as a transcriptional activator on the ABA-inducible gene expression under drought stress in *Arabidopsis thaliana* [66, 67]. Moreover, the bZIP families OG0000057 and OG0000069 presented high expression in all tissues. In family OG0000057, transcripts TRINITY_DN28492_c0_g1 and TRINITY_DN63679_c0_g1 were highly expressed in all tissues, while TRINITY_DN21593_c0_g1 was not expressed in leaf tissue. From family OG0000069, TRINITY_DN22238_c0_g1 and TRINITY_DN28034_c0_g1 were expressed in leaf, leaf sheave, bulb, and root tissues. Its up-regulation leads to alleviation of oxidative stress symptoms, high expression of ABA/stress-responsive genes in *Arabidopsis* [68], and increases tolerance to drought and salinity in both *Arabidopsis*, cotton [69] and rice [70].

Transcripts of OG0000005 and OG0000012 families are orthologs to the transcription factor NF-Y, a plant TF with numerous members. In the first family, three transcripts (TRINITY_DN34990_c2_g1, TRINITY_DN34990_c2_g2, and TRINITY_DN34990_c2_g4) had high expression in the flower tissues (female and male), whereas the transcript TRINITY_DN28214_c0_g1 (OG0000012) presented high expression in all tissues studied. Members of the NF-Y family are known to exhibit diverse patterns of organ and tissue-specific expression [71]. They are post-transcriptionally regulated by natural small interfering RNAs (nat-siRNAs) and/or micro-RNAs (miRNAs) [72] and they are targets of the miR169 family, one of the largest and most conserved miRNA families in plants [73]. Li et al. [74] observed that NFYA5 overexpressing transgenic *Arabidopsis* show less leaf water loss and more resistance to drought stress compared to wild type. Leyva-González et al. [75] proposed that the level of NF-Y factor expression is part of an adaptive response to environmental conditions. They observed that NF-YA2, 7 and 10 overexpressing transgenic *Arabidopsis* were dwarf and senescent with greater tolerance to many abiotic stresses. Moreover, wheat TaNF-YA10-1 overexpressing *Arabidopsis* showed significant tolerance of drought and greater root length, although increased sensitivity to salinity [76].

The superfamily AP2/ERF (APETALA2/ethylene responsive factor) is one of the largest transcription factors in plants. The transcripts of ERF (ethylene-responsive factor) family were grouped in OG0000796, OG0000112, OG0000209 families. It was observed high expression only in family OG0000112, which transcript TRINITY_DN33317_c0_g1 were expressed in the fruit tissue. Omidvar et al. [77] observed an up-regulation of this superfamily, EgAP2-1 (ERE-binding protein), in the mesocarp of oil palm fruits in response to ethylene and abscisic acid, suggesting a regulatory role of this TF in ripening and development of palm fruits. The effect of this TF on fruits improved seed germination [78] and in regulates ripening of pear fruits [79]. Transcript TRINITY_DN30070_c0_g1 similar to AP2/ERF (family OG0001372) had high expression only in the root tissue, as found in sesame [80] and male flower. In rice, ERF overexpressing plants showed a higher survival rate and seed fixation when the plant was exposed to stress during the reproductive stage [81]. Our data, confirmed by the literature, suggest a possible involvement of these TFs the adaptive response to drought stress in vegetative and reproductive tissues.

Furthermore, family OG0000123 is composed of three transcripts like HSFA1 (heat-stress transcription factor A-1). Only the transcript TRINITY_DN27166_c0_g1 had high expression in leaf sheath, bulb, root and fruit tissues. HSFA1 is the most studied TF in plants under heat stress condition with different numbers of members per plant; *Arabidopsis thaliana* has four members - A1A, 1B, 1D, and 1E [82]. It is considered the main regulator in response to heat stress in plants since Mishra et al. [83] observed reduced thermotolerance in knockdown tomato plants. Its primary function was confirmed by Liu; Liao; and Charng [84], showing that the A1A, A1B, and A1D isoforms are essential for regulating the expression of different HS responsive genes. The family to TF WRKY had no high-expressed transcript.

Under natural conditions, plants are exposed to both thermal and drought stress, a condition that is becoming increasingly common since climate change associated with global warming. It is important to consider both drought and temperature together, as the combination increases the negative effects on the plant [85].

3.4.3 Kinases

One of the categories of stress-induced proteins is those that regulate gene expression and the transduction of stress response signals. Among the orthogroups provide by DroughtBD database, 11 families were classified as kinases, of which seven had high expression in the tissues analyzed.

Including the highly expressed kinases in the tissues, four families were similar to the kinases mediated or activated by calcium. The eight transcripts of the OG0000003 family were related to calcium-dependent protein kinase (CDPK) and six of them highly expressed (TRINITY_DN24067_c0_g1, TRINITY_DN31407_c0_g1, TRINITY_DN35120_c5_g1, and TRINITY_DN35503_c1_g1 were expressed in all tissues; TRINITY_DN34134_c0_g1 in leaf sheath, bulb, root, and fruit; TRINITY_DN35503_c1_g4 in root and fruit tissues), being all tissues represented. Moreover, transcripts from families OG0000107 (TRINITY_DN16962_c0_g1, TRINITY_DN35781_c0_g1, TRINITY_DN35781_c0_g6), OG0002906 (TRINITY_DN35120_c3_g1), and OG0000156 (TRINITY_DN34658_c0_g1, TRINITY_DN35028_c0_g1, TRINITY_DN35373_c0_g1) were similar with CBL-interacting protein kinase (CIPK) and they had high expression in all tissues (OG0000107, OG0002906) and in leaf sheath, bulb, root, and fruit (OG0000156).

Transient Ca^{+2} flux occurs under stress and development conditions and stimulates directly in these proteins. CDPKs has a Ca^{+2} binding and kinase activity, consequently forming "sensor-response" modules [86]. They are up-regulated when plants are in a drought stress condition, as observed in soybean [87] and in corn CDPK overexpressing *Arabidopsis*, wherein plants increased the survival rate on stress condition [88]. Furthermore, the CIPK protein has an N-terminal kinase domain and a C-terminal NAF-binding domain CBL-Ca^{+2}. When Ca^{+2} levels increase, CIPK phosphorylates the target proteins. It regulates the movements of stomata water, increasing the sensitivity of guard cells to ABA and regulating K^{+} uptake by guard cells, as observed in knockout mutants for AtCBL1, AtCBL9 and AtCIPK23 [89], in transgenic tobacco overexpressing BdCIPK31 from *Brachypodium distachyon* [90], and in *Arabidopsis* overexpressing wheat TaCIPK23 [91]. Stomatal closure is essential to prevent water loss, but it should be negative when there are combined stresses, such as drought and high temperatures.

Transcripts from the family OG0000023 had high expression in all tissues analyzed (TRINITY_DN28647_c0_g1 – not expressed in root; TRINITY_DN32053_c0_g1 – expressed in all tissues; TRINITY_DN34069_c1_g1 – not expressed in female flower; TRINITY_DN34211_c0_g1 – expressed in all tissues) and similarity to stress-activated protein kinase (SAPK), signaling plays essential roles in osmotic response signal transduction. SAPK phosphorylates the ABA signaling bZIP transcription factor [92, 93] and it is regulated positively by drought stress conditions. SAPK function in drought response has already been observed in root by [94]. The authors proposed that SAPK can act by (i) reducing water loss via the accumulation of solutes, promoting stomatal closure and increasing expression of stress response genes (ii) inducing expression of antioxidant enzyme genes to scavenge ROS.

Additionally, the family OG0000646 had only the transcript TRINITY_DN30116_c0_g1 expressed in root and leaf tissues, being like glycerol kinase (GK). Glycerol kinase is a key regulator involved in the catabolism of glycerol to glycerol-3-phosphate and dihydroxyacetone. Mutants of *Arabidopsis thaliana* gli1 lack glycerol kinase accumulate glycerol and are more resistant to abiotic stresses such as hyperosmotic, salt, freezing and desiccation [95]. The transcript TRINITY_DN34267_c3_g2 from family OG0000978, similar to phosphoenolpyruvate carboxykinase (PEPCK), were highly expressed in leaf, sheath, bulb, root and flower tissues. It catalyzes the conversion of oxaloacetate into phosphoenolpyruvate, which is a key enzyme in the gluconeogenesis route. In *Arabidopsis thaliana*, the expression of the PEPCK promoter is mainly on the expression of the PEPCK promoter is mainly on the guiding cells and the leaf trichomes. Also, the pck1 mutants reduced drought tolerance and exhibited greater stomatal conductance and wider stomatal apertures compared to wild-type [96]; study with *Camellia sinensis* PEPCK showed it was overexpressed under drought stress condition [97]. The high expression of these genes can be explained by the need to maintain the stomatal osmoregulation to close the guard cells by the solute accumulation, in order to control the water loss.

The orthogroups OG0000119 (LRR receptor-like serine / threonine-protein kinase RPK2), OG0000213 (G-type lectin S-receptor-like serine/threonine-protein kinase AT5g24080), OG0000580 (inositol polyphosphate multikinase beta-like) and OG0000985 (histidine kinase 1-like) did not show high expression transcripts in the tissues analyzed.

4. DISCUSSION

Macaúba palm is a non-model palm tree with great potential for oil production for food, cosmetic and bioenergetic use. Its oldest fossil sites are dated to 11,200 years BC in Santarém, Pará - Brazil (as Acrocomia sp.), followed by Panama (8040 B.P.), and Mexico (6750 B.P.) [7]. The macaúba palm (*Acrocomia aculeata*) is a widespread palm tree in Brazil, being found from Pará state to the north of Paraná. It is present in degraded and disturbed areas, road margins, and rainy regions, but its greater occurrence is in opening areas with high solar incidence. In Brazil, its greatest distribution is in the Cerrado but also occurs in regions of more pronounced drought, such as arid and semi-arid regions. In some places, macaúba palm occurs in dense populations, which characterizes the species as oligarchic [98–101]

The macaúba palm is adapted in morphological and molecular level to different edaphoclimatic systems, since its predominant occurrence in Brazilian Cerrado, corroborating its climatic robustness.

The mechanisms for plant adaptation are influenced by CO_2 concentration, solar radiation, temperature, relative humidity, drought/rain periods, and soil fertility (SANTOS and CARLESSO, 1998). According to Osório et al. [102], species with drought tolerance mechanism have high phenotypic plasticity, which has already been described by other authors for the macaúba palm [2, 103, 104]. Among phenotypic adaptations, cuticle on the leaf surface [105], cuticular wax, opening efficiency, stomatal morphology [106] and petioles bases covering stipe increases plant survival in drought, high temperature and fire [2, 107–109]. Müller et al. [110] provided a good example of a study in a date palm, a palm tree adapted to the desert climate, which has a unique cuticle and stoma architecture, with stomata surrounded by a wax wall. The macaúba palm, when compared to *Syagrus coronata*, decreased its gas exchange by 92% in the dry period of a semi-arid region in Northeastern Brazil, in addition to the larger adaxial cuticle with higher stomatal density on the abaxial leaf surface [111].

Morphological and structural adaptations are associated with biochemical and molecular changes that allow the species survival in several environments. In this study, we present the first large-scale NGS data in macaúba palm species, providing valuable genome resources for studying evolution, gene expression, and genomic studies. We present an extensive multi-tissue transcriptome analysis and the annotation of transcripts showed many predicted genes that play important role in many biochemical roots of abiotic and biotic stress, providing a possible explanation about plasticity and rusticity of macaúba palm and about differences with oil palm.

There are two gene sets related to abiotic and biotic stresses: the first set contains proteins with a function in stress tolerance, such as chaperones, osmotins, mRNA binding proteins, key enzymes for the osmolytes biosynthesis, aquaporins, sugar and proline transporters, detoxifying enzymes and proteases. The second set is involved in signal transduction, comprising transcription factors, protein kinases, phosphatases, enzymes involved in phospholipid metabolism, and other signaling molecules, that play a key role in the plant's response to stress [112].

Here, we found 341 root – specific transcript with high expression, suggesting that macaúba palm displays high rusticity and tolerance to abiotic and biotic stress. Interestingly, most of the transcripts observed are related to signal transduction and defense mechanisms, such as serine/threonine kinases, mitogen-activated protein kinase (MAPK) and PR-proteins. When we observe family expansion between macaúba and oil palm, we can propose a molecular adjustment of macaúba palm to environmental adaptation. Four families were annotated as PPR proteins, two families were similar to heat-shock proteins and three families were like to protein kinase; six families were unique to macaúba and 11 of them had at least one TS transcripts, being all root-specific transcripts. The OG0000038 family was highly expressed were root-specific, which annotation showed similarity to ubiquitin, a regulatory protein that plays function in the most different biochemical processes.

Additionally, transcripts related to drought stress were grouped in 114 families containing at least one macaúba transcript. It was possible to analyze and suggest that the plant adjusts expressing regulatory proteins. Among the TS families, proteins from the 14-3-3 family and CAM had high expression. In addition, 13 families expressed stress-related transcription factors and 11 families were classified as kinases.

The macaúba palm is considered a Neotropical species incipiently domesticated and it was evolved a long time in different environments and edaphoclimatic conditions. Habitat changes require morphological and physiological changes, such as the evolution of root system, photosynthetic changes and reserve accumulation for the successful colonization in different ecosystems. Bacon et al. [113] reported the shifting of palm tree genus *Allagoptera* from tropical forest to dry habitats, which is correlated with morphological adaptations, although they did not find metabolic differences due to the recent origin of the species in Brazilian Cerrado. Lopes et al. [114] reported that chloroplastic genes related to photosynthesis, plastid gene expression and fatty acid biosynthesis in the macaúba palm are presumably under positive selection, but authors haven't correlated these sites with the habitats of palm species and the existence of specific selective pressures.

Our data suggest that it is probably more economical and efficient for the plant to regulate biological processes than to detoxify the plant and deal with environmental adaptation. Our data collected provides a new resource as well as an extended catalog of macaúba palm genes. Therefore, the data generated will be of significant support for the ongoing research in Acrocomia genera, particularly.

5. REFERENCES

1. Azevedo Filho J a., Colombo C a., Berton LHC. Macaúba: palmeira nativa como opção bioenergética. Pesqui Tecnol. 2012;9. http://www.aptaregional.sp.gov.br/index.php?option=com_docman&task=doc_view&gid=1275&Itemid=284.

2. Colombo CA, Berton LHC, Diaz BG, Ferrari RA. Macauba: a promising tropical palm for the production of vegetable oil. Oilseeds Fats Crop Lipids. 2018;25:9.

3. Steinfeld H, Gerber PJ, Wassenaar T, Castel V, Rosales M, De haan C. Livestock's long shadow - environmental issues and options. Food Agric Organ United Nations. 2006;3:1–377. doi:10.1007/s10666-008-9149-3.

4. Fitzherbert EB, Struebig MJ, Morel A, Danielsen F, Brühl CA, Donald PF, et al. How will oil palm expansion affect biodiversity? Trends Ecol Evol. 2008;23:538–45. doi:10.1016/j.tree.2008.06.012.

5. Rival A, Levang P. Palms of controversies: Oil palm and development challenges. 2014.

6. Sheil D, Casson A, Meijaard E, van Noordwijk M, Gaskell J, Sunderland-groves J, et al. The Impacts and Opportunities of Oil Palm in Southeast Asia: What Do We Know and What Do We Need to Know? 2009. http://hdl.handle.net/10535/5256.

7. Morcote-Ríos G, Bernal R. Remains of Palms (Palmae) at Archaeological Sites in the New World: A Review. Bot Rev. 2001;67:309–50. doi:10.1007/BF02858098.

8. Ciconini G, Favaro SP, Roscoe R, Miranda CHB, Tapeti CF, Miyahira MAMM, et al. Biometry and oil contents of Acrocomia aculeata fruits from the Cerrados and Pantanal biomes in Mato Grosso do Sul, Brazil. Ind Crops Prod. 2013;45 March 2016:208–214.

9. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10:57–63. doi:10.1038/nrg2484.

10. Unamba CIN, Nag A, Sharma RK. Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. Front Plant Sci. 2015;6 December. doi:10.3389/fpls.2015.01074.

11. Ho C-L, Kwan Y-Y, Choi M-C, Tee S-S, Ng W-H, Lim K-A, et al. Analysis and functional annotation of expressed sequence tags (ESTs) from multiple tissues of oil palm (Elaeis guineensis Jacq.). BMC Genomics. 2007;8:381. doi:10.1186/1471-2164-8-381.

12. Pootakham W, Uthaipaisanwong P, Sangsrakru D, Yoocha T, Tragoonrung S, Tangphatsornruang S. Development and characterization of single-nucleotide polymorphism markers from 454 transcriptome sequences in oil palm (Elaeis guineensis). Plant Breed. 2013;132:711–7. doi:10.1111/pbr.12095.

13. Shearman JR, Jantasuriyarat C, Sangsrakru D, Yoocha T, Vannavichit A, Tragoonrung S, et al. Transcriptome analysis of normal and mantled developing oil palm flower and fruit. Genomics. 2013;101:306–12. doi:10.1016/j.ygeno.2013.02.012.

14. Dussert S, Guerin C, Andersson M, Joët T, Tranbarger TJ, Pizot M, et al. Comparative transcriptome analysis of three oil palm fruit and seed tissues that differ in oil content and fatty acid composition. Plant Physiol. 2013;162:1337–58. doi:10.1104/pp.113.220525.

15. Yin Y, Zhang X, Fang Y, Pan L, Sun G, Xin C, et al. High-throughput sequencing-based gene profiling on multi-staged fruit development of date palm (Phoenix dactylifera, L.). Plant Mol Biol. 2012;78:617–26. doi:10.1007/s11103-012-9890-5.

16. Zhang G, Pan L, Yin Y, Liu W, Huang D, Zhang T, et al. Large-scale collection and annotation of gene models for date palm (Phoenix dactylifera, L.). Plant Mol Biol. 2012;79:521–36. doi:10.1007/s11103-012-9924-z.

17. Zeng Y, Yang T. RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. Plant Mol Biol Report. 2002;20:417. doi:10.1007/BF02772130.

18. Le Provost G, Herrera R, Paiva JA, Chaumeil P, Salin F, Plomion C. A micromethod for high throughput RNA extraction in forest trees. Biol Res. 2007;40:291–7.

19. Rezaian MA, Krake LR. Nucleic acid extraction and virus detection in grapevine. J Virol Methods. 1987;17:277–85. doi:https://doi.org/10.1016/0166-0934(87)90137-6.

20. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

21. Kopylova E, Noé L, Touzet H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.

22. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol. 2011;29:644–52. doi:10.1038/nbt.1883.

23. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525. http://dx.doi.org/10.1038/nbt.3519.

24. Koskinen P, Törönen P, Nokso-Koivisto J, Holm L. PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment. Bioinformatics. 2015;31:1544–52. doi:10.1093/bioinformatics/btu851.

25. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: A Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 2011;39 SUPPL. 1:225–9.

26. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015;16:157.

doi:10.1186/s13059-015-0721-2.

27. Librado P, Vieira FG, Rozas J. BadiRate: estimating family turnover rates by likelihood-based methods. Bioinformatics. 2012;28:279–81. doi:10.1093/bioinformatics/btr623.

28. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC135756/.

29. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3. doi:10.1093/bioinformatics/btu033.

30. Xiao SJ, Zhang C, Zou Q, Ji ZL. TiSGeD: A database for tissue-specific genes. Bioinformatics. 2010;26:1273–5.

31. Wang ZQ, Li GZ, Gong QQ, Li GX, Zheng SJ. OsTCTP, encoding a translationally controlled tumor protein, plays an important role in mercury tolerance in rice. BMC Plant Biol. 2015;15:1–15.

32. Berkowitz O, Jost R, Pollmann S, Masle J. Characterization of TCTP, the Translationally Controlled Tumor Protein, from Arabidopsis thaliana. Plant Cell Online. 2008;20:3430–47. doi:10.1105/tpc.108.061010.

33. Lee SB, Go YS, Bae H-J, Park JH, Cho SH, Cho HJ, et al. Disruption of Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein Gene Altered Cuticular Lipid Composition, Increased Plastoglobules, and Enhanced Susceptibility to Infection by the Fungal Pathogen Alternaria brassicicola. Plant Physiol. 2009;150:42–54. doi:10.1104/pp.109.137745.

34. Liu F, Zhang X, Lu C, Zeng X, Li Y, Fu D, et al. Non-specific lipid transfer proteins in plants: Presenting new advances and an integrated functional analysis. J Exp Bot. 2015;66:5663–81.

35. Chen C, Chen G, Hao X, Cao B, Chen Q, Liu S, et al. Plant Science CaMF2 , an anther-specific lipid transfer protein (LTP) gene , affects pollen development in Capsicum annuum L . Plant Sci. 2011;181:439–48. doi:10.1016/j.plantsci.2011.07.003.

36. Guo C, Ge X, Ma H. The rice OsDIL gene plays a role in drought tolerance at vegetative and reproductive stages. Plant Mol Biol. 2013;82:239–53.

37. Rafferty JB, Simon JW, Baldock C, Artymiuk PJ, Baker PJ, Stuitje AR, et al. Common themes in redox chemistry emerge from the X-ray structure of oilseed rape (Brassica napus) enoyl acyl carrier protein reductase. Structure. 1995;3:927–38.

38. Fawcett T, Simon WJ, Swinhoe R, Shanklin J, Nishida I, Christie WW, et al. Expression of mRNA and steady-state levels of protein isoforms of enoyl-ACP reductase from Brassica napus. Plant Mol Biol. 1994;26:155–63.

39. Budiani A, Santoso D, Aswindinnoor H, Suwanto A, Sudiatso S. Isolation and characterization of protein differentially expressed during oil palm mesocarp development. Menara Perkeb. 2002;70:1–11. http://sitp.rpn.co.id/uploads/riset/sawit/MP70-01-01.pdf.

40. Wong YC, Teh HF, Mebus K, Ooi TEK, Kwong Q Bin, Koo KL, et al. Differential gene expression at different stages of mesocarp development in high- and low-yielding oil palm. BMC Genomics. 2017;18:1–13.

41. Ruzlan N, Othman AM, Win W, Kua SF, Low JYS, Chandran AN, et al. METHODS FOR PREDICTING PALM OIL YIELD OF A TEST OIL PALM PLANT. 2018;:30.

42. Bourgis F, Kilaru A, Cao X, Ngando-Ebongue G-FG-F, Drira N, Ohlrogge JB, et al. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. Proc Natl Acad Sci. 2011;108:12527–32. doi:10.1073/pnas.1106502108.

43. Shiu S-H, Bleecker AB. Plant receptor-like kinase gene family: diversity, function, and signaling. Sci Signal. 2001;2001:13 P.

44. Patankar H V., Al-Harrasi I, Al-Yahyai R, Yaish M. Identification of Candidate Genes Involved in the Salt Tolerance of Date Palm (Phoenix dactylifera L .) Based. DNA Cell Biol. 2018;37:524–34.

45. Radwan O, Arro J, Keller C, Korban SS. RNA-Seq Transcriptome Analysis in Date Palm Suggests Multi-Dimensional Responses to Salinity Stress. Trop Plant Biol. 2015;8:74–86.

46. Van Loon LC, Van Strien EA. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol Mol Plant Pathol. 1999;55:85–97.

47. Sudisha J, Sharathchandra RG, Amruthesh KN, Kumar A, Shetty HS. Pathogenesis Related Proteins in Plant Defense Response. Plant Def. 2012; March:379–403.

48. Baker WJ, Dransfield J. Beyond *Genera Palmarum* : progress and prospects in palm systematics. Bot J Linn Soc. 2016;182:207–33. doi:10.1111/boj.12401.

49. Small ID, Peeters N. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci. 2000;25:45–7. doi:10.1016/S0968-0004(99)01520-0.

50. Manna S. An overview of pentatricopeptide repeat proteins and their applications. Biochimie.

2015;113:93–9. doi:10.1016/j.biochi.2015.04.004.

51. Eyüboglu B, Pfister K, Haberer G, Chevalier D, Fuchs A, Mayer KFX, et al. Molecular characterisation of the STRUBBELIG-RECEPTOR FAMILY of genes encoding putative leucine-rich repeat receptor-like kinases in Arabidopsis thaliana. BMC Plant Biol. 2007;7:1–24.

52. Chevalier D, Batoux M, Fulton L, Pfister K, Yadav RK, Schellenberg M, et al. STRUBBELIG defines a receptor kinase-mediated signaling pathway regulating organ development in Arabidopsis. Proc Natl Acad Sci. 2005;102:9074–9. doi:10.1073/pnas.0503526102.

53. Kwak SH, Schiefelbein J. The role of the SCRAMBLED receptor-like kinase in patterning the Arabidopsis root epidermis. Dev Biol. 2007;302:118–31.

54. Shiu S, Karlowski W, Pan R. Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell …. 2004;16 May:1220–34. doi:10.1105/tpc.020834.1.

55. Clark SE, Williams RW, Meyerowitz EM. The CLAVATA1 Gene Encodes a Putative Receptor Kinase That Controls Shoot and Floral Meristem Size in Arabidopsis. Cell. 1997;89:575–585. doi:10.1016/S0030-6657(08)70226-9.

56. Weller B, Zourelidou M, Frank L, Barbosa ICR, Fastner A, Richter S, et al. Dynamic PIN-FORMED auxin efflux carrier phosphorylation at the plasma membrane controls auxin efflux-dependent growth. Proc Natl Acad Sci. 2017;114:E887--E896. doi:10.1073/pnas.1614380114.

57. Kinoshita T, Doi M, Suetsugu N, Kagawa T, Wada M, Shimazaki K. phot1 and phot2 mediate blue light regulation of stomatal opening. Nature. 2001;414:656. http://dx.doi.org/10.1038/414656a.

58. Willige BC, Ahlers S, Zourelidou M, Barbosa ICR, Demarsy E, Trevisan M, et al. D6PK AGCVIII Kinases Are Required for Auxin Transport and Phototropic Hypocotyl Bending in *Arabidopsis*. Plant Cell. 2013;25:1674–88. doi:10.1105/tpc.113.111484.

59. Alter S, Bader KC, Spannagl M, Wang Y, Bauer E, Schön CC, et al. DroughtDB: An expert-curated compilation of plant drought stress genes and their homologs in nine species. Database. 2015;2015:1–7.

60. Roberts MR, Salinas J, Collinge DB. 14-3-3 Proteins and the Response To Abiotic and Biotic Stress. Plant Mol Biol. 2002;50:1031–9. doi:10.1023/A:1021261614491.

61. Yan J, He C, Wang J, Mao Z, Holaday SA, Allen RD, et al. Overexpression of the Arabidopsis 14-

3-3 protein GF14λ in cotton leads to a “stay-green” phenotype and improves stress tolerance under moderate drought conditions. Plant Cell Physiol. 2004;45:1007–14.

62. He Y, Wu J, Lv B, Li J, Gao Z, Xu W, et al. Involvement of 14-3-3 protein GRF9 in root growth and response under polyethylene glycol-induced water stress. J Exp Bot. 2015;66:2271–81.

63. Kudla J, Batistic O, Hashimoto K. Calcium Signals: The Lead Currency of Plant Information Processing. Plant Cell. 2010;22:541–63. doi:10.1105/tpc.109.072686.

64. Xu GY, Rocha PSCF, Wang ML, Xu ML, Cui YC, Li LY, et al. A novel rice calmodulin-like gene, OsMSR2, enhances drought and salt tolerance and increases ABA sensitivity in Arabidopsis. Planta. 2011;234:47–59.

65. Munir S, Liu H, Xing Y, Hussain S, Ouyang B, Zhang Y, et al. Overexpression of calmodulin-like (ShCML44) stress-responsive gene from Solanum habrochaites enhances tolerance to multiple abiotic stresses. Sci Rep. 2016;6 August:1–20. doi:10.1038/srep31772.

66. Lu X, Yang L, Yu M, Lai J, Wang C, McNeil D, et al. A novel Zea mays ssp. mexicana L. MYC-type ICE-like transcription factor gene ZmmICE1, enhances freezing tolerance in transgenic Arabidopsis thaliana. Plant Physiol Biochem. 2017;113:78–88. doi:10.1016/j.plaphy.2017.02.002.

67. Abe H, Urao T, Ito T, Seki M, Shinozaki K. Transcriptional Activators in Abscisic Acid Signaling. Society. 2003;15 January:63–78.

68. Zhang X, Wang L, Meng H, Wen H, Fan Y, Zhao J. Maize ABP9 enhances tolerance to multiple stresses in transgenic Arabidopsis by modulating ABA signaling and cellular levels of reactive oxygen species. Plant Mol Biol. 2011;75:365–78.

69. Liang C, Meng Z, Meng Z, Malik W, Yan R, Lwin KM, et al. GhABF2, a bZIP transcription factor, confers drought and salinity tolerance in cotton (Gossypium hirsutum L.). Sci Rep. 2016;6 October:1–14. doi:10.1038/srep35040.

70. Liu C, Mao B, Ou S, Wang W, Liu L, Wu Y, et al. OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. Plant Mol Biol. 2014;84:19–36.

71. Rí-podas C, Castaingts M, Clúa J, Blanco F, Zanetti ME. Annotation, phylogeny and expression analysis of the nuclear factor Y gene families in common bean (Phaseolus vulgaris). Front Plant Sci. 2015;5 January. doi:10.3389/fpls.2014.00761.

72. Combier J, Frugier F, Billy F De, Boualem A, El-yahyaoui F, Moreau S, et al. MtHAP2-1 is a key

transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula service MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. Genes Dev. 2006;20:3084–8.

73. Zanetti ME, Rípodas C, Niebel A. Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. Biochim Biophys Acta - Gene Regul Mech. 2017;1860:645–54.

74. Li W-X, Oono Y, Zhu J, He X-J, Wu J-M, Iida K, et al. The Arabidopsis NFYA5 Transcription Factor Is Regulated Transcriptionally and Posttranscriptionally to Promote Drought Resistance. Plant Cell Online. 2008;20:2238–51. doi:10.1105/tpc.108.059444.

75. Leyva-González MA, Ibarra-Laclette E, Cruz-Ramírez A, Herrera-Estrella L. Functional and Transcriptome Analysis Reveals an Acclimatization Strategy for Abiotic Stress Tolerance Mediated by Arabidopsis NF-YA Family Members. PLoS One. 2012;7.

76. Ma X, Li C, Wang M. Wheat NF-YA10 functions independently in salinity and drought stress. Bioengineered. 2015;6:245–7.

77. Omidvar V, Abdullah SNA, Ebrahimi M, Ho CL, Mahmood M. Gene expression of the oil palm transcription factor EgAP2-1 during fruit ripening and in response to ethylene and ABA treatments. Biol Plant. 2013;57:646–54.

78. Pirrello J, Jaimes-Miranda F, Sanchez-Ballesta MT, Tournier B, Khalil-Ahmad Q, Regad F, et al. Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. Plant Cell Physiol. 2006;47:1195–205.

79. Hao P-P, Wang G-M, Cheng H-Y, Ke Y-Q, Qi K-J, Gu C, et al. Transcriptome analysis unravels an ethylene response factor involved in regulating fruit ripening in pear. Physiol Plant. 2018;163:124–35.

80. Dossa K, Wei X, Li D, Fonceka D, Zhang Y, Wang L, et al. Insight into the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB subfamily under drought stress. BMC Plant Biol. 2016;16:1–16. doi:10.1186/s12870-016-0859-4.

81. Jin Y, Pan W, Zheng · Xiufang, Cheng X, Liu M, Ma H, et al. OsERF101, an ERF family transcription factor, regulates drought stress response in reproductive tissues. Plant Mol Biol. 2018;98:51–65. doi:10.1007/s11103-018-0762-5.

82. Guo M, Liu J-H, Ma X, Luo D-X, Gong Z-H, Lu M-H. The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. Front Plant Sci. 2016;7 February:114. doi:10.3389/fpls.2016.00114.

83. Mishra SK, Tripp J, Winkelhaus S, Tschiersch B, Theres K, Nover L, et al. In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes Dev. 2002;16:1555–67.

84. Liu HC, Liao HT, Charng YY. The role of class A1 heat shock factors (HSFA1s) in response to heat and other stresses in Arabidopsis. Plant, Cell Environ. 2011;34:738–51.

85. Dreesen FE, De Boeck HJ, Janssens IA, Nijs I. Summer heat and drought extremes trigger unexpected changes in productivity of a temperate annual/biannual plant community. Environ Exp Bot. 2012;79:21–30. doi:10.1016/j.envexpbot.2012.01.005.

86. Weinl S, Kudla J. The CBL – CIPK Ca 2 + -decoding signaling network : function and perspectives. New Phytol. 2009;184:517–28.

87. Hettenhausen C, Su G, He Y, Zhuang H, Sun T, Qi J, et al. Genome-wide identification of calcium-dependent protein kinases in soybean and analyses of their transcriptional responses to insect herbivory and drought stress. Sci Rep. 2016;6:1–14. doi:10.1038/srep18973.

88. Song C, Wang W. Calcium-dependent protein kinase gene ZmCPK12 from maize confers tolerance to drought and salt stresses in transgenic plants. Acta Physiol Plant. 2013;35:1659–66.

89. Cheong YH, Pandey GK, Grant JJ, Batistic O, Li L, Kim B, et al. Two calcineurin B-like calcium sensors , interacting with protein kinase CIPK23 , regulate leaf transpiration and root potassium uptake in Arabidopsis. 2007;:223–39.

90. Luo Q, Wei Q, Wang R, Zhang Y, Zhang F, He Y, et al. BdCIPK31 , a Calcineurin B-Like Protein-Interacting Protein Kinase , Regulates Plant Response to Drought and Salt Stress. Front Plant Sci. 2017;8 July:1–16.

91. Cui X, Du Y, Fu J, Yu T, Wang C, Chen M, et al. Wheat CBL-interacting protein kinase 23 positively regulates drought stress and ABA responses. BMC Plant Biol. 2018;18:1–13.

92. Joo J, Hab Y, Sang L, Song I. Overexpression of the rice basic leucine zipper transcription factor OsbZIP12 confers drought tolerance to rice and makes seedlings hypersensitive to ABA. 2014;:431–41.

93. Tang N, Zhang H, Li X, Xiao J, Xiong L. Constitutive Activation of Transcription Factor OsbZIP46. Plant Physiol. 2012;158:1755–68.

94. Lou D, Wang H, Liang G, Yu D. OsSAPK2 Confers Abscisic Acid Sensitivity and Tolerance to Drought Stress in Rice. Front Plant Sci. 2017;8:1–15.

95. Eastmond PJ. Glycerol-insensitive Arabidopsis mutants : gli1 seedlings lack glycerol kinase, accumulate glycerol and are more resistant to abiotic stress. Plant J. 2004;37:617–25.

96. Penfield S, Clements S, Bailey KJ, Gilday AD, Leegood RC, Gray JE, et al. Expression and manipulation of PHOSPHOENOLPYRUVATE CARBOXYKINASE 1 identifies a role for malate metabolism in stomatal closure. 2012;:679–88.

97. Das A, Das S, Mondal TK. Identification of Differentially Expressed Gene Profiles in Young Roots of Tea [Camellia sinensis (L .) O . Kuntze] Subjected to Drought Stress Using Suppression Subtractive Hybridization. Plant Mol Biol Report. 2012;30:1088–101.

98. Belén-camacho PDR, López I, García D, González M, Rodríguez US, Canoabo N, et al. Evaluación fisico-química de la semilla y del aceite de corozo (Acrocomia aculeata Jacq .). 2005;56 March 2003:311–6.

99. da Motta PEF, Curi N, de Oliveira-Filho AT, Gomes JBV. Ocorrência da macaúba em Minas Gerais : relação com atributos. Pesqui Agropecuária Bras. 2002;37:1023–31.

100. Clement CR, Lleras Pérez E, van Leeuwen J. O potencial das palmeiras tropicais no Brasil: acertos e fracassos das últimas décadas. Agrociencias. 2005;9:67–71.

101. Mota CS, Corrêa TR, Grossi JAS, Ribeiro A da S. Exploração sustentável da macaúba para produção de biodiesel : colheita , pós-colheita e qualidade dos frutos. Inf Agropecuário. 2011;32:41–51.

102. Osório ML, Gonçalves S, Coelho N, Osório J, Romano A. Morphological , physiological and oxidative stress markers during acclimatization and field transfer of micropropagated Tuberaria major plants. Plant Cell Tissue Organ Cult. 2013;115:85–97.

103. Pires TP, Souza E dos S, Kuki KN, Motoike SY. Ecophysiological traits of the macaw palm : A contribution towards the domestication of a novel oil crop. Ind Crops Prod. 2013;44:200–10.

104. Marques R, Carla C, Joaquim C, Azevedo A De, Henrique L, Berton C, et al. Non-biotic factors determining plasticity of the prospective oil-rich macauba palm (Acrocomia aculeata). Agrofor

Syst. 2017; Lorenzi 1992. doi:10.1007/s10457-017-0173-7.

105. Shao H, Chu L, Abdul C, Zhao C. Water-deficit stress-induced anatomical changes in higher plants ☆. 2008;331:215–25.

106. Chen Z, Chen G, Dai F, Wang Y, Hills A, Ruan Y, et al. Molecular Evolution of Grass Stomata. Trends Plant Sci. 2017;22:124–39. doi:10.1016/j.tplants.2016.09.005.

107. Vianna SA, Larry SMC. Leaf anatomy of Acrocomia (Arecaceae): an additional contribution to the taxonomic resolution of a genus with great economic potential. Plant Syst Evol. 2017;303:233–48.

108. Carvalho AL De, José E, Ferreira L, Marie J, Lima T. Comparações florísticas e estruturais entre comunidades de palmeiras em fragmentos de floresta primária e secundária da Área de Proteção Ambiental Raimundo Irineu Serra – Rio Branco , Acre , Brasil. Acta Amaz. 2010;40:657–66.

109. Herbert F. Commentary Learning from the experts : drought resistance in desert plants. New Phytol. 2017;216:5–7.

110. Müller HM, Schäfer N, Bauer H, Geiger D, Lautner S, Fromm J, et al. The desert plant Phoenix dactylifera closes stomata via nitrate-regulated SLAC1 anion channel. New Phytol. 2017;216:150–62.

111. Oliveira D, Medeiros M, Pereira S, Oliveira M, Frosi G, Arruda E. Ecophysiological leaf traits of native and exotic palm tree species under semi-arid conditions. 2016;:128–34.

112. Seki M, Kamei A, Yamaguchi-shinozaki K, Shinozaki K. Molecular responses to drought , salinity and frost : common and different paths for plant protection. Curr Opin Biotechnol. 2003;14:194–9.

113. Bacon CD, R MM, Jaramillo C, Antonelli A. Endemic palm species shed light on habitat shifts and the assembly of the Cerrado and Restinga floras Short running head Department of Biological and Environmental Sciences , University of Gothenburg , Herbario Nacional de Bolivia , Instituto de Ecología . Mol Phylogenet Evol. 2017. doi:10.1016/j.ympev.2017.03.013.

114. Lopes ADS, Túlio GP, Nimz T, Vieira L do N, Guerra MP, Nodari RO, et al. The complete plastome of macaw palm [Acrocomia aculeata (Jacq .) Lodd . ex Mart .] and extensive molecular analyses of the evolution of plastid genes in Arecaceae. Planta. 2018;247:1011–30. doi:10.1007/s00425-018-2841-x.

Additional file 1. TPM values of tissue-specific transcripts (SPM ≥ 0.95) root tissue. TPM values ≥ 0.95 in bold and pink-colored box.

Transcript ID	SPM	Annotation	Root		
			S 4	S 12	S 20
TRINITY_DN34234_c0_g1	1	CFEM domain.	19.74	**55.41**	**21.05**
TRINITY_DN35843_c1_g1	1	Transcriptional regulator ICP4; Provisional.	19.34	**33.72**	**41.46**
TRINITY_DN31296_c0_g1	0.97	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	19.13	**31.17**	**28.53**
TRINITY_DN35208_c2_g2	1	DNA polymerase III subunits gamma and tau; Validated.	18.77	**40.88**	**20.56**
TRINITY_DN3413_c0_g1	1	--	18.49	**21.27**	**24.46**
TRINITY_DN24265_c0_g1	1	Zinc finger, C2H2 type. The C2H2 zinc finger is a classical zinc finger domain.	18.35	**41.86**	**40.73**
TRINITY_DN35647_c0_g1	1	Transcriptional regulator ICP4; Provisional.	18.28	**35.2**	**30.63**
TRINITY_DN35080_c0_g1	0.99	Peptidase_C1, Papain family cysteine protease.	18.08	**32.88**	**117.18**
TRINITY_DN20382_c0_g1	1	--	17.81	**23.07**	**26.54**
TRINITY_DN27598_c0_g1	1	--	17	**39.95**	**26.11**
TRINITY_DN35227_c0_g1	1	DUF3712, Protein of unknown function (DUF3712).	16.93	**37.81**	**26.95**
TRINITY_DN19498_c0_g1	0.98	Cytochrome P450.	16.76	**26.42**	**20.98**
TRINITY_DN21367_c0_g1	0.97	ANK, ankyrin repeats.	15.88	**30.62**	**23.4**
TRINITY_DN17096_c0_g1	0.99	FR_SDR_e, flavonoid reductase (FR), extended (e) SDRs.	15.68	**46.71**	**28.78**
TRINITY_DN31963_c0_g1	1	--	15.58	**26.73**	**48.51**
TRINITY_DN35071_c0_g1	1	SUR7, SUR7/PalI family.	15.05	**23.66**	**21.22**
TRINITY_DN31025_c0_g1	1	AP2, DNA-binding domain in plant proteins such as APETALA2 and EREBPs.	14.9	**32.44**	**35.8**
TRINITY_DN23902_c0_g1	1	DUF4360, Domain of unknown function (DUF4360).	14.64	**38.35**	**31.29**
TRINITY_DN25769_c1_g1	0.99	ubiquitin, Ubiquitin family.	14.44	**40.07**	**29.66**
TRINITY_DN35799_c1_g1	1	bZIP_u1, Basic leucine zipper (bZIP) domain of bZIP transcription factors.	14.34	**27.35**	**29.15**
TRINITY_DN33654_c0_g1	1	GlnA, Glutamine synthetase [Amino acid transport and metabolism].	14.29	**43.22**	**23.67**
TRINITY_DN27516_c0_g2	1	ML, Domain involved in innate immunity and lipid metabolism.	14.05	**46.42**	**24.02**
TRINITY_DN35251_c0_g1	1	PTR2, POT family.	13.86	**47.6**	**24.41**
TRINITY_DN25689_c0_g1	1	ZF_C2H2, Zinc finger, C2H2 type. The C2H2 zinc finger is a classical zinc finger domain.	13.59	**21.6**	**29.96**

TRINITY_DN35142_c0_g1	0.98	Mito_carr, Mitochondrial carrier protein.	13.49	32.38	34.08
TRINITY_DN21380_c0_g1	1	UBQ, Ubiquitin homologues.	13.34	42.74	28.02
TRINITY_DN7019_c0_g2	1	GluR_Plant, Plant glutamate receptor domain.	12.15	45.83	30.66
TRINITY_DN12824_c0_g1	1	UreE_C, UreE urease accessory protein, C-terminal domain.	11.96	35.7	26.8
TRINITY_DN26731_c0_g1	1	RPE65, Retinal pigment epithelial membrane protein.	11.82	27.1	20.82
TRINITY_DN12005_c0_g1	0.97	p450, Cytochrome P450.	11.73	30.96	24.86
TRINITY_DN11303_c0_g1	1	Lig_chan, Ligand-gated ion channel.	11.21	52.56	30.38
TRINITY_DN33739_c0_g1	1	Transcription termination factor Rho; Provisional.	10.59	36.36	20.22
TRINITY_DN20654_c0_g1	1	Lycopene_cycl, Lycopene cyclase protein.	9.41	40.11	24.03
TRINITY_DN34878_c0_g1	1	p450, Cytochrome P450.	8.85	22.58	20.57
TRINITY_DN31034_c0_g1	1	DnaK, molecular chaperone DnaK; Provisional.	8.29	25.75	23.99
TRINITY_DN32717_c0_g1	1	DUF2638, Protein of unknown function (DUF2638).	7.95	39.67	20.88
TRINITY_DN35553_c0_g1	1	Peptidases_S8_3, Peptidase S8 family domain.	1147.5	1848.1	1082.8
TRINITY_DN29395_c0_g1	0.98	MIP, Major intrinsic protein. MIP (Major Intrinsic Protein) family proteins.	878.23	308.72	125.45
TRINITY_DN32252_c0_g1	1	--	808.06	787.18	682.98
TRINITY_DN29285_c0_g2	1	Phytocyanin, Phytocyanins are plant blue or type I copper proteins.	682.76	640.19	481.78
TRINITY_DN34368_c2_g1	1	B_lectin, Bulb-type mannose-specific lectin.	599.57	245.22	0
TRINITY_DN32126_c0_g1	1	PAR1, PAR1 protein.	589.33	794	563.17
TRINITY_DN35639_c0_g1	1	Cupin_1, Cupin.	502.71	445.7	354.29
TRINITY_DN31923_c0_g1	1	Ammonium_transporter, ammonium transporter. The Ammonium Transporter (Amt) Family (TC 2.A.49)	481.91	595.37	355.82
TRINITY_DN32518_c0_g2	1	B_lectin, Bulb-type mannose-specific lectin.	473.38	538.58	338.33
TRINITY_DN16187_c0_g1	1	LysM, Lysin motif.	438.61	347.37	207.83
TRINITY_DN19440_c0_g1	1	acpP, acyl carrier protein; Provisional.	425.28	650.19	462.76
TRINITY_DN31785_c0_g1	0.99	BURP, BURP domain.	366.75	145.24	71.63
TRINITY_DN20647_c0_g1	1	SGNH_plant_lipase_like, SGNH_plant_lipase_like.	338.19	440.3	285.26
TRINITY_DN34687_c0_g1	1	GH18_hevamine_XipI_class_III	329.04	237.57	166.06
TRINITY_DN29805_c0_g1	0.98	SCP_PR-1_like, SCP_PR-1_like: SCP-like extracellular protein domain, PR-1 like subfamily.	307.19	25.13	22.84
TRINITY_DN18590_c0_g1	1	SGNH_plant_lipase_like, SGNH_plant_lipase_like.	274.96	488.39	283.74
TRINITY_DN33416_c0_g1	1	p450, Cytochrome P450.	227.15	353.41	234.17
TRINITY_DN17182_c0_g1	1	MFS, The Major Facilitator Superfamily (MFS).	217.19	321.73	154.22
TRINITY_DN35665_c1_g1	1	Peptidases_S8_3, Peptidase S8 family domain.	215.47	258.63	149.73

TRINITY_DN30737_c0_g1	1	TauD, Taurine catabolism dioxygenase TauD, TfdA family.	208.59	304.67	205.37
TRINITY_DN25728_c0_g1	0.99	Alginate_lyase2, Alginate lyase.	201.86	113.1	182.31
TRINITY_DN26108_c0_g1	1	--	193.24	231.48	104.36
TRINITY_DN30915_c0_g1	0.99	Peptidase_S10, Serine carboxypeptidase.	192.03	342.42	220.09
TRINITY_DN34463_c0_g2	0.97	BURP, BURP domain.	185.93	122.95	127.49
TRINITY_DN33218_c0_g1	1	p450, Cytochrome P450.	183.17	174.72	31.68
TRINITY_DN34456_c0_g1	1	Sugar_tr, Sugar (and other) transporter.	177.19	245.27	129.25
TRINITY_DN21295_c0_g1	1	Phytocyanin, Phytocyanins are plant blue or type I copper proteins.	171.99	120.4	88.2
TRINITY_DN33652_c0_g1	1	OLE1, Fatty-acid desaturase [Lipid metabolism].	168.72	325.22	81.56
TRINITY_DN29550_c0_g1	1	--	159.79	417.35	308.1
TRINITY_DN23294_c0_g1	0.99	DUF4033, Domain of unknown function (DUF4033).	155.76	294.25	245.68
TRINITY_DN29715_c0_g1	1	TB2_DP1_HVA22, TB2/DP1, HVA22 family.	154.09	261.26	54.59
TRINITY_DN16053_c0_g1	1	CsbD, CsbD-like. CsbD is a bacterial general stress response protein.	153.73	90.75	5.89
TRINITY_DN32913_c0_g1	1	Alpha_tubulin, The alpha-tubulin family.	152.86	140.99	68.46
TRINITY_DN23243_c0_g1	0.97	Pat17_PNPLA8_PNPLA9_like, Patatin-like phospholipase.	140.41	20.56	6.67
TRINITY_DN22631_c1_g1	1	--	134.66	62.52	0.45
TRINITY_DN35187_c4_g3	1	UBQ, Ubiquitin homologues.	131.87	527.27	136.9
TRINITY_DN29952_c1_g1	1	UBQ, Ubiquitin homologues.	121.29	496.99	105.43
TRINITY_DN34070_c2_g1	1	YbiR_permease, Putative anion permease YbiR.	119.96	135.95	91.19
TRINITY_DN25531_c0_g1	0.97	Leucoanthocyanidin dioxygenase like protein.	116.86	46.77	12.23
TRINITY_DN34794_c0_g1	1	--	112.75	143.01	99.63
TRINITY_DN35505_c0_g1	1	Peptidases_S8_3, Peptidase S8 family domain, uncharacterized subfamily 3.	109.05	184.65	101.9
TRINITY_DN30877_c0_g2	1	AP2, DNA-binding domain in plant proteins such as APETALA2 and EREBPs.	107.25	138.56	182.45
TRINITY_DN34794_c1_g2	0.97	GH18_plant_chitinase_class_V, The class V plant chitinases have a glycosyl hydrolase family 18 (GH18) domain.	104.44	133.91	83.9
TRINITY_DN14946_c0_g1	0.97	HPS_like, HPS_like: Hydrophobic Protein from Soybean (HPS)-like subfamily.	103.49	110.17	111.73
TRINITY_DN33224_c0_g2	1	Secretory peroxidase.	101.63	40.35	15.92
TRINITY_DN3141_c0_g1	1	Ceramidase_alk, Neutral/alkaline non-lysosomal ceramidase, N-terminal.	101.31	109.05	56.23
TRINITY_DN29814_c0_g1	1	PRX_1cys, Peroxiredoxin (PRX) family, 1-cys PRX subfamily.	99.75	85.42	4.42
TRINITY_DN28903_c0_g1	0.98	PLAC8, PLAC8 family. This family includes the Placenta-specific gene 8 protein.	98.6	120.39	52.46
TRINITY_DN21742_c0_g1	1	Cupin_1, Cupin.	94.15	100.22	86.61

TRINITY_DN17745_c0_g1	0.98	Bet_v1-like, Ligand-binding bet_v_1 domain of major pollen allergen of white birch (Betula verrucosa).	93.33	69.44	41.82
TRINITY_DN30668_c0_g1	0.99	RPE65, Retinal pigment epithelial membrane protein.	91.74	369.2	253.61
TRINITY_DN28917_c0_g1	1	FRQ1, Ca2+-binding protein (EF-Hand superfamily).	90.86	54.34	21.73
TRINITY_DN35141_c0_g1	0.98	retinol-DH_like_SDR_c_like, retinol dehydrogenase (retinol-DH).	89.34	167.46	118.15
TRINITY_DN29764_c0_g1	1	--	89.28	102.65	35.54
TRINITY_DN29992_c0_g2	1	--	88.57	21.76	4.99
TRINITY_DN35187_c4_g2	1	UBQ, Ubiquitin homologues.	85.87	60.64	38.19
TRINITY_DN26832_c0_g1	1	Grp1_Fun34_YaaH, GPR1/FUN34/yaaH family.	85.78	37.44	3.44
TRINITY_DN316_c0_g1	1	cnd41_like, Chloroplast Nucleoids DNA-binding Protease.	85.76	124.52	56.24
TRINITY_DN27090_c1_g1	0.98	glucosyltransferase; Provisional.	85.54	77.63	42.72
TRINITY_DN29990_c0_g1	1	SYLF_SH3YL1_like, The SYLF domain (also called DUF500).	84.45	58.83	18.32
TRINITY_DN30832_c0_g1	0.96	Kelch, Kelch repeat. .	84.41	123.04	51.85
TRINITY_DN33274_c0_g1	1	Manganese transport protein MntH; Reviewed.	83.86	135.87	111.62
TRINITY_DN19606_c0_g1	1	--	81.68	47.79	0.24
TRINITY_DN35187_c4_g1	1	UBQ, Ubiquitin homologues.	80.65	88.46	31.19
TRINITY_DN35725_c1_g1	0.96	GST_C_Tau, C-terminal, alpha helical domain of Class Tau Glutathione S-transferases.	78.75	92.43	74.6
TRINITY_DN26156_c0_g1	1	PTR2, POT family.	76.93	203.03	100.45
TRINITY_DN32387_c0_g1	0.96	Phytoene synthase.	74.93	156.07	133.65
TRINITY_DN33963_c0_g1	1	TEF1, Translation elongation factor EF-1alpha (GTPase).	74.58	242.53	40.74
TRINITY_DN40308_c0_g1	1	--	73.75	33.03	0
TRINITY_DN35271_c0_g1	1	HSP70, Hsp70 protein.	73.08	187.7	22.9
TRINITY_DN24994_c0_g1	1	--	72.14	103.92	27.41
TRINITY_DN31629_c0_g1	1	Nucleic acid binding protein; Provisional.	70.57	80.89	28.66
TRINITY_DN30877_c0_g1	0.99	AP2, DNA-binding domain in plant proteins such as APETALA2 and EREBPs.	69.28	54.45	155.32
TRINITY_DN24609_c0_g1	1	DNA polymerase III subunits gamma and tau; Validated.	68.01	28.76	9.17
TRINITY_DN65542_c0_g1	0.98	p450, Cytochrome P450.	65.94	13.21	31.68
TRINITY_DN58002_c0_g1	1	Short chain dehydrogenase; Provisional.	65.53	75	3.36
TRINITY_DN24608_c0_g1	1	--	65.17	54.65	21.93
TRINITY_DN35111_c0_g2	1	GH18_narbonin .	63.56	62.34	83.54
TRINITY_DN26853_c0_g1	0.95	Caleosin, Caleosin related protein.	63.4	38.91	48.71

TRINITY_DN30469_c0_g1	1	--	62.88	141.19	14.72
TRINITY_DN86038_c0_g1	0.98	--	62.8	82.77	71.18
TRINITY_DN29992_c0_g1	1	--	62.8	70.37	5.65
TRINITY_DN24775_c0_g1	1	GH18_narbonin.	62.74	65.36	57.15
TRINITY_DN22083_c0_g1	1	Transcriptional regulator ICP4; Provisional.	61.99	137.39	53.54
TRINITY_DN26489_c0_g1	1	Alpha_tubulin, The alpha-tubulin family.	61.6	51.72	13.37
TRINITY_DN1085_c0_g1	0.99	Cupin_1, Cupin.	61.19	117.95	63.24
TRINITY_DN30814_c0_g1	1	GABARAP, Ubiquitin domain of GABA-receptor-associated protein.	60.51	123.49	29.22
TRINITY_DN13151_c0_g1	1	Grp1_Fun34_YaaH, GPR1/FUN34/yaaH family.	59.84	54.16	2.46
TRINITY_DN46546_c0_g1	1	--	59.46	68.67	27.97
TRINITY_DN20970_c0_g1	1	DUF588, Domain of unknown function (DUF588).	59.24	97.52	56.09
TRINITY_DN4499_c0_g1	1	--	58.63	56.02	25.02
TRINITY_DN32701_c0_g1	1	Rab11_like, Rab GTPase family 11 (Rab11)-like includes Rab11a, Rab11b, and Rab25. Rab11a.	57.84	60.85	18.59
TRINITY_DN34321_c0_g1	1	MATA_HMG-box, MATA_HMG-box, class I member of the HMG-box superfamily of DNA-binding proteins.	57.83	56.38	16.52
TRINITY_DN22846_c0_g1	0.97	Secretory peroxidase.	56.29	9.36	24.91
TRINITY_DN31839_c0_g1	1	ACTIN, Actin. ACTIN subfamily of ACTIN/mreB/sugarkinase/Hsp70 superfamily.	56.21	170.75	37.54
TRINITY_DN34096_c0_g1	0.99	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	55.38	76.86	44.18
TRINITY_DN29462_c0_g1	0.99	Cu-oxidase_3, Multicopper oxidase.	55.18	90.68	54.15
TRINITY_DN34185_c0_g1	1	RRM_HP0827_like.	54.86	73.94	25.5
TRINITY_DN33662_c0_g1	1	Beta_tubulin, The beta-tubulin family.	54.26	66.32	13.6
TRINITY_DN31799_c0_g2	1	p450, Cytochrome P450.	54.24	121.9	115.2
TRINITY_DN10293_c0_g1	0.99	Peptidase_C1, Papain family cysteine protease.	54.24	55.83	37.81
TRINITY_DN33153_c0_g1	0.95	--	53.39	107.59	64.87
TRINITY_DN32467_c0_g1	0.99	UDP-glycosyltransferase family protein.	53.29	44.72	31.36
TRINITY_DN31288_c0_g1	1	GlnA, Glutamine synthetase.	52.32	85.52	25.28
TRINITY_DN35333_c2_g1	0.98	DUF296, Domain of unknown function (DUF296).	52.28	53.34	46.07
TRINITY_DN32581_c0_g1	1	Histone, Core histone H2A/H2B/H3/H4.	51.79	96.32	21.34
TRINITY_DN27284_c0_g1	1	--	50.86	39.11	43.52
TRINITY_DN30242_c0_g1	1	DUF2298, Uncharacterized membrane protein (DUF2298).	50.06	33.09	52.47
TRINITY_DN27583_c0_g1	1	DUF3421, Protein of unknown function (DUF3421).	49.7	32.68	9.23
TRINITY_DN33089_c0_g1	1	Ornithine_aminotransferase, ornithine aminotransferase.	49.6	76.24	5.3
TRINITY_DN24604_c0_g1	1	Pro_isomerase, Cyclophilin type peptidyl-prolyl cis-trans isomerase/CLD.	49.51	90.53	9.15

TRINITY_DN32216_c0_g1	1	Malectin_like, Carbohydrate-binding protein of the ER.	48.14	66.49	47.55
TRINITY_DN26322_c0_g1	0.97	gibberellin 20-oxidase.	48	30.37	22.4
TRINITY_DN27440_c0_g1	1	Cu-Zn_Superoxide_Dismutase, Copper/zinc superoxide dismutase (SOD).	47.96	31.85	5.86
TRINITY_DN31488_c0_g1	1	NAM, No apical meristem (NAM) protein.	47.5	105.06	87.03
TRINITY_DN15672_c0_g1	1	Ammonium_transporter, ammonium transporter.	47.27	66.94	44.44
TRINITY_DN29052_c0_g1	1	Alginate_lyase2, Alginate lyase.	46.46	110.66	100.09
TRINITY_DN23337_c0_g1	1	STKc_MAP3K-like, Catalytic domain of Mitogen-Activated Protein Kinase (MAPK) Kinase Kinase-like Serine/Threonine kinases.	46.14	30.12	8.88
TRINITY_DN31577_c0_g1	0.99	PTR2, POT family.	45.85	60.64	36.86
TRINITY_DN35433_c0_g1	1	RRM4_I_PABPs, RNA recognition motif 4 in type I polyadenylate-binding proteins.	45.65	98.59	29.26
TRINITY_DN1914_c0_g1	1	DUF4228, Domain of unknown function (DUF4228).	45.52	44.11	18.32
TRINITY_DN35862_c0_g1	1	SOG2, RAM signalling pathway protein.	45.24	118.13	91.14
TRINITY_DN33288_c0_g1	1	--	44.57	68.61	56.49
TRINITY_DN32322_c1_g1	1	RRM_HP0827_like, RNA recognition motif in Helicobacter pylori HP0827 protein and similar proteins.	44.47	57.69	7.04
TRINITY_DN27995_c0_g1	1	Abhydro_lipase, Partial alpha/beta-hydrolase lipase region.	44.46	68.85	153.39
TRINITY_DN19886_c0_g1	1	GRAS, GRAS domain family.	44.3	84.04	43.8
TRINITY_DN31417_c0_g1	0.99	PPX1, Inorganic pyrophosphatase/exopolyphosphatase [Energy production and conversion].	43.85	64.25	35.29
TRINITY_DN35183_c0_g1	1	HSP90, Hsp90 protein.	42.93	89.23	13.82
TRINITY_DN24495_c0_g1	0.97	C2_SRC2_like, C2 domain.	42.85	106.9	59
TRINITY_DN27764_c0_g1	1	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	42.56	52.56	33.04
TRINITY_DN29690_c0_g1	1	60S ribosomal protein L34; Provisional.	42.27	41.88	6.93
TRINITY_DN31659_c0_g1	1	rps5p, 30S ribosomal protein S5P; Reviewed.	42.25	93.95	20.8
TRINITY_DN28128_c0_g1	1	p450, Cytochrome P450.	42.22	80.85	60.73
TRINITY_DN23104_c0_g1	1	Ribosomal_S25, S25 ribosomal protein.	41.71	129.08	13.08
TRINITY_DN27284_c2_g1	0.98	--	41.7	46.01	58.51
TRINITY_DN20411_c0_g1	0.96	Protein_white, The Eye Pigment Precursor Transporter (EPP) Family protein.	41.51	98.6	111.69
TRINITY_DN31766_c0_g1	1	Phosphoadenosine phosphosulfate reductase; Provisional.	41.31	47.68	13.29
TRINITY_DN26448_c0_g1	1	eIF1_SUI1, Eukaryotic initiation factor 1.	40.16	41.87	5.77
TRINITY_DN31672_c1_g1	1	DUF966, Domain of unknown function (DUF966).	39.98	44.8	40.63

TRINITY_DN27248_c0_g1	1	p450, Cytochrome P450.	39.65	81.49	43.5
TRINITY_DN28593_c0_g1	1	ATP-synt_C, ATP synthase subunit C.	39.34	92.3	17.92
TRINITY_DN26668_c0_g1	1	ZnuB, ABC-type Mn2+/Zn2+ transport systems, permease components.	39.34	38.7	23.36
TRINITY_DN32031_c0_g1	1	Sterol_MT_C, Sterol methyltransferase C-terminal.	39.32	90.34	17.93
TRINITY_DN32882_c0_g1	1	Asp, Eukaryotic aspartyl protease. Aspartyl (acid) proteases include pepsins, cathepsins, and renins.	39.2	78.24	9.04
TRINITY_DN35441_c1_g1	1	Plasma_membrane_ATPase, plasma-membrane proton-efflux P-type ATPase.	38.77	61.59	16.85
TRINITY_DN29025_c0_g1	1	rpl14p, 50S ribosomal protein L14P; Reviewed.	38.73	63.19	12.16
TRINITY_DN31766_c0_g2	1	phosphoadenosine phosphosulfate reductase.	38.61	21.32	11.19
TRINITY_DN28958_c0_g1	1	bZIP_GCN4, Basic leucine zipper (bZIP) domain of General control protein GCN4.	38.35	27.32	12.46
TRINITY_DN34243_c0_g1	1	TonB_N, TonB N-terminal region.	38.34	25.5	38.19
TRINITY_DN67125_c0_g1	1	DUF4228, Domain of unknown function (DUF4228).	38.2	49.95	39.34
TRINITY_DN25866_c0_g1	1	NDPk_I, Nucleoside diphosphate kinase Group I (NDPk_I)-like:	38.11	61.02	23.18
TRINITY_DN35111_c0_g3	1	GH18_narbonin.	38	46.45	34.61
TRINITY_DN32519_c0_g1	1	ZnF_GATA, Zinc finger DNA binding domain;	37.94	85.46	10.65
TRINITY_DN29702_c0_g1	1	Ran, Ras-related nuclear proteins (Ran)/TC4 family of small GTPases.	37.9	87.93	20.91
TRINITY_DN14916_c0_g1	1	Synaptobrevin, Synaptobrevin.	37.81	36.52	16.8
TRINITY_DN63449_c0_g1	0.99	Transcriptional regulator ICP4; Provisional.	36.59	59.61	45.2
TRINITY_DN27334_c0_g1	1	ADF_cofilin_like, Cofilin, Destrin, and related actin depolymerizing factors.	36.47	98.99	29.08
TRINITY_DN32746_c1_g1	1	ChaA, Ca2+/H+ antiporter [Inorganic ion transport and metabolism].	36.37	39.42	6.18
TRINITY_DN34259_c0_g5	0.99	Double_bond_reductase_like, Arabidopsis alkenal double bond reductase and leukotriene B4 12-hydroxydehydrogenase.	36.32	52.81	43.99
TRINITY_DN14333_c0_g1	0.96	--	36.32	45.73	29.74
TRINITY_DN28470_c0_g1	1	Skp1, Skp1 family, dimerization domain.	35.8	67.96	15.94
TRINITY_DN34995_c0_g1	1	SOBP, Sine oculis-binding protein.	35.75	119.64	116.86
TRINITY_DN4152_c0_g1	0.99	GRAS, GRAS domain family.	35.53	40.51	25.93
TRINITY_DN34400_c0_g1	1	TT_ORF1, TT viral orf 1. TT virus (TTV).	35.36	133.33	118.74
TRINITY_DN17544_c0_g1	1	--	34.93	70.64	34.13
TRINITY_DN15952_c0_g1	0.95	AP2, DNA-binding domain in plant proteins such as APETALA2 and EREBPs.	34.87	34.23	22.32
TRINITY_DN28883_c0_g1	1	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	34.82	55.51	35.67
TRINITY_DN34169_c0_g1	1	14-3-3, 14-3-3 protein.	34.69	77.61	19.65

TRINITY_DN8479_c0_g1	0.98	--	34.67	28.17	38.39
TRINITY_DN28997_c0_g1	1	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	33.95	23.94	18.88
TRINITY_DN33432_c0_g1	1	HLH, Helix-loop-helix domain.	33.56	37.06	13.08
TRINITY_DN65454_c0_g1	1	--	33.54	27.26	23.91
TRINITY_DN32351_c0_g1	1	Ubiquitin-protein ligase [Posttranslational modification, protein turnover, chaperones].	33.32	81.94	22.58
TRINITY_DN35495_c0_g1	1	MdlB, ABC-type multidrug transport system, ATPase and permease components [Defense mechanisms].	33.26	117.17	65.91
TRINITY_DN30654_c0_g1	1	Mito_carr, Mitochondrial carrier protein.	33.2	79.67	15.62
TRINITY_DN24378_c0_g1	1	TLP-PA, allergenic/antifungal thaumatin-like proteins: plant and animal homologs.	33.18	21.98	38.07
TRINITY_DN20568_c0_g1	1	40S ribosomal protein S15; Provisional.	33.06	58.96	11.43
TRINITY_DN17112_c0_g1	1	Cyt-b5, Cytochrome b5-like Heme/Steroid binding domain.	32.86	21.43	2.12
TRINITY_DN17954_c0_g1	1	PLAC8, PLAC8 family.	32.81	30.62	34.45
TRINITY_DN32205_c1_g1	1	Glucosyltransferase; Provisional.	32.52	70.93	43.77
TRINITY_DN30876_c0_g1	1	--	32.5	58.39	11.58
TRINITY_DN31634_c0_g1	1	WcaG, Nucleoside-diphosphate-sugar epimerases.	32.5	56.46	16.29
TRINITY_DN35720_c0_g1	1	DUF1421, Protein of unknown function (DUF1421).	32.45	86.02	55.53
TRINITY_DN19825_c0_g1	1	cnd41_like, Chloroplast Nucleoids DNA-binding Protease.	32.36	134.32	72.66
TRINITY_DN19906_c0_g1	1	PB1, PB1 domain. Phox and Bem1p domain, present in many eukaryotic cytoplasmic signalling proteins.	32.3	30.71	7.11
TRINITY_DN26182_c0_g1	1	C2, C2 domain.	32.08	34.96	8.37
TRINITY_DN30989_c0_g1	1	Ribosomal_S23, S12-like family, 40S ribosomal protein S23 subfamily.	32.06	49.2	13.74
TRINITY_DN20586_c0_g1	1	Ribosomal_L36e, Ribosomal protein L36e.	31.92	58.99	10.8
TRINITY_DN32606_c0_g1	1	60S_ribosomal_protein_L7-A, 60S ribosomal protein uL30, eukaryotic form.	31.9	53.89	10.92
TRINITY_DN26066_c0_g1	1	Arginase, Arginase family.	31.78	32.54	10.31
TRINITY_DN16483_c0_g1	1	GRAS, GRAS domain family.	31.69	46.84	39.21
TRINITY_DN31572_c0_g1	1	rpl10e, 50S ribosomal protein L10e; Reviewed.	31.57	73.06	12.43
TRINITY_DN32365_c0_g1	0.98	--	30.92	33.47	35.86
TRINITY_DN26294_c0_g1	1	UBQ, Ubiquitin homologues.	30.86	76.15	51.6
TRINITY_DN29693_c0_g1	1	60S ribosomal protein L5; Provisional.	30.57	53.28	12.53
TRINITY_DN34440_c0_g1	1	DUF1421, Protein of unknown function (DUF1421).	30.37	21.38	9.92
TRINITY_DN27915_c0_g1	1	DUF1421, Protein of unknown function (DUF1421).	30.21	21.83	9.03
TRINITY_DN35711_c1_g1	0.95	--	30.19	25.89	39.27

TRINITY_DN28798_c0_g1	1	GapA, Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase.	**29.62**	**48.73**	13.58
TRINITY_DN32433_c0_g1	1	ACTIN, Actin. ACTIN subfamily of ACTIN/mreB/sugarkinase/Hsp70 superfamily.	**29.35**	**66.5**	11.06
TRINITY_DN21087_c0_g1	0.99	double_bond_reductase_like, Arabidopsis alkenal double bond reductase and leukotriene B4 12-hydroxydehydrogenase.	**29.06**	**47.88**	**37.63**
TRINITY_DN32548_c0_g1	1	Bestrophin, Bestrophin, RFP-TM, chloride channel.	**28.82**	**66.33**	11.95
TRINITY_DN27773_c0_g1	1	--	**28.76**	**41.51**	**29.82**
TRINITY_DN43780_c0_g1	1	C2, C2 domain. The C2 domain was first identified in PKC.	**28.72**	**148.36**	2.96
TRINITY_DN30450_c0_g1	1	biotin_synthase, biotin synthase. Catalyzes the last step of the biotin biosynthesis pathway.	**28.71**	**51.45**	6.97
TRINITY_DN29127_c0_g1	1	p450, Cytochrome P450.	**28.19**	**23.23**	0.65
TRINITY_DN26869_c0_g1	1	RRM_RCAN_like, RNA recognition motif in regulators of calcineurin (RCANs) and similar proteins.	**28.15**	**35.12**	11.8
TRINITY_DN26385_c0_g1	1	Cyt-b5, Cytochrome b5-like Heme/Steroid binding domain.	**27.99**	**40.37**	10.76
TRINITY_DN29902_c0_g1	1	CE4_MrCDA_like, Catalytic NodB homology domain of Mucor rouxii chitin deacetylase and similar proteins.	**27.92**	**37.11**	4.24
TRINITY_DN25944_c0_g1	1	ACTIN, Actin. ACTIN subfamily of ACTIN/mreB/sugarkinase/Hsp70 superfamily.	**27.91**	**117.84**	**23.76**
TRINITY_DN29037_c0_g1	1	fructose-bisphosphate aldolase; Provisional.	**27.82**	**46.19**	7.44
TRINITY_DN27272_c0_g1	0.99	CBFD_NFYB_HMF, Histone-like transcription factor (CBF/NF-Y) and archaeal histone.	**27.8**	**20.48**	**29.29**
TRINITY_DN24244_c0_g1	1	Root_cap, Root cap.	**27.6**	**25.87**	8.35
TRINITY_DN6809_c0_g1	0.96	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	**27.59**	**35.55**	**44.33**
TRINITY_DN32538_c0_g1	1	RHO, Rho (Ras homology) subfamily of Ras-like small GTPases.	**27.55**	**61.5**	14.11
TRINITY_DN31269_c0_g1	0.99	p450, Cytochrome P450.	**27.38**	**71.63**	**42.91**
TRINITY_DN21897_c0_g1	1	Hypothetical_protein, oligosaccharide repeat unit polymerase.	**27.17**	**23.47**	5.77
TRINITY_DN29771_c0_g1	1	LDH_like, L-lactate dehydrogenase-like enzymes.	**27.1**	**42.74**	8.89
TRINITY_DN28488_c0_g1	1	MDH_glyoxysomal_mitochondrial, Glyoxysomal and mitochondrial malate dehydrogenases.	**27.07**	**63.62**	7.15
TRINITY_DN23519_c0_g1	1	Ribosomal_S26e, Ribosomal protein S26e.	**27.02**	**93.3**	10.75
TRINITY_DN25997_c0_g1	1	Histone, Core histone H2A/H2B/H3/H4.	**26.76**	**32.28**	8.98
TRINITY_DN31011_c0_g1	1	SET, SET domain.	**26.75**	**50.4**	7.39
TRINITY_DN31096_c0_g1	0.96	K_trans, K+ potassium transporter.	**26.67**	**20.93**	19.1

TRINITY_DN26428_c0_g1	1	--	**26.56**	**52.44**	**22.28**
TRINITY_DN30257_c0_g1	1	rpl2p, 50S ribosomal protein L2P; Validated.	**26.43**	**50.96**	12.23
TRINITY_DN27532_c0_g1	1	REB1, Myb superfamily proteins, including transcription factors and mRNA splicing factors.	**26.22**	**87.94**	**34.06**
TRINITY_DN31711_c0_g1	1	Arf, ADP-ribosylation factor family.	**26.06**	**59.07**	14.91
TRINITY_DN29021_c0_g1	1	60S ribosomal protein L27a; Provisional.	**25.93**	**40.94**	6.76
TRINITY_DN21241_c0_g1	1	HHT1, Histones H3 and H4.	**25.91**	**51.91**	4.65
TRINITY_DN31543_c0_g1	1	Sar1, Sar1 is an essential component of COPII vesicle coats.	**25.91**	**42.51**	9.97
TRINITY_DN24750_c0_g1	1	HHT1, Histones H3 and H4.	**25.87**	**49.06**	16.57
TRINITY_DN29550_c0_g2	1	--	**25.85**	**22.97**	16.13
TRINITY_DN25453_c0_g1	1	Ribosomal_L40e, Ribosomal L40e family.	**25.78**	**111.04**	14.61
TRINITY_DN32629_c0_g1	1	Peptidases_S8_PCSK9_ProteinaseK_like, Peptidase S8 family domain in ProteinaseK-like proteins.	**25.61**	**23.48**	6.81
TRINITY_DN31290_c0_g1	0.98	Mlo, Mlo family.	**25.58**	**39.36**	**28.57**
TRINITY_DN35521_c0_g1	1	Dynein_heavy, Dynein heavy chain and region D6 of dynein motor.	**25.51**	**52.97**	12.92
TRINITY_DN34138_c0_g1	1	Peptidase_S10, Serine carboxypeptidase.	**25.39**	**46.83**	7.62
TRINITY_DN28333_c0_g1	1	RPL30, Ribosomal protein L30E [Translation, ribosomal structure and biogenesis].	**25.35**	**60.22**	3.77
TRINITY_DN23116_c0_g1	1	AP2, DNA-binding domain in plant proteins such as APETALA2 and EREBPs.	**25.25**	**75.51**	**58.56**
TRINITY_DN31840_c0_g1	1	Rab8_Rab10_Rab13_like, Rab GTPase families 8, 10, 13.	**25.1**	**44.66**	11.44
TRINITY_DN33251_c0_g1	1	NOX_Duox_like_FAD_NADP.	**25.08**	**37.49**	8.09
TRINITY_DN24523_c0_g1	1	4CL, 4-Coumarate-CoA Ligase (4CL). 4-Coumarate:coenzyme A ligase.	**24.95**	**47.21**	**21.86**
TRINITY_DN17529_c0_g1	0.97	unnamed_protein_product, F-box protein interaction domain.	**24.85**	**56.23**	**34.97**
TRINITY_DN23105_c0_g1	1	--	**24.84**	**572.81**	**183.49**
TRINITY_DN34005_c0_g1	1	Chs3p, Chitin synthase III catalytic subunit.	**24.79**	**43.47**	10.05
TRINITY_DN35906_c0_g1	1	S_TKc, Serine/Threonine protein kinases, catalytic domain. Phosphotransferases.	**24.79**	**31.11**	2.83
TRINITY_DN29137_c0_g1	1	p450, Cytochrome P450.	**24.51**	**29.23**	5.67
TRINITY_DN34996_c0_g1	1	--	**24.37**	**26.58**	2.07
TRINITY_DN32296_c1_g1	1	RAS, Ras subfamily of RAS small GTPases. Similar in fold and function to the bacterial EF-Tu GTPase.	**24.32**	**70.07**	11.56
TRINITY_DN29071_c0_g1	1	FKBP_C, FKBP-type peptidyl-prolyl cis-trans isomerase.	**24.19**	**32.75**	5.19
TRINITY_DN23776_c0_g1	0.99	K-box, K-box region.	**24.13**	**22.15**	18.83
TRINITY_DN29008_c0_g1	1	60S ribosomal protein L13a; Provisional.	**24.03**	**44.19**	7.66
TRINITY_DN28979_c0_g1	1	--	**24.02**	**37.09**	8.93

TRINITY_DN34728_c0_g1	1	SufI, Putative multicopper oxidases.	24.02	34.42	18.59
TRINITY_DN27396_c0_g1	1	--	24.01	26.36	14.97
TRINITY_DN29981_c0_g1	1	RPL18A, Ribosomal protein L18E.	23.99	39.3	9.7
TRINITY_DN18261_c0_g1	0.96	DUF296, Domain of unknown function (DUF296).	23.85	30.78	30.71
TRINITY_DN33869_c0_g1	1	MFS, The Major Facilitator Superfamily (MFS).	23.82	36.53	9.72
TRINITY_DN31803_c0_g3	0.95	Nucleic acid binding protein; Provisional.	23.74	37.53	32.55
TRINITY_DN29933_c0_g1	1	--	23.73	47.25	9.17
TRINITY_DN27643_c0_g1	1	Grp1_Fun34_YaaH, GPR1/FUN34/yaaH family.	23.73	21.08	3.93
TRINITY_DN35263_c0_g1	1	Peptidases_S8_3, Peptidase S8 family domain, uncharacterized subfamily 3.	23.59	23.03	4.77
TRINITY_DN31565_c0_g1	1	p450, Cytochrome P450.	23.51	92.51	50.74
TRINITY_DN3560_c0_g1	0.98	alpha_CA_prokaryotic_like, Carbonic anhydrase alpha.	23.51	27.04	11.26
TRINITY_DN34540_c0_g1	1	CDC3, Septin family protein.	23.34	41.1	14.55
TRINITY_DN33761_c0_g1	1	Rab5_related, Rab-related GTPase family includes Rab5 and Rab22; regulates early endosome fusion.	23.07	39.57	8.3
TRINITY_DN18890_c0_g1	1	p450, Cytochrome P450.	22.96	48.21	41.61
TRINITY_DN30966_c0_g1	1	, Rab8_Rab10_Rab13_like, Rab GTPase families 8, 10, 13 (Rab8, Rab10, Rab13).	22.85	43.78	13.58
TRINITY_DN31805_c0_g1	1	STKc_IRAK, Catalytic domain of the Serine/Threonine kinases, Interleukin-1 Receptor Associated Kinases and related STKs.	22.83	42.76	21.33
TRINITY_DN22218_c0_g1	0.98	NB-ARC, NB-ARC domain.	22.76	23.06	14.21
TRINITY_DN84412_c0_g1	1	--	22.72	86.74	34.69
TRINITY_DN27141_c0_g1	1	C2, C2 domain.	22.7	27.04	7.04
TRINITY_DN21747_c0_g1	1	TB2_DP1_HVA22, TB2/DP1, HVA22 family.	22.67	27.64	4.89
TRINITY_DN5458_c0_g1	1	TRX_family, TRX family.	22.63	28.83	0
TRINITY_DN12525_c0_g1	0.99	PTR2, POT family (proton-dependent oligopeptide transport).	22.62	49.04	17.41
TRINITY_DN16720_c0_g1	1	GRAS, GRAS domain family.	22.54	64.66	65.2
TRINITY_DN33978_c0_g1	1	Ribosomal_S3Ae, Ribosomal S3Ae family.	22.53	43.29	11.49
TRINITY_DN34633_c0_g1	1	Transcriptional regulator ICP4; Provisional.	22.49	41.69	25.32
TRINITY_DN32282_c0_g1	1	HSP70, Hsp70 protein. Hsp70 chaperones help to fold many proteins.	22.48	32.4	5.54
TRINITY_DN23018_c0_g1	1	DUF4452, Domain of unknown function (DUF4452).	22.46	52.97	26.86
TRINITY_DN14954_c0_g1	1	Uncharacterized conserved protein [Function unknown].	22.42	22.48	1.64
TRINITY_DN33794_c4_g1	0.99	unnamed_protein_product, myb-like DNA-binding domain, SHAQKYF class.	22.32	27.84	24
TRINITY_DN29164_c0_g1	0.98	--	22.32	6	21.47
TRINITY_DN31734_c1_g1	1	60S ribosomal protein L32; Provisional.	22.1	26.93	4.09

TRINITY_DN29123_c0_g1	1	MdlB, ABC-type multidrug transport system, ATPase and permease components.	**22.05**	**51.59**	**28.06**
TRINITY_DN26470_c0_g1	1	Ribosomal_L19e_E, Ribosomal protein L19e, eukaryotic.	**21.89**	**65.55**	9.98
TRINITY_DN26168_c0_g1	1	rpl3p, 50S ribosomal protein L3P; Reviewed.	**21.7**	**59.17**	8.78
TRINITY_DN26607_c0_g1	1	zf-rbx1, RING-H2 zinc finger.	**21.69**	**27.45**	7.16
TRINITY_DN26829_c0_g1	1	Ribosomal_L44, Ribosomal protein L44.	**21.67**	**51.75**	11.42
TRINITY_DN34257_c0_g1	1	ARM, armadillo repeat. Armadillo (ARM)/beta-catenin-like repeats.	**21.6**	**39.19**	9.27
TRINITY_DN30361_c0_g1	1	--	**21.56**	**21.9**	3.73
TRINITY_DN27854_c0_g1	1	--	**21.48**	**21.16**	5.51
TRINITY_DN32072_c0_g1	1	DnaJ-class molecular chaperone with C-terminal Zn finger domain.	**21.46**	**29.18**	9.67
TRINITY_DN41011_c0_g1	1	Ornithine_aminotransferase, ornithine aminotransferase.	**21.38**	**20.82**	3.69
TRINITY_DN35535_c0_g1	1	Homeobox, Homeobox domain.	**21.33**	**30.02**	13.41
TRINITY_DN33615_c0_g1	1	HLH, Helix-loop-helix domain, found in specific DNA- binding proteins that act as transcription factors.	**21.19**	**33.86**	7.46
TRINITY_DN34776_c0_g1	1	bZIP, Basic leucine zipper (bZIP) domain of bZIP transcription factors	**21.18**	**37.82**	5.66
TRINITY_DN29436_c0_g1	1	Vma12, Endoplasmic reticulum-based factor for assembly of V-ATPase.	**21.11**	**47.56**	3.82
TRINITY_DN29712_c0_g1	1	Protein white, The Eye Pigment Precursor Transporter (EPP).	**21**	**40.08**	**25.89**
TRINITY_DN32611_c0_g1	1	EF1_GNE, EF-1 guanine nucleotide exchange domain. This family is the guanine nucleotide exchange domain of EF-1 beta and EF-1 delta chains.	**20.85**	**28.71**	6.95
TRINITY_DN29339_c0_g1	1	Ribosomal_L24e, Ribosomal protein L24e.	**20.75**	**53**	8.48
TRINITY_DN33535_c0_g1	1	Elongation factor EF-2; Reviewed.	**20.59**	**47.34**	5.31
TRINITY_DN30305_c0_g1	1	KOW_RPL6, KOW motif of Ribosomal Protein L6.	**20.47**	**61.34**	5.82
TRINITY_DN27450_c0_g1	1	rps9p, 30S ribosomal protein S9P; Reviewed.	**20.31**	**48.48**	8.16
TRINITY_DN30513_c0_g1	1	--	**327.34**	**210.72**	7.78

Additional file 2. Tissue-specific transcripts in expanded families of macauba palm with respective SPM values by tissue analyzed. SPM value ≥ 0.95 values in pink-colored box.

Family	Fruit pulp	Leaf	Leaf sheath	Bulb	Root	Female flower	Male flower
OG0000015							
TRINITY_DN16760_c0_g1	0.21	0.32	0.56	0.25	0.39	0.37	0.44
TRINITY_DN20595_c0_g1	0.34	0.43	0.28	0.32	0.47	0.32	0.44
TRINITY_DN21123_c0_g1	0.21	0.31	0.51	0.35	0.52	0.22	0.40
TRINITY_DN21123_c0_g2	0.43	0.29	0.37	0.41	0.53	0.21	0.30
TRINITY_DN23177_c0_g1	0.61	0.41	0.34	0.31	0.34	0.22	0.30
TRINITY_DN26488_c0_g1	0.48	0.42	0.21	0.34	0.52	0.22	0.34
TRINITY_DN30512_c0_g1	0.33	0.42	0.38	0.29	0.42	0.31	0.47
TRINITY_DN33294_c1_g1	0.45	0.32	0.43	0.31	0.40	0.31	0.39
TRINITY_DN33687_c0_g1	0.37	0.47	0.38	0.34	0.46	0.20	0.36
TRINITY_DN33687_c0_g2	0.26	0.46	0.42	0.32	0.54	0.21	0.33
TRINITY_DN35010_c1_g1	0.30	0.45	0.40	0.34	0.48	0.23	0.39
TRINITY_DN35097_c2_g1	0.20	0.52	0.41	0.35	0.40	0.28	0.39
TRINITY_DN35678_c2_g1	0.37	0.42	0.36	0.29	0.43	0.35	0.41
TRINITY_DN35722_c0_g1	0.49	0.23	0.37	0.42	0.42	0.31	0.34
TRINITY_DN35722_c0_g2	0.56	0.20	0.30	0.42	0.41	0.34	0.31
TRINITY_DN35914_c0_g1	0.33	0.43	0.39	0.35	0.43	0.28	0.41
TRINITY_DN35914_c0_g3	0.34	0.43	0.40	0.36	0.45	0.27	0.36
TRINITY_DN8915_c0_g1	0.45	0.13	0.15	0.36	0.33	0.46	0.56
OG0000024							
TRINITY_DN12992_c0_g1	0.17	0.42	0.53	0.19	0.68	0.12	0.11
TRINITY_DN13894_c0_g1	0.17	0.49	0.36	0.29	0.44	0.45	0.34
TRINITY_DN14263_c0_g1	0.15	0.42	0.47	0.34	0.67	0.09	0.11
TRINITY_DN15704_c0_g1	0.10	0.45	0.41	0.25	0.49	0.34	0.45
TRINITY_DN17326_c0_g1	0.79	0.19	0.33	0.26	0.34	0.12	0.17
TRINITY_DN17383_c0_g1	0.01	0.00	0.13	0.21	0.97	0.06	0.03
TRINITY_DN19679_c0_g1	0.28	0.48	0.47	0.38	0.40	0.32	0.26
TRINITY_DN20032_c0_g1	0.20	0.30	0.27	0.20	0.76	0.25	0.34
TRINITY_DN20759_c0_g1	0.06	0.38	0.49	0.20	0.60	0.17	0.44
TRINITY_DN21019_c0_g1	0.16	0.46	0.40	0.32	0.49	0.43	0.27
TRINITY_DN21432_c0_g1	0.23	0.44	0.19	0.20	0.78	0.16	0.17
TRINITY_DN2271_c0_g1	0.00	0.53	0.39	0.17	0.47	0.36	0.44
TRINITY_DN22952_c0_g1	0.11	0.49	0.41	0.25	0.46	0.33	0.44
TRINITY_DN24320_c0_g1	0.12	0.30	0.54	0.35	0.62	0.21	0.23

TRINITY_DN24320_c0_g2	0.54	0.35	0.32	0.27	0.38	0.24	0.46
TRINITY_DN25551_c0_g1	0.26	0.42	0.31	0.40	0.60	0.18	0.32
TRINITY_DN27540_c0_g1	0.19	0.49	0.44	0.16	0.53	0.19	0.42
TRINITY_DN29480_c0_g1	0.32	0.31	0.41	0.42	0.45	0.34	0.37
TRINITY_DN31794_c0_g1	0.22	0.35	0.44	0.24	0.59	0.38	0.30
TRINITY_DN34323_c0_g3	0.35	0.49	0.35	0.44	0.46	0.22	0.24
TRINITY_DN35207_c1_g1	0.11	0.30	0.32	0.39	0.67	0.29	0.32
TRINITY_DN35503_c1_g2	0.21	0.43	0.48	0.32	0.55	0.26	0.26
TRINITY_DN35767_c0_g1	0.39	0.43	0.34	0.35	0.41	0.27	0.44
TRINITY_DN35909_c1_g2	0.38	0.41	0.35	0.34	0.42	0.27	0.44
TRINITY_DN35916_c0_g1	0.24	0.47	0.36	0.31	0.45	0.32	0.42
TRINITY_DN35917_c1_g1	0.34	0.44	0.37	0.32	0.39	0.36	0.41
TRINITY_DN35920_c1_g1	0.26	0.47	0.39	0.30	0.45	0.31	0.41
TRINITY_DN37743_c0_g1	0.00	0.77	0.00	0.00	0.63	0.00	0.10
TRINITY_DN38348_c0_g1	0.19	0.30	0.18	0.24	0.86	0.10	0.18
TRINITY_DN41718_c0_g1	0.14	0.56	0.40	0.22	0.50	0.26	0.38
TRINITY_DN44299_c0_g1	0.15	0.49	0.26	0.24	0.72	0.23	0.21
TRINITY_DN50574_c0_g1	0.09	0.08	0.24	0.20	0.55	0.07	0.76
TRINITY_DN50745_c0_g1	0.27	0.31	0.24	0.09	0.40	0.39	0.67
TRINITY_DN5271_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN55745_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN6257_c0_g2	0.19	0.56	0.36	0.16	0.49	0.31	0.41
TRINITY_DN64334_c0_g1	0.00	0.00	0.00	0.04	1.00	0.00	0.00
TRINITY_DN64954_c0_g1	0.50	0.26	0.29	0.24	0.57	0.25	0.38
TRINITY_DN70651_c0_g1	0.00	0.14	0.12	0.13	0.56	0.16	0.78
TRINITY_DN72537_c0_g1	0.34	0.35	0.34	0.32	0.50	0.35	0.42
TRINITY_DN74835_c0_g1	0.00	0.73	0.16	0.16	0.64	0.00	0.00
TRINITY_DN78169_c0_g1	0.40	0.36	0.48	0.32	0.56	0.16	0.20
TRINITY_DN82862_c0_g1	0.12	0.47	0.23	0.19	0.80	0.06	0.18
TRINITY_DN84816_c0_g1	0.05	0.21	0.00	0.19	0.94	0.15	0.09
TRINITY_DN85037_c0_g1	0.33	0.56	0.49	0.32	0.41	0.12	0.20
TRINITY_DN87291_c0_g1	0.00	0.07	0.00	0.00	1.00	0.00	0.00
TRINITY_DN87569_c0_g1	0.78	0.12	0.31	0.18	0.39	0.15	0.28
OG0000038							
TRINITY_DN11382_c0_g1	0.32	0.18	0.36	0.30	0.73	0.25	0.21
TRINITY_DN11382_c0_g2	0.00	0.00	0.00	0.19	0.98	0.00	0.00
TRINITY_DN14629_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN16579_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN17682_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN19913_c0_g1	0.04	0.05	0.04	0.20	0.96	0.16	0.09
TRINITY_DN20716_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TRINITY_DN21380_c0_g1	0.00	0.04	0.00	0.00	1.00	0.00	0.00
TRINITY_DN22879_c0_g1	0.34	0.38	0.40	0.38	0.37	0.40	0.38
TRINITY_DN23914_c0_g1	0.39	0.36	0.38	0.38	0.41	0.36	0.36
TRINITY_DN25769_c1_g1	0.05	0.09	0.09	0.05	0.99	0.00	0.00
TRINITY_DN26294_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN26813_c1_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN32299_c2_g1	0.35	0.31	0.36	0.40	0.39	0.43	0.39
TRINITY_DN32299_c4_g1	0.41	0.36	0.39	0.36	0.40	0.36	0.37
TRINITY_DN32299_c4_g3	0.38	0.34	0.36	0.39	0.37	0.42	0.39
TRINITY_DN32299_c4_g4	0.41	0.38	0.39	0.38	0.40	0.33	0.34
TRINITY_DN33606_c3_g3	0.38	0.39	0.38	0.38	0.40	0.37	0.34
TRINITY_DN34075_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g2	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35187_c4_g3	0.00	0.00	0.00	0.00	1.00	0.01	0.03
TRINITY_DN35187_c5_g1	0.40	0.39	0.40	0.38	0.36	0.37	0.35
TRINITY_DN45547_c0_g1	0.35	0.33	0.37	0.38	0.38	0.42	0.40
TRINITY_DN7570_c0_g1	0.16	0.21	0.45	0.14	0.74	0.00	0.39
OG0000057							
TRINITY_DN11979_c0_g1	0.41	0.45	0.40	0.38	0.44	0.26	0.24
TRINITY_DN13737_c0_g1	0.32	0.40	0.39	0.32	0.49	0.41	0.29
TRINITY_DN18163_c0_g1	0.27	0.62	0.46	0.29	0.30	0.26	0.31
TRINITY_DN18831_c0_g1	0.25	0.40	0.38	0.40	0.51	0.32	0.34
TRINITY_DN28373_c0_g1	0.35	0.33	0.41	0.42	0.42	0.37	0.33
TRINITY_DN2987_c0_g1	0.29	0.45	0.35	0.44	0.43	0.36	0.28
TRINITY_DN37517_c0_g1	0.33	0.61	0.48	0.30	0.33	0.19	0.24
TRINITY_DN37878_c0_g1	0.33	0.40	0.36	0.35	0.46	0.36	0.39
TRINITY_DN38626_c0_g1	0.31	0.42	0.60	0.13	0.48	0.29	0.21
TRINITY_DN38746_c0_g1	0.39	0.41	0.35	0.45	0.39	0.36	0.26
TRINITY_DN39223_c0_g1	0.32	0.34	0.44	0.31	0.40	0.42	0.40
TRINITY_DN45326_c0_g1	0.38	0.67	0.39	0.18	0.26	0.27	0.29
TRINITY_DN46243_c0_g1	0.32	0.35	0.47	0.29	0.40	0.39	0.40
TRINITY_DN4843_c0_g1	0.29	0.47	0.39	0.33	0.42	0.37	0.34
TRINITY_DN48710_c0_g1	0.35	0.44	0.37	0.35	0.38	0.42	0.32
TRINITY_DN49355_c0_g1	0.38	0.53	0.41	0.33	0.37	0.28	0.30
TRINITY_DN5206_c0_g1	0.41	0.49	0.35	0.33	0.45	0.30	0.27
TRINITY_DN53969_c0_g1	0.36	0.28	0.46	0.34	0.38	0.44	0.36
TRINITY_DN5501_c0_g1	0.30	0.39	0.42	0.43	0.44	0.34	0.30
TRINITY_DN57837_c0_g1	0.36	0.43	0.44	0.33	0.41	0.35	0.28
TRINITY_DN57986_c0_g1	0.21	0.42	0.44	0.43	0.44	0.34	0.31
TRINITY_DN60009_c0_g1	0.12	0.55	0.43	0.31	0.49	0.36	0.17

TRINITY_DN62878_c0_g1	0.38	0.43	0.35	0.34	0.36	0.43	0.34
TRINITY_DN65689_c0_g1	0.37	0.48	0.52	0.30	0.38	0.27	0.23
TRINITY_DN66840_c0_g1	0.29	0.41	0.38	0.38	0.39	0.44	0.34
TRINITY_DN67038_c0_g1	0.19	0.46	0.41	0.37	0.55	0.28	0.24
TRINITY_DN69053_c0_g1	0.31	0.22	0.41	0.31	0.44	0.45	0.44
TRINITY_DN6912_c0_g1	0.36	0.60	0.40	0.32	0.34	0.28	0.24
TRINITY_DN7211_c0_g1	0.40	0.43	0.41	0.36	0.45	0.27	0.28
TRINITY_DN81045_c0_g1	0.15	0.55	0.47	0.32	0.46	0.29	0.24
TRINITY_DN82379_c0_g1	0.31	0.37	0.37	0.40	0.47	0.38	0.32
TRINITY_DN82692_c0_g1	0.34	0.41	0.37	0.33	0.39	0.42	0.36
TRINITY_DN83019_c0_g1	0.15	0.54	0.51	0.25	0.43	0.40	0.13
TRINITY_DN88570_c0_g1	0.42	0.27	0.38	0.46	0.45	0.29	0.32
TRINITY_DN89236_c0_g1	0.26	0.50	0.44	0.29	0.42	0.36	0.32
OG0000077							
TRINITY_DN24002_c0_g1	0.03	0.02	0.16	0.08	0.98	0.06	0.05
TRINITY_DN25795_c0_g1	0.26	0.44	0.32	0.28	0.65	0.26	0.26
TRINITY_DN28676_c0_g1	0.20	0.46	0.33	0.23	0.43	0.31	0.55
TRINITY_DN34206_c0_g1	0.29	0.43	0.44	0.35	0.52	0.25	0.29
TRINITY_DN34206_c0_g3	0.34	0.44	0.40	0.30	0.47	0.20	0.43
TRINITY_DN35909_c1_g6	0.27	0.35	0.26	0.32	0.40	0.35	0.59
TRINITY_DN4622_c0_g1	0.00	0.44	0.18	0.30	0.74	0.22	0.31
TRINITY_DN46286_c0_g1	0.12	0.38	0.28	0.24	0.53	0.35	0.54
TRINITY_DN50965_c0_g1	0.24	0.49	0.39	0.18	0.71	0.05	0.11
TRINITY_DN56262_c0_g1	0.44	0.36	0.26	0.39	0.65	0.18	0.09
OG0000193							
TRINITY_DN18647_c0_g1	0.34	0.23	0.31	0.41	0.40	0.50	0.39
TRINITY_DN2916_c0_g1	0.41	0.25	0.33	0.29	0.44	0.50	0.38
TRINITY_DN34833_c0_g1	0.43	0.23	0.38	0.42	0.45	0.37	0.34
TRINITY_DN47894_c0_g1	0.27	0.21	0.33	0.51	0.59	0.32	0.26
TRINITY_DN48994_c0_g1	0.42	0.33	0.16	0.44	0.43	0.50	0.24
TRINITY_DN50021_c0_g1	0.35	0.34	0.30	0.39	0.53	0.44	0.20
TRINITY_DN51080_c0_g1	0.52	0.10	0.09	0.53	0.56	0.27	0.19
TRINITY_DN663_c0_g1	0.20	0.19	0.40	0.42	0.48	0.41	0.43
TRINITY_DN68321_c0_g1	0.34	0.05	0.33	0.43	0.37	0.47	0.48
TRINITY_DN85527_c0_g1	0.30	0.08	0.29	0.41	0.42	0.53	0.43
TRINITY_DN9682_c0_g1	0.30	0.17	0.37	0.44	0.51	0.41	0.35
OG0000487							
TRINITY_DN19412_c0_g1	0.41	0.41	0.46	0.39	0.47	0.14	0.24
TRINITY_DN19657_c0_g1	0.39	0.35	0.42	0.42	0.40	0.31	0.35
TRINITY_DN22141_c0_g1	0.50	0.44	0.39	0.43	0.40	0.16	0.18
TRINITY_DN25685_c0_g1	0.42	0.39	0.44	0.41	0.43	0.21	0.29

TRINITY_DN32040_c0_g1	0.13	0.32	0.45	0.52	0.48	0.24	0.34
TRINITY_DN32040_c1_g1	0.48	0.43	0.45	0.42	0.41	0.11	0.17
TRINITY_DN32040_c1_g2	0.21	0.31	0.45	0.43	0.41	0.40	0.39
TRINITY_DN3945_c1_g1	0.43	0.31	0.33	0.56	0.46	0.18	0.22
TRINITY_DN79776_c0_g1	0.41	0.40	0.42	0.43	0.37	0.28	0.31
OG0000659							
TRINITY_DN16790_c0_g1	0.00	0.00	0.00	0.03	0.99	0.00	0.16
TRINITY_DN17015_c0_g1	0.00	0.58	0.44	0.41	0.44	0.18	0.29
TRINITY_DN18319_c0_g1	0.66	0.33	0.53	0.15	0.29	0.16	0.19
TRINITY_DN25697_c0_g1	0.00	0.00	0.00	0.01	0.99	0.00	0.15
TRINITY_DN58518_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN67071_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN81769_c0_g1	0.39	0.39	0.40	0.40	0.43	0.29	0.32
OG0000697							
TRINITY_DN124_c0_g1	0.42	0.41	0.42	0.42	0.44	0.19	0.28
TRINITY_DN15505_c0_g1	0.50	0.43	0.41	0.29	0.47	0.14	0.28
TRINITY_DN26500_c0_g1	0.60	0.00	0.06	0.00	0.00	0.54	0.59
TRINITY_DN26500_c0_g2	0.10	0.00	0.00	0.00	0.00	0.70	0.71
TRINITY_DN26500_c0_g3	0.42	0.37	0.38	0.34	0.36	0.34	0.42
TRINITY_DN6019_c0_g1	0.42	0.33	0.40	0.41	0.41	0.32	0.34
TRINITY_DN72431_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000748							
TRINITY_DN11486_c0_g1	0.25	0.58	0.46	0.35	0.37	0.27	0.22
TRINITY_DN20811_c0_g1	0.36	0.44	0.34	0.38	0.44	0.38	0.28
TRINITY_DN3182_c0_g1	0.30	0.41	0.44	0.33	0.44	0.38	0.33
TRINITY_DN33968_c0_g1	0.23	0.54	0.44	0.32	0.48	0.26	0.24
TRINITY_DN34609_c0_g2	0.26	0.40	0.38	0.36	0.43	0.39	0.38
TRINITY_DN36360_c0_g1	0.33	0.28	0.45	0.41	0.49	0.31	0.34
TRINITY_DN36612_c0_g1	0.38	0.42	0.42	0.37	0.45	0.32	0.25
TRINITY_DN43278_c0_g1	0.17	0.68	0.36	0.39	0.46	0.00	0.12
TRINITY_DN52161_c0_g1	0.30	0.43	0.42	0.37	0.47	0.33	0.29
TRINITY_DN53905_c0_g1	0.26	0.22	0.39	0.48	0.40	0.41	0.41
TRINITY_DN55797_c0_g1	0.24	0.65	0.39	0.31	0.37	0.25	0.28
TRINITY_DN57132_c0_g1	0.33	0.33	0.44	0.42	0.50	0.26	0.33
TRINITY_DN61213_c0_g1	0.26	0.62	0.31	0.33	0.52	0.20	0.19
TRINITY_DN62958_c0_g1	0.36	0.49	0.42	0.37	0.44	0.22	0.28
TRINITY_DN63307_c0_g1	0.24	0.70	0.38	0.33	0.38	0.14	0.17
TRINITY_DN64587_c0_g1	0.31	0.51	0.40	0.38	0.50	0.22	0.19
TRINITY_DN65856_c0_g1	0.34	0.59	0.42	0.37	0.45	0.12	0.06
TRINITY_DN72299_c0_g1	0.22	0.70	0.48	0.32	0.32	0.14	0.10
TRINITY_DN74635_c0_g1	0.35	0.50	0.46	0.37	0.41	0.20	0.26

OG0001176							
TRINITY_DN12752_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN15658_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN16748_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN19613_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN21803_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN21989_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN22432_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN23091_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN25199_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN27832_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN28824_c0_g1	0.01	0.18	0.14	0.10	0.97	0.00	0.00
TRINITY_DN51820_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN54299_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN55033_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN57180_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN57390_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN62765_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN65272_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN68960_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0001613							
TRINITY_DN11200_c0_g1	0.29	0.36	0.38	0.40	0.38	0.41	0.40
TRINITY_DN25915_c0_g1	0.36	0.28	0.42	0.40	0.36	0.43	0.37
TRINITY_DN34203_c0_g1	0.29	0.38	0.38	0.37	0.42	0.41	0.38
TRINITY_DN34203_c1_g3	0.37	0.30	0.40	0.36	0.44	0.39	0.37
TRINITY_DN50577_c0_g1	0.04	0.48	0.36	0.19	0.62	0.35	0.30
TRINITY_DN77511_c0_g1	0.08	0.55	0.35	0.22	0.61	0.33	0.21
OG0001853							
TRINITY_DN12421_c0_g1	0.00	0.00	0.15	0.44	0.44	0.55	0.55
TRINITY_DN23105_c0_g1	0.00	0.00	0.00	0.00	1.00	0.04	0.06
TRINITY_DN23321_c0_g1	0.09	0.04	0.08	0.41	0.47	0.61	0.48
TRINITY_DN32709_c0_g1	0.02	0.07	0.24	0.43	0.44	0.61	0.44
TRINITY_DN5214_c0_g1	0.04	0.02	0.14	0.44	0.40	0.59	0.53
TRINITY_DN87120_c0_g1	0.19	0.00	0.17	0.42	0.49	0.46	0.56
OG0002240							
TRINITY_DN21553_c0_g2	0.31	0.44	0.43	0.42	0.39	0.30	0.32
TRINITY_DN2944_c0_g1	0.27	0.61	0.44	0.36	0.32	0.24	0.27
TRINITY_DN3166_c0_g1	0.31	0.37	0.40	0.38	0.36	0.44	0.38
TRINITY_DN51101_c0_g1	0.42	0.46	0.35	0.35	0.50	0.19	0.27
TRINITY_DN54175_c0_g1	0.24	0.56	0.35	0.26	0.51	0.41	0.13
TRINITY_DN64072_c0_g1	0.35	0.39	0.44	0.23	0.40	0.40	0.40

TRINITY_DN67422_c0_g1	0.36	0.39	0.44	0.33	0.52	0.22	0.31
TRINITY_DN67547_c0_g1	0.27	0.46	0.48	0.36	0.40	0.32	0.29
TRINITY_DN68133_c0_g1	0.36	0.55	0.42	0.31	0.35	0.31	0.25
TRINITY_DN83728_c0_g1	0.54	0.58	0.44	0.16	0.30	0.16	0.20
OG0002627							
TRINITY_DN11021_c0_g1	0.25	0.48	0.39	0.43	0.49	0.28	0.21
TRINITY_DN14731_c0_g1	0.39	0.38	0.39	0.38	0.46	0.34	0.29
TRINITY_DN39337_c0_g1	0.26	0.49	0.36	0.48	0.51	0.17	0.21
TRINITY_DN58472_c0_g1	0.31	0.37	0.46	0.41	0.52	0.27	0.21
TRINITY_DN58913_c0_g1	0.21	0.46	0.37	0.41	0.43	0.38	0.33
TRINITY_DN72570_c0_g1	0.35	0.42	0.44	0.48	0.44	0.15	0.24
TRINITY_DN73485_c0_g1	0.39	0.40	0.43	0.38	0.51	0.24	0.22
TRINITY_DN87176_c0_g1	0.45	0.47	0.28	0.29	0.49	0.26	0.32
OG0002773							
TRINITY_DN16942_c0_g1	0.00	0.00	0.19	0.11	0.30	0.60	0.71
TRINITY_DN21585_c0_g1	0.52	0.20	0.40	0.37	0.46	0.33	0.26
TRINITY_DN23766_c0_g1	0.21	0.19	0.43	0.51	0.59	0.20	0.30
TRINITY_DN25296_c0_g1	0.45	0.22	0.38	0.40	0.48	0.36	0.31
TRINITY_DN26562_c0_g1	0.09	0.31	0.31	0.45	0.37	0.47	0.49
TRINITY_DN33266_c0_g1	0.51	0.24	0.32	0.42	0.52	0.28	0.24
TRINITY_DN35700_c0_g1	0.00	0.27	0.38	0.48	0.34	0.52	0.41
OG0003836							
TRINITY_DN40963_c0_g1	0.00	0.50	0.46	0.39	0.48	0.23	0.31
TRINITY_DN58644_c0_g1	0.28	0.61	0.33	0.49	0.35	0.22	0.16
TRINITY_DN74389_c0_g1	0.27	0.39	0.30	0.28	0.71	0.18	0.27
TRINITY_DN83117_c0_g1	0.00	0.60	0.21	0.29	0.67	0.21	0.17
OG0004289							
TRINITY_DN13368_c0_g1	0.52	0.07	0.49	0.57	0.40	0.02	0.00
TRINITY_DN17142_c1_g1	0.93	0.09	0.14	0.25	0.22	0.04	0.09
TRINITY_DN34407_c2_g5	0.41	0.37	0.35	0.35	0.38	0.39	0.39
TRINITY_DN42187_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN45384_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN69918_c0_g1	0.77	0.00	0.00	0.00	0.45	0.00	0.45
TRINITY_DN86414_c0_g1	0.48	0.32	0.34	0.40	0.42	0.30	0.36
OG0005442							
TRINITY_DN24928_c0_g1	0.37	0.46	0.22	0.26	0.53	0.24	0.43
TRINITY_DN33063_c0_g1	0.37	0.46	0.44	0.26	0.45	0.27	0.33
TRINITY_DN3646_c0_g1	0.45	0.45	0.33	0.19	0.42	0.21	0.48
TRINITY_DN49720_c0_g1	0.39	0.64	0.33	0.21	0.50	0.12	0.14
TRINITY_DN58898_c0_g1	0.45	0.38	0.51	0.31	0.50	0.14	0.16

TRINITY_DN72415_c0_g1	0.35	0.37	0.37	0.47	0.58	0.13	0.19
OG0005452							
TRINITY_DN14350_c0_g1	0.00	0.00	0.00	0.05	1.00	0.00	0.00
TRINITY_DN18058_c0_g1	0.41	0.26	0.38	0.43	0.41	0.39	0.34
TRINITY_DN18058_c0_g2	0.43	0.33	0.38	0.41	0.39	0.36	0.33
TRINITY_DN30150_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN33407_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN35218_c1_g4	0.40	0.33	0.38	0.40	0.40	0.38	0.35
TRINITY_DN73525_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0010763							
TRINITY_DN14129_c0_g1	0.44	0.38	0.35	0.35	0.39	0.37	0.36
TRINITY_DN17142_c0_g1	0.41	0.38	0.35	0.37	0.39	0.37	0.37
TRINITY_DN31034_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN37866_c0_g1	0.52	0.15	0.51	0.54	0.39	0.01	0.02
OG0010767							
TRINITY_DN16764_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN17000_c1_g1	0.36	0.48	0.43	0.34	0.45	0.22	0.29
TRINITY_DN9921_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN9921_c1_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
OG0012394							
TRINITY_DN23126_c0_g1	0.41	0.23	0.27	0.40	0.48	0.43	0.37
TRINITY_DN73669_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN7602_c0_g1	0.38	0.37	0.37	0.40	0.37	0.39	0.37
TRINITY_DN7786_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0012409							
TRINITY_DN20672_c0_g1	0.39	0.29	0.39	0.31	0.56	0.34	0.29
TRINITY_DN25641_c0_g1	0.00	0.00	0.00	0.00	1.00	0.00	0.00
TRINITY_DN25641_c0_g2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN70288_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Additional file 3. Individual transcripts in expanded families of macauba palm with respective PANNZER and CDD protein description, protein similarity (%) and TPM value by tissue. TPM value ≥ 20 values in blue-colored box.

OG0000015	CDD Description	Leaf			Leaf Sheath			Bulb			Root			Female Flower			Male Flower			Ripe fruits		
		S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	1 ♀	2 ♀	3 ♀	1 ♂	2 ♂	3 ♂	S 7	S 5	S 8
TRINITY_DN16760_c0_g1	RNase_HI_RT_Ty3.	1.37	0.57	0.57	1.19	1.32	0	0.86	0.13	0.41	3.37	0.73	0.14	1.33	0.19	0.67	0.12	0.85	1.97	0.1	0.96	0.34
TRINITY_DN20595_c0_g1	RT_LTR: Reverse transcriptases	12.55	1.94	5.2	9.98	2.2	0.23	2.89	1.21	5.41	6.1	6.01	8.17	4.51	2.84	1.25	0.34	5.34	8.39	1.3	5.64	3.19
TRINITY_DN21123_c0_g1	RT_LTR: Reverse transcriptases	4.46	0.5	0.13	2.04	0.94	0	0.59	0.28	0.57	4	0.95	0.48	1.11	0.31	0.33	0	0.67	0.85	0.32	0.4	0
TRINITY_DN21123_c0_g2	RT_LTR: Reverse transcriptases	3.44	0.82	0.83	2.92	1.17	0	1.33	0.78	1.49	3.38	2.03	0.45	0.94	0.49	0.56	0	0.88	0.91	2.5	0.25	1.43
TRINITY_DN23177_c0_g1	--	4.99	0.18	1.84	2.95	1.36	0.32	1.2	1.21	1.81	5.04	1.38	0.92	1.53	0.76	0.7	0.19	1.16	3.87	4.18	0.77	3.69
TRINITY_DN26488_c0_g1	--	2.55	2.09	3.29	0.89	2.71	0.3	1.76	1.81	2.95	3.94	4.53	3.69	0.98	1.65	0.72	0.24	1.86	1.8	3.47	3.33	1.53
TRINITY_DN30512_c0_g1	Transcriptional regulator ICP4.	15.06	4.81	7.82	13.03	6	0.31	2.23	3.4	6.09	6.19	8.17	7.65	4.16	4.11	2.36	1	10.22	14.07	1.93	9.62	4.66
TRINITY_DN33294_c1_g1	Retropepsin_like.	**63.17**	1.26	0.77	**28.04**	1.99	0.4	5.01	0.53	1.21	8.21	1.77	0.99	1.91	1.16	1.22	0.46	1.73	3.19	2.6	0.77	2.17
TRINITY_DN33687_c0_g1	RT_LTR: Reverse transcriptases	8.42	9.91	7.02	4.98	6.81	0.48	2.29	4.02	6.33	5.81	10.91	8	1.62	1.56	1.01	3.09	4.39	7.86	4.67	8.22	1.62
TRINITY_DN33687_c0_g2	RNase_HI_RT_Ty3.	7.91	11.02	7.87	6.41	9.52	0.14	2.04	3.69	7.05	5.44	12.18	11.88	1.95	1.66	0.96	2.33	3.89	4.17	2.4	4.65	1.77
TRINITY_DN35010_c1_g1	RT_LTR: Reverse transcriptases	**22.26**	7.03	8.99	**23.04**	6.92	0.51	4.6	2.49	7.75	10.63	7.44	13.39	4.92	2.26	1.84	0.51	6.33	7.18	2.96	9.1	3.79
TRINITY_DN35097_c2_g1	RNase_HI_RT_Ty3.	9.19	8.51	4.66	5.53	4.97	1.52	2.03	3.55	4.68	4.62	4.87	3.27	6.26	2.12	2.36	1.56	4.45	5.17	1.38	2.93	0.94
TRINITY_DN35678_c2_g1	RT_LTR: Reverse transcriptases	14.16	10.53	14.42	9.37	12.7	1.25	4.06	5.51	12.08	7.66	15.72	17.07	18.37	8.77	7.55	2.54	13.62	**26.16**	10.06	18.28	6.11
TRINITY_DN35722_c0_g1	RT_LTR: Reverse transcriptases	5.99	7.31	3.02	**27.4**	**22.02**	4.67	**72.75**	**33.52**	17.37	**39.89**	**32.58**	**22.05**	12.46	6.69	**26.91**	0.96	16.57	**67.74**	**132**	**57.9**	19.8
TRINITY_DN35722_c0_g2	Retrotrans_gag.	2.72	10.56	5.1	15.19	**48.78**	3.47	**44.54**	**72.73**	**35.12**	**22.84**	**72.25**	**41.76**	**21.39**	10.56	**35.23**	3.33	16.35	**66.12**	**445**	**165.5**	**30.02**

TRINITY_DN35914_c0_g1	RT_LTR: Reverse transcriptases	**27.73**	**23.76**	**30.65**	**20.11**	**24.61**	0.64	9.03	14.52	**26.92**	**20.05**	**29.34**	**31.95**	18.44	8.22	6.24	2.28	**24.1**	**29.08**	12.98	**20.07**	9.29
TRINITY_DN35914_c0_g3	Retropepsin_like.	**30.61**	**30.97**	**34.54**	**25.13**	**30.43**	2.31	10.83	17.89	**36.76**	**27.09**	**37.45**	**42.81**	**29.82**	8.17	6.79	2.83	18	**26.96**	14.57	**24.95**	6.61
TRINITY_DN8915_c0_g1	CHROMO.	0.91	3.88	1.28	1.5	7.35	1.22	7.03	8.36	8.35	4.67	6.85	10.65	16.39	9.1	17.78	2.82	**31.49**	**89.68**	**31.58**	15.59	11.13
OG0000024																						
TRINITY_DN12992_c0_g1	gag_pre-integrs.	1.88	1.93	2.17	3.08	2.86	0.75	0.61	0.6	1.26	4.63	5.45	4.63	0.45	0.36	0.29	0	0.57	0.31	0.55	1.93	0.44
TRINITY_DN13894_c0_g1	RVT_2. Reverse transcriptase.	3.04	7.25	14.48	2.49	3.8	6.69	0.85	2.54	7.81	5.68	4.49	**22.62**	8.51	1.83	5.99	3.31	2.35	**26.3**	1.04	0.86	1.07
TRINITY_DN14263_c0_g1	RVT_2. Reverse transcriptase.	4.58	1.93	1.95	3.99	2.41	0	0.61	1.43	3.78	3.79	5.38	4.69	0.27	0.27	0.47	0.33	0	0.47	2.87	0.17	0.49
TRINITY_DN15704_c0_g1	RVT_2. Reverse transcriptase.	6.62	13.63	6.23	5.46	6.86	2.95	1.72	2.09	3.17	7.94	11.22	7.99	4.05	2.7	3.57	3.79	6.65	8.35	0.99	0.55	0
TRINITY_DN17326_c0_g1	RVT_2. Reverse transcriptase.	1.31	1.2	2.82	3.12	3.32	0.35	1.96	2.04	7.93	3.43	3.19	5.26	1.14	0.66	0.31	5.39	0.94	1.08	**29.97**	**30.37**	0.16
TRINITY_DN17383_c0_g1	rve. Integrase core domain.	0	0	0	0	0.47	1.39	2.58	0.34	0.82	**59.22**	8.09	14.75	1.14	0.19	0.06	0.1	0.2	0	0.04	0.2	0
TRINITY_DN19679_c0_g1	RVT_2. Reverse transcriptase.	9.99	14.66	10.71	9.88	9.74	3.99	5.79	7.08	4.14	6.74	9.68	6.81	4.2	1.71	4.1	3.01	2.56	2.7	4.22	3.12	2.72
TRINITY_DN20032_c0_g1	RVT_2. Reverse transcriptase.	1.11	1.39	1.45	1.17	1.65	0.59	0.4	0.75	1.83	5.23	7.74	7.75	1.45	0.18	1.05	3.51	1.17	1.67	0.79	0.76	0.34
TRINITY_DN20759_c0_g1	--	1.36	3.04	2.84	4.62	6.74	3.05	1.06	0.88	10.08	4.39	7.29	9.13	2.46	0.56	0.83	4.33	3.72	1.09	2.2	0.24	0.23
TRINITY_DN21019_c0_g1	--	2.25	3.91	5.84	0.53	2.95	3.74	0.53	2.07	2.87	4.98	4.23	4.49	3.45	1.24	3.8	1.55	0.99	3.06	0.63	0.79	0.76
TRINITY_DN21432_c0_g1	rve. Integrase core domain.	1.02	3.28	2.32	0.68	3.14	0.1	0.57	0.73	3.27	6.06	7.33	7.56	0.55	0.64	0.24	0.51	1.11	0.58	0.88	2.14	0.67
TRINITY_DN2271_c0_g1	RVT_2. Reverse transcriptase.	5.9	15.39	8.81	4.43	5.09	2.43	1.05	0.62	1.68	6.45	8.82	6.12	3.75	4.01	3.78	3.28	5.54	8.43	1.42	0	0
TRINITY_DN22952_c0_g1	RVT_2. Reverse transcriptase.	7.77	16.96	11.8	6.96	8.89	7.59	1.7	3	2.62	8.91	16.6	9.75	4.62	4.32	6.67	6.44	8.74	13.41	1.44	0.74	0.3

TRINITY_DN24320_c0_g1	--	1.95	1.08	0.78	2.65	3.21	0	1.31	0.34	5.99	3.48	2.9	4.69	1.04	0.52	0.64	0.8	0.33	0.75	0.39	0.33	0
TRINITY_DN24320_c0_g2	RVT_2. Reverse transcriptase.	5.26	4.22	4.16	3.63	4.65	0.92	1.3	2.58	4.58	5.3	7.02	5.08	3.82	2.1	1.12	10.21	7.88	5.4	4.94	**118.7**	12.41
TRINITY_DN25551_c0_g1	rve. Integrase core domain.	2.65	3.87	2.34	1.58	2.8	0.09	0.9	2.46	5.51	5.31	9.39	3.63	0.76	1.21	0.74	0.63	3.81	1.65	1.2	1.54	0.15
TRINITY_DN27540_c0_g1	--	5.53	9.72	5.37	4.38	5.8	1.92	0.87	0.61	2.41	6.49	8.72	6.05	0.77	1.07	1.92	1.72	4.01	5.85	1.25	1.07	0.19
TRINITY_DN29480_c0_g1	Transcriptional regulator ICP4	8.19	12.75	8.68	18.47	**22.78**	5.49	18.65	**24.88**	**20.27**	**20.63**	**26.86**	**24.49**	11.22	9.62	10.64	2.66	14.36	13.6	6.08	8.98	8.91
TRINITY_DN31794_c0_g1	RVT_2. Reverse transcriptase.	3.05	2.41	3.65	4.59	5.12	1.13	1.33	1.59	3.88	6.31	9.09	9.09	4.97	1.43	3.56	1.89	2.22	3.92	1.35	3.41	0.21
TRINITY_DN34323_c0_g3	RVT_2. Reverse transcriptase.	10.58	7.05	3.27	**21.37**	3.41	1.65	15.12	5.36	3.02	**32.92**	6.03	3.01	3.76	1.43	1.56	2.67	1.74	1.8	3.4	4.51	1.12
TRINITY_DN35207_c1_g1	RVT_2. Reverse transcriptase.	2.17	3.8	1.36	2.32	5.08	1.56	2.32	3.46	6.54	11.77	12	11.54	1.89	1.96	2.58	2.2	6.51	2.4	0.76	0.5	0.26
TRINITY_DN35503_c1_g2	RVT_2. Reverse transcriptase.	4.3	5.71	2.89	5.32	6.68	2.35	1.03	2.46	2.6	7.53	13.28	6.23	2.19	1.77	1.36	1.74	3.53	1.29	1.3	3.74	0.37
TRINITY_DN35767_c0_g1	RVT_2. Reverse transcriptase.	**30.83**	19.2	**31.18**	15.07	**20.6**	1.07	7.87	15.54	**21.09**	15.31	**30.99**	**25.87**	16.83	7.94	4.17	2.89	**34.22**	**43.99**	**22.22**	**24**	2.1
TRINITY_DN35909_c1_g2	RNase_HI_RT_Ty1.	**38.81**	**27.6**	**46.82**	**22.17**	**27.45**	4.27	11.95	19.94	**26.66**	19.71	**43.92**	**40**	**20.16**	9.66	7.03	5.67	**47.84**	**51.7**	**29.56**	**30.5**	4.63
TRINITY_DN35916_c0_g1	RVT_2. Reverse transcriptase.	19.01	**31.1**	18.05	9.24	19.38	4.39	3.45	6.44	6.55	14.05	**28.93**	17.09	7.21	4.73	6.91	6.24	13.72	19.54	3.75	**21.37**	3.35
TRINITY_DN35917_c1_g1	RVT_2. Reverse transcriptase.	**28.51**	**24.22**	**29.17**	16.2	**21.75**	2.32	8.26	10.36	**20.8**	18.18	**28.1**	19.25	**29.27**	14.52	8.62	1.73	**22.21**	**32**	11.97	**34.75**	9.12
TRINITY_DN35920_c1_g1	RVT_2. Reverse transcriptase.	**58.23**	**95.74**	**76.82**	**37.03**	**48.67**	**24.83**	10.54	15.69	**36.21**	**48.38**	**61.09**	**63.51**	17.22	15.82	**24.84**	**23.12**	**43.7**	**50.82**	8.66	14.22	9.69
TRINITY_DN37743_c0_g1	RVT_2. Reverse transcriptase.	1.17	4.62	0.12	0	0.3	0	0.34	0	0	2.92	0.9	0	0	0.2	0	0	0.51	0.11	0	0	0
TRINITY_DN38348_c0_g1	RVT_2. Reverse transcriptase.	1.23	1.13	0.44	0.58	1.15	0.08	0.16	1.3	0.82	4.53	11.18	7.42	0.58	0.11	0.27	0.36	0.55	0.96	0.11	0.6	0.62
TRINITY_DN41718_c0_g1	--	6.19	13.5	8.7	4.11	5.93	2.59	1.4	0.83	1.61	3.5	11.23	6.59	1.87	1.85	2.11	3.09	3.64	7.32	1.52	0.79	0.3

TRINITY_DN44299_c0_g1	gag_pre-integrs.	1.46	2.66	2.58	0.95	3.41	0.48	0.21	0.86	3.31	3.41	7.78	5.48	2.05	0.83	0.36	0.58	1.18	0.75	1.89	0.47	0.45
TRINITY_DN50574_c0_g1	RVT_2. Reverse transcriptase.	4.17	0	0.13	0.5	0.44	0	0.62	0	0.36	2.15	0	1.33	0.11	0	0.12	2.18	**29.54**	1.39	0.33	0.14	0
TRINITY_DN50745_c0_g1	rve. Integrase core domain.	0.9	1.64	2.49	1.1	1.91	0	0.31	0.18	1.04	2.7	2.31	2.5	4.28	2.32	0.49	0.17	7.01	8.57	0.42	1.91	1.33
TRINITY_DN5271_c0_g1	--	0	0	0	0	0	0	0	0	0	4.66	0.89	0	0	0	0	0	0	0	0	0	0
TRINITY_DN55745_c0_g1	RVT_2. Reverse transcriptase.	0	0	0	0	0	0	0	0	0	2.82	1.12	0.46	0	0	0	0	0	0	0	0	0
TRINITY_DN6257_c0_g2	rve. Integrase core domain.	6.57	8.67	4.65	2.65	5.2	0.53	0.66	0.77	2.51	4.9	10.31	4.62	2.19	2.02	1.72	1.08	3.33	3.86	0.96	1.38	0
TRINITY_DN64334_c0_g1	RVT_2. Reverse transcriptase.	0	0	0	0	0	0	0.08	0.09	0.08	2.89	5.44	8.77	0	0	0	0	0	0	0	0	0.17
TRINITY_DN64954_c0_g1	RVT_2. Reverse transcriptase.	1.27	0.52	1.06	1.3	2.36	0	0.99	0.7	3.04	3.23	4.08	4.39	0.62	1.05	1.05	1.97	2.57	1.83	1.61	17.87	3.08
TRINITY_DN70651_c0_g1	rve. Integrase core domain.	6.47	0.35	0	0.99	0.29	0	0.82	0	0.31	2.67	0.37	2.2	0.88	0	0.41	4.15	**40.71**	2.68	0	0.18	0
TRINITY_DN72537_c0_g1	rve. Integrase core domain.	1.17	3.64	7.18	2.46	6.06	3.41	0.92	3.06	6.43	2.64	7.97	15.55	7.11	2.96	3.78	3.89	5.41	5.84	3.55	0.56	4.69
TRINITY_DN74835_c0_g1	rve. Integrase core domain.	1.82	4.14	0.7	0.26	0.8	0	0.26	0	0.75	2.27	1.49	0	0	0	0	0	0.44	0	0	0	0
TRINITY_DN78169_c0_g1	--	1	2.08	1.18	1.87	2.38	0.8	0.37	1.01	1.77	3.61	2.39	0.83	0.88	0.43	0.23	0.14	0.56	1.51	0.44	1.38	3.96
TRINITY_DN82862_c0_g1	RVT_2. Reverse transcriptase.	0.61	3.45	2.23	0.79	2.64	0.24	0.52	0.61	3.37	3.12	9.54	6.3	0.46	0.11	0.16	0.57	1.77	0.4	0.35	1.75	0.14
TRINITY_DN84816_c0_g1	rve. Integrase core domain.	0.46	0.25	0.43	0	0.98	0	0.16	0.37	0.76	3.89	5.09	2.99	0.78	0.28	0.1	0	0.27	0.16	0.14	0	0.09
TRINITY_DN85037_c0_g1	rve. Integrase core domain.	6.56	6.26	7.36	4.8	4.79	0.35	1.55	2.45	2.2	3.07	4.94	3.44	1.47	0.56	0.45	0.76	1.08	1.31	2.31	3.82	1.6
TRINITY_DN87291_c0_g1	RNase_HI_RT_Ty1.	0.46	0.17	0	0	0.07	0	0	0	0.15	3.5	8.55	10.14	0	0	0	0	0	0	0	0	0
TRINITY_DN87569_c0_g1	rve. Integrase core domain.	0.59	0.18	1.64	2.24	3	0	0.85	1	8.15	2.96	3.89	3.43	0.76	0.6	0.96	2.25	1.92	1.04	18.54	**28.78**	0.37
OG0000038																						
TRINITY_DN11382_c0_g1	Ubiquitin	0.52	0.84	0.36	1.37	1.79	0.63	1.02	0.66	1.19	4.66	2.71	9.82	1.11	0.1	0.81	0.62	0.64	1.27	1.22	1.14	0.85
TRINITY_DN11382_c0_g2	Ubiquitin.	0	0	0	0	0	0	0.45	0	0.65	6.12	2.19	9	0	0	0	0	0	0	0	0.13	0

TRINITY_DN14629_c0_g1	Ubiquitin	0	0	0	0	0.21	0	0.36	0	0	5.81	0.27	0.65	0	0	0	0	0	0	0	0	0
TRINITY_DN16579_c0_g1	Ubiquitin	0	0	0	0	0	0	0	0	0	12.81	39.89	7.59	0	0	0	0	0	0	0	0	0
TRINITY_DN17682_c0_g1	Ubiquitin	0	0	0	0	0	0	0	0	0	4.04	2.28	0.18	0	0	0	0	0	0	0	0	0
TRINITY_DN19913_c0_g1	Ubiquitin	0.12	0.11	0.12	0.22	0	0.1	0.65	0.63	0.52	30.56	9.24	0.72	0.48	0.28	0.67	0.36	0.24	0	0.1	0.24	0
TRINITY_DN20716_c0_g1	Ubiquitin	0	0	0	0	0	0	0.13	0	0	2.34	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN21380_c0_g1	Ubiquitin	0.16	0	0.44	0	0	0.88	0	0	0.26	13.34	42.74	28.02	0	0	0	0	0	0.14	0	0	0.15
TRINITY_DN22879_c0_g1	Ubiquitin	71.69	89.48	90.28	112.9	83.69	133.8	74.63	92.05	126	87.15	81.67	107.9	148.8	104	125.4	122.3	89.14	78.96	55.24	169.5	41.34
TRINITY_DN23914_c0_g1	Ubiquitin	1927	1878	1849	2647	2668	3344	3477	2466	2367	4691	4600	5413	1785	1180	2098	2303	1948	1629	3688	3056	5012
TRINITY_DN25769_c1_g1	Ubiquitin.	0	0.37	0.37	0.35	0.61	0.32	0.17	0.81	0.17	14.44	40.07	29.66	0	0	0	0	0	0	0	0.19	0.19
TRINITY_DN26294_c0_g1	Ubiquitin.	0	0	0	0	0	0	0	0	0.12	30.86	76.15	51.6	0	0	0	0	0	0	0	0	0
TRINITY_DN26813_c1_g1	Ubiquitin	0	0	0	0	0	0	0	0	0	19.75	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN32299_c2_g1	Ubiquitin	20.94	29.68	31.69	38.5	49.72	47.69	75.6	85.61	65.73	81.95	72.83	45.77	117.8	66.31	108.9	89.07	65.52	55.14	50.46	43.88	46.85
TRINITY_DN32299_c4_g1	Ubiquitin	757.8	573.9	963.5	1746	1486	928.6	1055	594	799.6	1755	1503	1461	883.8	277	745.2	856	807.6	880.1	2554	1947	1136
TRINITY_DN32299_c4_g3	Ubiquitin	400.7	418.1	488.6	602.6	654.7	538.3	1123	1088	997.6	804.8	719.5	649.5	1875	1485	2620	1635	1034	913.5	1334	929.5	750.4
TRINITY_DN32299_c4_g4	Ubiquitin.	404.3	490.2	523.4	529.1	565.2	412.5	508.9	443.2	537.9	670.7	641.7	807.5	400.8	126	225.5	326.1	221.8	229.1	773.7	1093	654.5
TRINITY_DN33606_c3_g3	Ubiquitin	105.6	116	111.4	155.4	100.8	76.8	98.1	103	161.8	125.8	111.5	128	136.5	67.43	89.06	109.4	58.39	57.77	92.53	278.6	36.48
TRINITY_DN34075_c1_g1	Ubiquitin.	0	0	0	0	0	0	0	0	0	15.01	0.21	0.89	0	0	0	0	0	0	0	0	0
TRINITY_DN35187_c4_g1	Ubiquitin	0	0	0	0	0	0	0	0	0	80.65	88.46	31.19	0	0	0.02	0	0.07	0	0	0	0
TRINITY_DN35187_c4_g2	Ubiquitin	0	0	0	0	0	0	0.12	0	0	85.87	60.64	38.19	0	0.11	0	0.13	0	0	0	0	0
TRINITY_DN35187_c4_g3	Ubiquitin	0	0	0	0	0	0	0	0	0.17	131.9	527.3	136.9	0.08	0	0.03	0.19	0	0.18	0	0	0
TRINITY_DN35187_c5_g1	Ubiquitin	1291	1881	2425	1718	2081	2921	1030	1403	1473	1109	1043	1762	2509	715	1269	1222	819.2	948.9	2090	3100	1997
TRINITY_DN45547_c0_g1	Ubiquitin	589.6	606.8	728.3	1383	1442	723.5	2449	1487	1153	2196	1678	1447	4253	1886	3454	4874	2413	1408	695.6	1410	934.6
TRINITY_DN7570_c0_g1	Ubiquitin.	0.58	0.35	0.54	0.17	1.47	3.24	0.33	0.98	0	3.45	1.91	9.19	0	0	0.52	1.1	1.32	1.2	0.45	0.37	0.36
OG0000057																						
TRINITY_DN11979_c0_g1	DYW_deaminase.	3.04	6.14	3.63	2.93	3.43	2.16	2.62	3.17	1.9	2.77	3.64	3.51	1.26	1.43	1.77	1.56	1.12	1.23	2.99	3.09	2.28
TRINITY_DN13737_c0_g1	DYW_deaminase.	0.84	2.44	4.51	2.23	2.28	3.81	1.44	2.2	1.65	3.44	4.44	2.96	2.86	2.37	2.52	1.32	1.41	1.47	1.77	1.67	1.08
TRINITY_DN18163_c0_g1	DYW_deaminase.	5.47	11.2	14.41	5.41	3.58	5.72	2.68	2.17	1.68	3.31	2.43	2.23	1.81	0.65	1.85	3.15	2.4	2.46	2	2.07	1.29
TRINITY_DN18831_c0_g1	DYW_deaminase.	4.09	4.66	3.95	4.99	3.72	1.95	4.05	4.15	4.38	7.27	6.04	8.04	3.3	1.09	2.75	1.16	6.17	3.06	1.75	3.55	1.32
TRINITY_DN28373_c0_g1	DYW_deaminase.	5.24	6.14	4.77	8.49	7.45	9.49	8.92	8.85	6.76	10.64	7.65	8.84	6.79	4.84	6.42	6.17	4.45	5.15	5.98	3.08	7.51

TRINITY_DN2987_c0_g1	DYW_deaminase.	1.62	4.32	3.18	2.07	2.21	1.29	2.94	3.08	4.09	3.62	2.75	2.89	2.24	1.9	2.11	1.3	1.86	1.39	1.52	1.25	1.89
TRINITY_DN37517_c0_g1	DYW_deaminase.	14.98	**32.52**	**32.43**	14.8	12.29	**21.28**	5.64	4.52	4.12	5.62	5.51	5.3	4.3	1.79	1.98	3.07	2.3	5.17	7.06	2.8	5.61
TRINITY_DN37878_c0_g1	DYW_deaminase.	0.93	2.89	3.65	1.04	2.41	3.79	2.96	1.61	2.29	3.83	1.04	4.59	3.15	2.41	2.07	2.76	2.7	2.91	2.05	2.3	1.1
TRINITY_DN38626_c0_g1	DYW_deaminase.	1.64	1.16	1.34	1.74	2.34	3.76	0.31	0.92	0.3	2.89	1.61	0.84	1.39	0.14	0.78	0.52	0.88	0.48	0.85	2.28	0
TRINITY_DN38746_c0_g1	DYW_deaminase.	0.57	1.29	1.62	1.04	0.95	1.54	1.52	0.98	1.45	2.34	0.84	1.18	1.61	0.65	1.06	0.7	0.77	0.4	0.92	2.35	1.21
TRINITY_DN39223_c0_g1	DYW_deaminase.	2.52	1.77	1.52	2.79	2.72	2.55	1.5	1.77	0.72	2.18	2.29	2.69	3.95	1.25	2.54	2.39	2.36	2.22	2.18	1.59	1.17
TRINITY_DN45326_c0_g1	DYW_deaminase.	**27.49**	16.02	17.57	5.63	4.56	4.19	0.78	1.22	1.73	2.11	1.18	4.52	2.77	2.28	1.9	3.56	2.63	1.72	4.31	4.93	1.67
TRINITY_DN46243_c0_g1	DYW_deaminase.	1.29	1.98	3.26	2.1	3.29	3.65	1.45	2.36	1.33	3.46	1.66	2.51	2.41	0.85	2.45	2.37	2.5	4.32	1.68	1.86	1.04
TRINITY_DN4843_c0_g1	DYW_deaminase.	2.18	5.07	5.53	3.45	3.14	7.05	2.3	2.84	2.61	4.33	4.16	3.34	3.31	2.33	3.24	2.79	2.25	2.78	2.04	4.04	1.32
TRINITY_DN48710_c0_g1	STKc_CDK9_like.	1.73	4.69	4.61	3.27	3.18	6.74	1.78	5.27	2.98	2.75	4.79	3.34	4.18	2.89	5.46	2.1	2.48	3.44	2.87	5.11	2.8
TRINITY_DN49355_c0_g1	DYW_deaminase.	4.57	5.4	5.46	2.71	3.25	3.41	2.18	2.86	2.17	2.64	2.2	4.27	2.24	1.7	1.52	2.49	1.8	1.89	2.79	2.91	1.26
TRINITY_DN5206_c0_g1	DYW_deaminase.	3.03	8.43	6.05	3.02	4.44	2.93	1.87	6.67	2.68	4.8	4.98	7.27	2.33	2.14	2.32	1.51	3.37	1.91	3.7	4.88	4.01
TRINITY_DN53969_c0_g1	DYW_deaminase.	1.05	2.44	1.59	2.08	3.77	5.03	1.89	2.13	2.76	2.59	2.35	2.76	5.51	1.81	3.45	2.72	2.33	2.44	2.45	2.86	1.59
TRINITY_DN5501_c0_g1	DYW_deaminase.	3.25	3.47	4.79	4.93	3.94	3.18	6.07	4.05	3.22	5.3	4.39	3.43	2.64	1.55	2.58	3.17	2.08	1.47	3.31	2.15	1.89
TRINITY_DN57837_c0_g1	DYW_deaminase.	1.23	3.55	5.58	3.4	3.74	5.72	2.92	1.66	2.2	3.46	1.62	3.2	2.28	2.41	2.57	0.88	1.7	1.9	3.9	1.19	2.56
TRINITY_DN57986_c0_g1	DYW_deaminase.	1.88	3.03	2.87	2.96	3.08	3.4	2.81	3.08	2.99	3.05	2.21	3.85	2.72	1.98	1.77	1.72	2.68	1.28	0.73	1.47	0.94
TRINITY_DN60009_c0_g1	DYW_deaminase.	2.37	4.05	5.18	2.06	2.58	3.54	1.53	1.2	2.44	3.32	2.92	3.42	3.18	0.67	1.95	1.69	0.29	0.65	2.18	0.43	0
TRINITY_DN62878_c0_g1	DYW_deaminase.	2.87	4.58	4.33	2.84	2.23	4.63	2.64	2.38	2.69	3.05	2.89	4.53	5.52	3.62	4.28	2.54	2.73	2.76	3.29	6.17	1.69
TRINITY_DN65689_c0_g1	DYW_deaminase.	4.46	5.93	5.7	6.85	3.76	8.93	3.09	2.26	1.79	4.05	3.42	1.79	1.9	0.47	3.91	2.06	0.42	1.47	3.22	3.17	4.47
TRINITY_DN66840_c0_g1	DYW_deaminase.	3.86	7.03	7.84	4.29	5.86	9.86	3.9	6.34	5.82	7.1	6.03	6	8.17	3.74	9.04	4.29	6.09	4.46	3.3	4.88	2.95
TRINITY_DN67038_c0_g1	DYW_deaminase.	0.98	2.49	1.71	1.43	2.49	1.04	0.28	1.22	1.71	3.27	2.26	2.02	0.84	0.58	1.12	0.1	1.28	0.67	0.51	4.21	0.5
TRINITY_DN69053_c0_g1	DYW_deaminase.	1.02	0.62	1.04	3.66	1.46	2.62	1.66	1.96	1.22	4.25	1.9	3.02	4.18	2.24	3.09	3.12	1.55	3	1.67	1.64	0.1
TRINITY_DN6912_c0_g1	DYW_deaminase.	11.9	**27.27**	**27.3**	7.79	9.14	8.27	3.98	7.09	4.94	4.59	5.61	6.17	3.93	3.72	5.93	2.94	4	2.16	6.66	4.31	7.27
TRINITY_DN7211_c0_g1	DYW_deaminase.	3.19	5.13	5.04	4.59	4.57	6.64	2.88	4.31	3.54	3.94	5.7	5.91	3.35	1.11	2.15	4.26	2.18	1.56	3.94	4.4	4.29
TRINITY_DN81045_c0_g1	DYW_deaminase.	3.96	4.65	3	2.91	3	2.59	1.97	1.52	1.02	3.21	2.77	2.85	1.35	0.33	1.64	0.92	1.71	1.01	0.48	1.78	0.57
TRINITY_DN82379_c0_g1	DYW_deaminase.	0.8	3.83	3.13	3.09	3.24	4.61	3.24	4.04	3.7	2.78	5.27	6.03	3.33	2.97	4.51	3.71	2.46	2.3	3.09	2.37	1.94
TRINITY_DN82692_c0_g1	DYW_deaminase.	0.97	2.95	3.34	2.53	2.21	2.58	1.78	2.43	2.08	2.48	2.8	3.44	3.92	2.21	3.21	2.09	3.46	2.39	2.21	2.37	1.79
TRINITY_DN83019_c0_g1	DYW_deaminase.	0.17	2.34	4.1	2.09	2.73	1.63	0.74	0.87	0.71	3.04	0.84	1.58	1.44	0.78	1.47	0.32	1.33	0.3	0.4	2.97	0.32
TRINITY_DN88570_c0_g1	DYW_deaminase.	0.58	1.71	1.2	2.89	0.88	1.95	2.73	2.05	3.45	2.7	3.27	2.17	2.99	0.55	1.32	1.51	1.55	1.78	2.8	2.37	1.2

TRINITY_DN89236_c0_g1	DYW_deaminase.	1.54	4.14	3.29	1.84	2.53	7.5	1.75	1.32	0.67	2.35	1.68	2.44	1.81	1.17	2.49	1.54	0.87	2.15	1.2	0.78	1.13
OG0000077																						
TRINITY_DN24002_c0_g1	RNase_HI_RT_Ty1.	0.05	0.09	0	0.49	0.75	0.22	0.61	0.19	0.23	6.02	11.14	16.45	0.76	0.04	0.16	0.18	0.14	0.04	0.07	0	0.22
TRINITY_DN25795_c0_g1	RNase_HI_RT_Ty1.	1.74	2.95	3.3	1.7	3.8	0.48	0.61	1.42	3.83	7.22	6.67	5	1.6	1.26	0.77	1.15	2.31	1.28	1.27	1.6	0.51
TRINITY_DN28676_c0_g1	RNase_HI_RT_Ty1.	8.68	12.33	9.12	4.23	7.78	1.72	2.2	2.24	6.68	5.28	10.63	7.88	4.22	3.76	2.89	1.58	14.84	**25.63**	1.71	1.89	1.36
TRINITY_DN34206_c0_g1	RNase_HI_RT_Ty1.	15.81	10.03	4.35	12.09	11.32	6.23	5.65	6.24	8.24	17.69	18.42	8.67	4.51	3.18	0.64	4.1	7.97	1.44	4.02	5.35	1.28
TRINITY_DN34206_c0_g3	RNase_HI_RT_Ty1.	6.75	5.24	16	5.62	4.12	6.87	3.53	1.23	3.01	13.06	3.16	7.81	5.69	0.94	1.52	6.62	4.42	12.58	5.41	3.84	2.15
TRINITY_DN35909_c1_g6	RNase_HI_RT_Ty1.	7.66	6.43	5.01	3.33	6.73	1.76	2.77	5.11	6.15	4.71	8.53	10.71	9.58	6.42	4.02	4.78	**27.44**	**29.3**	2.97	10.5	3.72
TRINITY_DN4622_c0_g1	RVT_2. Reverse transcriptase	1.08	0	1.28	1.21	0.35	0	0.67	0.16	2.88	2.46	1.07	2.69	0.54	0.24	0.44	0.22	0.68	1.37	0	0.15	0
TRINITY_DN46286_c0_g1	RNase_HI_RT_Ty1.	3.14	4.14	2.09	1.83	4.11	0.28	0.9	1.59	1.44	2.94	6.02	7.07	2.68	3.44	1.79	0.5	6.63	6.47	0.54	2.7	0.48
TRINITY_DN50965_c0_g1	RVT_2. Reverse transcriptase	3.91	1.61	1.28	2.09	1.15	0.1	0.43	0.26	1.77	3.07	2.73	3.21	0.58	0	0.1	0	0.25	0.89	0.59	0.97	0.35
TRINITY_DN56262_c0_g1	RVT_2. Reverse transcriptase	1.19	1.08	1.09	0.69	2.1	0	0.34	1.2	3.1	4.62	2.14	2.74	0.45	0.15	0.48	0	0.19	1.04	1.53	0.95	1.46
OG0000193																						
TRINITY_DN18647_c0_g1	--	0.75	1.31	1.64	2	2.14	2.43	2.08	4.02	3.46	3.28	3.06	3.39	5.65	5.06	5.17	3.17	4.47	2.94	3.08	1.89	2.45
TRINITY_DN2916_c0_g1	--	1.05	2.28	1.82	2.87	2.95	5.57	1.76	2.41	2.61	4.5	5.65	5.34	6.84	8.66	7.07	4.01	3.27	3.95	4.5	2.42	5
TRINITY_DN34833_c0_g1	--	2.77	3.56	3.18	9.6	11.31	8.39	8.98	12.52	14.27	12.41	15.09	17.23	10.53	5.48	8.79	6.49	7.26	7.08	17.82	2.59	13.21
TRINITY_DN47894_c0_g1	--	1.64	0.41	0.89	2.53	1.76	0.24	3.79	3.03	5.42	5.93	5.07	3.45	2.75	1.02	1.69	0.78	1.23	2.37	1.04	1.3	1.59
TRINITY_DN48994_c0_g1	--	0.97	1.9	1.28	0.48	2.32	0.11	1.32	2.11	1.95	2.34	1.92	1.29	1.28	2.43	2.4	0.79	0.54	1.22	1.83	1.88	1.41
TRINITY_DN50021_c0_g1	--	1.45	1.86	1.21	2.03	1.21	0	0.63	1.77	1.8	2.99	2.86	3.03	2.12	1.43	2.19	0.69	1.13	0.51	1.47	1.12	2.96
TRINITY_DN51080_c0_g1	--	0.23	0.32	0.21	0.2	0.62	0.19	0.9	1.89	3.47	2.07	3.1	1.21	1.88	0.44	0.72	0.33	0.45	1.65	2.72	1.35	1.83
TRINITY_DN663_c0_g1	--	2.49	2.49	3.5	11.17	12.52	14.37	9.82	14.09	15.09	**22.88**	16.37	**21.69**	13.76	9.73	16.96	9.24	16.18	15.38	2.58	1.83	9.99
TRINITY_DN68321_c0_g1	--	0	0.52	0.18	1.99	1.88	1.97	2.46	3.09	3.61	2.31	3.56	1.76	3.94	2.45	3.59	5.59	3.7	2.68	3.69	2.02	0.7
TRINITY_DN85527_c0_g1	--	0.19	0.44	0.35	1.92	1.45	2.75	3.47	3.11	4.27	4.75	3.3	3.77	7.58	4.58	6.12	10.01	3.08	3.89	5.06	2.03	0.97
TRINITY_DN9682_c0_g1	--	1.42	1.39	2.91	9.86	6.03	5.38	9.1	7.29	12.31	14.98	14.05	13.35	7.61	4.71	7.98	5.17	5.51	8.74	3.8	2.63	8.26
OG0000487																						
TRINITY_DN19412_c0_g1	--	6.01	5.75	6.97	7.6	8.61	0.39	5.26	5.21	10.4	5.08	8.12	8.86	0.37	0.96	1.28	1.66	3.59	2.05	5.64	9.08	5.83

TRINITY_DN19657_c0_g1	--	9.29	9.81	10.81	16.19	19.39	6.91	8.85	17.49	16.24	12.24	14.21	18.67	10.61	5.66	7.02	6.62	11.54	10.03	13	12.27	13.68
TRINITY_DN22141_c0_g1	--	6.8	4.63	5.52	4.42	6.54	0.43	2.58	5.43	6.36	4.24	5.36	4.62	0.28	0.96	1	0.26	1.94	1.12	9.15	7.53	4.78
TRINITY_DN25685_c0_g1	--	7.28	8.6	8.23	11.24	12.61	2.14	5.36	9.13	12.33	6.97	10.66	11.6	2.39	1.91	2.87	1.92	5.9	4.23	11.46	9.74	8.57
TRINITY_DN32040_c0_g1	--	4.84	1.27	1.68	2.97	4.12	0.06	2.54	3.93	6.04	5.04	3.4	3.09	1.1	0.77	1.13	0.4	1.81	2.01	0.51	4.11	0.43
TRINITY_DN32040_c1_g1	--	4.27	3.16	5.04	4.65	5.37	0	2.3	4.06	5.18	3.75	4.47	4.01	0.61	0.2	0.54	0.51	0.91	2.24	5.49	11.98	3.46
TRINITY_DN32040_c1_g2	--	12.82	6	6.38	17.15	**21.03**	7.28	15.65	14.74	18.27	16.48	12.57	12.89	12.44	7.8	12.34	8.55	14.03	11.59	2.23	6.68	2.82
TRINITY_DN3945_c1_g1	--	5.21	0.9	1.45	1.54	3.51	0	2.12	3.99	4.38	3.59	2.62	2.73	0.23	0.67	0.74	0.56	0.86	1.56	2.44	4.65	2.27
TRINITY_DN79776_c0_g1	--	10.62	12.1	11.9	14.35	19.2	9.02	8.68	15.25	**22.19**	9.6	16.73	9	6.25	2.95	5.01	3.69	6.37	7.64	12.63	14.34	8.18
OG0000659																						
TRINITY_DN16790_c0_g1	INPP5c_INPP5B. Catalytic inositol polyphosphate 5-phosphatase	0	0	0	0	0.05	0	0.11	0.07	0	5.16	12.66	6.75	0	0.1	0	0	9.45	0.4	0	0	0
TRINITY_DN17015_c0_g1	INPP5c_ScInp51p-like. Catalytic inositol polyphosphate 5-phosphatase	**36.92**	**30.96**	**25.66**	13.03	11.72	16.7	7.71	10.82	13.05	11.13	12.68	13.05	1.04	1.95	3.9	2.1	4.66	4.54	0	0	0.11
TRINITY_DN18319_c0_g1	INPP5c. Catalytic domain of inositol polyphosphate 5-phosphatases.	7.22	5.89	1.18	**21.43**	16.37	**42.31**	1.45	2.35	0.86	7.78	4.5	3.06	1.56	1.06	2.39	2.1	2.57	1.51	**44.7**	3.58	**54.02**
TRINITY_DN25697_c0_g1	INPP5c_INPP5B. Catalytic inositol polyphosphate 5-phosphatase	0.03	0	0	0	0	0	0.09	0	0.04	9.16	16.41	14.24	0	0	0.02	1.17	0.51	0.02	0	0	0
TRINITY_DN58518_c0_g1	INPP5c_ScInp51p-like. Catalytic inositol polyphosphate 5-phosphatase	0	0	0	0	0	0	0	0	0	2.8	3.34	1.15	0	0	0	0	0	0	0	0	0

TRINITY_DN67071_c0_g1	INPP5c_INPP5B. Catalytic inositol polyphosphate 5-phosphatase	0	0	0.32	0	0	0.28	0.3	0	0	2.78	1.72	1.41	0	0	0	1.16	0	0	0	0	0
TRINITY_DN81769_c0_g1	INPP5c. Catalytic domain of inositol polyphosphate 5-phosphatases.	9.03	14.18	11.2	11.39	14.97	6.41	7.36	14.4	11.73	10.31	19.37	14.86	5.69	5.13	4.21	4.93	8.62	6.81	11.11	13.51	7.7
OG0000697																						
TRINITY_DN124_c0_g1	--	3.03	3.31	1.88	3.3	3.74	1.81	3.2	3.23	3.44	3.68	3.13	3.5	0.93	0.63	1.14	1.58	1.77	1.6	9.05	3.22	3.22
TRINITY_DN15505_c0_g1	LRRNT_2. Leucine rich repeat N-terminal domain.	2.08	2.25	2.99	3.16	2.07	0.52	0.67	1.18	2.14	2.91	2.55	2.25	0.5	0.3	0.45	1.35	1.14	0.8	4.83	2.88	2.28
TRINITY_DN26500_c0_g1	STKc_IRAK.	8.15	0	0	4.18	0	0.12	1.62	0	0	4.35	0	0	1.67	2.65	1.68	1.93	2.81	1.28	2.03	2.94	0
TRINITY_DN26500_c0_g2	STKc_IRAK.	5.49	0	0	5.11	0	0	1.22	0	0	5.43	0	0	3.18	4.03	2.81	2.29	5.65	3.32	0.24	0.83	0
TRINITY_DN26500_c0_g3	STKc_IRAK.	4.16	5.32	5.76	5.28	5.7	12.51	4.14	4.49	5.29	4.89	3.9	7.8	9.4	4	4.46	4.82	7.09	8.87	13.66	4.41	7.26
TRINITY_DN6019_c0_g1	STKc_IRAK.	53.08	16.41	29.36	60.92	74.39	12.99	55.05	64.85	67.13	40.37	70.92	64.83	52.25	19.91	26.83	44.67	32.12	26.95	152.6	4.52	77.31
TRINITY_DN72431_c0_g1	STKc_IRAK.	0	0	0	0	0	0	0	0	0	1.86	0	0	0	0	0	0	0	0	0	0	0
OG0000748																						
TRINITY_DN11486_c0_g1	PPR. Pentatricopeptide repeat.	16.17	21.39	18.13	11.36	9.55	9.5	3.2	7.21	4.9	5.85	5.59	5.65	3	2.23	3.62	2.17	2.03	1.09	2.07	2.58	4.28
TRINITY_DN20811_c0_g1	PPR. Pentatricopeptide repeat.	3.12	3.36	3.08	3.89	2.03	1.33	1.83	3.28	2.39	3.24	2.86	3.35	2.48	1.27	2.59	1.48	1.85	1.28	2.25	1.29	2.58
TRINITY_DN3182_c0_g1	PPR. Pentatricopeptide repeat.	4.17	4.89	2.92	4.84	7.23	3.48	2.73	4.03	1.96	4.83	7.23	2.75	4.72	1.51	3.58	1.51	2.7	4.02	6.78	1.92	2.38
TRINITY_DN33968_c0_g1	PPR. Pentatricopeptide repeat.	8.62	13.24	9.91	6.03	5.34	6.16	2.11	3.26	3.17	4.86	7.37	8.22	2.32	1.29	2.22	1.58	1.94	2.12	1.83	5.04	1.6
TRINITY_DN34609_c0_g2	PPR. Pentatricopeptide repeat.	7.11	9.57	10.16	8.78	8.37	7.52	5.39	7.31	9.89	10.62	11.39	12.39	8.96	5.88	9.72	8.36	8.36	7.11	3.54	3.28	3.92
TRINITY_DN36360_c0_g1	PPR. Pentatricopeptide repeat.	2.84	1.46	1.59	3.55	4.59	0.2	2.55	3	3.26	2.89	5.59	4.28	2.46	0.75	1.83	1.52	2.16	3.26	2.11	0.95	2.74

TRINITY_DN36612_c0_g1	PPR. Pentatricopeptide repeat.	4.66	8.71	4.65	4.53	4.7	5.91	1.98	3.64	4.92	4.36	8.36	5.39	2.82	1.33	4.24	1.38	1.86	2.49	3.85	3.76	3.04
TRINITY_DN43278_c0_g1	PPR. Pentatricopeptide repeat.	3.83	6.24	5.01	1.58	4.74	0	1.39	2.86	1.83	2.64	2.38	1.86	0	0.46	0	0.19	0.39	0.36	0.78	0.58	0.37
TRINITY_DN52161_c0_g1	PPR. Pentatricopeptide repeat.	2.18	4.81	3.72	3.52	4.24	0.99	1.2	3.62	2.69	4.65	4.28	2.52	2.27	1.89	2.47	1.77	1.66	6.16	1.93	2.7	0.57
TRINITY_DN53905_c0_g1	PPR. Pentatricopeptide repeat.	2.48	3.53	2.71	9.3	11.86	4.31	9.86	20.56	16.22	7.82	11.28	9.66	10.57	7.63	10.68	10.14	10.69	5.32	7.09	0.75	3.58
TRINITY_DN55797_c0_g1	PPR. Pentatricopeptide repeat.	7.99	11.06	8.35	4.68	2.8	1.22	1.98	1.56	1.9	2.16	2.87	2.66	1.77	0.93	1.41	1.6	1.64	1.35	1.31	5.19	0.43
TRINITY_DN57132_c0_g1	PPR. Pentatricopeptide repeat.	1.17	1.56	3.23	2.51	2.73	1	1.08	2.28	2.37	3.11	2.92	5.3	1.24	0.48	1.07	1.54	2.54	1.11	1.53	5.29	0.33
TRINITY_DN61213_c0_g1	PPR. Pentatricopeptide repeat.	3.87	4.15	4.51	1.29	2.33	0.27	0.78	2.19	1.41	2.98	3.7	2.1	0.7	1.55	0.61	0.65	0.66	0.6	0.97	2.63	0.84
TRINITY_DN62958_c0_g1	PPR. Pentatricopeptide repeat.	5.49	9.65	5.76	4.03	6.36	0.61	2.32	3.22	4.21	4.41	6.54	4.34	0.81	1.39	1.38	1.92	4.22	1.61	3.06	3.9	0.53
TRINITY_DN63307_c0_g1	PPR. Pentatricopeptide repeat.	14.03	17.58	17.59	8.06	3.85	0.78	2.94	5.04	1.85	2.64	3.74	3.94	0.67	0.81	0.92	1.02	0.57	1.37	1.74	4.23	1.26
TRINITY_DN64587_c0_g1	PPR. Pentatricopeptide repeat.	2.16	4.98	4.64	2.88	3.38	1.15	2.23	2.63	3.32	2.69	4.53	4.82	1.1	1.09	1.05	0.82	1.26	0.89	1.12	5.14	1.86
TRINITY_DN65856_c0_g1	PPR. Pentatricopeptide repeat.	4.74	3.07	4.56	2.38	3.59	0.25	2.63	0.64	1.91	2.66	3.11	2.43	0.73	0.24	0.4	0.5	0	0.19	1.71	1.52	2.14
TRINITY_DN72299_c0_g1	PPR. Pentatricopeptide repeat.	20.69	15.64	15.8	6.01	7.53	1.21	1.88	3.58	2.64	3.62	1.49	2.72	1.03	0.77	0.77	0.48	0.33	0.89	1.44	1.95	0.62
TRINITY_DN74635_c0_g1	PPR. Pentatricopeptide repeat.	3.23	4.91	6.07	5.32	4.18	1.24	2.17	3.04	2.67	3.25	2.6	3.32	1.84	1	1.03	1.48	1.75	0.84	2.51	3.12	0.66
OG0001176																						
TRINITY_DN12752_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	3.53	5.79	3.49	0	0	0	0	0	0	0	0	0

TRINITY_DN15658_c0_g1	CYSTM. Cysteine-rich TM module stress tolerance.	0	0	0	0	0	0	0	0	0	4.44	3.06	3.8	0	0	0	0	0	0	0	0	0
TRINITY_DN16748_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	3.18	0.67	0	0	0	0	0	0	0	0	0	0
TRINITY_DN19613_c0_g1	--	0	0	0	0	0	0	0	0	0	5.19	5.71	1.75	0	0	0	0	0	0	0	0	0
TRINITY_DN21803_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	5.1	2.87	1.55	0	0	0	0	0	0	0	0	0
TRINITY_DN21989_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	3.84	1.5	0.41	0	0	0	0	0	0	0	0	0
TRINITY_DN22432_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	4.93	8.56	7.12	0	0	0	0	0	0	0	0	0
TRINITY_DN23091_c0_g1	TLD.	0	0	0	0	0	0	0.08	0	0	3.11	2.39	0.73	0	0	0	0	0	0	0	0	0
TRINITY_DN25199_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	4.84	5.51	1.47	0	0	0	0	0	0	0	0	0
TRINITY_DN27832_c0_g1	--	0	0	0	0	0	0	0	0	0	3.61	4.79	2.09	0	0	0	0	0	0.07	0	0	0
TRINITY_DN28824_c0_g1	BTB. Broad-Complex. Tramtrack and Bric a brac. Domain in Broad-Complex.	0.39	0.04	0.43	0.3	0	1.67	0.38	0	0.2	5.1	8.41	3.77	0	0	0	0	0	0	0.02	0	0.07
TRINITY_DN51820_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	2.81	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN54299_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	4	7.16	0.07	0	0	0	0	0	0	0	0	0
TRINITY_DN55033_c0_g1	--	0	0	0	0	0	0	0	0	0	2.95	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN57180_c0_g1	--	0	0	0	0	0	0	0	0	0	2.7	6.03	1.25	0	0	0	0	0	0	0	0	0
TRINITY_DN57390_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	2.55	5.28	0.79	0	0	0	0	0	0	0	0	0
TRINITY_DN62765_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	7.08	5.74	0.19	0	0	0	0	0	0	0	0	0
TRINITY_DN65272_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	3.61	8.39	1.08	0	0	0	0	0	0	0	0	0
TRINITY_DN68960_c0_g1	TLD.	0	0	0	0	0	0	0	0	0	2.91	4.66	0.47	0	0	0	0	0	0	0	0	0
OG0001613																						
TRINITY_DN11200_c0_g1	STKc_phototropin_like.	**29.31**	**27.91**	**28.88**	**30.24**	**33.29**	**49.65**	**33.75**	**41.46**	**60.67**	**30.75**	**34.93**	**49.86**	**47.28**	**39.92**	**72.28**	**40.75**	**36.7**	**46.9**	11.44	14.23	19.84
TRINITY_DN25915_c0_g1	STKc_phototropin_like.	11.05	18.41	**24.79**	**47.76**	**90.19**	**93.36**	**46.39**	**101.7**	**69.27**	**27.59**	**45.07**	**45.86**	**99.46**	**68.15**	**103.5**	**50.86**	**49.15**	**61.79**	**44.02**	10.73	**60**
TRINITY_DN34203_c0_g1	STKc_phototropin_like.	16.02	16.61	19.85	17.58	15.64	**21.83**	15.49	16.03	19.01	**22.94**	**26.49**	**23.05**	**21.77**	18.27	**24.47**	13.44	16.38	19.4	7.94	10.42	8.37
TRINITY_DN34203_c1_g3	STKc_phototropin_like.	7	10.75	**20.41**	**26.06**	**22.39**	**38.65**	15.53	**38.03**	18.01	**25.44**	**36.09**	**40.43**	**21.33**	**23.18**	**34.51**	**26.74**	19.04	15.92	**34.72**	19.09	5.05
TRINITY_DN50577_c0_g1	STKc_phototropin_like.	7.54	5.07	5	2.82	2.3	3.54	0.73	1.5	1.05	7.19	9.16	9.62	3.9	2.68	2.39	2.34	1.1	2.11	0.16	0.14	2.48
TRINITY_DN77511_c0_g1	STKc_phototropin_like.	5.19	5.73	4.78	3.11	2.14	2.2	1.04	1.46	0.52	4.91	7.61	6.53	3.19	1.4	1.94	1.64	0.67	1.02	0.31	0	2.35
OG0001853																						

TRINITY_DN12421_c0_g1	--	0.14	0	0	0.49	0.63	0	2.28	2.12	11.62	4.44	2.2	2.26	2.78	3.39	4.73	0.66	3.39	9.59	0	0.4	0
TRINITY_DN23105_c0_g1	--	0	0.79	0	0	0	0	0	0.18	0	**24.84**	**572.8**	**183.5**	0.94	0.13	0.21	1.24	0.34	0.38	0	0	0.48
TRINITY_DN23321_c0_g1	--	0.15	0	0.14	0.33	0.75	0	3.36	2.95	12.55	5.15	4.22	3.27	7.57	6.21	7.93	2.32	4.46	14.42	0.47	0.37	0.28
TRINITY_DN32709_c0_g1	MIP-T3. Microtubule-binding protein MIP-T3.	0.8	0.45	0.32	2.52	3.44	0.29	8.91	7.75	18.99	15.91	9.63	5.89	15.43	**28.45**	**24.57**	9.38	13.03	5.24	0.11	0	0.17
TRINITY_DN5214_c0_g1	--	0	0.1	0.1	0.76	0.91	0.09	4.87	4.52	17.22	5.74	4.01	3.52	8.97	10.03	14.91	4.02	7.56	18.3	0.17	0.52	0
TRINITY_DN87120_c0_g1	--	0.5	0	0	0.58	0.63	0	1	2.01	13.24	3.1	1.96	2.68	0.64	2.38	3.03	1.26	3.38	5.25	0.64	0.64	0
OG0002240																						
TRINITY_DN21553_c0_g2	PPR. Pentatricopeptide repeat.	2.55	4.64	6.48	3.2	5	4.41	2.58	4.78	4.14	3.08	3.54	6.06	2.3	1.02	3.11	1.78	2.55	3.83	2.56	2.38	1.45
TRINITY_DN2944_c0_g1	PPR. Pentatricopeptide repeat.	10.25	6.73	10.28	7.48	3.32	4.68	3.12	3.52	3.11	3.95	2.52	1.74	1.39	1.94	1.63	0.72	2.05	1.86	1.4	3.35	1.95
TRINITY_DN3166_c0_g1	PPR. Pentatricopeptide repeat.	3.56	4.49	6.6	4.33	5.17	7.95	5.04	4.68	3.96	3.58	4.11	5.49	7.29	5.18	6.29	4.68	4.58	4.77	3.79	1.91	3.05
TRINITY_DN51101_c0_g1	PPR. Pentatricopeptide repeat.	3.26	1.27	4.44	2.03	2.59	1.11	2.01	3.62	1.03	3.78	3.21	4.13	0.48	0.94	0.84	1.33	1.81	1.1	2.77	3.15	1.15
TRINITY_DN54175_c0_g1	PPR. Pentatricopeptide repeat.	1.2	2.18	1.79	0.91	1.36	0.72	0.39	0.61	1.11	3.22	1.47	1.56	1.38	0.57	1.13	0.14	0.73	0.26	0.46	1.16	0.55
TRINITY_DN64072_c0_g1	PPR. Pentatricopeptide repeat.	2.28	1.76	3.55	1.68	2.93	4.75	1.96	0.89	1.01	2.46	2.76	2.03	2.83	2.4	2.27	1.83	3.07	2.47	1.91	2.88	0.65
TRINITY_DN67422_c0_g1	PPR. Pentatricopeptide repeat.	0.87	1.42	1.75	2.11	1.71	1.52	0.6	2.46	1.14	2.27	2.56	2.2	0.13	0.92	0.65	0	1.01	1.22	1.61	1.17	1.28
TRINITY_DN67547_c0_g1	PPR. Pentatricopeptide repeat.	4.5	8.7	5.8	6.43	5.19	7.6	3.57	4.57	2.08	3.03	4.43	4.37	2.76	3.69	2.81	1.54	3.33	2.38	2.93	2.09	1.17
TRINITY_DN68133_c0_g1	PPR. Pentatricopeptide repeat.	8.33	13.94	12.84	6.31	6.25	7.43	2.57	7.16	3.44	3.13	4.37	5.55	3.82	2.68	3.37	1.98	2.29	4.15	4.63	2.88	4.52
TRINITY_DN83728_c0_g1	PPR. Pentatricopeptide repeat.	2.81	9.09	6.21	3.42	5.01	2.24	0.73	1.28	0.35	3.42	1.24	1.79	0.75	0.64	1.32	0.4	0.96	1.11	**21.68**	2.98	5.32
OG0002627																						

TRINITY_DN11021_c0_g1	PPR. Pentatricopeptide repeat.	3.74	2.87	4.53	2.57	2.54	1.25	3.05	1.79	3.4	3.86	0.97	5	1.96	0.75	1.43	0.56	0.96	1.04	1.22	1.9	0.91
TRINITY_DN14731_c0_g1	PPR. Pentatricopeptide repeat.	1.62	2.21	2.24	2.32	2.11	3.04	2.19	2.23	3.25	3.79	1.48	3.08	1.95	0.79	1.87	1.21	2.14	1.43	3.23	1.12	2.35
TRINITY_DN39337_c0_g1	PPR. Pentatricopeptide repeat.	2.45	6.01	3.64	2.13	2.72	0.9	1.14	3.44	5.59	5.29	2.78	3.92	0.72	0.71	2.54	0.89	0.92	1.33	2.48	0.18	1.22
TRINITY_DN58472_c0_g1	PPR. Pentatricopeptide repeat.	1.36	2.75	2.78	4.2	4.12	1.92	2.6	3.67	3.23	3.39	5.2	5.93	3.01	1.26	1.63	1.14	1.91	0.93	1.99	2.04	0.98
TRINITY_DN58913_c0_g1	PPR. Pentatricopeptide repeat.	4.06	7.05	5.84	3.76	3.18	4.85	3.39	5.27	4.61	4.98	5.62	4.55	3.99	3.18	5.81	3.73	3.09	2.46	3.54	1.47	0.94
TRINITY_DN72570_c0_g1	PPR. Pentatricopeptide repeat.	1.32	2	2.97	2.04	2.44	2.1	3.66	2.52	2.17	3.15	2.16	1.19	1.24	0.44	0.47	0.83	0.86	1.16	1.48	1.55	0.54
TRINITY_DN73485_c0_g1	PPR. Pentatricopeptide repeat.	1.42	1.13	1.63	1.7	0.81	1.55	0.3	1.44	1.31	2.82	0.87	2.05	0.68	0.27	2.05	0.34	0.69	0.62	1.78	1.37	0
TRINITY_DN87176_c0_g1	PPR. Pentatricopeptide repeat.	2.68	3.16	4.21	1.37	2.16	1.26	1.77	1.44	1.43	3.4	3.42	4.01	1.21	0.48	1.49	1.65	2.15	0.56	4.9	2.44	2.92
OG0002773																						
TRINITY_DN16942_c0_g1	STKc_IRAK.	0.06	0	0	0.79	0.34	0	0.88	0.17	0.19	5.67	0.61	0.39	0.44	1.59	2	0.43	2.04	2.15	0.44	0	0
TRINITY_DN21585_c0_g1	STKc_IRAK.	2.68	1.82	0.82	9.99	5.83	6.92	11.37	5.29	5.84	10.87	7.57	9.72	4.61	3.96	4.41	1.83	5.25	2.77	13.53	15.28	3.67
TRINITY_DN23766_c0_g1	STKc_IRAK.	0.98	0.84	0.9	2.67	4.7	3.18	5.5	4.36	4.25	5.91	5.92	9.44	2.59	0.19	0.94	1.63	1.67	3.57	0.1	3.67	1.02
TRINITY_DN25296_c0_g1	STKc_IRAK.	7.16	9.22	**39.28**	**24.82**	**50.07**	**69.44**	**164.6**	**60.59**	**32.11**	**156.9**	**60.66**	**145.3**	**75.47**	10.66	**40.85**	10.45	**34.32**	**23.08**	**119.1**	**92.68**	**105.8**
TRINITY_DN26562_c0_g1	STKc_IRAK.	**23.77**	13.7	3.82	**23.4**	13.47	5.06	**50.44**	**41.04**	**46.12**	**22.5**	**28.81**	14.73	**55.16**	**57.95**	**68.89**	**28.5**	**70.44**	**77.21**	1.11	0.09	3.13
TRINITY_DN33266_c0_g1	LRR_RI.	5.18	6.48	**23.78**	14	**27.14**	5.23	**85.79**	**35.29**	17.24	**89.78**	**38.13**	**82.98**	16.12	2.78	9.76	3.93	10.23	6.64	**75.43**	**73.62**	**51.26**
TRINITY_DN35700_c0_g1	STKc_IRAK.	19.53	7.32	8.97	**25.63**	**46.69**	1.49	**57.82**	**103.5**	**64.08**	18.69	**26.7**	16.89	**74.63**	**99.24**	**85.1**	7.15	**43.29**	**32.78**	0	0	0.31
OG0003836																						
TRINITY_DN40963_c0_g1	NB-ARC. NB-ARC domain.	3.61	7.8	3.68	3.15	5.48	0.3	2.29	2.31	2.67	3.03	6.18	3.3	1.02	1.59	0.67	1.62	2.22	0.34	0	0	0.18
TRINITY_DN58644_c0_g1	NB-ARC. NB-ARC domain.	6.84	5.92	5.36	3.57	1.82	0.41	1.18	3.82	3.95	2.08	2.19	1.78	1.05	1.04	0.69	0.65	1.33	0.45	0.66	1.82	1.42
TRINITY_DN74389_c0_g1	NB-ARC. NB-ARC domain.	0.97	1.03	1.33	0.7	0.98	0	0.42	0.65	0.66	2.56	1.43	2.79	0	0.37	0.43	0	0.63	1.28	0.62	0.31	3.57

TRINITY_DN83117_c0_g1	NB-ARC. NB-ARC domain.	3.73	1.7	3.28	2.81	0.13	0.68	1.02	0	2.8	4.13	0.67	8.43	0.65	0.39	0.91	0.32	0.5	0.75	0	0	0.16
OG0004289																						
TRINITY_DN13368_c0_g1	Hsp70	0.83	0.5	0.51	13.72	15.18	3.45	**22.54**	**25.05**	**21.52**	8.19	**30.31**	6.61	0	0.14	0.22	0	0	0	8.29	**265.5**	16.87
TRINITY_DN17142_c1_g1	Hsp70	0.07	0.3	0.12	0.29	0.2	0.11	0.35	0.54	0.38	11.38	0.33	0.15	0.41	0.05	0	0.69	0.13	0	2.37	4.55	1.54
TRINITY_DN34407_c2_g5	Hsp70	**322.7**	**382**	**483**	**313.4**	**200.7**	**507.4**	**364.3**	**309.9**	**258.9**	**496.1**	**422.2**	**479.5**	**1080**	**562.7**	**394.3**	**575.3**	**569.7**	**297.5**	**1161**	**311.7**	**777.4**
TRINITY_DN42187_c0_g1	Hsp70	0	0	0	0	0	0	0	0	0	2.17	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN45384_c0_g1	Hsp70	0	0	0	0	0	0	0	0	0	3.01	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN69918_c0_g1	Hsp70	0	0	0	0	0	0	0	0	0.47	4.29	0.19	0	0.88	0	0	0.54	0.19	0	0.44	0	0.35
TRINITY_DN86414_c0_g1	Hsp70	**37.72**	**38.09**	**54.21**	**51.56**	**49.43**	10.2	**74.51**	**94.45**	**97.38**	**116.4**	**170.7**	**112.7**	**51.32**	**31.92**	**31.85**	**67.55**	**62.59**	**55.62**	**234.5**	**346.5**	**132.8**
OG0005442																						
TRINITY_DN24928_c0_g1	--	5.32	4.71	4.48	1.32	5.07	0.73	1.44	1.69	2.88	6.11	8.68	6.32	3.5	1.5	1.02	0.87	4.14	8.3	5.66	3.08	1.54
TRINITY_DN33063_c0_g1	rve. Integrase core domain.	**22.27**	**20.75**	16.43	18.45	13.35	17.31	3.27	4.55	7.04	9.86	19.22	**21.14**	**31.56**	1.8	4.87	**20.54**	4.75	8.17	10.44	12.41	9.84
TRINITY_DN3646_c0_g1	Chromo domain.	5.11	7.08	8.1	3.56	5.43	1.14	1.41	1.04	3.88	3.64	10.27	6.15	5.81	1.55	1.6	1.55	8.56	12.98	7.14	10.46	1.51
TRINITY_DN49720_c0_g1	--	3.76	4.28	1.44	1.23	1.43	0	0.94	0.32	0.65	2.93	1.7	2.35	0.36	0	0.34	0	0.91	0.41	1.58	3.63	0.87
TRINITY_DN58898_c0_g1	--	3.19	1.79	2.43	4.12	5.63	2.88	1.74	1	3.55	3.73	4.05	5.68	0.8	0.47	0.56	0.23	0.66	1.08	1.86	5.15	3.18
TRINITY_DN72415_c0_g1	Chromo domain.	1.94	1.09	2.06	1.95	7.92	0	2.82	2.87	3.81	4.33	3.52	6.9	0.57	0.45	0.24	0.42	0.73	1.05	2.55	0.29	1.79
OG0005452																						
TRINITY_DN14350_c0_g1	Elongation factor 1-alpha.	0	0	0	0	0	0	0	0.07	0.12	3.02	0.14	2.44	0	0	0	0	0	0	0	0	0.06
TRINITY_DN18058_c0_g1	EF1_alpha_III. Domain III of Elongation Factor 1.	3.98	7.15	5.55	14.15	15.24	12.61	**21.87**	**24.44**	18.67	**39.78**	17.93	17.48	16.64	15.16	15.38	10.8	12.95	10.01	18.73	19.26	13.98
TRINITY_DN18058_c0_g2	EF1_alpha_III. Domain III of Elongation Factor 1.	**120**	**280.1**	**247.7**	**523.8**	**512.7**	**367.4**	**893.4**	**842.5**	**726**	**631.5**	**695.6**	**689.7**	**411.1**	**380.8**	**468.4**	**281.7**	**227.1**	**196.4**	**1317**	**1475**	**460.4**
TRINITY_DN30150_c0_g1	Elongation factor 1-alpha.	0	0	0	0	0	0	0.39	0	0	10.37	0.64	4.98	0	0	0	0	0	0	0	0	0
TRINITY_DN33407_c1_g1	Elongation factor 1-alpha.	0	0	0	0	0	0	0.05	0	0	6.95	3.98	0.07	0	0	0	0	0	0	0	0	0
TRINITY_DN35218_c1_g4	Elongation factor 1-alpha.	**75.31**	**120.2**	**123.6**	**235.8**	**297.2**	**167.6**	**356.2**	**490.1**	**327.2**	**342.2**	**435.8**	**333.7**	**265.8**	**198.4**	**250.2**	**228.2**	**159.8**	**131.7**	**349.4**	**383.4**	**250.2**

TRINITY_DN73525_c0_g1	Elongation factor 1-alpha.	0	0	0	0	0	0	0	0	0	2.71	0.6	0.56	0	0	0	0	0	0	0	0	0
OG0010763																						
TRINITY_DN14129_c0_g1	dnaK.	**46.04**	**53.07**	**66.96**	**41.68**	**23.14**	**41.35**	**44.03**	**38.27**	**36.53**	**69.1**	**49.1**	**64.45**	**113.3**	**48.92**	**37.56**	**65.27**	**45.83**	**28.72**	**163.7**	**34.45**	**102.5**
TRINITY_DN17142_c0_g1	dnaK.	**325**	**377.9**	**485.9**	**324.1**	**195.5**	**264.4**	**322**	**322.9**	**270.8**	**481.6**	**450.8**	**493.8**	**774.1**	**343.1**	**254.1**	**409.7**	**334**	**200.8**	**1186**	**348.3**	**672.2**
TRINITY_DN31034_c0_g1	dnaK.	0	0	0	0	0	0	0.07	0	0	8.29	**25.75**	**23.99**	0	0	0	0.04	0	0	0	0	0
TRINITY_DN37866_c0_g1	dnaK.	1.1	0.54	1.28	13.7	11.26	1.82	11.24	19.49	12.82	5.78	**20.69**	3.45	0.08	0	0.05	0.09	0	0.18	4.25	**184.7**	11.33
OG0010767																						
TRINITY_DN16764_c0_g1	ZnF_C3H1. zinc finger.	0	0	0	0	0	0	0	0	0	3.05	3.8	0.26	0	0	0	0	0	0	0	0	0
TRINITY_DN17000_c1_g1	ZnF_C3H1. zinc finger.	13.41	**22.68**	**23.41**	15.86	**20.27**	4.36	5.39	10.88	8.4	12.59	17.68	18.77	2.6	3.45	3.32	2.97	5.64	9.8	9.88	**20.29**	9.12
TRINITY_DN9921_c0_g1	ZnF_C3H1. zinc finger.	0	0	0	0	0	0	0	0	0	5.97	16.57	3.77	0	0	0	0	0	0	0	0	0
TRINITY_DN9921_c1_g1	ZnF_C3H1. zinc finger.	0	0	0	0	0	0	0	0	0	3.3	5.52	0.44	0	0	0	0	0	0	0	0	0
OG0012394																						
TRINITY_DN23126_c0_g1	GapA.	6.66	7.6	7.39	13.12	11.34	8.45	**88.19**	16.99	**41.21**	**202.7**	**89.24**	**70.83**	**93.28**	**52.55**	**53.27**	**76.95**	**30.84**	**27.3**	**46.66**	**2054**	7.78
TRINITY_DN73669_c0_g1	GapA.	0	0	0	0	0	0	0	0	0	7.23	0	0.27	0	0	0	0	0	0	0	0	0
TRINITY_DN7602_c0_g1	GapA.	**581.6**	**542.7**	**584.5**	**623**	**575.1**	**753.6**	**1138**	**1105**	**1005**	**699.6**	**613.2**	**496.4**	**882.9**	**662**	**1359**	**863.1**	**672.1**	**666.2**	**690.9**	**408.8**	**1100**
TRINITY_DN7786_c0_g1	GapA.	0	0	0	0	0	0	0	0	0	5.9	0	0	0	0	0	0	0	0	0	0	0
OG0012409																						
TRINITY_DN20672_c0_g1	Elongation factor EF-2	2.74	3.21	5.67	7.07	5.85	4.21	3.73	4.66	2.99	17.67	15.14	15.67	7.59	1.55	4.43	5.77	3.21	2.29	5.98	**28.71**	0.94
TRINITY_DN25641_c0_g1	EF2_snRNP_III. Domain III of Elongation Factor 2 (EF2).	0	0	0	0	0	0	0	0	0	3.69	0	0.17	0	0	0	0	0	0	0	0	0
TRINITY_DN25641_c0_g2	Elongation factor EF-2	0	0	0	0	0	0	0	0	0	5.71	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN70288_c0_g1	FusA. Translation elongation factors (GTPases).	0	0	0	0	0	0	0	0	0	4.61	0	0	0	0	0	0	0	0	0	0	0

Additional file 4. Tissue-specific expression of expanded family OG0000038 in macauba palm transcripts with respective TPM value by tissue. TPM value ≥ 20 values in pink-colored box.

Expanded Family	Leaf			Leaf Sheath			Bulb			Root			Female Flower			Male Flower			Ripe fruits		
	S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	1 ♀	2 ♀	3 ♀	1 ♂	2 ♂	3 ♂	S 7	S 5	S 8
OG0000038																					
TRINITY_DN11382_c0_g2	0	0	0	0	0	0	0.45	0	0.65	6.12	2.19	9	0	0	0	0	0	0	0	0.13	0
TRINITY_DN14629_c0_g1	0	0	0	0	0.21	0	0.36	0	0	5.81	0.27	0.65	0	0	0	0	0	0	0	0	0
TRINITY_DN16579_c0_g1	0	0	0	0	0	0	0	0	0	12.81	**39.89**	7.59	0	0	0	0	0	0	0	0	0
TRINITY_DN17682_c0_g1	0	0	0	0	0	0	0	0	0	4.04	2.28	0.18	0	0	0	0	0	0	0	0	0
TRINITY_DN19913_c0_g1	0.12	0.11	0.12	0.22	0	0.1	0.65	0.63	0.52	**30.56**	9.24	0.72	0.48	0.28	0.67	0.36	0.24	0	0.1	0.24	0
TRINITY_DN21380_c0_g1	0.16	0	0.44	0	0	0.88	0	0	0.26	13.34	**42.74**	**28.02**	0	0	0	0	0	0.14	0	0	0.15
TRINITY_DN25769_c1_g1	0	0.37	0.37	0.35	0.61	0.32	0.17	0.81	0.17	14.44	**40.07**	**29.66**	0	0	0	0	0	0	0	0.19	0.19
TRINITY_DN26294_c0_g1	0	0	0	0	0	0	0	0	0.12	**30.86**	**76.15**	**51.6**	0	0	0	0	0	0	0	0	0
TRINITY_DN34075_c1_g1	0	0	0	0	0	0	0	0	0	15.01	0.21	0.89	0	0	0	0	0	0	0	0	0
TRINITY_DN35187_c4_g1	0	0	0	0	0	0	0	0	0	**80.65**	**88.46**	**31.19**	0	0	0.02	0	0.07	0	0	0	0
TRINITY_DN35187_c4_g2	0	0	0	0	0	0	0.12	0	0	**85.87**	**60.64**	**38.19**	0	0.11	0	0.13	0	0	0	0	0
TRINITY_DN35187_c4_g3	0	0	0	0	0	0	0	0	0.17	**131.87**	**527.27**	**136.9**	0.08	0	0.03	0.19	0	0.18	0	0	0

Additional file 5. Drought stress families with macauba palm transcripts and respective SPM values by tissue and analyzed. TPM value ≥ 20 in pink-colored box.

	Leaf			Leaf Sheath			Bulb			Root			Female Flower			Male Flower			Ripe fruits		
	S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	S 4	S 12	S 20	1 ♀	2 ♀	3 ♀	1 ♂	2 ♂	3 ♂	S 7	S 5	S 8
OG0000003	*PREDICTED_ calcium-dependent protein kinase 26 [Elaeis guineensis]*																				
TRINITY_DN17405_c0_g2	16.09	19.24	13.67	18.85	11.1	**57.77**	13.29	6.63	6.49	9.19	16.02	8.69	8.33	8.34	2.85	3.55	5.02	3.37	**35.41**	0.09	**144.69**
TRINITY_DN24067_c0_g1	**26.03**	**31.88**	**30.6**	**45.86**	**50.49**	**60.59**	**55.4**	**64.65**	**54.71**	**87.08**	**107**	**97.09**	**44.47**	**33.72**	**56.58**	**35.28**	**50.8**	**40.81**	**32.84**	15.63	**34.79**
TRINITY_DN31407_c0_g1	**24.52**	**32.74**	**36.63**	**38.27**	**42.43**	**43.02**	**45.43**	**56**	**59.75**	**51.88**	**53.43**	**61.69**	**53.33**	**44.67**	**57.96**	**28.82**	**36.9**	**29.02**	**77.12**	**64.11**	**63.83**
TRINITY_DN34134_c0_g1	18.69	15.64	15.42	**23.57**	**24.09**	9.22	**23.4**	**33.51**	**36.38**	**26.21**	**30.17**	**23.95**	17.19	12.42	19	14.86	18.75	15.53	**22.23**	14.78	**24.94**
TRINITY_DN35120_c5_g1	**26.11**	**42.13**	**30.39**	**60.16**	**68.23**	**76.35**	**88.11**	**89.66**	**66.06**	**78**	**94.23**	**89.4**	**66.74**	**20.13**	**37.35**	**42.32**	**32.99**	**31.45**	**125.9**	**70.8**	**122.79**
TRINITY_DN35503_c1_g1	19.91	**24.54**	**21.33**	**30.9**	**29.53**	**35.46**	**28.97**	**32.29**	**41.84**	**46.78**	**46.48**	**45.96**	**34.78**	15.62	**27.87**	**23.72**	**23.43**	19.6	**25.47**	**85.4**	**25.54**
TRINITY_DN35503_c1_g4	9.01	11.95	14.44	14.88	14.45	**23.76**	16.89	14.45	14.57	**46.29**	**53.09**	**47.93**	19.7	9.56	13.56	11.64	9.16	6.15	**26.89**	**50.37**	**27.26**
TRINITY_DN82979_c0_g1	1.98	0.46	0.36	1.23	0.9	0.41	2.93	2.53	1.73	4.55	4.4	3.93	2.26	2.11	4.59	1.34	18.95	**45.95**	1.67	0.22	3.56
OG0000004	*E3 ubiquitin-protein ligase At3g02290 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN22468_c1_g1	3.32	4.84	6.34	7.15	7.45	8.96	5.56	6.5	7.56	9.94	11.82	9.33	3.35	1.33	3.06	7.25	3.14	1.34	1.8	2.01	1.45
TRINITY_DN22468_c2_g1	12.7	14.87	14.48	16.59	12.68	**25.05**	5.72	5.84	5.4	13.71	15.3	14.34	2.93	2.74	7.55	2.13	1.81	1.72	12.88	7.63	19.06
TRINITY_DN24198_c1_g1	4.8	8.03	6.84	12.66	13.21	18.38	11.78	8.23	11.43	18.32	17.35	**21.61**	13.98	10.81	15.69	17.48	13.81	10.02	8.98	5.47	11.31
OG0000005	*nuclear transcription factor Y subunit B-4 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN34990_c2_g1	4.6	6.58	6.07	7.24	9.01	11.02	19.88	19.05	**24.28**	18.9	16.44	14.91	**32.8**	**25**	**41.32**	**22.44**	**25.64**	**26.82**	2.83	3.84	3.65
TRINITY_DN34990_c2_g2	6.64	14.01	12.97	12.6	13.48	**20.49**	**27.25**	18.07	**26.83**	17.59	17.09	**20.96**	**23.12**	18.32	**26.56**	**31.32**	**23.81**	18.44	**51.73**	13.55	**48.85**
TRINITY_DN34990_c2_g4	3.91	5.22	6.78	9.07	7.24	14.04	14.22	17.3	19.96	19.52	17.88	16.85	**37.47**	**23.15**	**45.66**	**25.01**	**45.32**	**40.36**	5.56	5.08	3.2
OG0000008	*PREDICTED_beta-glucosidase 12-like [Elaeis guineensis]*																				
TRINITY_DN18360_c0_g1	0.28	16.25	0.61	4.11	10.52	3.33	13.81	18.18	11.27	6.57	10.43	11.81	4.1	4	2.8	11.32	8.34	2.36	**27.18**	13.95	**70.31**
TRINITY_DN27655_c0_g1	**93.68**	16.5	15.43	**145.8**	**68.67**	17.16	**79.44**	**60**	**65.83**	**159.5**	**76.26**	**52.9**	**68.28**	**24.84**	**28**	**74.77**	**52.33**	17.61	14.47	0.2	9.42
TRINITY_DN28112_c0_g1	0	0	0	0.11	3.58	0.02	0.46	0.46	0.46	13.72	1.07	0.52	0.04	0	0	1.58	0	0	0	0.05	0.14
TRINITY_DN35708_c4_g1	**27.26**	1.09	1.96	**55.01**	**89.99**	**27.46**	**68.84**	**65.44**	**78.07**	**209.8**	**120.8**	**122.7**	**27.67**	**32.62**	**57.52**	**40.85**	**49.85**	**44.91**	0.41	0.05	**20.26**
TRINITY_DN47933_c0_g1	0	0	0	0.11	0	0	1.02	0.86	5.78	3.86	0.45	0.15	0.6	1.19	0.67	0	0.57	0.69	0	0	0
TRINITY_DN9444_c0_g1	4.48	1.82	2.17	6.02	1.4	10.11	1.38	0.2	1.67	14.7	**34.71**	**47.67**	0.47	0.35	0.93	1.34	12.35	1.89	0	0.05	0.07

OG0000009	*PREDICTED_ alpha_alpha-trehalose-phosphate synthase [UDP-forming] 1-like isoform X1 [Elaeis guineensis]*																				
TRINITY_DN35704_c1_g1	79.85	72.83	60.35	35.61	31.9	48.05	30.2	41.95	24.84	31.87	49.57	31.19	19.27	16.22	17.41	46.86	28.32	27.09	2.96	2.13	35.6
OG0000011	*PREDICTED_ L-ascorbATe peroxidase_ cytosolic [Elaeis guineensis]*																				
TRINITY_DN24915_c0_g1	0	0	0	0	0	0	0	0.04	0	9.98	0	0.05	0	0	0	0	0	0	0	0	0
TRINITY_DN32804_c1_g1	112.7	104.6	171.2	257.7	200.7	305.7	572.5	318.6	429	512.2	413.4	442.4	753.5	487.4	737.6	705.9	426.8	315.5	257.2	327.9	211.93
OG0000012	*PREDICTED_ nuclear transcription factor Y subunit A-7 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN28214_c0_g1	35.34	35.51	24.82	47.97	42.66	36.81	24.08	23.95	31.37	31.83	35.97	41	32.71	19.24	29.19	37.9	27.61	26.17	21.5	39.2	20.69
OG0000015	*PREDICTED_ probable protein phosphatase 2C 6 [Elaeis guineensis]*																				
TRINITY_DN34032_c0_g1	27.62	27.35	10.96	25.49	15.07	8.96	20.95	13.12	11.87	23.3	19.93	20.89	27.89	10.62	15.6	28.02	14.39	17.95	67.28	13.01	44.56
TRINITY_DN34032_c0_g2	23.47	22.4	20.23	27.2	21.82	11.62	22.83	27.15	19.13	23.19	22.02	23.7	28.02	16.39	44.29	22.5	21	18.36	7.48	4.13	3.35
TRINITY_DN34791_c2_g1	76.06	93.81	88.03	102.6	76.5	118.8	96.83	90.61	98.41	52.17	44.56	43.61	71.81	43.38	71.01	58.37	58.01	44.9	30.24	6.63	13.79
TRINITY_DN34791_c2_g4	93.92	103.1	117.8	149.7	139.1	135.5	118.7	133.3	97.25	94.95	95.99	92.26	102	57.75	125.1	89.41	91.69	92.11	56.86	22.43	45.83
OG0000017	*PREDICTED_ calmodulin*																				
TRINITY_DN16570_c1_g1	87.41	93.01	117.9	46.77	16.9	224.1	66.44	17.63	39.53	137.6	45.17	99.8	19.33	2.98	7.62	4.33	1.7	1.91	0.97	6.98	28.87
TRINITY_DN16580_c0_g1	0	0	0.1	0	0.16	0.17	0	0.19	0.35	10.87	0.1	0.24	0.41	0.31	0.08	0.2	0	0.03	0.25	0	0.1
TRINITY_DN16580_c0_g1	0	0.05	0	0	0	0	0	0	0	90.86	54.34	21.73	0	0	0	0	0.11	0	0	0	0
TRINITY_DN33665_c0_g1	271.6	408	387.1	402.7	410.6	662.7	631.9	635.1	780.9	473.4	463.6	480.8	953.3	417.2	693.8	763.5	470.4	408.3	1148	726.4	902.54
TRINITY_DN63959_c0_g1	0.95	1.01	0.51	0.69	0.96	2.14	1.16	1.04	0.26	3.85	2.57	8.32	1.58	0.66	1	0.15	0	0.14	0.18	0	0.44
TRINITY_DN9572_c0_g1	32.7	37.62	35.05	78.18	130.5	134	227.2	401.1	272.6	187.1	200.2	189.6	298.9	141.2	291.6	330.3	157.5	168.3	372.9	282.2	345.56
OG0000018	*PREDICTED_ABC transporter G family member 45-like isoform X1*																				
TRINITY_DN16806_c0_g1	1.18	1.43	0.56	1.34	1.48	0.07	4.4	0.72	3.13	10.94	3.27	3.34	1.9	0.44	0.68	1.02	0.72	1.11	1.56	197.9	0.3
TRINITY_DN22152_c0_g1	0.95	0.37	0.19	1.3	0.56	0.75	2.27	0.89	0.61	12.16	1.73	1.41	2.6	2.06	2.1	0.77	2.31	0.84	0.32	0.07	1
TRINITY_DN22700_c0_g1	0.24	0.26	0.34	1.63	4.32	0.72	5.27	0.93	6.02	15.01	19.02	19.71	0.25	0.08	0.07	0.26	0.52	0.15	0.08	0	0.05
TRINITY_DN29259_c0_g1	2.48	3.62	1.65	4.21	4.37	4.54	12.01	2.58	6.63	25.73	6.99	12.39	15.98	4.48	5.84	6.39	6.13	4.68	3.85	290.5	1.36
TRINITY_DN2976_c0_g1	0	0	0	0	0	0	0	0	0	4.19	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN31780_c2_g1	1.35	2.19	1.86	3.18	8.46	1.76	8.05	7.5	4.16	33.93	70.89	42.69	1.4	0.84	1.18	1.1	3	1.62	0.91	0.34	1.29
TRINITY_DN31836_c2_g1	0.29	0.26	0.29	0.57	0.46	0.15	0.46	0.22	0.64	8.6	5.42	3.26	0.36	0.24	0.55	0.25	0.4	0.27	0.08	0.07	0.26
TRINITY_DN35048_c1_g1	8.79	13.7	8.24	15.02	15.97	17.21	14.09	10.57	13.31	20.38	23.8	30.28	7.2	4.4	8.77	6.14	11.21	11.76	20.11	88.41	18.98
TRINITY_DN35048_c1_g2	19.76	27.34	71.7	5.82	4.14	32.06	2.99	1.51	2.36	26.97	16	19.29	7.34	1	3.33	2.8	2.19	2.69	1.01	1.77	2.52

OG0000023	*PREDICTED_ serine/threonine-protein kinase SAPK2 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN28647_c0_g1	26.67	28.78	22.2	37.36	33.45	40.14	19.37	33.54	27.21	34.51	37.35	25.01	52.96	38.12	76.65	29.23	39.16	36.31	1.49	9.06	1.15
TRINITY_DN32053_c0_g1	23.48	31.95	30.28	55.59	74.38	34.05	59.5	52.58	47.61	60.3	64.99	61.76	35.99	20.8	31.32	46.03	35.83	28.31	45.72	75.3	54.51
TRINITY_DN34069_c1_g1	29.74	27.61	24.95	60.8	56.59	24.4	45.26	33.82	31.05	37.64	32.1	33.36	18.55	13.02	19.86	27.83	21.15	18.93	62.63	21.68	25.69
TRINITY_DN34211_c0_g1	131.5	172.5	167.2	247.3	265.4	190.8	186.3	247.5	199.3	192.5	214.4	197.5	83.95	80.93	148.3	130.3	132.1	136.3	113	39.54	195.51
OG0000024	*PREDICTED_ 14-3-3-like protein B [Elaeis guineensis]*																				
TRINITY_DN32377_c0_g1	0	0	0	0	0	0	0	0	0	19.55	49.44	9.21	0	0	0	0	0	0.06	0	0	0
TRINITY_DN34169_c0_g1	0	0.05	0	0	0	0.02	0.05	0	0	34.69	77.61	19.65	0	0.04	0.02	0	0	0	0	0	0
TRINITY_DN44954_c0_g1	0	0	0	0	0	0	0	0	0	6.8	2.68	0	0	0	0	0	0	0	0	0	0
TRINITY_DN56448_c0_g1	0	0	0	0	0	0	0	0	0	5.14	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN58591_c0_g1	0	0	0	0	0	0	0	0	0	3.92	0	0	0	0	0	0	0	0	0	0	0
OG0000027	*PREDICTED_ plasma membrane ATPase-like [Elaeis guineensis]*																				
TRINITY_DN18901_c0_g1	0.2	0.23	0.13	0.61	0.36	0.29	0.29	0.2	0.14	5.19	4.69	3.97	0.28	0.13	0.2	1.55	0.38	0.47	0.37	0.03	0.65
TRINITY_DN34220_c0_g1	9.14	7.73	8.05	10.23	11.59	3.9	7.16	7.94	8.81	90.01	108.4	60.92	8.91	4.15	8.31	9.27	13.25	14.67	12.67	5.03	14.24
TRINITY_DN35114_c0_g1	0.44	0.52	3.01	7.46	6.79	18.41	14.95	9.99	7.33	24.98	25.98	29.05	14.36	6.4	16.18	7.09	5.57	10.8	3.71	1.66	3.89
TRINITY_DN35114_c0_g2	47.94	30	37.95	51.95	27.51	50.33	94.72	42.39	43.61	137.7	96.59	118.3	30.75	20.37	23.83	9.86	24.63	13.37	0.75	0.49	1.46
TRINITY_DN35822_c0_g1	32.35	54.44	82.12	67.31	99.8	83.44	161.5	101.4	79.13	335.9	285.9	295.4	24.15	8.31	21.55	20.68	18.93	18.03	69.5	1094	28.55
TRINITY_DN35822_c0_g2	54.1	51.64	46.02	88.4	83.53	32.69	114.5	157.2	116	108	108	107.4	56.6	30.75	55.67	72.5	66.52	80.69	40.13	20.23	17.51
OG0000028	*PREDICTED_ GPCR-type G protein COLD1 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN36198_c0_g1	38.33	42.64	35.1	47.11	51.79	54.88	47.69	66.16	65.85	56.56	47.68	49.81	33.57	13.28	31.73	40.53	28.68	39.13	26.7	33.84	34.87
OG0000034	*PREDICTED_ respirATory burst oxidase homolog protein E isoform X2 [Elaeis guineensis]*																				
TRINITY_DN24271_c0_g1	0	0	0	0	0.18	0.19	3.63	0.67	1.63	21.93	6.97	9.34	0.62	0.04	0.03	0.65	0	0	0.21	0.05	2.79
TRINITY_DN28914_c1_g1	4.82	0.13	0.15	9.53	10.21	0.86	20.98	5.03	6.9	28.84	19.74	20.62	5.19	0.95	1.2	6.2	20.7	27.07	0.33	0.02	0.45
TRINITY_DN32025_c0_g1	3.14	1.34	1.38	7.09	6.02	2.97	9.34	2.04	5.59	21.25	10.04	17.08	2.49	0.36	0.71	1.95	8.99	5.25	0.18	0.06	0.33
TRINITY_DN35726_c1_g1	3.08	1.35	3.17	6.99	26.05	8.77	37.93	18.98	21.06	22.17	14.2	54.9	22.73	4.6	6.9	12.7	14.04	15.97	174.5	40.9	224.37
TRINITY_DN35726_c1_g2	9.66	7.84	4.76	98.55	139.8	15.26	63.51	30.83	45.6	85.02	58.09	76.95	17.61	13.28	25.72	27.44	72.6	53.33	11.72	1.89	226.72
OG0000036	*PREDICTED_ putATive chloride channel-like protein CLC-g isoform X1 [Elaeis guineensis]*																				
TRINITY_DN23259_c0_g1	50.91	24.65	15.24	128.1	79.61	97.4	107.3	77.62	135.9	32.77	30.39	43.9	37.62	17.57	24.71	40.14	54.76	42.49	52.31	9.93	54.74
TRINITY_DN27635_c0_g1	7.38	10.68	11.58	5.88	10.39	17.63	4.92	9.7	13.59	9.22	11.93	15.89	3.53	3.65	4.17	3.06	5.85	5.54	9.32	12.15	2.95

TRINITY_DN27648_c0_g1	0.45	2.01	0.76	1.06	2.86	0.75	4.71	2.13	18.42	10.71	4.72	9.2	11.83	11.33	10.78	5.78	9.94	5.05	1.23	0.31	0.92
TRINITY_DN27648_c1_g1	**65.36**	**71.26**	**28.18**	**61.15**	**30.83**	**22.91**	**29.33**	**20.9**	17.17	**36.61**	**32.8**	**45.8**	11.68	6.7	16.22	8.46	11.87	15.16	4.28	**26.9**	8.83
TRINITY_DN33770_c0_g1	0.73	0.27	0.12	4.54	2.09	1.33	2.87	0.42	0.97	8.04	3.59	2.16	0.04	0.06	0.07	0.17	0.79	0.2	0	0.03	0
OG0000041	*PREDICTED_ S-type anion channel SLAH2-like isoform X1 [Elaeis guineensis]*																				
TRINITY_DN30378_c0_g2	0.15	0.27	0.45	0.25	0.39	3.01	0.74	0.57	0.35	7.43	1.02	2.38	**48.52**	2.69	9.6	3.8	1.54	1.19	3.12	**30.28**	0.39
TRINITY_DN31821_c0_g1	3.29	5.77	3.33	3.37	1.82	2.21	4.22	6.83	12.61	9.9	11.45	15.67	14.34	3.35	5.21	**56.8**	**25.87**	11.41	3.29	0.17	0.38
OG0000043	*PREDICTED_ aquaporin PIP2-4 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN24828_c0_g1	0	0	0	0	0	0	0	0	0	3.97	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN35053_c3_g1	**81.74**	**58.4**	**64.05**	**175.3**	**131.1**	**96.5**	**687.9**	**207.5**	**350**	**1047**	**645.5**	**423.5**	**260.4**	**109.5**	**181.8**	**145.1**	**163**	**134.1**	**82.58**	5.04	**94.31**
OG0000046	*PREDICTED_ protein argonaute 1A-like [Elaeis guineensis]*																				
TRINITY_DN27842_c0_g1	0.54	0.93	0.28	2.16	1.24	0.22	4.03	**21.18**	**55.53**	8.09	6.21	3.96	**38.75**	**65.14**	**98.46**	**25.47**	**45.01**	**45.88**	0.05	0	0.06
TRINITY_DN34013_c0_g1	0.39	0.81	0.89	3.97	3.67	0.62	2.89	8.52	**24.57**	9.33	17	13.28	**25.78**	**34.82**	**53.96**	6.24	**23.39**	15.3	0.23	0	0.25
TRINITY_DN35016_c0_g1	3.38	8.56	5.56	8.09	11.4	14.39	11.02	8.49	8.74	17.06	**26.25**	**24.11**	**24.07**	11.22	16.93	12.61	**22.07**	10.8	**21.48**	0.28	**89.13**
TRINITY_DN35738_c0_g1	**37.76**	19.95	**21.98**	**45.18**	**51.76**	18.03	**39.77**	**62.26**	**73.74**	**56.14**	**60.1**	**54.62**	**51.13**	**41.88**	**61.47**	**70.4**	**83.31**	**68.01**	**45.91**	14.65	**48.65**
OG0000052	*PREDICTED_ transcription factor MYC2-like [Elaeis guineensis]*																				
TRINITY_DN2149_c0_g1	**59.76**	**70.8**	**140.2**	**203.8**	**411.9**	**226.5**	**291.6**	**320.7**	**291.7**	**193.5**	**191.6**	**359.5**	**112.1**	**61.16**	**117.4**	**66.73**	**85.63**	**94.12**	**142.2**	**71.22**	**151.61**
TRINITY_DN22244_c0_g1	9.4	7.69	13.43	15.29	19.9	2.15	**40.99**	**29.69**	**21.82**	**24.01**	11.87	19.3	9.12	4.11	11.43	2.44	5.09	3.31	1.74	9.73	0.21
TRINITY_DN27638_c0_g2	15.04	**25.75**	**24.23**	**54.52**	**67.3**	**47.32**	**81.13**	**73.65**	**65.33**	**63.77**	**64.97**	**119**	**48.05**	**32.43**	**54.24**	12.25	**26.62**	**22.67**	**28.92**	**116.1**	17.47
TRINITY_DN31421_c0_g1	12.67	10.02	**20.11**	**25.6**	**29.51**	7.25	**66.64**	**50.58**	**29.66**	**32.05**	15.92	**32.24**	**22.28**	14.71	**42**	4.67	11.36	8.6	3.01	14.9	1.18
TRINITY_DN31421_c0_g2	4.69	9.59	8.73	8.05	**29.74**	12	**38.9**	**23.88**	**23.35**	**34.09**	16.43	**52.39**	14.41	5.4	**27.18**	5.04	9.99	8.69	**66.98**	13.63	4.6
TRINITY_DN31421_c1_g1	**27.7**	**37.59**	**71.91**	**100.1**	**187.1**	**47.34**	**146.3**	**154**	**154.9**	**101.8**	**105.6**	**176.8**	**39.42**	19.36	**34.71**	**21.49**	**31.66**	**30.45**	**66.62**	**53.47**	**74.34**
TRINITY_DN76297_c0_g1	0	0	0	0	0	0	0	0	0	3.51	0	0	0	0	0	0	0	0	0	0	0
OG0000053	*PREDICTED_ phosphoinositide phospholipase C 2 [Elaeis guineensis]*																				
TRINITY_DN35011_c0_g1	**28.42**	**28.3**	**23.72**	**48.24**	**27.1**	**30.84**	12.28	11.04	18.16	**36.84**	**22.95**	**29.07**	**39.07**	19.43	**24.81**	**27.57**	17.68	14.18	17.74	1.81	**43.48**
TRINITY_DN35011_c0_g2	0.53	0.45	0.22	1.35	2.57	1.11	5.39	8.28	7.67	15.8	**22.58**	16.4	13.11	8.3	14.09	9.43	7.28	7.92	0.16	0.81	0.22
TRINITY_DN35011_c0_g3	**89.99**	**87.71**	**94.69**	**88.16**	**85.4**	**118**	**29.55**	**40.29**	**48.07**	**55.92**	**50.97**	**55.04**	**58.16**	19.32	**38.67**	**51.29**	**45.12**	**35.6**	**26.51**	18.93	**33.81**
OG0000054	*PREDICTED_ delta-1-pyrroline-5-carboxylATe synthase [Elaeis guineensis]*																				
TRINITY_DN35030_c1_g1	**27.92**	**30.02**	**25.19**	**30.19**	**36.95**	19.53	**23.81**	**33.72**	**33.29**	**34.98**	**46**	**39.14**	**32.16**	**22.08**	**41.6**	**37.62**	**28.32**	**26.9**	**50.88**	**37.59**	**41.83**
OG0000057	*PREDICTED_ bZIP transcription factor 53 [Elaeis guineensis]*																				

TRINITY_DN21593_c0_g1	8.43	17.96	17	24.75	43.4	26.41	101.8	116	65.77	77.97	96.16	118.2	50.25	20.54	38.83	27.95	33.27	24.54	220.6	13.92	216.9
TRINITY_DN28492_c0_g1	38.76	33.59	37.41	72.72	93.95	11.23	247.1	308.4	366.8	97.06	156.8	147.6	140.6	70.23	105	121.1	173.9	109.1	40.23	53.76	95.59
TRINITY_DN28492_c0_g2	0.5	0.2	0.31	2.08	2.23	2.34	3.25	10.74	28.01	14.78	13.49	7.24	10.64	12.46	28.41	14.21	32.32	25.75	0	0	0
TRINITY_DN63679_c0_g1	27.93	78.45	35.36	52.31	56.33	57.03	94.97	97.26	33.42	116.6	111.1	173.4	45.2	11.86	30.71	141	70.8	44.97	127.9	7.51	125.49
TRINITY_DN82183_c0_g1	0	0	0	0	0	0	0	0	0	2.76	0	0	0	0	0	0	0	0	0	0	0
OG0000063	*PREDICTED_ homeobox-leucine zipper protein ROC8-like isoform X2 [Elaeis guineensis]*																				
TRINITY_DN27172_c0_g1	70.54	61.06	69.77	47.02	83.25	14.44	10.59	20.95	37.84	5.78	8.7	6.32	44.92	26.24	55.74	46.89	60.77	54.83	0.46	0.18	0.06
OG0000064	*PREDICTED_ probable aquaporin PIP1-2 [Elaeis guineensis]*																				
TRINITY_DN29280_c0_g2	0	0	0	0	0	0	0	0	0	9.06	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN35003_c1_g1	533.2	416.3	326.3	856.9	868.4	268.8	1333	959.3	1095	1039	759	651.6	1465	832.8	1480	1195	985.4	956.2	338.7	233.8	249.9
OG0000065	*PREDICTED_ transcription repressor MYB5-like [Elaeis guineensis]*																				
TRINITY_DN18669_c0_g1	4.09	0.19	0.56	37.39	56.19	1.14	22.71	14.1	19.57	28.51	15.1	15.37	4.55	0.62	2.26	6.6	5.51	4.87	0.79	0	5.67
TRINITY_DN22101_c0_g1	5.29	0.1	0	24.59	12.89	2.42	7.94	1.71	1.39	3.37	0.11	0.91	1.73	0.51	1.39	3.52	1.2	1.98	0	0.22	1.87
TRINITY_DN24569_c0_g1	0	0	0.42	7.79	8.44	41.11	11.16	9.19	3.69	4.4	5.42	7.03	1.76	0.78	2.38	0.22	0.67	2.63	0	0	0
TRINITY_DN26933_c0_g1	0	0.69	0	0.11	0.1	1.72	1.86	0.39	1.57	13.17	3.5	7.19	0	0	0	0	0	0	0	0	0
TRINITY_DN35293_c0_g1	0	0	0	0.23	2.37	0.52	15.8	4.88	3.01	39.32	15.63	16.7	0.1	0	0	0	0	0	0	0	0
TRINITY_DN6643_c0_g1	0	0	0	2.49	5.74	2.44	6.34	4.35	6.24	5.17	5.84	3.07	9.89	3.72	4.96	13.58	3.38	4.15	0.32	0	0
TRINITY_DN78240_c0_g1	0	0.27	0	0.65	0.57	0.24	1.8	0.46	0.74	7.28	1.18	2.25	0	0	0	0	0	0	0	0	0
OG0000069	*PREDICTED_ bZIP transcription factor TRAB1 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN17434_c0_g1	8.39	12.84	11.52	17.99	16.11	5.3	10.87	16.37	12.26	13.69	20.13	14.05	2.31	1.86	2.86	2.56	2.16	2.76	5.03	24.28	3.11
TRINITY_DN22238_c0_g1	22.01	23.21	13.76	43.85	36.29	10.2	31.36	43.63	29.96	39.78	48.46	38.6	6.69	3.64	6.7	9.06	8.52	10.64	14.81	29.82	13.91
TRINITY_DN28034_c0_g1	26.14	35.05	33.16	47.6	47.19	52.42	33.28	42.21	34.91	41.26	54.04	48	14.55	12.49	24.95	18.46	17.24	18.33	13.02	51.33	14.65
OG0000090	*PREDICTED_ abscisic acid 8'-hydroxylase 1 [Elaeis guineensis]*																				
TRINITY_DN32057_c0_g1	30.94	36.49	43.66	15.67	28.63	107.5	11.51	4.19	21.88	35.52	12.11	98	4.04	0.96	2.08	3.08	0.9	0.42	542.6	23.78	166.9
OG0000094	*PREDICTED_ uncharacterized protein LOC109504732 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN15141_c0_g1	24	32.91	21.85	32.21	36.22	36.66	24.47	32.35	27.27	30.18	33.76	31.1	21.36	15.48	27.98	25.1	23.76	23.87	28.66	23.29	26.32
TRINITY_DN35068_c2_g1	57.88	10.07	5.8	66	14.85	11.14	41.95	15.28	11.53	54.08	16.54	10.44	29.13	17.13	24.28	25.64	32.92	26.95	38.86	4.28	8.84
OG0000107	*PREDICTED_ CBL-interacting protein kinase 18-like [Elaeis guineensis]*																				
TRINITY_DN16962_c0_g1	71.85	30.13	32.1	506	202.8	110.5	373.2	291.4	303.8	101	67.17	76.46	58.59	36.91	49.74	34.74	53.71	31.44	215.3	42	92.26

TRINITY_DN35781_c0_g1	**45.79**	**57.24**	**42.07**	**151.6**	**104**	**62.11**	**130.4**	**120.9**	**151.7**	**54.85**	**67.06**	**81.99**	**186.9**	**42.28**	**66.37**	**103.3**	**68.12**	**38.37**	**258.6**	**319.5**	**109.85**
TRINITY_DN35781_c0_g6	**76.75**	**86.17**	**95.92**	**223.9**	**232**	**220**	**305.1**	**280.2**	**316**	**221.9**	**255.4**	**341.6**	**119.9**	**67.51**	**139.9**	**218.7**	**120**	**115.1**	**287.8**	**52.14**	**360.2**
OG0000110	*PREDICTED_ protein ECERIFERUM 3-like isoform X2 [Elaeis guineensis]*																				
TRINITY_DN45525_c0_g1	**73.89**	**37.23**	**32.04**	**140.9**	**117**	12.03	6.27	9.71	19.02	2.78	4.51	2.65	16.25	13.4	**44.94**	**313.8**	**78.33**	**88.42**	0	0	0
TRINITY_DN748_c0_g1	**97.79**	**38.29**	**45.03**	**177.8**	**177**	**26.44**	5.03	6.97	13.59	3.76	4.09	4.65	**28.87**	19.78	**71.43**	**522.2**	**109.9**	**126.9**	0	0	0
TRINITY_DN81466_c0_g1	**116.9**	**51.93**	**53.24**	**252.8**	**263.6**	**70.86**	5.23	8.76	11.03	2.6	5.98	3.92	**53.47**	**37.08**	**148.9**	**1005**	**182.8**	**233.6**	0	0	0
OG0000112	*PREDICTED_ ethylene-responsive transcription factor ERF003-like [Elaeis guineensis]*																				
TRINITY_DN33317_c0_g1	3.64	1.33	0	13.3	14.67	1.41	**22.82**	**32.88**	6.69	10.11	9.82	4.69	17.23	1.34	12.77	**35.12**	6.27	10.62	**342.7**	**236.3**	**309.73**
OG0000117	*PREDICTED_ ABC transporter G family member 22 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN18534_c0_g1	**115.7**	**191.4**	**125.8**	**55.33**	**46.91**	**34.69**	4.5	4.55	5.17	7.54	2.76	9.39	8.42	6.08	9.88	7.66	17.13	14.76	0.45	**72.83**	0.25
OG0000119	*PREDICTED_ LRR receptor-like serine/threonine-protein kinase RPK2 [Elaeis guineensis]*																				
TRINITY_DN26288_c0_g1	0.32	0.06	0	3.27	4.91	0.13	8.07	10.98	13.72	9.94	9.99	7.26	11.37	5	7.82	4.83	5.06	4.95	1.53	0.07	0.87
TRINITY_DN35712_c0_g1	5.23	6.14	8.5	12.54	14.14	8.82	9.22	10.34	11.94	12.69	11.89	16.74	10.94	4.22	9.65	4.73	6.76	7.21	7.83	7.1	7.61
TRINITY_DN47823_c0_g1	2.78	1.78	3.51	8.47	10.39	9.26	7.54	8.35	11.38	11.12	10.64	12.91	12.41	8.28	11.07	7.67	10.56	6.62	3.63	1.79	2.09
TRINITY_DN81687_c0_g1	1.16	0.85	1.69	3.1	4.17	0.41	3.61	4.22	5.72	4.66	5.06	5.25	1.41	0.87	1.45	0.8	2.56	1.64	2.17	2.29	0.71
OG0000123	*PREDICTED_ heat stress transcription factor A-1 [Elaeis guineensis]*																				
TRINITY_DN27166_c0_g1	5.05	6.42	8.69	**43.11**	**57.59**	18.22	**55.37**	**78.86**	**70.72**	**38.51**	**50.62**	**43.77**	17.23	13.92	**24.68**	**22.13**	15.67	15.08	**90.87**	**46.17**	**82.49**
TRINITY_DN32888_c4_g4	7.98	13.04	6.88	18.52	19.3	**21.9**	15.75	**27.5**	14.78	13.91	19.33	**21.64**	11.08	5.77	11.86	8.95	10.79	19.34	8.42	11.09	11.3
TRINITY_DN32888_c4_g5	5.92	8.3	8.39	10.37	12.46	5.93	10.58	15.24	10.72	7.5	10.35	11.33	5.72	3.72	5.56	5.38	6.2	2.69	5.53	13.51	5.18
OG0000138	*PREDICTED_ small nuclear ribonucleoprotein E [Elaeis guineensis]*																				
TRINITY_DN28076_c0_g2	**31.95**	**44.31**	**43.44**	**42.9**	**58.49**	**65.19**	**43.69**	**56.04**	**58.67**	**38.12**	**41.57**	**35.53**	**113.3**	**78.13**	**139.5**	**98.26**	**62.23**	**54.78**	**53.04**	**78.41**	**47.17**
OG0000139	*PREDICTED_ probable phospholipid hydroperoxide glutathione peroxidase [Elaeis guineensis]*																				
TRINITY_DN20565_c0_g1	9.49	10.84	11.79	**23.38**	10.78	**20.43**	**22.05**	12.88	17.77	19.15	10.23	9	13.53	7.36	9.09	14.11	7.72	4.95	**22.85**	6.39	**47.95**
TRINITY_DN28768_c0_g1	**141.6**	**151.8**	**136.6**	**132.5**	**137.9**	**119.6**	**91.89**	**85.46**	**82.45**	**110**	**73.38**	**64.24**	**55.01**	**23.32**	**74.14**	**82.71**	**48.34**	**53.46**	**220.2**	**379.9**	**230.99**
OG0000145	*PREDICTED_ zinc finger A20 and AN1 domain-containing stress-associated protein 8 [Elaeis guineensis]*																				
TRINITY_DN23122_c0_g1	**37.38**	**51.25**	**43.09**	**80.45**	**80.29**	**60.44**	**102.7**	**100.1**	**120.6**	**64.89**	**66.4**	**86.38**	**64.59**	**29.64**	**48.73**	**78.17**	**59.85**	**44.62**	**190.5**	**188.6**	**150.57**
TRINITY_DN27472_c0_g1	**107.1**	**184.8**	**124.7**	**283.7**	**341.6**	**353.1**	**360.5**	**367.3**	**389.5**	**224.4**	**242.7**	**281.2**	**267.3**	**173.3**	**311.5**	**360.3**	**210.3**	**216.3**	**431.8**	**149.9**	**458.57**
OG0000149	*PREDICTED_ 1-aminocyclopropane-1-carboxylate synthase-like [Elaeis guineensis]*																				
TRINITY_DN12825_c0_g1	0	0.38	0.68	0.73	1.67	3	0.36	0.32	0.17	2.94	0.1	3.27	3.94	0.16	0.31	0	0.92	0.92	0	2.32	0

TRINITY_DN56910_c0_g1	0	0.27	0.05	0	0	0.08	0	0	0	4.87	0.24	1.02	0	0	0	0	0.1	0	2.36	**219.8**	0.45
OG0000156	*PREDICTED_ CBL-interacting protein kinase 32-like isoform X1 [Elaeis guineensis]*																				
TRINITY_DN34658_c0_g1	**87.94**	**67.37**	**48**	**81.91**	**75.14**	**70.14**	**48.61**	**47.47**	**54.36**	**87.41**	**102.4**	**74.3**	**28.17**	19.66	**23.81**	**26.16**	**45.34**	**28.98**	18.7	18.74	**25.37**
TRINITY_DN35028_c0_g1	**81.39**	**89.49**	**62.45**	**243**	**256.6**	**133.4**	**87.24**	**101.2**	**112.6**	**138.7**	**105.1**	**127.9**	**334.7**	**121.2**	**223.8**	**275.9**	**142.3**	**143.7**	**370.2**	**312.2**	**349.74**
TRINITY_DN35373_c0_g1	**107.2**	**101.9**	**50.86**	**98.07**	**134**	**37.53**	**82.13**	**88.29**	**75.48**	**103.3**	**59.44**	**61.18**	**61.1**	**70.55**	**79.84**	**58.51**	**89.72**	**55.87**	**59.64**	**32.01**	**105.48**
OG0000187	*PREDICTED_ LOW-QUALITY PROTEIN_ indole-3-acetaldehyde oxidase-like [Elaeis guineensis]*																				
TRINITY_DN29615_c0_g1	11.6	**23.38**	**25.21**	**25.88**	**21.8**	**49.16**	11.76	11	6.62	**38.68**	**44.63**	**45.2**	3.71	3.54	7.47	4.29	4.87	2.87	10.37	1.46	**24.81**
OG0000207	*PREDICTED_ histone deacetylase HDT2 [Elaeis guineensis]*																				
TRINITY_DN34680_c0_g1	**29.6**	**49.98**	**32.44**	**51.89**	**82.93**	19.02	**97.63**	**173.1**	**93.76**	**90.54**	**117.5**	**61.96**	**122.6**	**97.48**	**125.9**	**97.78**	**53.67**	**45.11**	**172.3**	**84.91**	**222.33**
TRINITY_DN34680_c0_g2	**39.53**	**82.72**	**25.64**	**81.74**	**150.1**	**38.19**	**179.4**	**333.4**	**167.2**	**150.2**	**202.6**	**103.8**	**240**	**176**	**237.9**	**160.1**	**116.3**	**98.67**	**246.9**	**70.25**	**461.53**
OG0000209	*PREDICTED_ ethylene-responsive transcription factor ERF027-like [Elaeis guineensis]*																				
TRINITY_DN24582_c0_g1	0	0	0.12	0	2.66	0.32	1.54	0	2.89	10.63	0.78	1.45	3.8	0.45	2.92	1.25	0.71	0.12	0.26	0	0.49
TRINITY_DN80402_c0_g1	1.3	2.84	18.87	0.45	**24.6**	**80.04**	2.01	0.53	2.14	3.64	0	**21.14**	7.27	0	0.42	0.98	0	0	8.05	0.5	1.08
OG0000213	*PREDICTED_ G-type lectin S-receptor-like serine/threonine-protein kinase AT5g24080 [Elaeis guineensis]*																				
TRINITY_DN19269_c1_g1	0.02	0	0.04	0	0	0.05	0.17	0.04	0.02	15.91	**24.22**	18.22	0	0.03	0	0.37	0	0	0	0	0.04
TRINITY_DN32893_c0_g2	0.43	0.1	0.05	1.25	0.52	0.3	1.22	0.22	1.57	11.25	3.64	**22.74**	0.06	0.04	0.04	0	0.05	0.41	0	3.04	0.05
OG0000226	*PREDICTED_ glutathione S-transferase U17-like [Elaeis guineensis]*																				
TRINITY_DN35725_c2_g1	**466**	**448.3**	**369.9**	**260.8**	**161.5**	**150.5**	**80.94**	**65.04**	**55.67**	**37.09**	18.67	**28.55**	**59.76**	**28.36**	**51.74**	**155.2**	**67.85**	**55.37**	1.05	1	6.95
TRINITY_DN62888_c0_g1	2.61	8.36	4.06	4.11	10.46	13.43	**34.31**	10.92	6.81	**66.22**	**27.06**	**26.38**	1.46	1.21	4.57	0.73	0.95	1.27	9.15	**170.3**	**204.56**
OG0000262	*PREDICTED_ protein FAR-RED IMPAIRED RESPONSE 1-like [Elaeis guineensis]*																				
TRINITY_DN31063_c0_g1	7.75	6.13	6.91	7.86	8.63	6.32	5.66	9.06	7.74	6.43	8.11	8.97	8.26	3.58	5.34	5.74	5.19	3.92	7.43	7.53	5.65
OG0000277	*PREDICTED_ NADP-dependent malic enzyme isoform X2 [Elaeis guineensis]*																				
TRINITY_DN27748_c0_g1	**160.5**	**304.1**	**396**	**189.8**	**129.6**	**183.5**	**106**	**72.54**	**69.69**	**185.3**	**205.6**	**233.5**	**55**	**39.05**	**63.42**	**57.44**	**65.98**	**59.77**	**193.8**	**230.6**	**148.67**
TRINITY_DN27852_c0_g1	**166.7**	**69.32**	**84.76**	**129.4**	**96.22**	**51.12**	**160.9**	**90.49**	**53.06**	**143.3**	**94.37**	**74.3**	**306.4**	**140.3**	**273.3**	**229.2**	**182.9**	**105.5**	**109.4**	**131.8**	**487.48**
OG0000282	*PREDICTED_ protein farnesyltransferase subunit beta isoform X3 [Elaeis guineensis]*																				
TRINITY_DN33437_c0_g1	**21.21**	**24.36**	16.79	18.21	18.82	**20.61**	11.55	16.07	14.35	17.46	**23.41**	18.33	**36.57**	**20.31**	**30.58**	18.7	16.94	14.11	**23.72**	13.31	16.47
OG0000305	*PREDICTED_ syntaxin-61 [Elaeis guineensis]*																				
TRINITY_DN20536_c0_g2	5.05	7.73	6.59	9.08	11.19	9.58	8.26	14.28	8.97	10.27	13.21	13.98	6.34	4.88	7.16	7.95	7.71	11.88	**21.37**	**33.7**	11.18
TRINITY_DN33913_c0_g1	**25.43**	**25.28**	**23.7**	**48.49**	**61.63**	**57.88**	**79.77**	**76.83**	**69.62**	**66.27**	**57.59**	**49.06**	**72.18**	**43.93**	**76.82**	**91.1**	**89.34**	**85**	**69.2**	**52.25**	**106.3**

OG0000309	*PREDICTED_ pyrophosphate-energized vacuolar membrane proton pump [Elaeis guineensis]*																				
TRINITY_DN35773_c1_g1	77.89	76.42	48.14	128.4	138.3	85.36	275	235.2	253.1	223.7	181.9	174	114.8	43.36	59.17	120.8	103.9	61.6	161.9	65.13	78.72
OG0000312	*PREDICTED_ probable protein phosphatase 2C 59 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN33850_c0_g1	54.66	66.61	63.87	49.02	44.29	48.75	26.12	22.34	26.26	30.88	33.93	36.64	16.88	9.89	17.47	26.55	27.97	24.88	6.92	10.26	12.98
TRINITY_DN34401_c0_g1	152	163.5	119.1	210.1	206.3	157.5	94.45	123.3	114.9	120.5	115.3	114.4	82.33	52.28	96.8	173.1	156.8	130.9	118.3	141.5	166.07
OG0000315	*PREDICTED_ transcription factor MYB108-like [Elaeis guineensis]*																				
TRINITY_DN20129_c0_g1	0.95	6.58	0.76	0.93	2.96	2.36	13.15	22.81	15.82	7.2	2.56	7.39	4.27	2.16	4.27	2.7	6.57	6.11	0.18	0	0
TRINITY_DN28160_c0_g1	0.1	0.43	0.67	2.21	4.26	0.99	5.36	1.19	0.86	12.97	2.69	19	0.56	0	0.05	0.06	0	0	2.66	3.31	0
TRINITY_DN9553_c0_g1	1.3	4.47	5.58	0.73	4.45	6.14	4.26	2.37	0.89	6.74	0.83	3.86	8.48	0	2.61	1.02	0	0.54	1.13	120.2	0.32
OG0000318	*PREDICTED_ histone deacetylase 19-like [Elaeis guineensis]*																				
TRINITY_DN35795_c1_g1	33.13	30.56	40.85	35.17	25.34	27.19	28.89	29.85	39.75	43.81	37.91	42.17	41.68	21.48	29.32	31.01	32.1	26.59	115.8	187.3	82.24
OG0000344	*PREDICTED_ dehydration-responsive element-binding protein 2B isoform X1 [Elaeis guineensis]*																				
TRINITY_DN18242_c0_g1	11.75	28.94	26.14	18.37	12.71	29.8	15.83	10.08	13.05	23.1	22.01	24.54	13.08	5.84	11.26	16.95	8.09	7.38	30.33	134.3	45.16
TRINITY_DN34839_c0_g1	31.45	39.1	49.35	52.61	53.99	68.01	43.23	39.04	65.17	36.41	39.53	77.29	49.37	23.3	27.06	41.76	28.82	23.98	35.44	73.44	28.78
OG0000348	*PREDICTED_ galactinol synthase 2-like [Elaeis guineensis]*																				
TRINITY_DN34784_c4_g1	53.6	82.77	59.12	24.21	16.25	25.8	12.58	6.94	8.18	75.85	60.09	50.37	20.19	2.77	5.95	10.08	6.78	4.08	0.26	0.59	0.93
OG0000357	*PREDICTED_ uncharacterized protein LOC105045662 [Elaeis guineensis]*																				
TRINITY_DN33626_c0_g1	10.74	11.56	10.25	14.89	17.21	16.39	11.11	9.83	14.1	16.95	16.88	21.46	12.33	7.58	11.9	10.69	13.84	23.92	11.61	16.34	13.12
OG0000363	*PREDICTED_ homeobox-leucine zipper protein HOX8 [Elaeis guineensis]*																				
TRINITY_DN30488_c0_g1	0.06	0.41	0.28	15.33	28.11	0.31	39.95	39.4	28.48	18.54	44.28	35.66	39.71	30.39	109.8	28.67	17.92	20.68	0.46	1.51	0.6
TRINITY_DN30488_c0_g2	68.11	58.87	43.38	202.1	212	112.3	227.8	216.9	130.7	151.9	235	228.2	35.46	43.9	64.2	28.41	57.55	19.61	0.74	0.15	25.32
OG0000371	*PREDICTED_ probable E3 ubiquitin-protein ligase XERICO [Elaeis guineensis]*																				
TRINITY_DN73320_c0_g1	0	0	0	0	0	0	0	0	0	2.45	0	0	0	0	0	0	0	0	0	0	0
OG0000410	*PREDICTED_ potassium channel KOR1 [Elaeis guineensis]*																				
TRINITY_DN7448_c0_g1	11.98	12.12	8.76	16.78	7.65	33.05	4.12	2.85	2.5	5.08	4.9	7.88	1.68	1.82	2.05	0.69	4.82	19.37	0.58	0.24	5.54
OG0000412	*PREDICTED_ zinc finger protein 1-like [Elaeis guineensis]*																				
TRINITY_DN17984_c0_g1	0	0	0	0	0	0	0	0	0	11.65	0	0	0	0	0	0	0	0	0	0	0
TRINITY_DN31498_c0_g1	21.79	100.1	232.6	86.4	372.2	608.4	125	31.46	49.77	255	75.18	618.6	78.85	11.27	22.64	16.53	29.55	5.04	243.8	88.61	92.07
TRINITY_DN76639_c0_g1	0	0	0	0	0	0	0	0	0	7.02	0	0	0	0	0	0	0	0	0	0	0

OG0000420	*PREDICTED_ UDP-glucose 4-epimerase GEPI48 [Elaeis guineensis]*																				
TRINITY_DN33433_c0_g1	**26.44**	**27**	**25.67**	**24.69**	16.85	**24.6**	17.48	7.17	9.97	**41.11**	**22.8**	**22.6**	15.29	5.28	9.02	**32.98**	16.64	12.49	**28.4**	**27.67**	11.09
OG0000438	*PREDICTED_ 9-cis-epoxycarotenoid dioxygenase_ chloroplastic [Elaeis guineensis]*																				
TRINITY_DN28946_c0_g1	4.91	6.74	3.84	7.48	2.15	0.87	3.81	6.09	2.5	7.66	10.95	10.05	1.3	0.26	0.6	0.42	0.29	1.93	0.39	**58.62**	1.19
TRINITY_DN30486_c0_g1	2.59	4.81	4.7	5.62	1.52	0.22	2.61	4.48	1.37	6.77	9.75	8.1	1.22	0.26	0.28	0.2	0.34	0.55	0.7	**45.85**	0.38
OG0000466	*PREDICTED_ probable calcium-binding protein CML31 [Elaeis guineensis]*																				
TRINITY_DN47677_c0_g1	**37.44**	**164.7**	**379.9**	**87.01**	**319.7**	**1147**	**69.04**	**23.9**	**42.48**	**141.8**	**33.77**	**351.2**	**54.93**	7.68	11.22	13.05	7.28	0.68	**56.99**	**176.7**	**216.9**
TRINITY_DN65773_c0_g1	12.12	**54.73**	**60.6**	9.19	**76.48**	**149**	5.76	1.59	4.05	**29.6**	2.88	**87.07**	12.83	0.48	1.97	2.28	0.41	0	4.86	16.73	4.44
OG0000476	*PREDICTED_ potassium channel K AT3-like [Elaeis guineensis]*																				
TRINITY_DN32477_c2_g1	7.25	6.91	7.91	9.3	10.55	2.55	**29.52**	**21.1**	15.03	**29.1**	**28.46**	**45.69**	1.54	3.11	3.79	1.16	4.38	6.36	0.04	0.52	0.8
OG0000484	*PREDICTED_ annexin D2 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN33177_c0_g1	9.98	5.27	6.58	19.09	**45.4**	8.53	**779**	**1191**	**726.5**	**145.6**	**167.3**	**102.9**	**714.8**	**287.4**	**538.3**	**610.3**	**275.1**	**340.8**	**203.6**	2.79	**105.46**
TRINITY_DN33177_c0_g4	**84.07**	**25.06**	**21.71**	**209.3**	**473.1**	**27.74**	**410.2**	**708.4**	**757.5**	**189.1**	**341.1**	**156.9**	**353**	**183.1**	**377.7**	**256.2**	**176.9**	**306.9**	4.53	4.52	5.5
OG0000501	*PREDICTED_ dehydration-responsive element-binding protein 1G [Elaeis guineensis]*																				
TRINITY_DN23652_c0_g1	0.74	3.19	**29.89**	3.69	**70.05**	**140.2**	8.58	1.62	5.93	12.07	0.54	**82.44**	1.65	0.79	0.27	0.64	0.12	0	**45.91**	0	**33.82**
OG0000559	*PREDICTED_ protein arginine N-methyltransferase 5 [Elaeis guineensis]*																				
TRINITY_DN25497_c0_g1	15.91	19.74	19.14	**25.07**	**27.27**	**23.43**	**27.89**	**35.9**	**34.54**	**34.92**	**42.27**	**39.16**	**33.99**	**29.24**	**39.82**	**23.74**	**32.25**	**29.92**	**30.23**	15.25	**22.88**
OG0000574	*PREDICTED_ probable tyrosine-protein phosphatase AT1g05000 [Elaeis guineensis]*																				
TRINITY_DN27651_c0_g1	6.14	12.94	7.99	2.57	3.75	6.48	5.86	6.16	7.74	5.97	5.28	4.8	**33.89**	9.51	**30.3**	16.57	8.32	15.76	**81.99**	**95.11**	**24.17**
TRINITY_DN27651_c0_g2	6.25	4.33	5.52	6.69	8.63	2.48	6.24	11.28	11.72	6.9	7.45	5.64	**27.49**	15.32	**37.5**	**22.02**	16.54	**53.6**	10.52	8.64	16.11
OG0000580	*PREDICTED_ inositol polyphosphate multikinase beta-like [Elaeis guineensis]*																				
TRINITY_DN24747_c0_g1	6.29	7.83	8.27	10.6	9.05	13.23	12.14	13.18	13.89	10.7	13.36	14.01	10.4	7.13	10.37	6.31	7.37	7.14	14.35	10.21	9.54
OG0000590	*PREDICTED_ aldehyde dehydrogenase family 3 member H1 [Elaeis guineensis]*																				
TRINITY_DN24671_c0_g1	**368**	**499.4**	**449.9**	**139**	**102.9**	**165.8**	**79.62**	**58.88**	**48.69**	**169.1**	**148.8**	**140.9**	**49.74**	**26.1**	**59.82**	**62.03**	**37.32**	17.61	2.69	3.26	9.97
OG0000595	*PREDICTED_ poly [ADP-ribose] polymerase 2 [Elaeis guineensis]*																				
TRINITY_DN36004_c0_g1	1.16	0.87	3.01	0.83	0.88	4.78	2.64	3.04	6.3	3.12	3.36	3.15	9.42	6.93	15.79	4.01	4.57	3.24	0.48	0.66	0.13
OG0000617	*PREDICTED_ annexin D4 isoform X2 [Elaeis guineensis]*																				
TRINITY_DN25175_c0_g1	**50.78**	**27.24**	**22.18**	**137.3**	**110**	15.61	**39.62**	9.51	9.06	**72.97**	15.12	**23.25**	17.79	12.18	**36.16**	**72.09**	**41.45**	**31.17**	**24.69**	12.32	17.66
OG0000623	*PREDICTED_ guanine nucleotide-binding protein alpha-1 subunit-like [Elaeis guineensis]*																				

TRINITY_DN32765_c0_g1	8.7	5.26	4.09	14.14	15.34	9.09	12.8	11.32	20.04	18.28	13.25	16.64	23.15	15.91	20.56	19.29	20.74	14.64	4.15	0.4	5.79
OG0000638	*PREDICTED_ nuclear cap-binding protein subunit 2 [Elaeis guineensis]*																				
TRINITY_DN18003_c0_g1	25.38	28.16	27.2	35.68	42.66	44.72	39.75	59.43	75.5	42.04	38.79	49.78	69.45	47.99	70.21	46.86	49.49	36.98	45.27	32.43	38.15
OG0000646	*PREDICTED_ glycerol kinase [Elaeis guineensis]*																				
TRINITY_DN30116_c0_g1	27.82	51.51	50.73	22	18.6	22.23	13.76	9.4	5.13	84.56	75.3	58.01	11.19	4.54	12.12	17.62	19.91	14.34	7.27	2.86	31.39
OG0000682	*PREDICTED_ histone deacetylase 6-like [Elaeis guineensis]*																				
TRINITY_DN33114_c0_g1	33.7	46.33	36.32	32.4	36.64	31.58	30.59	45.87	30.34	29.96	40.63	39.02	17.8	16.42	26.36	15.78	24.55	40.95	26.06	15.98	31.12
OG0000705	*PREDICTED_ squalene monooxygenase [Elaeis guineensis]*																				
TRINITY_DN28003_c0_g1	87.58	48.58	70.02	63.08	89.68	26.95	60.78	56	97.88	44.12	40.08	38.21	67.97	32.25	54.37	92.81	66.93	67.14	37.29	9.57	49.35
OG0000720	*PREDICTED_ RING-H2 finger protein ATL78-like [Elaeis guineensis]*																				
TRINITY_DN12789_c0_g1	0.91	1.12	2.16	5.49	5.76	31.01	4.88	6.07	9.39	5.55	9.89	19.21	0.86	0	0.62	0.36	0.83	1.04	0.58	0	0.1
OG0000721	*PREDICTED_ ABC transporter G family member 25-like [Elaeis guineensis]*																				
TRINITY_DN34098_c1_g1	5.36	2.05	2.76	13.26	7.67	3.02	18.35	25.08	15.01	17.76	13.65	11.61	15.29	8.21	11.8	7.96	9.48	10.24	1.86	0.4	1.56
OG0000761	*PREDICTED_ double-stranded RNA-binding protein 8 [Elaeis guineensis]*																				
TRINITY_DN15102_c0_g1	9.48	15.8	14.61	14.58	14.43	20.39	12.27	12.45	13.26	12.11	14.06	13.31	16.73	10.1	17.21	13.36	12.34	11.88	12.56	17.63	9.91
OG0000763	*PREDICTED_ protein OVEREXPRESSOR OF CATIONIC PEROXIDASE 3 [Elaeis guineensis]*																				
TRINITY_DN1508_c0_g1	48.38	61.08	65.86	41.32	44.38	45.71	50.09	54.06	37.69	67.67	55.2	54.42	31.04	19.33	39.19	32	28.95	31.8	27.26	45.17	75.39
OG0000796	*PREDICTED_ ethylene-responsive transcription factor ERF017-like [Elaeis guineensis]*																				
TRINITY_DN16697_c0_g1	2.69	9.07	13.25	13.47	46.28	31.65	14.36	4.3	9.54	29.69	5.96	118.4	10.69	0.97	1	1.28	0.98	0.06	2.76	0.13	2.24
TRINITY_DN6144_c0_g1	2.54	2.16	4.22	5.02	21.23	6.35	4.67	1.03	2.66	18.57	3.17	64.12	5.21	2.96	3.88	1.61	0.33	0.45	0.39	0	0.16
OG0000815	*PREDICTED_ poly [ADP-ribose] polymerase 1-like isoform X1 [Elaeis guineensis]*																				
TRINITY_DN35514_c0_g1	20.43	19.35	19.22	20.57	20.99	14.57	14.4	23.43	31.5	23.27	30.18	33.18	34.64	29.37	45.17	25.63	25.14	23.46	12.36	10.88	18.39
OG0000817	*PREDICTED_ 3-ketoacyl-CoA thiolase 2_ peroxisomal [Elaeis guineensis]*																				
TRINITY_DN35269_c1_g2	13.97	12.79	14.51	37.99	44.64	23.72	80.71	24.44	32.21	166.4	111.8	114.5	48.58	22.09	48.84	47.11	43.87	34.52	8.99	12.23	15.11
OG0000847	*PREDICTED_ glutathione S-transferase DHAR2 isoform X1 [Elaeis guineensis]*																				
TRINITY_DN32172_c0_g1	104.9	64.01	70.95	275.6	319.2	206.8	516.5	430.7	272.8	329.6	311.9	271.5	131.9	61.1	127.3	153.9	128.7	172.2	339.3	334.9	317.46
OG0000850	*PREDICTED_ omega-3 fatty acid desaturase_ chloroplastic [Elaeis guineensis]*																				
TRINITY_DN34006_c0_g1	110.4	86.69	169.8	39.43	70.34	80.47	34.87	19.8	15.27	27.4	9.75	61	52.09	15.23	20.58	20.12	14.35	6.91	211.2	54.08	225.15
OG0000856	*PREDICTED_ uncharacterized protein LOC105044901 [Elaeis guineensis]*																				

TRINITY_DN35032_c0_g1	**203**	**243.2**	**169.6**	**147.1**	**134.9**	**182.5**	**58.92**	**89.02**	**57.45**	**52.79**	**76.22**	**74.77**	**52.36**	**32.56**	**60.39**	**91.56**	**68.55**	**66.96**	**88.62**	**87.73**	**98.95**
OG0000868	*PREDICTED_ NAC domain-containing protein 68 [Elaeis guineensis]*																				
TRINITY_DN35062_c5_g3	11.25	**32.78**	**128**	10.75	**38.7**	**151**	**31.83**	11.04	7.25	**57.37**	17.98	**85.83**	**53.91**	0.94	**30.56**	4.45	0.99	0.45	**25.45**	**565.2**	4.46
OG0000919	*PREDICTED_ molybdenum cofactor sulfurase [Elaeis guineensis]*																				
TRINITY_DN29766_c0_g1	7.99	10.63	8.42	11.03	8.4	8.51	4.28	4.38	4.29	14.07	14.53	15.4	5.16	4	5.88	7.03	6.66	6.19	8.82	5.09	5.92
OG0000932	*PREDICTED_ transcription factor MYB44-like [Elaeis guineensis]*																				
TRINITY_DN33427_c1_g2	19.29	**25.98**	**139.4**	**27.83**	**120.5**	**139.8**	**26.82**	14.55	**122.5**	**106.3**	**40.32**	**225.3**	**26.44**	14.82	12.82	11.67	8.71	6.97	**51.17**	**49.99**	**39.95**
OG0000934	*PREDICTED_ omega-6 fatty acid desaturase_ endoplasmic reticulum isozyme 2-like [Elaeis guineensis]*																				
TRINITY_DN29082_c0_g1	0	0	0	0	0	0	0	0	0	12.79	**46.68**	7.56	0	0	0	0	0	0	0	0	0
OG0000963	*PREDICTED_ G-protein coupled receptor 1 [Elaeis guineensis]*																				
TRINITY_DN73492_c0_g1	11.28	8.97	9.59	14.36	19.06	16.04	16.89	14.21	17.37	**21.1**	16.76	18.9	10.09	6.28	10.38	11.62	10.89	10.17	11.08	15.72	14.7
OG0000978	*PREDICTED_ phosphoenolpyruvate carboxykinase [ATP] [Elaeis guineensis]*																				
TRINITY_DN34267_c3_g2	**46.26**	**35.62**	**46.31**	**61.33**	**61.71**	12.76	**156.1**	**362.2**	**190.1**	**174.2**	**293.5**	**101.3**	**94.18**	**57.92**	**119.5**	**178.6**	**268.3**	**254.9**	**29.6**	8.98	10.73
OG0000985	*PREDICTED_ histidine kinase 1-like [Elaeis guineensis]*																				
TRINITY_DN32093_c0_g1	1.09	0.25	0.49	4.81	2.38	1.63	5.52	1.57	1.48	9.72	3.78	2.32	0.72	0.53	0.72	1.99	1.12	0.71	0.05	0.1	0.08
OG0001023	*PREDICTED_ chlorophyll a-b binding protein CP24 10A_ chloroplastic-like [Elaeis guineensis]*																				
TRINITY_DN29717_c0_g2	**3697**	**1437**	**1712**	**1587**	**639.2**	**585.9**	**86.35**	**85.46**	0.76	8.45	3.15	11.05	**41.19**	**22.11**	19.27	**49.26**	**30.94**	12.27	0.06	0	0.55
OG0001046	*PREDICTED_ omega-6 fatty acid desaturase_ chloroplastic isoform X1 [Elaeis guineensis]*																				
TRINITY_DN26075_c0_g1	**351.4**	**333.9**	**275.7**	**278.9**	**233.5**	**145.2**	**100.3**	**161.6**	**72.78**	**50.8**	**71.88**	**70.12**	**41.84**	**21.02**	**47.13**	**49.84**	**47.45**	**55.89**	**120**	**44.22**	**151.9**
OG0001076	*PREDICTED_ zinc finger A20 and AN1 domain-containing stress-associated protein 6-like [Elaeis guineensis]*																				
TRINITY_DN16529_c0_g1	**78.56**	**101.6**	**92.48**	**50.09**	**38.57**	**106.4**	**31.18**	**29.38**	**34.78**	**26.25**	**26.42**	**30.47**	**33.64**	**23.1**	**31.03**	**30.76**	**32.05**	**23.88**	**70.33**	**78.74**	**22.68**
OG0001144	*PREDICTED_ probable purine permease 4_ partial [Elaeis guineensis]*																				
TRINITY_DN20091_c0_g1	0.17	0.23	0.35	3.02	2.01	1.77	7.88	0.87	1.94	9.88	8.8	7.32	1.28	2.57	4.48	1.02	6.57	14.17	0	0	0
OG0001311	*PREDICTED_ abscisic stress-ripening protein 2-like [Elaeis guineensis]*																				
TRINITY_DN27408_c1_g1	**31.62**	**28.11**	**32.19**	**43.49**	**83.2**	**73.13**	**184**	**43.04**	**80.09**	**501.8**	**273.4**	**326.7**	**282**	6.18	**42.02**	**99.47**	**33.38**	12.03	**1736**	**3410**	**226.66**
TRINITY_DN301_c0_g1	0	0	0	0	0	0	0	0	0	8.19	0	0	0	0	0	0	0	0	0	0	0
OG0001372	*PREDICTED_ AP2-like ethylene-responsive transcription factor AT1g79700 [Elaeis guineensis]*																				
TRINITY_DN30070_c0_g1	4.24	1.85	1.1	10.08	9.44	1.74	4.32	0.98	2.39	**39.27**	**44.38**	**30.48**	10.22	3.15	5.59	**25**	**35.94**	18.79	0	0.06	0.06
OG0001690	*PREDICTED_ nuclear cap-binding protein subunit 1 [Elaeis guineensis]*																				

TRINITY_DN25443_c0_g1	17.25	20.75	20.38	21.11	19.98	23.67	16.05	21.21	26.25	30.28	30.4	35.6	28.44	25.48	41.78	22.92	24.07	23.02	13.45	15.4	18.4
OG0001792	*PREDICTED_ abscisic acid receptor PYL8 [Elaeis guineensis]*																				
TRINITY_DN35530_c2_g1	114	130.7	126.1	126.6	136.6	103.6	143.7	124.3	134.4	109.2	89.27	120.4	107.2	99.9	127.2	95.36	97.35	58.74	55.51	55.07	52.04
OG0001847	*PREDICTED_ choline monooxygenase_ chloroplastic isoform X1 [Elaeis guineensis]*																				
TRINITY_DN27028_c0_g1	4	3.08	3.79	2.74	1.98	4.77	2.85	2.36	2.63	8.83	9.53	10.56	3.64	2.48	4.66	3.76	2.38	4.11	3.07	1.2	3.23
OG0001919	*PREDICTED_ spermidine synthase 1-like [Elaeis guineensis]*																				
TRINITY_DN31477_c0_g1	48.19	51.05	67.39	46.62	41.33	100.7	53.96	48.84	43.96	48.29	61.32	56.3	70.66	24.71	39.05	83.34	39.6	32.03	166.1	90.27	90.74
OG0001983	*PREDICTED_ ABC transporter C family member 14-like [Elaeis guineensis]*																				
TRINITY_DN35691_c0_g1	100	138.3	191.3	89.14	88.54	111.2	25.58	25.1	19.72	90.55	135	116.9	16.75	11.36	22.98	9.53	26.57	24.43	1.91	11.28	2.85
OG0002008	*PREDICTED_ glu S.griseus protease inhibitor-like [Elaeis guineensis]*																				
TRINITY_DN27479_c0_g1	27.4	10	7.05	7.84	3.42	12.14	2.86	2.08	1.56	114.4	100.9	99.88	1.21	0.72	1.7	2.62	8.99	1.67	0.12	0	1.31
OG0002138	*PREDICTED_ LOW QUALITY PROTEIN_ elongator complex protein 1 [Elaeis guineensis]*																				
TRINITY_DN25638_c0_g1	7.57	10.55	9.27	12.69	15.17	9.8	8.17	13.08	11.47	11.39	14.95	16.01	9.23	6.86	9.81	7.13	7.98	6.18	7.81	13.98	5.01
OG0002152	*PREDICTED_ myb-related protein 306 [Elaeis guineensis]*																				
TRINITY_DN58516_c0_g1	43.14	41.3	29.94	33.15	45.57	25.98	5.31	13.73	17.25	5.2	3.81	1.68	15.12	2.84	6.06	3.79	4.59	1.38	0	0	0
OG0002211	*PREDICTED_ chaperone protein ClpD2_ chloroplastic [Elaeis guineensis]*																				
TRINITY_DN33551_c0_g1	48.79	66.89	62.16	29.77	19.83	31.57	11.53	10.76	10.47	21.25	20.81	24.49	16.7	7.16	14.23	22.05	15.1	12.29	82.27	133.8	43.44
OG0002303	*PREDICTED_ putative disease resistance protein AT5g47280 [Elaeis guineensis]*																				
TRINITY_DN6834_c0_g1	29.24	48.87	38.4	35.79	57.08	43.44	76.63	65.46	28.97	27.21	39.75	38.37	25.94	9.83	21.78	25.69	25.05	15.45	103.8	278.8	30.63
OG0002779	*PREDICTED_ E3 ubiquitin-protein ligase SDIR1 [Elaeis guineensis]*																				
TRINITY_DN32852_c0_g1	26.32	33.68	27.43	58.37	78.49	61.26	66.26	70.89	65.43	39.77	55.95	54.7	94.92	37.03	96.04	110.4	65.95	85.38	82.31	57.24	61.75
OG0002791	*PREDICTED_ probable WRKY transcription factor 70 [Elaeis guineensis]*																				
TRINITY_DN29063_c0_g1	1.69	16.26	3.8	1.3	9.67	39.46	8.44	4.59	2.9	12.52	11.45	13.95	19.24	0.33	14.19	8.28	0.79	0.77	2.81	23.77	13.57
OG0002802	*PREDICTED_ aldehyde dehydrogenase 22A1-like isoform X1 [Elaeis guineensis]*																				
TRINITY_DN35365_c0_g2	33.72	19.52	19.5	40.58	32.91	21.76	32.29	57.42	71.81	31.82	36.19	37.39	56.16	44.66	58.52	32.72	49.24	44.67	27.43	6.9	25.19
OG0002906	*PREDICTED_ CBL-interacting protein kinase 19-like [Elaeis guineensis]*																				
TRINITY_DN35120_c3_g1	10.62	18.96	95.07	37.99	142.3	360.3	55.59	36.59	39.16	60.16	33.94	170.1	7.52	3.74	4.03	4.4	10.88	3.81	144.7	60.54	177.56

Additional file 6. Drought stress families with macauba palm respective SPM values by tissue and analyzed. SPM value ≥ 0.95 values in blue-colored box.

	SPM Value						
	Leaf	Leaf sheath	Bulb	Root	Female flower	Male flower	Fruit pulp
OG0000003	*PREDICTED_ calcium-dependent protein kinase 26*						
TRINITY_DN17405_c0_g2	0.42	0.44	0.30	0.34	0.33	0.22	0.53
TRINITY_DN24067_c0_g1	0.34	0.38	0.39	0.45	0.37	0.36	0.34
TRINITY_DN31407_c0_g1	0.35	0.37	0.40	0.39	0.39	0.33	0.41
TRINITY_DN34134_c0_g1	0.34	0.39	0.43	0.40	0.35	0.34	0.38
TRINITY_DN35120_c5_g1	0.32	0.39	0.41	0.41	0.33	0.32	0.44
TRINITY_DN35503_c1_g1	0.34	0.38	0.39	0.43	0.37	0.35	0.36
TRINITY_DN35503_c1_g4	0.33	0.36	0.35	0.50	0.34	0.30	0.43
TRINITY_DN82979_c0_g1	0.09	0.16	0.31	0.42	0.29	0.74	0.24
OG0000004	*E3 ubiquitin-protein ligase At3g02290*						
TRINITY_DN22468_c1_g1	0.37	0.45	0.42	0.50	0.30	0.30	0.22
TRINITY_DN22468_c2_g1	0.45	0.47	0.31	0.45	0.23	0.17	0.43
TRINITY_DN24198_c1_g1	0.30	0.39	0.37	0.44	0.40	0.40	0.34
OG0000005	*nuclear transcription factor Y subunit B-4*						
TRINITY_DN34990_c2_g1	0.27	0.32	0.42	0.40	0.49	0.45	0.21
TRINITY_DN34990_c2_g2	0.32	0.32	0.40	0.35	0.38	0.39	0.47
TRINITY_DN34990_c2_g4	0.24	0.31	0.39	0.39	0.49	0.50	0.24
OG0000008	*PREDICTED_beta-glucosidase 12-like*						
TRINITY_DN18360_c0_g1	0.08	0.28	0.46	0.41	0.27	0.38	0.57
TRINITY_DN27655_c0_g1	0.29	0.43	0.43	0.45	0.34	0.41	0.24
TRINITY_DN28112_c0_g1	0.00	0.13	0.46	0.88	0.00	0.00	0.06
TRINITY_DN35708_c4_g1	0.12	0.43	0.46	0.52	0.38	0.41	0.04
TRINITY_DN47933_c0_g1	0.00	0.00	0.67	0.35	0.49	0.43	0.00
TRINITY_DN9444_c0_g1	0.26	0.44	0.19	0.80	0.09	0.24	0.01
OG0000009	*PREDICTED_α-trehalose-phosphate synthase [UDP-forming] 1*						
TRINITY_DN35704_c1_g1	0.49	0.41	0.39	0.40	0.33	0.38	0.16
OG0000011	*PREDICTED_ L-ascorbATe peroxidase_ cytosolic*						
TRINITY_DN24915_c0_g1	0.00	0.00	0.00	1.00	0.00	0.00	0.00
TRINITY_DN32804_c1_g1	0.31	0.36	0.39	0.39	0.43	0.39	0.36
OG0000012	*PREDICTED_ nuclear transcription factor Y subunit A-7*						
TRINITY_DN28214_c0_g1	0.39	0.41	0.35	0.40	0.37	0.37	0.34
OG0000015	*PREDICTED_ probable protein phosphatase 2C 6*						

TRINITY_DN34032_c0_g1	0.41	0.34	0.32	0.38	0.34	0.36	0.47
TRINITY_DN34032_c0_g2	0.40	0.39	0.40	0.40	0.42	0.39	0.21
TRINITY_DN34791_c2_g1	0.41	0.42	0.42	0.35	0.39	0.37	0.25
TRINITY_DN34791_c2_g4	0.38	0.41	0.40	0.38	0.38	0.37	0.32
OG0000017							
TRINITY_DN16570_c1_g1	0.51	0.43	0.41	0.51	0.24	0.12	0.23
TRINITY_DN16580_c0_g1	0.00	0.35	0.41	0.51	0.63	0.07	0.22
TRINITY_DN16580_c0_g1	0.00	0.00	0.00	1.00	0.00	0.00	0.00
TRINITY_DN33665_c0_g1	0.36	0.36	0.39	0.37	0.39	0.37	0.41
TRINITY_DN63959_c0_g1	0.32	0.32	0.34	0.75	0.33	0.06	0.08
TRINITY_DN9572_c0_g1	0.26	0.36	0.41	0.38	0.41	0.37	0.43
OG0000018	*PREDICTED_ABC transporter G family member 45-like*						
TRINITY_DN16806_c0_g1	0.29	0.32	0.53	0.55	0.19	0.26	0.35
TRINITY_DN22152_c0_g1	0.17	0.30	0.34	0.53	0.60	0.32	0.15
TRINITY_DN22700_c0_g1	0.06	0.26	0.50	0.82	0.02	0.06	0.01
TRINITY_DN29259_c0_g1	0.25	0.33	0.40	0.52	0.38	0.39	0.31
TRINITY_DN2976_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN31780_c2_g1	0.21	0.29	0.44	0.77	0.16	0.20	0.13
TRINITY_DN31836_c2_g1	0.13	0.19	0.19	0.93	0.15	0.12	0.04
TRINITY_DN35048_c1_g1	0.32	0.40	0.37	0.45	0.30	0.35	0.43
TRINITY_DN35048_c1_g2	0.61	0.35	0.22	0.55	0.27	0.24	0.19
OG0000023	*PREDICTED_ serine/threonine-protein kinase SAPK2*						
TRINITY_DN28647_c0_g1	0.38	0.41	0.38	0.40	0.45	0.41	0.10
TRINITY_DN32053_c0_g1	0.34	0.40	0.39	0.41	0.34	0.36	0.40
TRINITY_DN34069_c1_g1	0.37	0.45	0.39	0.39	0.33	0.34	0.36
TRINITY_DN34211_c0_g1	0.38	0.41	0.40	0.40	0.33	0.37	0.35
OG0000024	*PREDICTED_ 14-3-3-like protein B*						
TRINITY_DN32377_c0_g1	0.00	0.00	0.00	1.00	0.00	0.00	0.00
TRINITY_DN34169_c0_g1	0.00	0.00	0.00	1.00	0.01	0.00	0.00
TRINITY_DN44954_c0_g1	0.00	0.00	0.00	1.00	0.00	0.00	0.00
TRINITY_DN56448_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN58591_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000027	*PREDICTED_ plasma membrane ATPase-like*						
TRINITY_DN18901_c0_g1	0.10	0.17	0.10	0.93	0.10	0.21	0.17
TRINITY_DN34220_c0_g1	0.30	0.33	0.30	0.61	0.30	0.36	0.35
TRINITY_DN35114_c0_g1	0.07	0.36	0.40	0.55	0.46	0.35	0.26
TRINITY_DN35114_c0_g2	0.40	0.43	0.41	0.52	0.35	0.29	0.06
TRINITY_DN35822_c0_g1	0.36	0.39	0.41	0.51	0.28	0.27	0.38
TRINITY_DN35822_c0_g2	0.36	0.40	0.43	0.42	0.36	0.39	0.27
OG0000028	*PREDICTED_ GPCR-type G protein COLD1*						
TRINITY_DN36198_c0_g1	0.37	0.40	0.42	0.39	0.35	0.37	0.35

OG0000034	*PREDICTED_ respirATory burst oxidase homolog protein E*						
TRINITY_DN24271_c0_g1	0.00	0.07	0.38	0.92	0.02	0.00	0.08
TRINITY_DN28914_c1_g1	0.03	0.43	0.38	0.57	0.15	0.57	0.05
TRINITY_DN32025_c0_g1	0.19	0.43	0.42	0.64	0.12	0.41	0.04
TRINITY_DN35726_c1_g1	0.17	0.28	0.38	0.39	0.26	0.34	0.64
TRINITY_DN35726_c1_g2	0.23	0.48	0.40	0.46	0.31	0.42	0.27
OG0000036	*PREDICTED_ putATive chloride channel-like protein CLC-g*						
TRINITY_DN23259_c0_g1	0.31	0.44	0.45	0.34	0.31	0.37	0.39
TRINITY_DN27635_c0_g1	0.41	0.41	0.40	0.43	0.26	0.32	0.39
TRINITY_DN27648_c0_g1	0.13	0.16	0.39	0.52	0.57	0.43	0.15
TRINITY_DN27648_c1_g1	0.50	0.41	0.37	0.43	0.30	0.30	0.27
TRINITY_DN33770_c0_g1	0.12	0.55	0.33	0.75	0.03	0.09	0.00
OG0000041	*PREDICTED_ S-type anion channel SLAH2-like*						
TRINITY_DN30378_c0_g2	0.07	0.10	0.14	0.38	0.74	0.29	0.44
TRINITY_DN31821_c0_g1	0.28	0.22	0.39	0.47	0.34	0.62	0.06
OG0000043	*PREDICTED_ aquaporin PIP2-4*						
TRINITY_DN24828_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN35053_c3_g1	0.30	0.36	0.43	0.47	0.38	0.36	0.32
OG0000046	*PREDICTED_ protein argonaute 1A-like*						
TRINITY_DN27842_c0_g1	0.06	0.12	0.45	0.29	0.61	0.56	0.01
TRINITY_DN34013_c0_g1	0.10	0.26	0.38	0.45	0.60	0.47	0.03
TRINITY_DN35016_c0_g1	0.27	0.36	0.32	0.45	0.41	0.37	0.44
TRINITY_DN35738_c0_g1	0.30	0.37	0.40	0.39	0.38	0.41	0.37
OG0000052	*PREDICTED_ transcription factor MYC2-like*						
TRINITY_DN2149_c0_g1	0.32	0.41	0.43	0.40	0.36	0.34	0.38
TRINITY_DN22244_c0_g1	0.36	0.43	0.52	0.46	0.35	0.22	0.15
TRINITY_DN27638_c0_g2	0.32	0.40	0.43	0.42	0.39	0.32	0.34
TRINITY_DN31421_c0_g1	0.33	0.42	0.50	0.44	0.40	0.29	0.18
TRINITY_DN31421_c0_g2	0.31	0.35	0.44	0.48	0.37	0.31	0.36
TRINITY_DN31421_c1_g1	0.33	0.41	0.45	0.42	0.32	0.31	0.38
TRINITY_DN76297_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000053	*PREDICTED_ phosphoinositide phospholipase C 2*						
TRINITY_DN35011_c0_g1	0.40549	0.41547	0.31049	0.4086	0.39026	0.35145	0.35183
TRINITY_DN35011_c0_g2	0.07	0.17	0.43	0.57	0.52	0.43	0.04
TRINITY_DN35011_c0_g3	0.43	0.43	0.36	0.38	0.35	0.37	0.32
OG0000054	*PREDICTED_ delta-1-pyrroline-5-carboxylATe synthase*						
TRINITY_DN35030_c1_g1	0.36	0.37	0.38	0.40	0.38	0.36	0.40
OG0000057	*PREDICTED_ bZIP transcription factor 53*						
TRINITY_DN21593_c0_g1	0.27	0.31	0.43	0.43	0.34	0.31	0.50

TRINITY_DN28492_c0_g1	0.30	0.35	0.47	0.41	0.38	0.39	0.33
TRINITY_DN28492_c0_g2	0.05	0.21	0.43	0.47	0.46	0.58	0.00
TRINITY_DN63679_c0_g1	0.32	0.36	0.41	0.42	0.31	0.38	0.43
TRINITY_DN82183_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000063	*PREDICTED_ homeobox-leucine zipper protein ROC8-like*						
TRINITY_DN27172_c0_g1	0.48	0.44	0.35	0.23	0.43	0.46	0.02
OG0000064	*PREDICTED_ probable aquaporin PIP1-2*						
TRINITY_DN29280_c0_g2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN35003_c1_g1	0.34	0.39	0.40	0.38	0.42	0.39	0.32
OG0000065	*PREDICTED_ transcription repressor MYB5-like*						
TRINITY_DN18669_c0_g1	0.07	0.61	0.51	0.47	0.20	0.31	0.10
TRINITY_DN22101_c0_g1	0.03	0.82	0.31	0.20	0.27	0.34	0.06
TRINITY_DN24569_c0_g1	0.00	0.58	0.60	0.48	0.26	0.13	0.00
TRINITY_DN26933_c0_g1	0.00	0.05	0.41	0.91	0.00	0.00	0.00
TRINITY_DN35293_c0_g1	0.00	0.12	0.52	0.84	0.00	0.00	0.00
TRINITY_DN6643_c0_g1	0.00	0.33	0.52	0.48	0.47	0.43	0.00
TRINITY_DN78240_c0_g1	0.00	0.33	0.40	0.86	0.00	0.00	0.00
OG0000069	*PREDICTED_ bZIP transcription factor TRAB1*						
TRINITY_DN17434_c0_g1	0.43	0.48	0.44	0.46	0.20	0.22	0.30
TRINITY_DN22238_c0_g1	0.39	0.45	0.43	0.46	0.25	0.28	0.34
TRINITY_DN28034_c0_g1	0.40	0.44	0.40	0.44	0.31	0.33	0.31
OG0000090	*PREDICTED_ abscisic acid 8'-hydroxylase 1*						
TRINITY_DN32057_c0_g1	0.43	0.40	0.30	0.42	0.13	0.08	0.60
OG0000094	*PREDICTED_ uncharacterized protein LOC109504732*						
TRINITY_DN15141_c0_g1	0.37	0.41	0.38	0.39	0.35	0.36	0.38
TRINITY_DN35068_c2_g1	0.32	0.37	0.37	0.38	0.43	0.44	0.31
OG0000107	*PREDICTED_ CBL-interacting protein kinase 18-like*						
TRINITY_DN16962_c0_g1	0.29	0.45	0.48	0.37	0.33	0.30	0.38
TRINITY_DN35781_c0_g1	0.32	0.39	0.41	0.35	0.35	0.35	0.46
TRINITY_DN35781_c0_g6	0.32	0.39	0.41	0.40	0.35	0.35	0.41
OG0000110	*PREDICTED_ protein ECERIFERUM 3-like*						
TRINITY_DN45525_c0_g1	0.43	0.56	0.28	0.16	0.34	0.53	0.00
TRINITY_DN748_c0_g1	0.42	0.57	0.23	0.18	0.37	0.53	0.00
TRINITY_DN81466_c0_g1	0.40	0.55	0.23	0.16	0.40	0.55	0.00
OG0000112	*PREDICTED_ ethylene-responsive transcription factor ERF003-like*						
TRINITY_DN33317_c0_g1	0.10	0.32	0.38	0.29	0.31	0.29	0.69
OG0000117	*PREDICTED_ ABC transporter G family member 22*						
TRINITY_DN18534_c0_g1	0.63	0.50	0.22	0.28	0.29	0.36	0.05
OG0000119	*PREDICTED_ LRR receptor-protein kinase RPK2*						
TRINITY_DN26288_c0_g1	0.01	0.31	0.53	0.51	0.46	0.38	0.13

TRINITY_DN35712_c0_g1	0.32	0.42	0.39	0.43	0.38	0.33	0.35
TRINITY_DN47823_c0_g1	0.24	0.42	0.40	0.45	0.45	0.39	0.20
TRINITY_DN81687_c0_g1	0.23	0.41	0.49	0.53	0.26	0.28	0.34
OG0000123	*PREDICTED_ heat stress transcription factor A-1*						
TRINITY_DN27166_c0_g1	0.21	0.41	0.46	0.41	0.31	0.30	0.47
TRINITY_DN32888_c4_g4	0.31	0.43	0.40	0.43	0.35	0.35	0.35
TRINITY_DN32888_c4_g5	0.39	0.42	0.43	0.42	0.33	0.32	0.32
OG0000138	*PREDICTED_ small nuclear ribonucleoprotein E*						
TRINITY_DN28076_c0_g2	0.35	0.38	0.37	0.34	0.44	0.38	0.37
OG0000139	*PREDICTED_ probable phospholipid hydroperoxide glutathione peroxidase*						
TRINITY_DN20565_c0_g1	0.35	0.43	0.41	0.34	0.33	0.31	0.45
TRINITY_DN28768_c0_g1	0.41	0.40	0.37	0.35	0.33	0.33	0.45
OG0000145	*PREDICTED_ zinc finger A20 and AN1 domain-containing stress-associated protein 8*						
TRINITY_DN23122_c0_g1	0.33	0.38	0.40	0.37	0.34	0.36	0.46
TRINITY_DN27472_c0_g1	0.33	0.39	0.40	0.37	0.38	0.36	0.41
OG0000149	*PREDICTED_ 1-aminocyclopropane-1-carboxylate synthase-like*						
TRINITY_DN12825_c0_g1	0.17	0.52	0.15	0.73	0.14	0.35	0.00
TRINITY_DN56910_c0_g1	0.03	0.00	0.00	0.50	0.00	0.00	0.86
OG0000156	*PREDICTED_ CBL-interacting protein kinase 32-like*						
TRINITY_DN34658_c0_g1	0.42	0.43	0.39	0.44	0.32	0.34	0.29
TRINITY_DN35028_c0_g1	0.33	0.41	0.34	0.36	0.40	0.37	0.43
TRINITY_DN35373_c0_g1	0.40	0.40	0.39	0.36	0.37	0.36	0.36
OG0000187	*PREDICTED_ LOW QUALITY PROTEIN_ indole-3-acetaldehyde oxidase-like*						
TRINITY_DN29615_c0_g1	0.44	0.45	0.34	0.53	0.21	0.23	0.33
OG0000207	*PREDICTED_ histone deacetylase HDT2*						
TRINITY_DN34680_c0_g1	0.30	0.34	0.39	0.39	0.41	0.34	0.44
TRINITY_DN34680_c0_g2	0.29	0.34	0.40	0.39	0.42	0.37	0.42
OG0000209	*PREDICTED_ ethylene-responsive transcription factor ERF027-like*						
TRINITY_DN24582_c0_g1	0.00	0.14	0.47	0.45	0.69	0.27	0.12
TRINITY_DN80402_c0_g1	0.33	0.80	0.27	0.38	0.09	0.00	0.18
OG0000213	*PREDICTED_ G-type lectin S-receptor-like serine/threonine-protein kinase*						
TRINITY_DN19269_c1_g1	0.01	0.00	0.01	1.00	0.00	0.00	0.00
TRINITY_DN32893_c0_g2	0.04	0.16	0.30	0.94	0.01	0.02	0.02
OG0000226	*PREDICTED_ glutathione S-transferase U17-like*						
TRINITY_DN35725_c2_g1	0.54	0.45	0.37	0.30	0.35	0.38	0.06

TRINITY_DN62888_c0_g1	0.22	0.33	0.34	0.46	0.12	0.09	0.70
OG0000262	*PREDICTED_ protein FAR-RED IMPAIRED RESPONSE 1-like*						
TRINITY_DN31063_c0_g1	0.38	0.40	0.40	0.40	0.34	0.33	0.39
OG0000277	*PREDICTED_ NADP-dependent malic enzyme*						
TRINITY_DN27748_c0_g1	0.44	0.40	0.33	0.41	0.31	0.32	0.41
TRINITY_DN27852_c0_g1	0.35	0.36	0.35	0.36	0.44	0.41	0.38
OG0000282	*PREDICTED_ protein farnesyltransferase subunit beta*						
TRINITY_DN33437_c0_g1	0.39	0.38	0.34	0.37	0.43	0.36	0.36
OG0000305	*PREDICTED_ syntaxin-61*						
TRINITY_DN20536_c0_g2	0.32	0.37	0.36	0.42	0.31	0.34	0.49
TRINITY_DN33913_c0_g1	0.30	0.37	0.40	0.37	0.39	0.41	0.39
OG0000309	*PREDICTED_ pyrophosphate-energized vacuolar membrane proton pump*						
TRINITY_DN35773_c1_g1	0.35	0.39	0.44	0.41	0.33	0.37	0.35
OG0000312	*PREDICTED_ probable protein phosphatase 2C 59 isoform X1*						
TRINITY_DN33850_c0_g1	0.46	0.43	0.37	0.39	0.32	0.37	0.27
TRINITY_DN34401_c0_g1	0.39	0.41	0.37	0.37	0.34	0.39	0.38
OG0000315	*PREDICTED_ transcription factor MYB108-like*						
TRINITY_DN20129_c0_g1	0.15	0.26	0.62	0.46	0.36	0.43	0.00
TRINITY_DN28160_c0_g1	0.11	0.36	0.24	0.80	0.01	0.00	0.40
TRINITY_DN9553_c0_g1	0.49	0.49	0.35	0.45	0.37	0.12	0.22
OG0000318	*PREDICTED_ histone deacetylase 19-like*						
TRINITY_DN35795_c1_g1	0.36	0.34	0.35	0.38	0.35	0.35	0.49
OG0000344	*PREDICTED_ dehydration-responsive element-binding protein 2B*						
TRINITY_DN18242_c0_g1	0.42	0.37	0.33	0.40	0.32	0.28	0.48
TRINITY_DN34839_c0_g1	0.38	0.41	0.39	0.38	0.35	0.35	0.37
OG0000348	*PREDICTED_ galactinol synthase 2-like*						
TRINITY_DN34784_c4_g1	0.54	0.43	0.29	0.54	0.26	0.27	0.06
OG0000357	*PREDICTED_ uncharacterized protein LOC105045662*						
TRINITY_DN33626_c0_g1	0.35	0.41	0.35	0.41	0.36	0.38	0.38
OG0000363	*PREDICTED_ homeobox-leucine zipper protein HOX8*						
TRINITY_DN30488_c0_g1	0.03	0.37	0.49	0.47	0.49	0.40	0.06
TRINITY_DN30488_c0_g2	0.36	0.47	0.47	0.48	0.33	0.30	0.05
OG0000371	*PREDICTED_ probable E3 ubiquitin-protein ligase XERICO*						
TRINITY_DN73320_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000410	*PREDICTED_ potassium channel KOR1*						
TRINITY_DN7448_c0_g1	0.52	0.58	0.27	0.37	0.21	0.36	0.09
OG0000412	*PREDICTED_ zinc finger protein 1-like*						

TRINITY_DN17984_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TRINITY_DN31498_c0_g1	0.39	0.50	0.33	0.47	0.27	0.24	0.38
TRINITY_DN76639_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0000420	*PREDICTED_ UDP-glucose 4-epimerase GEPI48*						
TRINITY_DN33433_c0_g1	0.42	0.41	0.30	0.40	0.29	0.36	0.43
OG0000438	*PREDICTED_ 9-cis-epoxycarotenoid dioxygenase_ chloroplastic*						
TRINITY_DN28946_c0_g1	0.48	0.31	0.42	0.65	0.13	0.09	0.21
TRINITY_DN30486_c0_g1	0.53	0.28	0.39	0.67	0.07	0.09	0.16
OG0000466	*PREDICTED_ probable calcium-binding protein CML31*						
TRINITY_DN47677_c0_g1	0.44	0.50	0.32	0.43	0.21	0.18	0.44
TRINITY_DN65773_c0_g1	0.55	0.59	0.22	0.47	0.15	0.05	0.24
OG0000476	*PREDICTED_ potassium channel K AT3-like*						
TRINITY_DN32477_c2_g1	0.35	0.39	0.52	0.57	0.24	0.28	0.07
OG0000484	*PREDICTED_ annexin D2 isoform X2*						
TRINITY_DN33177_c0_g1	0.15	0.23	0.50	0.37	0.47	0.44	0.35
TRINITY_DN33177_c0_g4	0.24	0.40	0.49	0.39	0.44	0.42	0.13
OG0000501	*PREDICTED_ dehydration-responsive element-binding protein 1G*						
TRINITY_DN23652_c0_g1	0.22	0.65	0.29	0.39	0.09	0.02	0.54
OG0000559	*PREDICTED_ protein arginine N-methyltransferase 5*						
TRINITY_DN25497_c0_g1	0.33	0.36	0.40	0.41	0.40	0.38	0.35
OG0000574	*PREDICTED_ probable tyrosine-protein phosphatase*						
TRINITY_DN27651_c0_g1	0.30	0.21	0.27	0.25	0.47	0.38	0.60
TRINITY_DN27651_c0_g2	0.28	0.30	0.37	0.31	0.50	0.47	0.36
OG0000580	*PREDICTED_ inositol polyphosphate multikinase beta-like*						
TRINITY_DN24747_c0_g1	0.34	0.38	0.41	0.42	0.38	0.33	0.38
OG0000590	*PREDICTED_ aldehyde dehydrogenase family 3 member H1*						
TRINITY_DN24671_c0_g1	0.53	0.43	0.35	0.43	0.34	0.31	0.13
OG0000595	*PREDICTED_ poly [ADP-ribose] polymerase 2*						
TRINITY_DN36004_c0_g1	0.21	0.17	0.38	0.39	0.64	0.44	0.11
OG0000617	*PREDICTED_ annexin D4 isoform X2*						
TRINITY_DN25175_c0_g1	0.37	0.53	0.26	0.36	0.33	0.42	0.33
OG0000623	*PREDICTED_ guanine nucleotide-binding protein alpha-1 subunit-like*						
TRINITY_DN32765_c0_g1	0.27	0.40	0.38	0.42	0.45	0.44	0.24
OG0000638	*PREDICTED_ nuclear cap-binding protein subunit 2*						
TRINITY_DN18003_c0_g1	0.33	0.37	0.40	0.37	0.42	0.38	0.36

OG0000646	*PREDICTED_ glycerol kinase*						
TRINITY_DN30116_c0_g1	0.48	0.38	0.28	0.52	0.30	0.35	0.25
OG0000682	*PREDICTED_ histone deacetylase 6-like*						
TRINITY_DN33114_c0_g1	0.40	0.39	0.38	0.41	0.33	0.36	0.37
OG0000705	*PREDICTED_ squalene monooxygenase*						
TRINITY_DN28003_c0_g1	0.40	0.39	0.39	0.35	0.38	0.40	0.34
OG0000720	*PREDICTED_ RING-H2 finger protein ATL78-like*						
TRINITY_DN12789_c0_g1	0.20	0.50	0.52	0.63	0.13	0.16	0.03
OG0000721	*PREDICTED_ ABC transporter G family member 25-like*						
TRINITY_DN34098_c1_g1	0.22	0.36	0.50	0.45	0.43	0.40	0.16
OG0000761	*PREDICTED_ double-stranded RNA-binding protein 8*						
TRINITY_DN15102_c0_g1	0.39	0.39	0.36	0.37	0.40	0.36	0.37
OG0000763	*PREDICTED_ protein OVEREXPRESSOR OF CATIONIC PEROXIDASE 3*						
TRINITY_DN1508_c0_g1	0.41	0.38	0.39	0.40	0.34	0.35	0.38
OG0000796	*PREDICTED_ ethylene-responsive transcription factor ERF017-like*						
TRINITY_DN16697_c0_g1	0.38	0.57	0.39	0.56	0.11	0.11	0.19
TRINITY_DN6144_c0_g1	0.29	0.46	0.30	0.69	0.37	0.09	0.03
OG0000815	*PREDICTED_ poly [ADP-ribose] polymerase 1-like*						
TRINITY_DN35514_c0_g1	0.36	0.37	0.38	0.41	0.43	0.39	0.31
OG0000817	*PREDICTED_ 3-ketoacyl-CoA thiolase 2_ peroxisomal*						
TRINITY_DN35269_c1_g2	0.28	0.38	0.37	0.50	0.41	0.40	0.27
OG0000847	*PREDICTED_ glutathione S-transferase DHAR2*						
TRINITY_DN32172_c0_g1	0.30	0.40	0.43	0.40	0.34	0.35	0.41
OG0000850	*PREDICTED_ omega-3 fatty acid desaturase_ chloroplastic*						
TRINITY_DN34006_c0_g1	0.46	0.41	0.29	0.32	0.30	0.26	0.52
OG0000856	*PREDICTED_ uncharacterized protein LOC105044901*						
TRINITY_DN35032_c0_g1	0.45	0.42	0.34	0.36	0.33	0.36	0.38
OG0000868	*PREDICTED_ NAC domain-containing protein 68*						
TRINITY_DN35062_c5_g3	0.42	0.43	0.29	0.48	0.41	0.08	0.39
OG0000919	*PREDICTED_ molybdenum cofactor sulfurase*						
TRINITY_DN29766_c0_g1	0.40	0.40	0.30	0.49	0.32	0.36	0.34
OG0000932	*PREDICTED_ transcription factor MYB44-like*						
TRINITY_DN33427_c1_g2	0.34	0.49	0.34	0.48	0.28	0.23	0.40
OG0000934	*PREDICTED_ omega-6 fatty acid desaturase_ endoplasmic reticulum isozyme 2-like*						
TRINITY_DN29082_c0_g1	0.00	0.00	0.00	1.00	0.00	0.00	0.00
OG0000963	*PREDICTED_ G-protein coupled receptor 1*						
TRINITY_DN73492_c0_g1	0.33	0.40	0.41	0.42	0.34	0.35	0.39

OG0000978	*PREDICTED_ phosphoenolpyruvate carboxykinase [ATP]*						
TRINITY_DN34267_c3_g2	0.32	0.34	0.44	0.43	0.38	0.46	0.21
OG0000985	*PREDICTED_ histidine kinase 1-like*						
TRINITY_DN32093_c0_g1	0.16	0.50	0.39	0.65	0.22	0.31	0.03
OG0001023	*PREDICTED_ chlorophyll a-b binding protein CP24 10A_ chloroplastic-like*						
TRINITY_DN29717_c0_g2	0.62	0.54	0.37	0.19	0.26	0.29	0.00
OG0001046	*PREDICTED_ omega-6 fatty acid desaturase_ chloroplastic*						
TRINITY_DN26075_c0_g1	0.47	0.44	0.37	0.34	0.30	0.31	0.38
OG0001076	*PREDICTED_ zinc finger A20 and AN1 domain-containing stress-associated protein 6-like*						
TRINITY_DN16529_c0_g1	0.45	0.39	0.34	0.33	0.34	0.34	0.42
OG0001144	*PREDICTED_ probable purine permease 4_ partial*						
TRINITY_DN20091_c0_g1	0.06	0.30	0.30	0.62	0.35	0.55	0.00
OG0001311	*PREDICTED_ abscisic stress-ripening protein 2-like*						
TRINITY_DN27408_c1_g1	0.27	0.33	0.34	0.45	0.29	0.27	0.58
TRINITY_DN301_c0_g1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OG0001372	*PREDICTED_ AP2-like ethylene-responsive transcription factor*						
TRINITY_DN30070_c0_g1	0.17	0.39	0.20	0.62	0.31	0.54	0.01
OG0001690	*PREDICTED_ nuclear cap-binding protein subunit 1*						
TRINITY_DN25443_c0_g1	0.37	0.37	0.37	0.41	0.40	0.38	0.33
OG0001792	*PREDICTED_ abscisic acid receptor PYL8*						
TRINITY_DN35530_c2_g1	0.39	0.39	0.40	0.38	0.38	0.37	0.33
OG0001847	*PREDICTED_ choline monooxygenase_ chloroplastic*						
TRINITY_DN27028_c0_g1	0.37	0.31	0.30	0.55	0.36	0.37	0.33
OG0001919	*PREDICTED_ spermidine synthase 1-like*						
TRINITY_DN31477_c0_g1	0.38	0.37	0.37	0.39	0.35	0.35	0.43
OG0001983	*PREDICTED_ ABC transporter C family member 14-like*						
TRINITY_DN35691_c0_g1	0.50	0.45	0.33	0.48	0.29	0.33	0.14
OG0002008	*PREDICTED_ glu S.griseus protease inhibitor-like*						
TRINITY_DN27479_c0_g1	0.40	0.37	0.19	0.78	0.13	0.22	0.02
OG0002138	*PREDICTED_ LOW QUALITY PROTEIN_ elongator complex protein 1*						
TRINITY_DN25638_c0_g1	0.36	0.41	0.39	0.43	0.36	0.33	0.34
OG0002152	*PREDICTED_ myb-related protein 306*						
TRINITY_DN58516_c0_g1	0.57	0.54	0.41	0.24	0.30	0.24	0.00
OG0002211	*PREDICTED_ chaperone protein ClpD2_ chloroplastic*						
TRINITY_DN33551_c0_g1	0.47	0.38	0.28	0.35	0.31	0.31	0.50
OG0002303	*PREDICTED_ putative disease resistance protein*						

TRINITY_DN6834_c0_g1	0.37	0.38	0.42	0.37	0.31	0.32	0.46
OG0002779	*PREDICTED_ E3 ubiquitin-protein ligase SDIR1*						
TRINITY_DN32852_c0_g1	0.31	0.38	0.38	0.37	0.42	0.41	0.38
OG0002791	*PREDICTED_ probable WRKY transcription factor 70*						
TRINITY_DN29063_c0_g1	0.27	0.41	0.30	0.46	0.48	0.10	0.47
OG0002802	*PREDICTED_ aldehyde dehydrogenase 22A1*						
TRINITY_DN35365_c0_g2	0.31	0.37	0.42	0.38	0.42	0.40	0.34
OG0002906	*PREDICTED_ CBL-interacting protein kinase 19-like*						
TRINITY_DN35120_c3_g1	0.31	0.51	0.38	0.42	0.17	0.17	0.51

Bazzo *et al. BMC Plant Biology* (2018) 18:276
https://doi.org/10.1186/s12870-018-1509-9

BMC Plant Biology

CHAPTER 2:

RESEARCH ARTICLE **Open Access**

Development of novel EST-SSR markers in the macaúba palm (*Acrocomia aculeata*) using transcriptome sequencing and cross-species transferability in Arecaceae species

Bárbara Regina Bazzo[1], Lucas Miguel de Carvalho[1], Marcelo Falsarella Carazzolle[1],
Gonçalo Amarante Guimarães Pereira[1]* and Carlos Augusto Colombo[2]

Abstract

Background: The macaúba palm is a novel feedstock for oil production suitable for multiple uses, including as biodiesel and in the food and cosmetic industries. As an efficient alternative, the macaúba palm has limited genomic resources, particularly expressed sequence tag (EST) markers. We report a comprehensive set of validated EST-simple sequence repeat (SSR) markers by using transcriptome sequencing, its application in genetic diversity analysis and cross transferability in other palm trees with environmental and economic importance.

Results: In this study, a total of 418 EST-SSRs were identified to be unique for one transcript and region; 232 EST-SSRs were selected, with trinucleotide repeats being the most frequent motif, representing 380 (90.9%), followed by composited (4.5%), di- (3.6%), and hexanucleotides (3.6%). A total of 145 EST-SSRs (62.5%) were validated for consistent amplification in seventeen macaúba palm samples, and 100 were determined to be polymorphic with PIC values ranging from 0.25 to 0.77. Genetic diversity analysis was performed with the 20 most informative EST-SSR markers showing a distinct separation of the different groups of macaúba palm. Additionally, these 145 markers were transferred in six other palm species resulting in transferability rates of 99% (144) in *Acrocomia intumescens*, 98% (143) in *Acrocomia totai*, 80.7% (117 EST-EST) in African oil palm (*Elaeis guineensis*) and peach palm (*Bactris gasipaes*) samples, 70% (102) in the juçara palm (*Euterpe edulis*) and 71.7% (104) in the hat palm (*Sabal causiarum*). Analysis of genetic distance showed a high separation in accordance with geographic location, establishing distinct groups by genera.

Conclusions: The EST markers identified in our study are a valuable resource and provide a genomic tool for genetic mapping and further genetic studies, as well as evaluation of co-location between QTLs and functionally associated markers.

Keywords: RNA-Seq, Genic molecular markers, Transcript sequences, Genetic diversity, Polymorphism, Macaúba palm

---

* Correspondence: goncalo@unicamp.br
[1]Institute of Biology, Laboratory of Genomic and Expression, State University of Campinas, Campinas, Brazil
Full list of author information is available at the end of the article

Background

The macaúba palm (*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart., 2n = 30) is a native palm tree from America, belonging to the Arecaceae family. It is an arboreous oleiferous species, perennial, heliophilous, monoecious, with a single stem between 4 and 15 m tall and 20–30 cm in diameter [1]. The species is monoecious, and its genome size is 5.81 pg, distributed in 15 pairs of chromosomes (2n = 30) with a base composition of AT = 58.3% [2]. *Acrocomia aculeata* has been documented to inhabit areas from north Florida, Mexico and West Indies to south Paraguay and north Argentina; it is considered the most widespread palm in Brazil, and it can be found especially in the Midwest and Southeast region [3]. It grows naturally in large populations, is adapted to different ecosystems and can be used to rehabilitate degraded pastures or in agroforestry systems.

Its fruit, a product of great economic value, contains a large amount of pulp oil with oil content values up to 75% for the pulp and 65% for oil on the dry basis [4]. It is globose, with a fibrous mesocarp, strongly adhered endocarp, and a large endosperm, with up to four seeds per fruit. Mesocarp oil is rich in oleic acid with high oxidative stability and operability at low temperatures [5, 6]. The oil extracted from the endosperm is rich in short-chained saturated fatty acids, primarily lauric acid, constituting a valuable source for pharmaceutical and cosmetic use [7–9].

The macaúba palm has a great production potential for oil, similar to the African oil palm (*Elaeis guineensis*), whose global production reached 69 million of tons in 2017, representing 34% of global oil production [10]. The productivity of selected native plants of macaúba may reach 5000 kg of oil per ha^{-1} year^{-1} [11, 12], making it an inexpensive feedstock for oil production. Moreover, the macaúba palm is considered a novel oil-feedstock crop with a potential role as an environmentally and socially co-beneficial feedstock in South America. Its large-scale cultivation should be conducted in permanent protection areas, pastures, disturbed areas, and tilled land.

Despite its incipient domestication and current use based on its extractive character, macaúba has several uses, with multiple products from its exploitation including food, cosmetics, animal feed, and biofuels [12]. Currently, the renewed interest in this novel feedstock increased commercial interest and can lead to the propagation of plants without agronomic quality, which would render their competitiveness unfeasible.

Precaution should be taken in the use of non-domesticated feedstock species because possible biological variations and genetic diversity is observed in the macaúba in different environments of occurrence [13]. In this context, molecular markers are considered an essential tool to identify and to select superior plants for adoption in large-scale commercial crops, establishment of core collections, creation of seed garden, and initiation of breeding studies [14, 15].

Molecular markers are widely used to track loci and genome regions in several crop-breeding programmes. Furthermore, they can accelerate the generation of new varieties and allow for the connection of phenotypic characters with the genomic loci responsible for them [16, 17]. Molecular marker selection (MMS) is considered a simple and rapid technique, thus accelerating breeding time and developing segregating populations for several generations [16].

Microsatellites or single sequence repeats (SSRs), primarily genic SSRs (or EST-SSRs), are widely employed in palm tree studies with commercial interest [14, 18–21]. In addition to being functional, EST-SSR can lead to a gain or loss of gene function via frameshift mutation or other changes in the amino acid sequence [22]; this is useful for marker-assisted selection, especially when the markers reside in the genes responsible for a phenotypic trait. Assessing the performance and genetic diversity of the natural material is important for understanding the genetic structure and consequently for guiding breeding programmes to develop superior genotypes.

Cross-transferability is a dominant feature of EST-SSR markers among distantly related species and can shed light on the evolution of plant genomes, changes in gene location, and genome organization, whereas the genomic SSRs are not suitable for this purpose. EST-SSR transferability provides a cost-effective source of markers for related species, which is important for taxa with low microsatellite frequencies or for those whose microsatellites are difficult to isolate. Among Arecaceae plants, few species have SSR available, and its cross-transferability is an alternative for species with less information [23, 24].

Despite its increasing promotion in Brazil and the availability of molecular study data, only few SSR markers have been determined to be useful and validated for *A. aculeata*. The present study involves the first transcriptome sequencing of 8 *A. aculeata* tissues using the Illumina Hiseq2500 sequencing platform. The goal of this study was to provide a polymorphic set of genic microsatellite markers, which will allow for the improvement of the understanding of the genetic diversity, genotype characterization and genetic structure of *A. aculeata*, Acrocomia genera, and other environmental and commercial important palm species. Additionally, these markers will be useful in modern *A. aculeata* breeding programmes.

Results

Sequencing and reference assembly of Illumina paired-end reads

The cDNA libraries were sequenced using Illumina/HISEQ2000 (Illumina Inc. San Diego, CA, USA), producing millions of 100-bp paired-end reads. For each individual, the sequenced reads from all tissues were grouped and submitted to reference-based transcriptome assembly using the African oil palm genome as a reference. We decided to assemble the individuals separately to avoid the chimeric combination of EST-SSRs that may vary in size among the individuals. The average alignment rate was 71%, ranging from 64.6% (root samples from Amparo and Pedreira) to 80.2% (endosperm sample from Ibituruna). The transcriptome assemblies for the individuals i1, i2 and i3 were performed using vegetative tissues (merging leaves, roots, leaf sheaths and bulbs) that generated 60.323, 61.093 and 61.316 transcript sequences, respectively. The transcriptome assemblies of the female and male flowers resulted in 59.141 and 58.452 transcripts for the individuals i4, i5, and i6, respectively. With respect to fruit libraries (mesocarp and endosperm), the assembly produced 26.447, 28.749, 23.861, 29.847, 26.191, and 26.086 for the individuals i7, i8, i9, i10, i11, and i12, respectively. The minimum average size for all transcripts identified was 200 bp.

Frequency and distribution of EST-SSR

We identified 85.014 redundant EST-SSRs from 455.051 transcript sequences from all individuals. Among them, 7.492 EST-SSRs were chosen based on the criteria of flanking primers that generate PCR product size ranging from 100 to 500 bp and located inside the same exon. Of these, a total of 418 non-redundant EST-SSRs were selected as representative of each locus, of which sixty-three (15%) were exclusive to vegetative tissues, twenty-three (5.5%) to flower tissues, and twenty-eight (6.7%) to fruit. The number of repeats by SSR motif ranged from 5 to 13 repeats, with 5 and 6 being the most frequent (Table 1). Regarding the location in the African oil palm genome, chromosomes 1 and 2 harboured the majority of the markers, with 38 EST-SSRs in each, followed by chromosome 3 (36 EST-SSRs) and chromosome 4 (23).

The trinucleotide repeats exhibit the highest frequency of occurrence (380–90.91%), followed by composite repeats, dinucleotides, and hexanucleotides (4.55, 3.59 and 3.59%, respectively) (Table 1). Among the trinucleotide motifs, the most frequent motifs are GAG (7.18%, 30), CCT (5.02%, 21), GCC (4.78%, 20), GGA (4.55%, 19), CAG (3.83%, 16), GGC, AAG, AGG, GCC, CGG, GGT and TCC (see additional file 1). The major dinucleotide EST-SSR motif observed in the macaúba palm is CT/AG, comprising 66% of these motifs.

Development and validation of genic-SSR markers

To prevent amplicon size deviations, only EST-SSRs inside one exon were chosen for validation. A minimum distance of 1 Mpb between each marker was considered for validation to reduce the chance of linkage equilibrium among the markers. A total of 481 EST-SSR primers flanking unique sequences were designed, which were designated Acro01 to Acro418 (Acro = 'Acrocomia').

Table 1 Number and frequency of the 418 EST-SSRs identified by the tissue from which they were derived, number of motif repeats and SSR motif in the macaúba palm transcriptome data set

	Number of EST-SSRs	Frequency (%)
Tissue		
Flower (female and male)	23	5.50
Fruit	28	6.70
Vegetative tissue	63	15.07
Vegetative tissue/flower	60	14.35
Vegetative tissue/fruit	21	5.02
Fruit/flower	19	4.55
Vegetative tissue/flower/fruit	204	48.80
Number of repeats		
5	197	47.13
6	109	26.08
7	48	11.48
8	21	5.02
9	12	2.87
10	2	0.48
11	4	0.96
12	5	1.20
13	1	0.24
SSR Motif		
Dinucleotide	15	3.59
Trinucleotide	380	90.91
Hexanucleotide	15	3.59
Compound	19	4.55

A subset of 232 EST-SSRs was selected according to the trinucleotide motif, amplicon size (from 100 to 500 bp), and melting temperature of both forward and reverse primers. They were selected to monitor polymorphisms in seventeen samples of macaúba palm from different geographic regions. Additionally, the annotation of these 232 EST-SSRs was performed according to the African oil palm genome.

Of the 232 EST-SSRs tested, 145 (62.5%) were successfully amplified in the genomic DNA, producing clear PCR amplicons with the expected sizes. In total, the 145 EST-SSR markers generated 476 markers bands, and all markers were used for further analysis. The mean number of alleles per loci was 3.28, ranging from 1 to 11 alleles, and 87 (60%) markers had three or more alleles. The effective number of alleles per locus (Ne), expected heterozygosity (He), and observed heterozygosity (Ho) ranged from 1 to 5.02 (Acro205), 0 to 0.828 (Acro201), and 0 to 1.0 (Acro16, Acro64, Acro172), respectively. In addition, Shannon's information index (I) values ranged from 0 to 1.93 (Acro205), probability of identity (PI) ranged from 0.063 to 1.0, and PIC values ranged from 0 to 0.777/0.776 (Acro205/Acro201). Among these, 39 exhibited high PIC values, ranging from 0.77 (Acro205) to 0.50 (Acro220); 61 exhibited medium PIC value, ranging from 0.49 (Acro125) to 0.25 (Acro213); 31 with low PIC value, ranging from 0.24 (Acro33) to 0.05 (Acro124); and 14 with null PIC value (monomorphic markers), according to Botstein et al. [25] (see Additional file 2). From all loci analysed, 124 private alleles were detected in 74 EST-SSRs, with Acro205 being the marker with more private alleles (7 private alleles) (Table 2). The 145 novel designed EST-SSR primers, annealing temperature, product size, and corresponding primer pair sequences are listed in the additional file 3.

Based on PIC values and probability of identity (PI) for all EST-SSRs, the most informative EST-SSR markers were selected to verify efficiency in the genetic analysis of the samples (additional file 4). Principal coordinate analysis (PCoA) was performed on the genotype data of 17 samples of macaúba palm based on the Nei distance [26]. The first and second axes explained 27.17 and 18.04% of the variation observed, respectively (Fig. 1). The PCoA results revealed three distinct genetic groups in accordance with their geographic location, containing samples from Itapira and Jaguariúna city from São Paulo State/Brazil (Group 1); Rifaina from São Paulo State/Brazil, Serra da Canastra, and Capitólio from Minas Gerais State/Brazil (Group 2); and Luz (Minas Gerais State/Brazil) (Group 3).

Cross-transferability

The 145 EST-SSR primers from macaúba palm were examined for cross transferability in six species belonging to the Arecaceae family, including *Acrocomia totai* and *Acrocomia intumescens*. We observed a frequency of cross transferability of 99% in *Acrocomia intumescens* (144 EST-SSRs), 98% for *Acrocomia totai* (143), 80.7% (117 EST-EST) for African oil palm (*Elaeis guineensis*) and peach palm (*Bactris gasipaes*) samples, 70% (102) for the juçara palm (*Euterpe edulis*), and 71.7% (104) for the hat palm (*Sabal causiarum*).

Considering all parameters of descriptive genetics, the number of alleles ranged from 0 (when the EST-SSR was not transferred in the palm species) to 8 alleles (*Acrocomia totai*, EST-SSR 42). The juçara palm presented the lower means, and *Acrocomia totai*, the highest means for an effective number of alleles per locus (Ne) (1.193/1.936), mean expected heterozygosity (He) (0.189/0.367), mean observed heterozygosity (Ho) (0.102/0.204), and Shannon's information index (I) (0.328/0.629) (see in Additional file 5).

The principal coordinate analysis (PCoA) performed on all sample with the 145 EST-SSRs (Nei distance [26]) explained 50.95% of the variance in the first and second axes with clear distinction of species and genera, which were structured in different groups (1 to 5) (Fig. 2). Non-amplified EST-SSR loci were considered as missing data.

Discussion

As the next-generation DNA sequencing is becoming quicker and cheaper, vast amounts of sequence data are currently being generated exponentially, including a large number of ESTs from different plant species. Our study is the first to provide effective and useful markers from transcriptome analysis of 8 distinct tissues of macaúba palm of different individuals using next-generation sequencing, from which we identified and validated 145 macaúba palm EST-SSRs.

The macaúba palm EST database harboured highly abundant SSR sites. These EST-SSR markers proved to be useful for prior population structure analysis and will facilitate macaúba palm breeding, as well as studies with other palm trees with economic and environmental value upon transfer of these markers to these trees. In this study, 85.014 redundant EST-SSRs were identified in all samples, with a frequency of 1/1 Mbp, to prevent the validation of more than one locus in linkage equilibrium, according to assembly in the African oil palm genome. Although the density of the markers was lower than that obtained in other studies with pigeon pea (1/8.4 Kbp) [27], robusta coffee (1/3.4 Kbp) [28], caston bean (1/1.77 Kbp) [29], *Crambe abyssinica* (1/11.1 kb) [30] and *Cocos nucifera* (1/7.59 Kbp) [31], several useful markers were validated for use from the macaúba palm and other Arecaceae species. The density of SSR in EST depends on the search tool used, criteria used to identify SSRs in the database or redundancy in the set of SSR, which can lead to the identification of multiple markers at the same locus [15].

To date, in the macaúba palm, a small number of molecular markers have been available and validated [32], and they have been used in recent studies [13, 33–36]. Recently, 221 SSRs were identified in the chloroplast in a high frequency of occurrence (total mean of 1/1.75 Kb) located in intergenic spacers, coding regions, introns and tRNA genes; however, they have not been validated for the macaúba palm [37].

Our data revealed that trinucleotide repeats were the most frequent EST-SSRs, as observed by Zhao et al. [18] and Xia et al. [31] in the date palm and coconut palm, respectively. This suggests a result of selection and evolution because tri- and hexa-SSRs do not change the coding frame of the gene regions. Other repetitions change the frame and induce negative mutation when there is variation in SSR length [29].

Table 2 Private alleles and frequency for each EST-SSR detected in macaúba palm samples

EST-SSR	Allele size	Frequency	EST-SSR	Allele size	Frequency
Acro2	229	0.059	Acro162	136	0.059
Acro10	112	0.029	Acro164	246	0.059
Acro10	115	0.029	Acro165	157	0.294
Acro14	258	0.059	Acro168	233	0.063
Acro14	264	0.059	Acro168	290	0.063
Acro15	106	0.059	Acro171	449	0.147
Acro29	266	0.529	Acro172	252	0.029
Acro32	145	0.118	Acro178	350	0.029
Acro34	339	0.059	Acro178	356	0.029
Acro36	203	0.300	Acro183	383	0.235
Acro36	176	0.033	Acro183	395	0.059
Acro36	185	0.033	Acro183	404	0.059
Acro36	221	0.033	Acro187	201	0.038
Acro39	401	0.235	Acro188	206	0.059
Acro39	350	0.118	Acro188	209	0.029
Acro41	254	0.029	Acro189	277	0.088
Acro49	314	0.067	Acro189	291	0.059
Acro49	386	0.033	Acro189	307	0.059
Acro52	358	0.118	Acro191	243	0.059
Acro52	346	0.059	Acro192	358	0.059
Acro58	422	0.059	Acro193	300	0.059
Acro62	147	0.423	Acro195	397	0.067
Acro62	240	0.077	Acro196	253	0.063
Acro62	246	0.077	Acro198	280	0.067
Acro63	375	0.063	Acro199	349	0.059
Acro64	181	0.029	Acro199	352	0.059
Acro69	400	0.029	Acro201	354	0.179
Acro82	164	0.059	Acro201	342	0.036
Acro84	124	0.088	Acro201	369	0.036
Acro92	409	0.735	Acro201	417	0.036
Acro93	189	0.059	Acro203	215	0.206
Acro97	305	0.176	Acro203	167	0.118
Acro99	262	0.235	Acro203	260	0.029
Acro101	289	0.706	Acro205	120	0.094
Acro102	277	0.059	Acro205	126	0.063
Acro103	384	0.059	Acro205	171	0.063
Acro108	236	0.059	Acro205	180	0.031
Acro111	166	0.107	Acro205	183	0.031
Acro111	169	0.107	Acro205	192	0.031
Acro111	154	0.071	Acro205	195	0.031
Acro116	385	0.059	Acro206	338	0.250
Acro118	324	0.059	Acro206	362	0.063
Acro118	339	0.059	Acro206	365	0.063

Table 2 Private alleles and frequency for each EST-SSR detected in macaúba palm samples *(Continued)*

EST-SSR	Allele size	Frequency	EST-SSR	Allele size	Frequency
Acro118	378	0.059	Acro206	385	0.063
Acro118	358	0.029	Acro208	409	0.471
Acro125	223	0.059	Acro208	418	0.059
Acro126	350	0.059	Acro209	213	0.529
Acro130	283	0.059	Acro210	242	0.133
Acro136	160	0.059	Acro210	278	0.033
Acro137	146	0.033	Acro212	242	0.176
Acro142	393	0.067	Acro212	170	0.059
Acro143	126	0.059	Acro212	113	0.029
Acro144	175	0.059	Acro212	152	0.029
Acro144	211	0.059	Acro213	239	0.059
Acro146	173	0.088	Acro213	236	0.029
Acro147	273	0.125	Acro217	105	0.059
Acro153	348	0.500	Acro218	142	0.059
Acro153	345	0.346	Acro218	109	0.029
Acro153	342	0.077	Acro223	212	0.059
Acro153	306	0.038	Acro227	369	0.059
Acro153	351	0.038	Acro227	390	0.059
Acro159	354	0.059	Acro228	259	0.265

The most frequent motifs were GAG/CAG, CCT/CCG and GGA/GGC, which encode glutamic acid, proline, and glycine, respectively. Among the proteins containing 10 or more repetitions of single amino acid, glutamine, alanine, glycine, glutamic acid, and serine repeats were more frequent than other amino acids [38, 39]. Katti et al. [40] observed that codon repeats corresponding to small hydrophilic amino acids are possibly more easily tolerated, and selection pressure possibly eliminates codon repeats encoding hydrophobic and basic amino acids. Consequently, the high level of occurrence of these motifs is substantial because amino acids produced by them are observed to a high extent in proteins.

Of the 418 EST-SSRs identified in different transcripts of all tissue samples, 145 (62.5%) markers produced clear bands. This rate is higher than that reported in the rubber tree (50%) [41], alfalfa (30%) [42], and adzuki bean (59.2%) [43] but lower than that reported in the castor bean (81.2%) [29] and mung bean (65%) [44]. The 87 primer pairs failed to generate the expected amplicon size or non-amplification PCR product, which may be due to the presence of introns and indels, since we have used the oil palm genome as a reference rather than the macaúba genome, lack of specificity or assembly error.

Although genic SSRs generally exhibits PIC values lower than that of genomic markers in population and

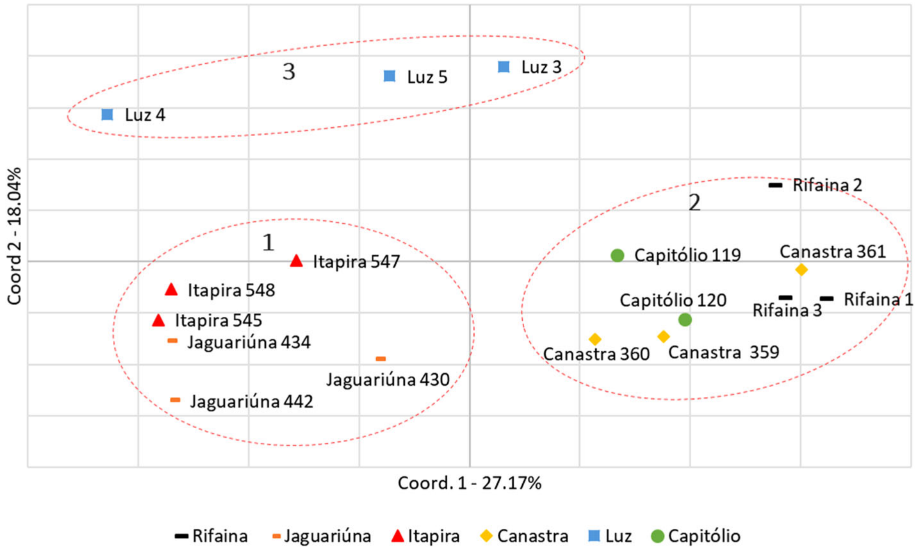

Fig. 1 ***Acrocomia aculeata*** **PCoA analysis.** Two-dimensional plot of principal coordinate analysis (PCoA) based on the Nei distance of *Acrocomia aculeata* samples from different locations for the 20 more discriminating EST-SSRs

diversity studies, we determined markers with high PIC values (> 0.5) [25], which were effective in discriminating the samples of the macaúba palm. The PIC values of our markers are higher than the values of the markers published by Nucci et al. [32] for the macaúba palm and are more valuable for marker-assisted selection and other applications.

The transferable nature of EST-SSR markers within related species or genera extends their usefulness in plant breeding and genetic studies, being limited in those species that have no available EST data set. We have achieved a high cross-species transferability of EST-SSRs found in the macaúba palm to other palm genera. It is advantageous to save time and cost for developing SSR markers for species that have not been largely studied. We obtained a transferability rate of 99% in *Acrocomia intumescens* (144 EST-SSRs), 98% for *Acrocomia totai* (143), and a high level of transferability in other palm trees.

The EST-SSR markers developed from *A. aculeata* offer a feasible solution for both correlational research of other related species that lack molecular markers and genetic studies in the *Acrocomia* genera. As demonstrated using the PCoA (Fig. 2), the Acrocomia genera was clustered into the same groups, and other palm species were clustered individually.

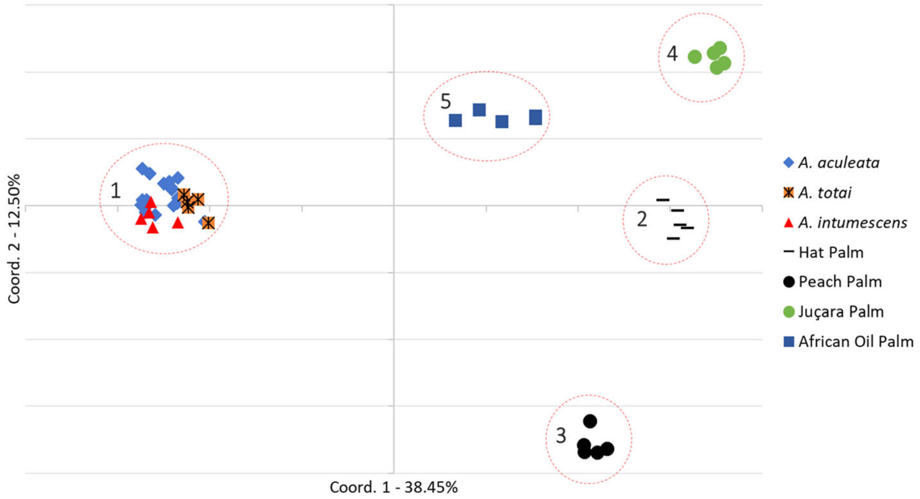

Fig. 2 Arecaceae family PCoA analysis. Two-dimensional plot of principal coordinate analysis (PCoA) based on the Nei distance of Arecaceae samples for 145 EST-SSR data

In the Arecaceae family, Zaki et al. [23] observed a high transferability level in Arecaceae members; 100% of *Elaeis oleifera* genomic SSRs were transferred to *Elaeis guineensis*, which are members of the same genre, and transferability in *Euterpe oleracea* (acai palm) was 72.7%, in *Oenocarpus multicaulis* was 63.6%, in *Jessinia bataua* was 54.5%, in *Ptychosperma macarthurii* (Macarthur palm) was 54.5%, in *Dictyosperma album* (Princess palm) was 45.5% and in *Cyrtostachys renda* (Lipstick *Palm*) was 45.5%. We obtained the same frequency of cross-transferability at the genera level as using genic SSRs (99 and 98%, respectively, for *A. totai and A. intumescens*/100% *Elaeis* sp.). Additionally, Mengistu et al. [24] observed that 44% of the markers developed by Zaki et al. [23] successfully amplified the genomic DNA in *A. aculeata*, of which 26% were polymorphic. Compared to genic SSRs, EST-SSRs are highly transferable at the genus, tribe and subfamily levels because of their location in conserved genic regions.

The strategy of obtaining EST-SSRs from reference-based transcriptome proved to be efficient since the SSRs identified were useful in separating the species and individuals within the *Acrocomia aculeata* species, as revealed in the principal coordinate analyzes. It is possible that some of EST-SSRs were lost and they could be capable of discriminate species of macauba palm and African oil palm, but since there is no macauba genome, the African oil palm reference was essential and useful to map regions with variable EST-SSRs.

We identified 418 EST-SSRs in all tissues; of these, 145 were amplified in *A. aculeata* samples, and a total of 100 polymorphic primer pairs were successfully amplified fragments, thus revealing abundant polymorphism between 17 *A. aculeata* samples. Additionally, of those 145 EST-SSRs, 144 were transferable in *Acrocomia intumescens*, 143 in *Acrocomia totai,* 117 in the African oil palm (*Elaeis guineensis*) and peach palm (*Bactris gasipaes*), 106 in the juçara palm, and 105 in the hat palm (*Sabal causiarum*), indicating that these newly developed EST-SSRs can be used with confidence in future population genetic studies of the 6 related species.

Conclusion

This study has identified the wide occurrence of microsatellites in *Acrocomia aculeata*. The use of reference-based transcriptomic data analysis of different tissues of the macaúba palm for microsatellite development has been shown to be promising, and we were able to increase the number of useful EST-SSRs as a valuable sequence resource in both *A. aculeata* and the Acrocomia genus. The EST-SSRs reported in this study can potentially be a useful genomic tool in addition to other published SSR, as they provide a potential resource for association mapping of genera-related species. These EST-SSR markers have proven to be useful for both genetic mapping and population structure analysis, facilitating crop breeding of the macaúba palm, as well as studies with other palm trees with economic and environmental value.

Methods

Plant material and RNA isolation

The present study was performed using the tissues of leaves, leaf sheaves, roots, bulbs, fruit (mesocarp and endosperm) and male and female flowers from the macaúba palm (*Acrocomia aculeata*) for RNA isolation and transcriptome sequencing.

Vegetative tissues were collected from eight-month-old seedlings from native plants of Dourado, Amparo and Pedreira from São Paulo State/Brazil. The plants were acclimated for one month in a greenhouse. The male and female flowers were collected from adult plants at the Experimental Unit Santa Elisa – IAC (Agronomic institute of São Paulo/ Campinas/ Sao Paulo State/Brazil). The flower bunches were removed from the base to avoid damage to these materials. The fruit tissues (mesocarp and endosperm) were collected from Santo Antônio de Posse, Amparo from São Paulo State/Brazil, and Ibituruna from Minas Gerais State/Brazil at two fruit development times (developing fruit and ripe fruit).

RNA was isolated using the lithium chloride method [45, 46]; for fruit tissues, RNA was isolated using the perchlorate protocol [47]. Total RNA was treated with RNase-free DNase I (Takara, Kyoto, Japan) for 30 min at 37 °C to remove residual DNA. The RNA quality was verified using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and RNase-free agarose gel electrophoresis. The concentration of the total RNA was further quantified using an RNA NanoDrop (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Illumina sequencing, data filtering and reference-based assembly

The mRNA libraries were synthesized using the TruSeq Stranded mRNA Library Preparation Kit and sequenced using the Illumina paired-end sequencing technology (HiSeq2000). Prior to assembly, the 100 bp paired-end reads were submitted to quality filtering and adapter trimming using the Trimmomatic software version 0.36 [48]. All reads with more than 10% of bases with a poor-quality score ($Q < 20$) or non-coding RNA, as well as ambiguous sequences containing an excess of “N” nucleotide calls or adaptor contamination, were removed. Subsequently, the trimmed reads were mapped against the African oil palm genome [49] using TopHat version 2.0.13 [50] configured to allow up to two mismatches and three indels. The reference-based transcriptome assembly for each individual was performed using Cufflinks version 2.2.1 [51], merging all respective tissues.

EST-SSR search and primer design

The MISA software [52] was employed for scanning EST-SSRs in the transcriptome assembly for each individual. The parameters were adjusted for the identification of perfect mono-, di-, tri-, tetra-, penta- and hexanucleotide motifs with a minimum of 10, 6, 5, 5, 5 and 5 repeats, respectively, and with a maximum distance of 100 bp between two SSRs.

SSR marker primer pairs were designed from the flanking sequences, using the PRIMER3 software [53], with the major parameters for primer lengths of 16–22 bases, GC content of 40–60%, annealing temperature of 50–60 °C, and PCR product size of 100 to 500 bp. To avoid amplicon size deviation generated by the presence of introns during DNA amplification, only PCR products located inside single exons were considered.

For the selection of unique and non-redundant SSRs in one transcript and region, the package BEDTools Version 2.26.0 was applied [54], with an intersection option that checked the overlap between all generated SSRs per individual. The chosen representative SSR was selected based on the nucleotide repeat size, melting temperature of forward and reverse primer, and motif type.

Marker validation

Seventeen *Acrocomia aculeata* (macaúba palm) plants were selected according to the geographic location for polymorphism investigation of the EST-SSRs (Table 3). DNA isolation from young leaf tissues was conducted according to the CTAB DNA extraction protocol [55]. The quality and quantity of DNA were evaluated on a 1% agarose gel using the NanoVue™ Plus Spectrophotometer (GE Healthcare). Contamination with phenol/carbohydrates and proteins was measured based on optical density A260/A230 and A260/A280, respectively.

PCR reactions were performed in a 15 μL total volume containing 20 ng of template DNA, 2.0 μL of each forward and reverse primers (5 μM/μL), 3 μL of Hot Start PCR Master Mix (2X) and 8.2 μL of ultrapure water. PCR amplifications were performed in a thermal cycler (T100 - Bio-Rad) as follows: initial denaturation at 94 °C for 2 min, followed by 30 cycles at 94 °C for 1 min, 55–58 °C (depending on the primers requirement) for 1 min, 72 °C for 1 min and a final extension at 72 °C for 10 min. The amplification products were separated by capillary electrophoresis using a Fragment Analyzer™ 96-capillary Automated CE System (Advanced analytical Technologies, Ames, IA, USA) using the DNF-905 double-stranded DNA Reagent Kit (Advanced Analytical Technologies, Ames, IA, USA). For this analysis, 5 μL of each amplification product was diluted in 19 μL of buffer and placed in 96-well microplates.

Table 3 Macaúba palm samples used for EST-SSR validation and genetic analysis

Species	Plant	Local
Acrocomia aculeata	Plant 1	Rifaina -SP
Acrocomia aculeata	Plant 2	Rifaina -SP
Acrocomia aculeata	Plant 3	Rifaina -SP
Acrocomia aculeata	Plant 430	Jaguariúna - SP
Acrocomia aculeata	Plant 434	Jaguariúna - SP
Acrocomia aculeata	Plant 442	Jaguariúna - SP
Acrocomia aculeata	Plant 545	Itapira - SP
Acrocomia aculeata	Plant 547	Itapira - SP
Acrocomia aculeata	Plant 548	Itapira - SP
Acrocomia aculeata	Plant 359	Serra da Canastra
Acrocomia aculeata	Plant 360	Serra da Canastra
Acrocomia aculeata	Plant 361	Serra da Canastra
Acrocomia aculeata	Plant 3	Luz – MG
Acrocomia aculeata	Plant 4	Luz – MG
Acrocomia aculeata	Plant 5	Luz – MG
Acrocomia aculeata	Plant 119	Capitólio - MG
Acrocomia aculeata	Plant 120	Capitólio - MG

Cross-species SSR transferability

It is well known that the genic regions are highly conserved and provide a cost-effective source of markers for related species, which is especially important for taxa with low microsatellite frequencies or from which microsatellites are difficult to isolate.

In this study, we investigated the transferability of genic SSRs with five plants of six different species, which have been selected based on leaf samples. *Acrocomia totai, Acrocomia intumescens*, the hat palm (*Sabal causiarum*), the juçara palm (*Euterpe edulis*) and the peach palm (*Bactris gasipaes*) have been collected from the palm trees of the Botanical Garden of the Agronomic Institute of Campinas. Samples of the African oil palm (*Elaeis guineensis*) were collected from the germplasm (genotypes BRS C2328, BRS C2528, BRS C7201, BRS C2001 and Manicoré hybrid). Genomic DNA isolation and PCR amplification were performed as described above.

Genetic diversity analysis

Estimates of expected heterozygosity (He), observed heterozygosity (Ho), effective number of alleles (Ne), polymorphism information content (PIC), probability of identity (PI), private alleles per locus, and Shannon's information index (I) were calculated using the software GenAlEx 6.5 [56, 57]. Principal Coordinate analysis (PCoA) was conducted using the software package GenAlEx based on Nei distance [26].

Additional files

Additional file 1: Table S1. Number and frequency of SSR motif and type of motif of 418 EST-SSRs identified in the macaúba palm transcriptome data. (XLSX 11 kb)

Additional file 2: Table S2. Descriptive statistics of 145 EST-SSR markers validated in the macaúba palm. (XLSX 20 kb)

Additional file 3: Table S3. Validation of 145 EST-SSR primer sequences in macaúba palm samples and their transfer to palm tree samples. (XLSX 24 kb)

Additional file 4: Table S4. Twenty EST-SSR that are more informative, primer sequence, and genetic statistics. (XLSX 12 kb)

Additional file 5: Table S5. EST-SSR transferred in *Acrocomia totai*, *Acrocomia intumescens*, hat palm, peach palm, juçara palm, and African oil palm. (XLSX 41 kb)

Abbreviations

EST- SSR: *Expressed Sequence Tag - Simple* Sequence Repeat; He: Expected heterozygosity; Ho: Observed heterozygosity; I: Shannon's information index; Ne: Effective number of alleles per locus; PCoA: Principal coordinate analysis; PI: Probability of identity; PIC: Polymorphism information content

Acknowledgements

Not applicable.

Funding

This work has been supported by The São Paulo Research Foundation, FAPESP (grant 2014/07265–2), the Center for Computational Engineering and Sciences - FAPESP/Cepid (grant 2013/08293–7), and the National Council for Scientific and Technological Development, CNPq (grant 458045/2014–4).

Availability of data and materials

The dataset supporting the conclusions of this article are included within the article (and its additional files).

The RNA-seq data has been deposited in Sequence Read Archives Database (https://www.ncbi.nlm.nih.gov/sra/) under accession number PRJNA489676.

Authors' contributions

BRB developed and screened the EST-SSR markers and performed the molecular biology studies and statistical genetic analyses; LMC and MFC performed the bioinformatics analyses and data mining analyses; and GAGP and CAC designed and coordinated the study. BRB wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Institute of Biology, Laboratory of Genomic and Expression, State University of Campinas, Campinas, Brazil. [2]Agronomic Institute of São Paulo (IAC), Genetic Resources Center, Campinas, SP, Brazil.

Received: 5 April 2018 Accepted: 29 October 2018

References

1. Lorenzi H, Noblick L, Kahn F, Ferreira E. *Flora*. Brasileira: *Arecaceae* (Palms). 1st ed. Nova Odessa: Instituto Plantarum; 2010.
2. Abreu IS, Carvalho CR, Carvalho GMA, Motoike SY. First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae) a promising plant for biodiesel production. Aust J Bot. 2011;59:149–55.
3. Henderson A, Galeano-Garces G, Bernal R. Field guide to the palms of the Americas. Princepton New Jersey: Princeton University Press; 1995.
4. Ciconini G, Favaro SP, Roscoe R, Miranda CHB, Tapeti CF, Miyahira MAMM, et al. Biometry and oil contents of *Acrocomia aculeata* fruits from the Cerrados and Pantanal biomes in Mato Grosso do Sul. Brazil Ind Crops Prod. 2013;45:208–14.
5. Berton LHC. Avaliação de populações naturais , estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (*Acrocomia aculeata*); 2013.
6. Silva JDCE, Barrichelo LEG, Brito JO. Endocarpos De Babaçu E De Macaúba Comparados A Madeira De *Eucalyptus grandis* Para A Produção De Carvão Vegetal. Ipef. 1986;34:31–4.
7. Beltrão NEM, Oliveira MIP. Oleaginosas potenciais do Nordeste para a produção de biodiesel. 2007. https://www.infoteca.cnptia.embrapa.br/bitstream/doc/277417/1/DOC177.pdf. Accessed 15 Jan 2018.
8. Bora PS, Rocha RVM. Macaiba palm: Fatty and amino acids composition on fruits. Cienc y Tecnol Aliment. 2004;4:158–62.
9. Coimbra MC, Jorge N. Fatty acids and bioactive compounds of the pulps and kernels of Brazilian palm species, guariroba (*Syagrus oleraces*), jerivá (*Syagrus romanzoffiana*) and macaúba (*Acrocomia aculeata*). J Sci Food Agric. 2012;92:679–84.
10. USDA. United States Department of Agriculture. 2018. https://apps.fas.usda.gov/psdonline/circulars/oilseeds.pdf Accessed 10 Feb 2018.
11. Evaristo AB, Grossi JAS, de CO CA, Pimentel LD, Motoike SY, Kuki KN. Actual and putative potentials of macauba palm as feedstock for solid biofuel production from residues. Biomass Bioenergy. 2016;85:18–24.
12. Colombo CA, Berton LHC, Diaz BG, Ferrari RA. Macauba: a promising tropical palm for the production of vegetable oil. Oilseeds Fats Crop Lipids. 2018;25:9.
13. Lanes ÉCM, Motoike SY, Kuki KN, Nick C, Freitas RD, Jones FA. Molecular characterization and population structure of the macaw palm, *Acrocomia aculeata* (Arecaceae), ex situ germplasm collection using microsatellites markers. J Hered. 2015;106:102–12.
14. Palliyarakkal MK, Ramaswamy M, Vadivel A. Microsatellites in palm (Arecaceae) sequences. Bioinformation. 2011;7:347–51.
15. Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 2005;23:48–55.
16. Gupta PK, Varshney RK. Cereal genomics: an overview. In: Gupta PK, Varshney RK, editors. Cereal genomics. Dordrecht: Springer Netherlands; 2005. p. 1–18.
17. Khan F. Molecular markers: an excellent tool for genetic analysis. J Mol Biomark Diagn. 2015;6:1–2.
18. Zhao Y, Williams R, Prakash C, He G. Identification and characterization of gene-based SSR markers in date palm (*Phoenix dactylifera* L.). BMC Plant Biol. 2013;12:237.
19. Singh R, Zaki NM, Ting N-C, Rosli R, Tan S-G, E-TL L, et al. Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity. Biologia (Bratisl). 2008;63:227–35.
20. Ting NC, Zaki NM, Rosli R, Low ETL, Ithnin M, Cheah SC, et al. SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. J Genet. 2010;89:135–45.
21. Aberlenc-Bertossi F, Castillo K, Tranchant-Dubreuil C, Chérif E, Ballardini M, Abdoulkader S, et al. In silico mining of microsatellites in coding sequences of the date palm (Arecaceae) genome, characterization, and transferability. Appl Plant Sci. 2014;2(1):1–5.
22. Li YC, Korol AB, Fahima T, Nevo E. Microsatellites within genes: structure, function, and evolution. Mol Biol Evol. 2004;21:991–1007.
23. Zaki NM, Singh R, Rosli R, Ismail I. *Elaeis oleifera* genomic-SSR markers: exploitation in oil palm germplasm diversity and cross-amplification in arecaceae. Int J Mol Sci. 2012;13:4069–88.
24. Mengistu FG, Motoike SY, Caixeta ET, Cruz CD, Kuki KN. Cross-species amplification and characterization of new microsatellite markers for the macaw palm, *Acrocomia aculeata* (Arecaceae). Plant Genet Resour Characterisation Util. 2016;14:163–72.
25. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980;32:314–31.

26. Nei M. Genetic distance between populations. Am Nat. 1972;106:283–92.
27. Dutta S, Kumawat G, Singh BP, Gupta DK, Singh S, Dogra V, et al. Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus Cajan*(L.) Millspaugh]. BMC Plant Biol. 2011;11:17.
28. Hendre PS, Aggarwal RK. Development of genic and genomic SSR markers of robusta coffee (*Coffea canephora* Pierre ex a. Froehner). PLoS One. 2014;9:1–34.
29. Qiu L, Yang C, Tian B, Yang J-B, Liu A. Exploiting EST Databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). BMC Plant Biol. 2010;10:278.
30. Qi W, Lin F, Liu Y, Huang B, Cheng J, Zhang W, et al. High-throughput development of simple sequence repeat markers for genetic diversity research in *Crambe abyssinica*. BMC Plant Biol. 2016;16:1–11.
31. Xia W, Xiao Y, Liu Z, Luo Y, Mason AS, Fan H, et al. Development of gene-based simple sequence repeat markers for association analysis in *Cocos nucifera*. Mol Breed. 2014;34:525–35.
32. Nucci SM, Azevedo-Filho JA, Colombo CA, RHG P, Coelho RM, Mata TL, et al. Development and characterization of microsatellites markers from the macaw. Mol Ecol Resour. 2008;8:224–6.
33. Mengistu F, Motoike S, Cruz C. Molecular characterization and genetic diversity of the macaw palm ex situ germplasm collection revealed by microsatellite markers. Diversity. 2016;8:20.
34. Lanes ÉCM, Motoike SY, Kuki KN, Resende MDV, Caixeta ET. Mating system and genetic composition of the macaw palm (*Acrocomia aculeata*): implications for breeding and genetic conservation programs. J Hered. 2016;107:527–36.
35. Abreu AG, Priolli RHG, Azevedo-Filho JA, Nucci SM, Zucchi MI, Coelho RM, et al. The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). Genet Mol Biol. 2012;35:116–21.
36. Neiva DS, Melo Júnior AF, Oliveira DA, Royo VA, Brandão MM, Menezes EV. *Acrocomia emensis* (Arecaceae) genetic structure and diversity using SSR molecular markers. Genet Mol Res. 2016;15:1–11.
37. De SA, Túlio L, Pacheco G, Nimz T, Vieira N, Guerra MP. The complete plastome of macaw palm [*Acrocomia aculeata* (Jacq .) Lodd . ex Mart.] and extensive molecular analyses of the evolution of plastid genes in Arecaceae. Planta. 2018. https://doi.org/10.1007/s00425-018-2841-x.
38. Katti MV, Sami-Subbu R, Ranjekar PK, Gupta VS. Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. Protein Sci. 2000;9:1203–9.
39. Asp T, Frei UK, Didion T, Nielsen KK, Lübberstedt T. Frequency, type, and distribution of EST-SSRs from three genotypes of *Lolium perenne*, and their conservation across orthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon*, and *Oryza sativa*. BMC Plant Biol. 2007;7:1–12.
40. Katti M, Ranjekar P, Gupta V. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol. 2001;18:1161–7.
41. Feng SP, Li WG, Huang HS, Wang JY, Wu YT. Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). Mol Breed. 2009;23:85–97.
42. Wang Z, Yan H, Fu X, Li X, Gao H. Development of simple sequence repeat markers and diversity analysis in alfalfa (*Medicago sativa* L.). Mol Biol Rep. 2013;40:3291–8.
43. Chen H, Liu L, Wang L, Wang S, Somta P, Cheng X. Development and validation of EST-SSR markers from the transcriptome of adzuki bean (*Vigna angularis*). PLoS One. 2015;10:1–14.
44. Chen H, Wang L, Wang S, Liu C, Blair MW, Cheng X. Transcriptome sequencing of mung bean (*Vigna radiate* L.) genes and the identification of EST-SSR markers. PLoS One. 2015;10:1–15.
45. Zeng Y, Yang T. RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. Plant Mol Biol Report. 2002;20:417.
46. Le Provost G, Herrera R, Paiva JA, Chaumeil P, Salin F, Plomion C. A micromethod for high throughput RNA extraction in forest trees. Biol Res. 2007;40:291–7.
47. Rezaian MA, Krake LR. Nucleic acid extraction and virus detection in grapevine. J Virol Methods. 1987;17:277–85.
48. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
49. Singh R, Ong-Abdullah M, Low E-TL, Manaf MAA, Rosli R, Nookiah R, et al. Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. Nature. 2013;500:335–9.
50. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105–11.
51. Trapnell C. Williams B a, Pertea G, Mortazavi a, Kwan G, van Baren MJ, et al. transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. Nat Biotechnol. 2011;28:511–5.
52. Thiel T, Michalek W, Varshney R, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet. 2003;106:411–22.
53. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012;40:1–12.
54. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2.
55. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. Focus. 1990;12:13–5.
56. Peakall R, Smouse PE. GenALEx 6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol Notes. 2006;6:288–95.
57. Peakall R, Smouse PE. GenALEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012;28:2537–9.

ADDITIONAL FILES

Additional file 1: Table S1. (.xls format) Number and frequency of SSR motif and type of motif of 418 EST-SSRs identified in the macaúba palm transcriptome data.

SSR Motif	Type of Motif	Number	Frequency (%)
Dinucleotide	CT	8	1.91
	AG	2	0.48
	TC	3	0.72
	TG	1	0.24
	AC	1	0.24
Trinucleotide	ATA/ATT/CAA/GTA/TAA/GTT/TTG/TGT	1	0.24
	CAT/TCG	2	0.48
	ATC/CGA/CGT/TGA	3	0.72
	ACC/AGA/GTG/TCA	4	0.96
	ATG/CAC	5	1.2
	GAT	6	1.44
	CGC/GAA/GCG/TGC	7	1.67
	AGC/CCA/CTT/GCA/GCT/TCT/TGG/TTC	9	2.15
	CTG	10	2.39
	CTC	12	2.87
	CGG/GGT/TCC	13	3.11
	AAG/AGG/GCC	14	3.35
	GGC	15	3.59
	CAG	16	3.83
	GGA	19	4.55
	CCG	20	4.78

	CCT	21	5.02
	GAG	30	7.18
Hexanucleotide	CCACCG / CCGTCT / GCCGAT / GGTGGA	1	0.24
Compound	(ACC)5tccaccaccaccacaaccggtgtctgtgacaccagcaaggccaatcccacaaagtg(CAC)5	1	0.24
	(GAG)5atcgtggtttacaatccggcgaaggttggacgtgttgggttagggtttcgtggtggt(GAG)5	1	0.24
	(GCC)7aataagggaaaaagaatatcctcctcctttctgttggctatcaaaggcagcctacgca(AGC)5	1	0.24
	(GCA)5ataccaagggcaatctttccagcaccaacatgaaggaggacatcactaccaacct(CAG)5	1	0.24
	(TG)8(AG)9 / (TGA)6(GGT)5 / (GAT)5c(GGA)8	1	0.24
	(CCT)5ccggcgtcgccgagcctcccatctccaacgccctcatggccgccctcaagcgcgcccaggccaaccagcgccgtggctgcccgg(AGC)5	1	0.24
	(TAC)7(TAA)6 / (TCA)6(GCA)5	1	0.24
	(CCT)6(CAT)5ctgaggaggagactttgttcgggtcttgtttgttcatgagtccccactgttggaagag tttgagctcatttatcaactttgcatggggatcgtcgtcg(TCT)6	1	0.24
	(TCC)5tcttcttcttctcctcgatccgttaacacgcggcctccgccggcggatatggtggcggagaccgac(GAG)5	1	0.24
	(CCT)5(CAT)5(CGT)5 / (CTC)7(CGC)6 / (CGC)6a(GCC)5 / (GGC)7(GGA)6 / (AAG)5gacaaga(AGC)5	1	0.24
	(CAT)5cccaccagcgccaccaccacctccaccgagaccaaccacattgttccccatgctcggag(CGC)5	1	0.24
	(CGC)5tgccctctccttcctcttcttcctcaaaccccctcatcccctccaaccacttctttcctt(CCG)6	1	0.24

Additional file 2: Table S2. (.xls format) Descriptive statistics of 145 EST-SSR markers validated in the macaúba palm.

SSR	Na	Ne	I	Ho	He	PIC	PI
Acro2	6	2.398	1.238	0.353	0.583	0.555	0.202
Acro3	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro6	3	1.273	0.444	0.000	0.215	0.204	0.628
Acro10	5	1.456	0.701	0.059	0.313	0.301	0.484
Acro14	3	1.273	0.444	0.000	0.215	0.204	0.628
Acro15	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro16	2	2.000	0.693	1.000	0.500	0.375	0.375
Acro17	2	1.841	0.649	0.000	0.457	0.352	0.399
Acro20	3	1.441	0.588	0.353	0.306	0.285	0.503
Acro22	2	1.882	0.662	0.000	0.469	0.359	0.392
Acro23	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro24	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro25	3	2.612	1.024	0.000	0.617	0.544	0.220
Acro27	2	1.262	0.362	0.000	0.208	0.186	0.649
Acro29	3	2.206	0.869	0.471	0.547	0.449	0.304
Acro30	3	1.438	0.578	0.000	0.304	0.281	0.507
Acro32	4	2.604	1.115	0.000	0.616	0.555	0.209
Acro33	2	1.410	0.466	0.000	0.291	0.248	0.545
Acro34	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro36	8	3.879	1.595	0.533	0.742	0.703	0.106
Acro37	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro38	3	2.206	0.869	0.000	0.547	0.449	0.304
Acro39	6	4.738	1.668	0.471	0.789	0.759	0.074
Acro41	2	1.061	0.133	0.059	0.057	0.055	0.891
Acro42	5	3.729	1.410	0.176	0.732	0.683	0.120
Acro44	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro45	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro46	2	2.000	0.693	0.059	0.500	0.375	0.375
Acro47	4	2.429	1.038	0.000	0.588	0.513	0.245
Acro48	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro49	4	1.420	0.626	0.067	0.296	0.282	0.510
Acro50	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro52	6	2.979	1.395	0.235	0.664	0.631	0.146
Acro54	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro55	2	1.882	0.662	0.000	0.469	0.359	0.392
Acro57	2	1.679	0.594	0.438	0.404	0.323	0.437
Acro58	3	1.273	0.444	0.000	0.215	0.204	0.628
Acro62	5	3.485	1.400	0.231	0.713	0.669	0.127
Acro63	2	1.133	0.234	0.000	0.117	0.110	0.786
Acro64	7	3.284	1.443	1.000	0.696	0.646	0.143
Acro69	3	1.908	0.750	0.059	0.476	0.387	0.363

Acro71	4	2.276	1.005	0.000	0.561	0.495	0.259
Acro72	2	1.192	0.298	0.059	0.161	0.148	0.717
Acro73	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro74	3	2.688	1.035	0.235	0.628	0.550	0.217
Acro76	3	1.273	0.444	0.118	0.215	0.204	0.628
Acro77	3	1.556	0.656	0.000	0.357	0.325	0.445
Acro80	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro82	3	2.206	0.869	0.000	0.547	0.449	0.304
Acro84	4	1.449	0.644	0.176	0.310	0.293	0.493
Acro87	3	1.841	0.804	0.235	0.457	0.411	0.341
Acro88	4	1.865	0.885	0.000	0.464	0.425	0.326
Acro89	3	2.109	0.846	0.118	0.526	0.436	0.314
Acro91	3	2.165	0.859	0.059	0.538	0.444	0.308
Acro92	2	1.637	0.578	0.176	0.389	0.314	0.449
Acro93	6	3.959	1.563	0.000	0.747	0.715	0.097
Acro94	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro97	2	1.410	0.466	0.118	0.291	0.248	0.545
Acro98	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro99	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro100	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro101	3	1.841	0.804	0.000	0.457	0.411	0.341
Acro102	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro103	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro105	2	1.969	0.685	0.000	0.492	0.371	0.379
Acro106	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro108	4	2.701	1.150	0.353	0.630	0.575	0.192
Acro109	3	1.615	0.678	0.000	0.381	0.340	0.424
Acro110	2	1.438	0.483	0.125	0.305	0.258	0.530
Acro111	6	4.780	1.670	0.714	0.791	0.761	0.073
Acro113	2	1.133	0.234	0.000	0.117	0.110	0.786
Acro116	3	1.438	0.578	0.000	0.304	0.281	0.507
Acro118	7	3.124	1.422	0.118	0.680	0.634	0.148
Acro121	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro122	3	1.795	0.753	0.000	0.443	0.384	0.369
Acro124	2	1.061	0.133	0.059	0.057	0.055	0.891
Acro125	4	2.276	1.005	0.000	0.561	0.495	0.259
Acro126	6	3.853	1.510	0.353	0.740	0.700	0.107
Acro130	3	1.795	0.753	0.000	0.443	0.384	0.369
Acro136	3	1.615	0.678	0.000	0.381	0.340	0.424
Acro137	3	1.145	0.291	0.067	0.127	0.123	0.767
Acro138	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro142	3	1.316	0.485	0.000	0.240	0.227	0.591
Acro143	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro144	4	2.079	0.955	0.118	0.519	0.466	0.284
Acro145	4	2.340	1.045	0.647	0.573	0.515	0.241
Acro146	3	2.014	0.848	0.118	0.503	0.437	0.313

Acro147	4	1.585	0.734	0.063	0.369	0.344	0.423
Acro148	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro149	3	2.240	0.876	0.000	0.554	0.452	0.300
Acro151	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro152	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro153	5	2.641	1.162	0.385	0.621	0.555	0.210
Acro155	2	1.710	0.606	0.000	0.415	0.329	0.428
Acro156	2	1.133	0.234	0.125	0.117	0.110	0.786
Acro158	2	1.262	0.362	0.000	0.208	0.186	0.649
Acro159	2	1.125	0.224	0.000	0.111	0.105	0.797
Acro160	3	1.438	0.578	0.000	0.304	0.281	0.507
Acro162	3	1.273	0.444	0.000	0.215	0.204	0.628
Acro164	5	2.651	1.197	0.000	0.623	0.567	0.198
Acro165	3	2.964	1.093	0.412	0.663	0.589	0.188
Acro166	2	1.778	0.630	0.059	0.438	0.342	0.412
Acro167	3	1.471	0.602	0.000	0.320	0.294	0.488
Acro168	4	1.488	0.689	0.000	0.328	0.313	0.467
Acro170	3	2.109	0.846	0.000	0.526	0.436	0.314
Acro171	5	3.341	1.391	0.294	0.701	0.662	0.128
Acro172	5	2.964	1.251	1.000	0.663	0.602	0.174
Acro174	2	1.562	0.546	0.000	0.360	0.295	0.475
Acro178	3	1.127	0.264	0.059	0.112	0.109	0.791
Acro180	2	1.710	0.606	0.000	0.415	0.329	0.428
Acro182	2	1.262	0.362	0.235	0.208	0.186	0.649
Acro183	7	3.635	1.557	0.059	0.725	0.690	0.110
Acro187	4	2.432	1.069	0.462	0.589	0.534	0.224
Acro188	5	1.456	0.701	0.059	0.313	0.301	0.484
Acro189	7	3.360	1.476	0.412	0.702	0.659	0.132
Acro190	7	3.122	1.437	0.438	0.680	0.642	0.140
Acro191	5	3.705	1.426	0.588	0.730	0.685	0.118
Acro192	4	2.141	1.006	0.000	0.533	0.491	0.260
Acro193	4	1.889	0.916	0.000	0.471	0.439	0.312
Acro195	2	1.142	0.245	0.000	0.124	0.117	0.774
Acro196	2	1.133	0.234	0.000	0.117	0.110	0.786
Acro197	2	1.800	0.637	0.000	0.444	0.346	0.407
Acro198	3	1.510	0.628	0.000	0.338	0.309	0.467
Acro199	4	1.651	0.790	0.000	0.394	0.370	0.391
Acro201	9	4.962	1.857	0.500	0.798	0.776	0.063
Acro203	6	3.853	1.493	0.765	0.740	0.700	0.108
Acro205	11	5.020	1.939	0.250	0.801	0.777	0.063
Acro206	6	3.556	1.488	0.125	0.719	0.680	0.117
Acro208	4	2.964	1.199	0.353	0.663	0.605	0.172
Acro209	3	2.206	0.869	0.000	0.547	0.449	0.304
Acro210	5	2.922	1.226	0.333	0.658	0.594	0.181
Acro211	2	1.867	0.657	0.067	0.464	0.357	0.395
Acro212	6	2.876	1.323	0.176	0.652	0.612	0.161

Acro213	4	1.360	0.573	0.176	0.265	0.253	0.552
Acro214	2	1.133	0.234	0.000	0.117	0.110	0.786
Acro215	2	1.841	0.649	0.000	0.457	0.352	0.399
Acro216	1	1.000	0.000	0.000	0.000	0.000	1.000
Acro217	3	1.438	0.578	0.000	0.304	0.281	0.507
Acro218	5	2.535	1.162	0.118	0.606	0.552	0.209
Acro220	3	2.388	0.956	0.824	0.581	0.500	0.256
Acro221	4	3.778	1.357	0.235	0.735	0.686	0.119
Acro223	3	1.966	0.808	0.000	0.491	0.415	0.335
Acro225	5	4.031	1.456	0.063	0.752	0.709	0.105
Acro227	7	4.446	1.681	0.000	0.775	0.744	0.082
Acro228	3	2.014	0.848	0.706	0.503	0.437	0.313

Additional file 3: Table S3. (.xls format) Validation of 145 EST-SSR primer sequences in macaúba palm samples and their transfer to palm tree samples.

EST-SSR	Tissue	Oil palm chromosome	Motif	Size	Forward primer	Reverse primer	Annealing temperature	Amplicon size	Gene function - oil palm annotation
Acro2	veget., flower,fruit	Chromosome 1	(CCT)6	18	ATCTCCTTGTACCTGAGGATG	AGGAGAAGGAGAAGAAGCTG	58	244	uncharacterized LOC105032745
Acro3	veget., flower	Chromosome 1	(GAA)6	18	TCGAGTAGAAGTCGAGAGAGA	GTTTGTTTCTTGTGCTGTTG	55	191	peroxidase 19
Acro6	veget., flower,fruit	Chromosome 1	(TCC)5	15	GCTAATAAAGAAGTCCGGAAG	CCTGTACTTAAACTCCCCTGT	55	159	uncharacterized LOC105046976
Acro10	veget.,flower,fruit	Chromosome 1	(CCG)5	15	TGTTGTATGTGGTAGCGTTG	AACTTCGACCTATTGCTCTTC	55	133	probable protein phosphatase 2C 64
Acro14	veget.,flower,fruit	Chromosome 1	(ATG)5	15	CACCTTTATCGAATCACTGAC	GAATCGAGGTTGAAAGAAGAG	55	264	H/ACA ribonucleoprotein complex non-core subunit NAF1-like
Acro15	flower ,fruit	Chromosome 1	(CTC)6	18	AAACTAGCAGGCGAATAGAAG	GTAATTATGATGAGGACCACG	55	100	uncharacterized LOC105047798
Acro16	flower, fruit	Chromosome 1	(GCC)8	24	GTCATATGGCTGGTGAGATT	GTTCCTTCTCTTGGTGGAAT	55	270	probable polygalacturonase
Acro17	flower	Chromosome 1	(GGC)5	15	GTGAGAATGGAAATGAGACAG	GAGCTCCTCTAGATACCTCCA	55	198	uncharacterized LOC105044230
Acro20	veget.,flower,fruit	Chromosome 2	(CCT)8	24	CCACCCTTAAGTTCATCTTCT	GACTGTTGGTGTTAAGGTTCA	55	294	protein LTV1 homolog
Acro22	veget.,flower	Chromosome 2	(TGG)5	15	TTGGAGTAGGGGTAGGTTTAC	CTGAGCCAGATGTCCAAC	58	382	uncharacterized LOC105036089
Acro23	veget.,flower	Chromosome 2	(TCC)5	15	CCACAAGTTATCACAGGACAT	ACCTCTACAACCTCCATGAAT	55	162	transcription factor HEC3-like
Acro24	veget.,flower	Chromosome 2	(CAG)5	15	GGCATATTGATCAGGCTAGTA	CATTATCTCCCACTAGGCTCT	55	145	uncharacterized LOC105038148
Acro25	veget.,flower,fruit	Chromosome 2	(GAG)5	15	ACAAGCTCATATGTCTTGGAG	TCCTTCTCCTATCTCTCCTTC	58	378	heat stress transcription factor A-3-like
Acro27	veget.,flower,fruit	Chromosome 2	(AAG)7	21	GCAAGAGGAAACAATGTGTAG	TCTTCACCACATTCTCATAGG	55	296	uncharacterized LOC105038454
Acro29	veget.,fruit	Chromosome 2	(TCG)6	18	GGTCAACAACTCTAACAGCAC	TACATCGAGACATCAAGTCCT	55	251	serine/threonine-protein kinase-like protein CCR4
Acro30	veget.,fruit	Chromosome 2	(TCT)5	15	CTTCTTCATCCAAATACTCCC	CGAGATCTTGAGTCATTTGTC	55	245	uncharacterized LOC105038729
Acro32	veget.,flower,fruit	Chromosome 2	(CAG)5	15	GAACCTCTACGTCAGCCC	ATCAGCAGAACCTCTCTCTTC	58	159	uncharacterized LOC105039591
Acro33	veget.,flower,fruit	Chromosome 2	(CGC)5	15	TCCCTTACTACTCTCCACTCC	ATTTCACCATCTACGACGAC	58	275	squamosa promoter-binding-like protein 9

Acro34	flower	Chromosome 2	(GAG)7	21	ACCCATCTCTCTACGATCTTC	CCTAACGTCTCCTCTCGTAAT	55	326	autophagy-related protein 13
Acro36	fruit	Chromosome 2	(CCT)5	15	GCAGCTAGGGTTTCTCTTG	ATGACCATGTAGTCAGTCCAG	55	180	protein lojap-related, mitochondrial
Acro37	veget.,flower,fruit	Chromosome 3	(ATG)6	18	CTGGTATACAAATATGGGCAG	CAGTTTGCTGTGGATAGACAT	55	178	uncharacterized LOC105040260
Acro38	veget.,flower,fruit	Chromosome 3	(CTT)5	15	AACATATACGCCAAGGTCTTC	TATCTAGGGTTTAATCGGACC	58	389	DEAD-box ATP-dependent RNA helicase 39
Acro39	veget.,flower	Chromosome 3	(GCG)7	21	GATGTTATGCTCAACTCCATC	GTACCATAGCTTCCTCCACTT	55	381	leucine-rich repeat extensin-like protein 5
Acro41	veget.,fruit	Chromosome 3	(CCG)5	15	ATAAACCCTAGAGTTGGCATC	GTCTTCGAGGACCTCTTCTT	58	252	putative E3 ubiquitin-protein ligase LIN-1
Acro42	veget.,flower,fruit	Chromosome 3	(GGC)6	18	CTCCAAGAATCCCTTCTCTC	CCTGATAGTACTCCTGCCTCT	55	175	uncharacterized LOC105040871
Acro44	veget.,flower	Chromosome 3	(CTG)5	15	ACTTGAGGACGAAGTCGAG	TCACCAACATATACCGTAAGC	58	375	PHD finger protein MALE MEIOCYTE DEATH 1
Acro45	veget.,flower,fruit	Chromosome 3	(CAG)5	15	TACTATCCGGAGAGAGGTTCT	ATCTGAGAATCATAGGGATGG	55	296	leucine-rich repeat extensin-like protein 2
Acro46	veget.,flower,fruit	Chromosome 3	(ATC)5	15	CAGATTATAGCACAGCTGGAG	AGTGACTTGAAGCTCATGTTG	55	398	uncharacterized LOC105041692
Acro47	veget.,flower	Chromosome 3	(GTG)8	24	AAATATCCAGAGAGCCTGTTC	ATACATCAACCTCTACAGCCA	58	383	uncharacterized LOC105041760
Acro48	veget.,fruit	Chromosome 3	(CAC)5	15	CTGATACTGGTACCACCCATA	GTAGAAGTGTCCGTCAGTGAT	55	263	zinc finger protein 8-like
Acro49	veget.,fruit	Chromosome 3	(TTC)5	15	ATCACTGACGGACACTTCTAC	CTGTGGAGAAGTTGATGAAGA	58	375	zinc finger protein 8-like
Acro50	veget.,flower	Chromosome 3	(GGC)5	15	ACGAAGACTATTACTACGCCC	GTTCATGTATGTCGGCTTCT	58	291	L-type lectin-domain containing receptor kinase S.6
Acro52	flower,fruit	Chromosome 3	(GAG)6	18	GAGATAAGAGAGCCGAGAAAG	AAAAATCCTCCTCCTCGTT	58	319	ethylene-responsive transcription factor ERF023-like
Acro54	fruit	Chromosome 3	(CCA)5	15	TCTTTCCATCAGAATAGCGT	GTACAGATTGAGATGCCTGAG	58	161	EG45-like domain containing protein
Acro55	fruit	Chromosome 3	(TCC)7	21	ATCACCAAAGCCAACCTC	GAAAACCAAGTAAGGTGGAGT	58	184	chloroplast processing peptidase-like
Acro57	veget., flower	Chromosome 4	(CTC)5	15	AGAAAGAGCAAGAAGAAGAGG	GAGATGTGTTGCTTGAGACC	58	268	uncharacterized LOC105042713
Acro58	veget.,flower,fruit	Chromosome 4	(TCT)6	18	CTCAACACTAGCGTAACGAAT	GTGAAAGCTATGATGATGGAG	58	317	nucleolin-like
Acro62	veget.	Chromosome 4	(CCT)5	15	TCTATTATTCCCTGCACCAC	ACTGCTCGAATAGGATTAGGT	58	132	protein GPR107-like
Acro63	veget.,flower,fruit	Chromosome 4	(CGG)6	18	GTGGGAGAACATCAACAAGTA	ATCTGTATAGTGCCTGACCCT	58	348	trihelix transcription factor GTL1-like
Acro64	veget.,flower	Chromosome 4	(CTG)12	36	GTATGGATGTCGTCGTTGAT	GACTATGGTAATGGACCAACA	55	168	zinc finger protein ZAT9-like

Acro69	fruit	Chromosome 4	(GGC)6	18	CCAGAAGACCCAATCATAGTA	GATTCTCCTCTGCATAGCC	58	388	probable alpha-1,6-mannosyltransferase MNN10
Acro71	veget., flower	Chromosome 5	(CGA)5	15	AGCAGCTTGAATACTCATCAC	ATACCTCATCGTAGACCCATT	55	303	protein trichome birefringence-like 33
Acro72	veget.,flower,fruit	Chromosome 5	(GGA)6	18	GATGTAGATGAGAACGACAGC	CCTCCTAAACAAGCATCTCTT	55	244	uncharacterized LOC105044928
Acro73	veget.,flower,fruit	Chromosome 5	(CGT)5	15	AAGAACGTCTACAGGGGAAT	CATCAAGGCCTTATCCAAC	55	363	ethylene-responsive transcription factor RAP2-3-like
Acro74	veget.,flower,fruit	Chromosome 5	(GCG)5	15	GTGAGTACATGTAGCCCTGAG	GAATCTGTAGCTAAGGAGGCT	55	226	pollen-specific leucine-rich repeat extensin-like protein 1
Acro76	veget.,flower	Chromosome 5	(GAA)7	21	TAAGATCTTGATCGTGTCACC	AGGGTCTTCTGATATGCTCTC	55	262	uncharacterized LOC105045915 [
Acro77	veget., flower	Chromosome 5	(TCA)7	21	TCTACTTCAGGTTCACCAAGA	CTAGACCTGTGTCGTCAGAAC	55	395	uncharacterized LOC105045867
Acro80	veget.,flower,fruit	Chromosome 5	(TCC)9	27	ACTTGTCCAACCCTAATGAAC	GGTTTATCTCCTCCCACTCT	55	344	probable anion transporter 1, chloroplastic
Acro82	veget.	Chromosome 6	(CAC)7	21	TGCTGCCACTACCTATATCAT	GGATGGTGTCGTAAGTGTAGA	55	179	subtilisin-like protease
Acro84	veget.,flower,fruit	Chromosome 6	(GAG)6	18	CTGCTGCAGTAAAGAGAGAGA	GTTGTAGCCATGATAGGAAGA	55	121	calcium-dependent protein kinase isoform 2-like
Acro87	veget.,flower,fruit	Chromosome 6	(CCG)8	24	CAGACTCCTCATACGACCC	GAACCCCACTACAACAAGAT	58	238	cytidine deaminase 1-like
Acro88	flower	Chromosome 6	(TGC)7	21	GGTTGTCTATTGATGAGTCCA	CAAATGAATATGCTTCACCC	58	181	uncharacterized LOC105046428
Acro89	flower,fruit	Chromosome 6	(TCC)8	24	CCTTGTAATTCTGACCTTCCT	CGTCTTCATCTCAGTCTTCTG	55	119	uncharacterized protein At4g00950
Acro91	fruit	Chromosome 6	(CGC)6	18	CTTCCCTTTCTCGTTTTCTC	AGACGTTAATGTCCTTCATCC	55	338	fatty-acid-binding protein 1
Acro92	fruit	Chromosome 6	(AGG)5	15	GCACGTATACTCCATTGATGA	CATTCCCTCCTCAACCTC	55	394	9-cis-epoxycarotenoid dioxygenase, chloroplastic-like
Acro93	veget.,flower,fruit	Chromosome 7	(GCA)5	15	ATCTAGTCTCAGCCTGACCA	GAAGGGGAGAGAGAGGAGT	58	213	uncharacterized LOC105048419
Acro94	veget.	Chromosome 7	(AAG)7	21	TTCCTTCCACCTTCTCTAAAC	AGGGTTTATGATGATAGCGA	58	109	RING-H2 finger protein ATL44-like
Acro97	veget.,flower,fruit	Chromosome 7	(GCT)12	36	ATTCTTCCCCTCTAAGAACCT	CTCAATTCCCTGCAACTATAC	55	302	uncharacterized LOC105049151
Acro98	veget.,flower	Chromosome 7	(AGA)5	15	TCCAGAGTTTAGAGGTGACAA	GCATTCTTCTTCCTCTACCAT	55	226	uncharacterized LOC105049157
Acro99	veget.,flower	Chromosome 7	(CGC)5	15	ATAACTACTTCACATCGCAGC	CTCTTGAAGTAGCCCCTGT	58	265	leucine-rich repeat extensin-like protein 2
Acro100	veget.,flower,fruit	Chromosome 7	(CTG)6	18	CTTGACCTTGTATCCCCTAAG	GCAGCAGAAGAGAAGAGAAA	55	369	uncharacterized LOC105049239
Acro101	flower, fruit	Chromosome 7	(GGA)5	15	GTACATGTCAAGAACGACGAT	TCTCTTCTTCTCCCTCTCTTC	58	283	thioredoxin F, chloroplastic-like

Acro102	flower, fruit	Chromosome 7	(CCA)5	15	GGCTAAGATCATTAATGGGAC	GGACCATACCAATTTCCTTAG	55	268	protein CHUP1, chloroplastic-like
Acro103	veget.	Chromosome 8	(AGC)11	33	GTTGAGAGGGAAGAAGGACT	GTCGACTACATCCAGAAAGG	58	369	uncharacterized acetyltransferase At3g50280-like
Acro105	veget., flower	Chromosome 8	(CGG)5	15	CTGGTAGGGGATATCGTAGTC	GCCACCTCATCTTCTACATC	55	194	uncharacterized LOC105049445
Acro106	veget., flower,fruit	Chromosome 8	(CGG)5	15	CACCACCAATACTCCAACTC	ATAAATTCCGGTAGAGTCCC	55	239	type I inositol 1,4,5-trisphosphate 5-phosphatase 13-like
Acro108	flower	Chromosome 8	(GCT)6	18	GTTATCTGATGTTTTCTCCCC	GAGCAAGTCAACAATTCTACC	55	242	transcription factor RF2a-like
Acro109	veget.,fruit	Chromosome 9	(AGG)5	15	AACTTCCTGGTAGGCAAACT	TTCAACCTTCTCTTCCACTCT	55	289	desiccation-related protein PCC13-62-like
Acro110	veget.,flower	Chromosome 9	(CTT)5	15	GAGAGCAACAGCATAATTGTC	CTTCAGAGGCAGGAGTTAGAT	55	246	endoglucanase 12-like
Acro111	veget.	Chromosome 9	(ATG)5	15	CAGAAAAGAGACCAGTTGATG	GTCACAGATTCTAAGCATTGG	55	181	probable LRR receptor-like serine/threonine-protein kinase At3g47570
Acro113	veget.	Chromosome 9	(CCG)5	15	GTTATTCCTCATGGCTGATG	ACATATGAGTAATGGAGGGCT	58	196	LOB domain-containing protein 25-like
Acro116	veget., flower,fruit	Chromosome 10	(GCC)6	18	CTCTTCACGCTCACAATTAAC	AGAAGAAGAAGGAGAGGATGA	55	340	probable inactive receptor kinase At5g67200
Acro118	veget.,flower	Chromosome 10	(AGC)6	18	GATGCTAATATACCGCACTTG	TAGTAGTCTGACTGAGGCTCG	55	336	uncharacterized LOC105052962
Acro121	veget., flower,fruit	Chromosome 10	(CAT)9	27	CATTACCACTGGTGTCATTCT	TGCTTGCTACTGATGATACCT	55	303	nucleolar protein 14
Acro122	veget.,flower	Chromosome 10	(ACC)5	15	AAAGGATCTGGAGACTCTGAA	ACGAGTTCTCCTATCGTTGAT	55	343	protein LONGIFOLIA 2-like
Acro124	flower,fruit	Chromosome 10	(CGC)5	15	AAAGACAACTGGAGAGACACA	GCTCTAATCCATCTAGGGTTC	55	229	uncharacterized LOC105052490
Acro125	veget.,flower,fruit	Chromosome 11	(GTA)5	15	GTGTTCGATAGCAGTATCGTC	TTACAGCCATAGTAAAGTCGC	55	196	nucleolin-like
Acro126	veget.,flower,fruit	Chromosome 11	(TTC)9	27	CCTATGCTGCTAGGTAATTCA	GTTGAGTCTGAAAGGGAGAAG	55	305	H/ACA ribonucleoprotein complex subunit 4-like
Acro130	veget.,flower,fruit	Chromosome 11	(TCT)6	18	GAACTCCTTGATCTTCTCCTC	CGTTAGGGATAAGAAAGTCGT	58	307	uncharacterized mitochondrial protein ymf11
Acro136	veget.	Chromosome 12	(GCA)5	15	GATTTCCACCATCTCCATC	CTCCAGTGTTGTTGTTGATGT	58	163	putative DNA-binding protein ESCAROLA
Acro137	veget.	Chromosome 12	(CAA)5	15	CAATACCTCCATTGTCTACCA	GGTTGTGTTCTCTGTCAACAT	55	146	trihelix transcription factor GTL1-like
Acro138	veget.	Chromosome 12	(AAG)11	33	GAACTCTGGAATCTCCAACTC	ACTTCTCTAGTGGTTGAAGGG	55	153	WD repeat-containing protein

									RUP2
Acro142	veget.,flower,fruit	Chromosome 12	(GGA)6	18	GTAGACGAGGTAGATGCTGC	GCCTCTAAGACCTCCTCCT	58	321	lysM domain receptor-like kinase 3
Acro143	flower	Chromosome 12	(TGC)6	18	CCTCTTTTCTTCTTCTTGCTC	ACTGGGCTTATACTGATCTCC	55	132	UDP-galactose/UDP-glucose transporter 4-like
Acro144	flower	Chromosome 12	(AGG)6	18	AAACTAGGTGTGGTGACAGAA	ACTCCTCTGAGATAATAAAGCG	55	217	ABC transporter B family member 1-like
Acro145	flower	Chromosome 12	(CAG)5	15	CAGGAAGGACTGACAATGATA	CTACACTACTCCCGTAGCTGA	55	372	transcription factor RAX2-like
Acro146	veget.,flower	Chromosome 13	(GGT)6	18	AAGTAGATCAGGTGGTGACAA	GGAAGTCGTCCAGATTTTAAC	55	176	leucine-rich repeat receptor-like serine/threonine-protein kinase BAM1
Acro147	veget.,flower,fruit	Chromosome 13	(TCC)6	18	GTCCTTGCTCTATTTCATCCT	CTCCATCTTTTGTTACCAGTG	55	273	uncharacterized LOC105055922
Acro148	veget.,flower,fruit	Chromosome 13	(CCG)5	15	ATGCTCCCTTATGATCTTCTC	AACCTCAACCTCTTCTCCAC	58	375	flavonoid 3'-monooxygenase-like
Acro149	veget.	Chromosome 13	(CGG)9	27	GATCTTCTCCTCCTCAATTTC	CGATACCCAAGATCAAGAAG	55	335	protein diaphanous homolog 1-like
Acro151	veget.,flower,fruit	Chromosome 13	(GCC)6	18	ATGTTCCGCTTCTAAGGTATC	AGTATACTTCACCAAGCTCCC	55	281	uncharacterized aarF domain-containing protein kinase At1g79600, chloroplastic
Acro152	veget.,flower,fruit	Chromosome 13	(AGG)5	15	ACTGAGAAGGCAAGTCATACA	TCAGGGATCATAACACTCATC	55	186	probable arabinosyltransferase ARAD1
Acro153	veget.,flower,fruit	Chromosome 13	(CAG)6	18	CATGTAGCAGCTCAAGCTAAC	AAGTGAGGATGATCTCTGGAT	55	258	uncharacterized LOC105056687
Acro155	veget.	Chromosome 13	(TCC)7	21	CTGTTCTTCCCTTGGTCTC	CCGGAGATATAGAGGGACTTA	55	397	probable inactive leucine-rich repeat receptor-like protein kinase At5g20690
Acro156	veget., flower, fruit	Chromosome 14	(GCA)8	24	GGCTCTCAGAAGAAGATGAGT	CAGTAAGAAACCTCTCCCATT	55	174	two-component response regulator-like PRR37
Acro158	veget.,flower,fruit	Chromosome 14	(CCG)5	15	GTCCATTTTATTACCTCCAGC	ACCAGCTCTCCTTCCTTAAC	55	368	zinc finger CCCH domain-containing protein 53-like
Acro159	veget.,flower,fruit	Chromosome 14	(TCT)7	21	GTAGATTATATGCCAAGCCTTC	CATCAACACACTGAGGTCTCT	55	352	UPF0496 protein At5g66675-like
Acro160	veget., flower,fruit	Chromosome 14	(GAG)5	15	GTTCATCCTGACAAGAACAAG	TATGAAATTAGTCCCTCCTCC	55	351	chaperone protein dnaJ 49
Acro162	veget.,flower,fruit	Chromosome 14	(CAG)6	18	CAACACGACTATCTACCATCC	AGATCGAGGTAAACCATTCTC	55	123	regulatory protein NPR5-like
Acro164	flower	Chromosome 14	(GGA)5	15	ACTTCGAGGTTCTCCACC	GCATCATGTGCTTGAACTC	58	291	probable calcium-binding protein CML27

Acro165	fruit	Chromosome 14	(GAG)7	21	AGACTGAAGAGACCGAGTTTT	TGTACCACACTCATGCTCC	58	148	F-box protein At3g54460
Acro166	veget.,flower,fruit	Chromosome 15	(GAA)6	18	AACCTTACAATCTCCGTAACC	TATCTCATCACTCTCGAGTCC	58	328	putative pentatricopeptide repeat-containing protein At5g06400, mitochondrial
Acro167	veget.,flower,fruit	Chromosome 15	(AGG)7	21	CATATCGCCTACAACTACCAG	ACCACCATGTTCTTGAACTC	55	308	uncharacterized LOC105061580
Acro168	veget.,flower,fruit	Chromosome 15	(CTG)6	18	AACTTCTTGAACTCCTCATCC	CACCTTCTCTCCCTCTCTTC	58	239	probable zinc metallopeptidase EGY3, chloroplastic
Acro170	veget.,flower,fruit	Chromosome 15	(CTT)7	21	GATCCACCAAGAGAAAAGACT	CTTCAAGATGACAAAGAGGGT	55	137	uncharacterized TPR repeat-containing protein At1g05150-like
Acro171	veget.,flower,fruit	Chromosome 15	(GAG)6	18	AGTCTTCTTACCGTCTGCTCT	CAGATATTCCATCCACAAGAC	55	398	zeaxanthin epoxidase, chloroplastic
Acro172	veget.,flower,fruit	Chromosome 15	(CTT)5	15	GCAGAGACTTCACCTCTTCTT	GTTCAAGTCCTTGTCTTCCTT	55	264	BRI1 kinase inhibitor 1-like
Acro174	veget.,flower,fruit	Chromosome 15	(TCA)5	15	GGTATGTCTTCAAGAAGGGTT	TGCATATCATCAGAGAGGAAG	55	193	translation initiation factor eIF-2B subunit beta
Acro178	veget., flower	Chromosome 16	(AAG)5	15	GAGGATGGAGTAGATTTCGTT	TCTATCAAGCAAGCATCTCTC	55	329	uncharacterized LOC105058955
Acro180	veget., flower,fruit	Chromosome 16	(CTT)6	18	CTCTTCTTCTTCGATCTGGAC	TCTCTCCCTCAAGGATTATCT	55	175	BUD13 homolog
Acro182	flower, fruit	Chromosome 16	(GCG)5	15	GGGTCTACAAGGACAAGAGTC	ACTGATGCCCAGTTATCATC	55	224	U11/U12 small nuclear ribonucleoprotein 31 kDa protein
Acro183	veget., flower,fruit	unplaced	(ATC)5	15	GAGATCTCTTCATCACTCCCT	TCCATGACTGATAGAGGTGAC	58	350	probable alpha-glucosidase Os06g0675700
Acro187	fruit	unplaced	(GCG)5	15	CTATAGAATTTGGAGCAGCAG	AACAGTATAGGCCTCCTGTG	58	216	splicing factor 3B subunit 4-like
Acro188	veget.,flower,fruit	unplaced	(GCC)5	15	GTCTCCTTCTTACCCTTCTCC	TCCATGTACAAGAAGATCCAC	55	236	CBL-interacting protein kinase 6
Acro189	veget.,flower,fruit	unplaced	(CTG)7	21	CCTCCCATCAATAGGTCTAAC	ATCAGGTGATACACAAGCATC	58	274	uncharacterized LOC105060244
Acro190	veget.,flower,fruit	unplaced	(AGC)5	15	AAGAAGAAAGATGAGGGTGAC	TACTCACTTGACCTTGTTTGG	55	157	veget.ative cell wall protein gp1
Acro191	veget., flower,fruit	unplaced	(GCT)7	21	AAGAGAGCGATGATTTCTTG	TATTCTCTCCGCCTCTTCTAT	55	242	sufE-like protein, chloroplastic
Acro192	veget.,flower,fruit	unplaced	(ATG)5	15	TGACAAGGATAGGACTGAGAA	AATCCTATATACATCCCTCGC	55	307	uncharacterized LOC105061170
Acro193	veget.,flower,fruit	unplaced	(GGT)5	15	GATGGTCAGATTTGTACTTGG	ATCTGTACTGCAGGAACTTCA	55	198	FAS1 domain-containing protein SELMODRAFT_448915-like
Acro195	veget.,flower,fruit	unplaced	(CTC)5	15	GTGTTCCTCCTGTTTCTGATA	TCTGACTGTTAGAAAACAGGG	55	385	uncharacterized membrane protein At1g75140-like

Acro196	veget., flower,fruit	unplaced	(GGA)7	21	AGAGTACTCAGTCTTGGGCTT	CAAGAACCCTAACCCTAATTC	55	259	uncharacterized LOC105061463
Acro197	veget.,flower,fruit	unplaced	(GGA)5	15	AAAGTCGACTCCTCTCTTGTT	CAGCAATTCTTCAGTGGATAG	55	381	probable dimethyladenosine transferase
Acro198	veget.	unplaced	(TGG)5	15	ACACGTCTTCATGCTCTAATG	TGAGTTCCTTCCTAAATCCTC	55	281	galactose oxidase-like
Acro199	veget.,flower,fruit	unplaced	(CCG)5	15	GGTCTTCGCTCTATCTCCTT	AGATAGCTCAGTAAAGGGAGG	55	355	uncharacterized LOC105032215
Acro201	veget.,flower	unplaced	(CAG)5	15	GGAGCTAAAGATGAGGAGAAG	CGGATGGGTGTAGAAGTTATT	55	390	uncharacterized LOC105032506
Acro203	flower	unplaced	(AGG)5	15	AGGGTGCCTATTTTGAGATAC	GATGACGTCTCCCTCCTC	58	224	flavonoid 3',5'-hydroxylase 1-like
Acro205	flower	unplaced	(CAC)6	18	AATGCTGTAATAGCCGGAG	GAGAGGAACTACGGGATCTT	55	162	glucan endo-1,3-beta-glucosidase 11-like
Acro206	fruit	unplaced	(CCA)7	21	TATCGTCTCCTCCACCCT	CGTAGGAGAGAATGACGGT	58	341	pentatricopeptide repeat-containing protein At4g36680, mitochondrial-like
Acro208	veget., flower,fruit	unplaced	(TGC)5	15	GGCAACTAAACTCATTGACAG	CAGCAGTCTAACCAGCATTAT	55	367	histone acetyltransferase HAC1-like
Acro209	flower, fruit	unplaced	(GGA)7	21	GTATCCCCTACGAGAAGAAGA	AAGAAGAAGAAGAACCACCC	55	189	protein LIGHT-DEPENDENT SHORT HYPOCOTYLS 6-like
Acro210	veget.	unplaced	(CTG)6	18	ACCAAGAGAAGGATGGATG	ATCTGGAGGAGTACGAGAAGA	55	218	agamous-like MADS-box protein AGL62
Acro211	veget., flower,fruit	unplaced	(CCT)6	18	GACTTTCTTTCCTCCTCCTC	GATTTGAAGGAGGAGGTGTAG	58	203	U11/U12 small nuclear ribonucleoprotein 48 kDa protein
Acro212	veget.	unplaced	(CCA)5	15	GGTTTCTAAGAGGGAGTTTGA	CAGGTGAAGAAGAATTGAGTG	58	206	umecyanin-like
Acro213	fruit	unplaced	(TGC)5	15	CCCCCTTCTACTCTTCTTGA	GTGTAAATAATGTCTCGCTGG	55	221	AP2-like ethylene-responsive transcription factor TOE3
Acro214	flower,fruit	unplaced	(GAG)5	15	GTTTTAGGAAGAGGAGAGGTG	GTGGCAATACTTCCTCGTT	58	133	pentatricopeptide repeat-containing protein At4g38150-like
Acro215	veget.	unplaced	(AAG)5	15	GGTTCTCCTCCTTCCTTCT	GAGCTTAAGAGCATGACAAAG	55	197	separase-like
Acro216	veget., flower,fruit	unplaced	(GAG)5	15	GGTAGGAAAAGTAGTGGCATT	ACAAGTATTTTCTCCAGGCTC	55	302	golgin candidate 2
Acro217	veget., flower	unplaced	(CCT)8	24	ACCAATAAATCCACGTTGTC	GAAGAAAAATGGGAGCTACAG	58	106	F-box protein SNE-like
Acro218	veget.,flower	unplaced	(CCT)7	21	ACTGCAAGAACCCTCCTC	GATTGTTTGACTAGCAGCAAC	58	145	WUSCHEL-related homeobox 4
Acro220	flower	unplaced	(CTC)5	15	CCCAGATACAGCAGATAACAG	CATGTGGTGTTGGAGTATAGG	55	209	WUSCHEL-related homeobox 3-like

Acro221	flower,fruit	unplaced	(GGA)5	15	GATGATGGAAACAGCTCG	CCTCAAAACCCTACCAAAA	55	126	amidophosphoribosyltransferase, chloroplastic-like
Acro223	veget., flower,fruit	unplaced	(GAT)5	15	TCCTGGAGGAGACTTACTAGG	CTTTGTTGACTAGTGCATTCC	55	212	uncharacterized LOC105035882
Acro225	veget.,flower,fruit	unplaced	(GTG)5	15	AGTCTTGAACTCATTCAGCAG	TACTTCAGAGTCTTAGCCACG	58	216	magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase, chloroplastic
Acro227	veget.	unplaced	(GGT)6	18	CAGAGACGACATAGGAATCTG	GCTGCTCTATTTCAACTCAAG	55	393	pentatricopeptide repeat-containing protein At3g12770-like
Acro228	veget.	unplaced	(CCA)7	21	AGCTATGCCACCTCACTG	TGTTTAGGGAGTCAAGTCTCA	55	247	probable leucine-rich repeat receptor-like protein kinase At1g35710

Additional file 4: Table S4. (.xls format) Twenty EST-SSR that are more informative, primer sequence, and genetic statistics.

	PIC	PI	He	Ho	Tissue	Motif	Motif lenght	Forward primer	Reverse primer	Ta	Amplicon size
Acro205	0.777	0.063	0.827	0.250	flower	(CAC)6	18	AATGCTGTAATAGCCGGAG	GAGAGGAACTACGGGATCTT	55	114 - 195
Acro201	0.776	0.063	0.828	0.500	veget.,flower	(CAG)5	15	GGAGCTAAAGATGAGGAGAAG	CGGATGGGTGTAGAAGTTATT	55	312 - 441
Acro111	0.761	0.073	0.820	0.714	veget.	(ATG)5	15	CAGAAAAGAGACCAGTTGATG	GTCACAGATTCTAAGCATTGG	55	154 - 187
Acro39	0.759	0.074	0.813	0.471	veget.,flower	(GCG)7	21	GATGTTATGCTCAACTCCATC	GTACCATAGCTTCCTCCACTT	55	350 - 401
Acro227	0.744	0.082	0.799	0.000	veget.	(GGT)6	18	CAGAGACGACATAGGAATCTG	GCTGCTCTATTTCAACTCAAG	55	348 - 417
Acro93	0.715	0.097	0.770	0.000	veget.,flower,fruit	(GCA)5	15	ATCTAGTCTCAGCCTGACCA	GAAGGGGAGAGAGAGGAGT	58	150 - 213
Acro225	0.709	0.105	0.776	0.063	veget.,flower,fruit	(GTG)5	15	AGTCTTGAACTCATTCAGCAG	TACTTCAGAGTCTTAGCCACG	58	216 - 285
Acro36	0.703	0.106	0.768	0.533	fruit	(CCT)5	15	GCAGCTAGGGTTTCTCTTG	ATGACCATGTAGTCAGTCCAG	55	180 - 221
Acro126	0.700	0.107	0.763	0.353	veget.,flower,fruit	(TTC)9	27	CCTATGCTGCTAGGTAATTCA	GTTGAGTCTGAAAGGGAGAAG	55	302 - 350
Acro203	0.700	0.108	0.763	0.765	flower	(AGG)5	15	AGGGTGCCTATTTTGAGATAC	GATGACGTCTCCCTCCTC	58	164 - 227
Acro183	0.690	0.110	0.747	0.059	veget., flower,fruit	(ATC)5	15	GAGATCTCTTCATCACTCCCT	TCCATGACTGATAGAGGTGAC	58	350 - 404
Acro221	0.686	0.119	0.758	0.235	flower,fruit	(GGA)5	15	GATGATGGAAACAGCTCG	CCTCAAAACCCTACCAAAA	55	126 - 141
Acro191	0.685	0.118	0.752	0.588	veget., flower,fruit	(GCT)7	21	AAGAGAGCGATGATTTCTTG	TATTCTCTCCGCCTCTTCTAT	55	230 - 245
Acro42	0.683	0.120	0.754	0.176	veget.,flower,fruit	(GGC)6	18	CTCCAAGAATCCCTTCTCTC	CCTGATAGTACTCCTGCCTCT	55	133 - 184
Acro206	0.680	0.117	0.742	0.125	fruit	(CCA)7	21	TATCGTCTCCTCCACCCT	CGTAGGAGAGAATGACGGT	58	338 - 371
Acro62	0.669	0.127	0.742	0.231	veget.	(CCT)5	15	TCTATTATTCCCTGCACCAC	ACTGCTCGAATAGGATTAGGT	58	132 - 246
Acro171	0.662	0.128	0.722	0.294	veget.,flower,fruit	(GAG)6	18	AGTCTTCTTACCGTCTGCTCT	CAGATATTCCATCCACAAGAC	55	398 - 449
Acro189	0.659	0.132	0.724	0.412	veget.,flower,fruit	(CTG)7	21	CCTCCCATCAATAGGTCTAAC	ATCAGGTGATACACAAGCATC	58	274 - 307
Acro64	0.646	0.140	0.717	1.000	veget.,flower	(CTG)12	36	GTATGGATGTCGTCGTTGAT	GACTATGGTAATGGACCAACA	55	159 - 180
Acro190	0.642	0.143	0.702	0.438	veget.,flower,fruit	(AGC)5	15	AAGAAGAAAGATGAGGGTGAC	TACTCACTTGACCTTGTTTGG	55	157 - 187

PIC	Polymorphism Information Content
He	Expected heterozygosity
Veget.	Vegetative tissues (root, leaves, bulb)
PI	Probability of Identity

Additional file 5: Table S5. (.xls format) EST-SSR transferred in *Acrocomia totai*, *Acrocomia intumescens*, hat palm, peach palm, juçara palm, and African oil palm.

	Acrocomia totai						*Acrocomia intumescens*					
	Transferabity	Na	Ne	I	Ho	He	Transferability	Na	Ne	I	Ho	He
Acro2	Yes	2	1.220	0.325	0.200	0.180	Yes	2	1.923	0.673	0.000	0.480
Acro3	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro6	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro10	Yes	4	3.333	1.280	0.800	0.700	Yes	2	1.923	0.673	0.000	0.480
Acro14	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro15	Yes	2	1.220	0.325	0.200	0.180	Yes	1	1.000	0.000	0.000	0.000
Acro16	Yes	2	2.000	0.693	1.000	0.500	Yes	2	2.000	0.693	1.000	0.500
Acro17	Yes	2	1.923	0.673	0.000	0.480	Yes	1	1.000	0.000	0.000	0.000
Acro20	Yes	2	1.923	0.673	0.400	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro22	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro23	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro24	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro25	Yes	2	1.600	0.562	0.500	0.375	Yes	1	1.000	0.000	0.000	0.000
Acro27	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro29	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro30	Yes	4	2.778	1.168	0.800	0.640	Yes	2	1.923	0.673	0.000	0.480
Acro32	Yes	5	3.125	1.359	0.400	0.680	Yes	1	1.000	0.000	0.000	0.000
Acro33	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro34	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro36	Yes	6	4.167	1.609	0.600	0.760	Yes	2	1.923	0.673	0.000	0.480
Acro37	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro38	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.471	0.500	0.000	0.320
Acro39	Yes	4	3.571	1.332	0.400	0.720	Yes	1	1.000	0.000	0.000	0.000
Acro41	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro42	Yes	8	7.143	2.025	0.800	0.860	Yes	3	2.381	0.943	0.600	0.580
Acro44	Yes	3	1.515	0.639	0.200	0.340	Yes	2	1.471	0.500	0.000	0.320
Acro45	Yes	4	2.381	1.089	0.400	0.580	Yes	2	1.923	0.673	0.000	0.480
Acro46	Yes	3	2.174	0.898	0.200	0.540	Yes	1	1.000	0.000	0.000	0.000
Acro47	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.923	0.673	0.000	0.480
Acro48	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.000	0.560
Acro49	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro50	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro52	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro54	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro55	Yes	3	2.778	1.055	0.000	0.640	Yes	3	2.273	0.950	0.000	0.560
Acro57	Yes	3	1.515	0.639	0.400	0.340	Yes	1	1.000	0.000	0.000	0.000
Acro58	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro62	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro63	Yes	2	1.220	0.325	0.200	0.180	Yes	2	1.471	0.500	0.000	0.320
Acro64	Yes	6	5.000	1.696	0.800	0.800	Yes	5	4.167	1.505	1.000	0.760
Acro69	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro71	Yes	3	1.852	0.802	0.600	0.460	Yes	2	1.471	0.500	0.000	0.320

Acro72	Yes	3	2.381	0.943	0.200	0.580	Yes	1	1.000	0.000	0.000	0.000
Acro73	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro74	Yes	2	1.471	0.500	0.400	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro76	Yes	3	1.852	0.802	0.400	0.460	Yes	2	1.471	0.500	0.000	0.320
Acro77	Yes	2	1.385	0.451	0.333	0.278	Yes	1	1.000	0.000	0.000	0.000
Acro80	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro82	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro84	Yes	5	4.545	1.557	0.600	0.780	Yes	2	1.923	0.673	0.000	0.480
Acro87	Yes	3	2.778	1.055	0.000	0.640	Yes	2	1.471	0.500	0.000	0.320
Acro88	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro89	Yes	2	1.220	0.325	0.200	0.180	Yes	2	1.471	0.500	0.000	0.320
Acro91	Yes	3	2.632	1.030	0.600	0.620	Yes	2	2.000	0.693	0.200	0.500
Acro92	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro93	Yes	2	1.923	0.673	0.000	0.480	Yes	3	2.778	1.055	0.000	0.640
Acro94	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro97	Yes	2	1.471	0.500	0.400	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro98	Yes	2	1.280	0.377	0.250	0.219	Yes	2	1.471	0.500	0.000	0.320
Acro99	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro100	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro101	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro102	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro103	Yes	1	1.000	0.000	0.000	0.000	Yes	4	2.381	1.089	0.200	0.580
Acro105	Yes	2	2.000	0.693	0.000	0.500	Yes	1	1.000	0.000	0.000	0.000
Acro106	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro108	Yes	2	1.471	0.500	0.400	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro109	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro110	Yes	2	1.923	0.673	0.000	0.480	Yes	3	2.778	1.055	0.400	0.640
Acro111	Yes	3	2.273	0.950	0.400	0.560	Yes	3	2.174	0.898	0.200	0.540
Acro113	Yes	2	1.923	0.673	0.000	0.480	Yes	5	3.125	1.359	0.400	0.680
Acro116	Yes	4	2.381	1.089	0.200	0.580	Yes	3	2.778	1.055	0.000	0.640
Acro118	Yes	3	2.381	0.943	0.600	0.580	Yes	2	1.923	0.673	0.000	0.480
Acro121	Yes	2	1.923	0.673	0.000	0.480	Yes	1	1.000	0.000	0.000	0.000
Acro122	Yes	2	1.923	0.673	0.000	0.480	Yes	1	1.000	0.000	0.000	0.000
Acro124	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro125	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro126	Yes	6	4.545	1.643	0.800	0.780	Yes	2	1.923	0.673	0.000	0.480
Acro130	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro136	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro137	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro138	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro142	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro143	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro144	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro145	Yes	2	2.000	0.693	1.000	0.500	Yes	4	2.941	1.194	1.000	0.660
Acro146	Yes	6	4.167	1.609	0.800	0.760	Yes	3	2.273	0.950	0.000	0.560
Acro147	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro148	Yes	4	3.571	1.332	0.000	0.720	Yes	2	1.471	0.500	0.000	0.320

Acro149	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.923	0.673	0.000	0.480
Acro151	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro152	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro153	Yes	6	4.545	1.643	1.000	0.780	Yes	2	1.923	0.673	0.000	0.480
Acro155	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro156	Yes	3	2.381	0.943	0.600	0.580	Yes	2	1.471	0.500	0.000	0.320
Acro158	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro159	Yes	3	2.174	0.898	0.800	0.540	Yes	2	1.471	0.500	0.000	0.320
Acro160	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro162	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro164	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.778	1.055	0.000	0.640
Acro165	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro166	Yes	3	2.778	1.055	0.000	0.640	Yes	3	2.273	0.950	0.000	0.560
Acro167	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro168	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.600	0.562	0.000	0.375
Acro170	Yes	2	1.923	0.673	0.000	0.480	Yes	3	1.852	0.802	0.200	0.460
Acro171	Yes	2	1.471	0.500	0.000	0.320	Yes	4	3.571	1.332	0.000	0.720
Acro172	Yes	3	2.632	1.030	1.000	0.620	Yes	2	1.471	0.500	0.000	0.320
Acro174	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro178	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro180	Yes	3	1.852	0.802	0.200	0.460	Yes	2	1.471	0.500	0.000	0.320
Acro182	Yes	2	1.724	0.611	0.600	0.420	Yes	1	1.000	0.000	0.000	0.000
Acro183	Yes	3	1.852	0.802	0.600	0.460	Yes	3	2.778	1.055	0.000	0.640
Acro187	Yes	3	1.515	0.639	0.400	0.340	Yes	3	2.273	0.950	0.000	0.560
Acro188	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro189	Yes	4	2.381	1.089	0.400	0.580	Yes	3	2.273	0.950	0.000	0.560
Acro190	Yes	4	3.571	1.314	1.000	0.720	Yes	2	1.724	0.611	0.600	0.420
Acro191	Yes	4	2.778	1.168	0.800	0.640	Yes	2	1.280	0.377	0.250	0.219
Acro192	Yes	3	2.778	1.055	0.000	0.640	Yes	1	1.000	0.000	0.000	0.000
Acro193	Yes	3	2.273	0.950	0.000	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro195	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro196	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro197	Yes	3	1.852	0.802	0.200	0.460	Yes	2	1.471	0.500	0.000	0.320
Acro198	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro199	Yes	6	5.333	1.733	0.500	0.813	Yes	2	1.471	0.500	0.000	0.320
Acro201	Yes	3	2.667	1.040	0.000	0.625	Yes	2	1.923	0.673	0.000	0.480
Acro203	Yes	5	4.167	1.505	0.800	0.760	Yes	1	1.000	0.000	0.000	0.000
Acro205	Yes	4	2.941	1.221	0.200	0.660	Yes	5	3.846	1.471	0.800	0.740
Acro206	Yes	3	2.273	0.950	0.000	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro208	Yes	4	2.381	1.089	0.200	0.580	Yes	2	1.923	0.673	0.000	0.480
Acro209	Yes	4	2.778	1.168	0.800	0.640	Yes	1	1.000	0.000	0.000	0.000
Acro210	Yes	5	4.545	1.557	0.400	0.780	Yes	4	2.381	1.089	0.200	0.580
Acro211	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.273	0.950	0.400	0.560
Acro212	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro213	Yes	5	3.125	1.359	0.400	0.680	Yes	3	2.273	0.950	0.000	0.560
Acro214	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro215	Yes	3	2.273	0.950	0.000	0.560	Yes	4	2.941	1.194	0.200	0.660

Acro216	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.220	0.325	0.200	0.180
Acro217	Yes	2	2.000	0.693	0.600	0.500	Yes	2	1.923	0.673	0.000	0.480
Acro218	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro220	Yes	2	2.000	0.693	1.000	0.500	Yes	2	2.000	0.693	1.000	0.500
Acro221	Yes	4	2.778	1.168	0.600	0.640	Yes	3	2.778	1.055	0.000	0.640
Acro223	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro225	Yes	3	2.778	1.055	0.000	0.640	Yes	2	1.923	0.673	0.000	0.480
Acro227	Yes	3	2.667	1.040	0.000	0.625	Yes	4	3.571	1.332	0.000	0.720
Acro228	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.174	0.898	0.800	0.540

	Hat palm						Peach palm					
	Transferability	Na	Ne	I	Ho	He	Transferability	Na	Ne	I	Ho	He
Acro2	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.471	0.500	0.000	0.320
Acro3	Yes	3	2.273	0.950	0.000	0.560	Yes	2	2.000	0.693	0.600	0.500
Acro6	Yes	2	1.800	0.637	0.000	0.444	Yes	2	1.471	0.500	0.000	0.320
Acro10	Yes	2	1.600	0.562	0.500	0.375	Yes	2	1.471	0.500	0.000	0.320
Acro14	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro15	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro16	Yes	2	1.923	0.673	0.800	0.480	Yes	2	2.000	0.693	1.000	0.500
Acro17	Yes	3	2.273	0.950	0.000	0.560	No	0	0.000	0.000	0.000	0.000
Acro20	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro22	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro23	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro24	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro25	Yes	3	2.273	0.950	0.000	0.560	No	0	0.000	0.000	0.000	0.000
Acro27	Yes	4	2.941	1.194	0.200	0.660	Yes	1	1.000	0.000	0.000	0.000
Acro29	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.000	0.560
Acro30	Yes	3	1.852	0.802	0.200	0.460	Yes	3	2.273	0.950	0.800	0.560
Acro32	Yes	2	1.600	0.562	0.000	0.375	Yes	6	5.000	1.696	0.400	0.800
Acro33	Yes	3	2.273	0.950	0.000	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro34	Yes	4	3.333	1.280	0.600	0.700	Yes	2	1.923	0.673	0.800	0.480
Acro36	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro37	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro38	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.471	0.500	0.000	0.320
Acro39	Yes	2	1.471	0.500	0.000	0.320	Yes	4	3.571	1.332	0.000	0.720
Acro41	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro42	Yes	3	2.273	0.950	0.000	0.560	Yes	3	2.778	1.055	0.000	0.640
Acro44	Yes	4	3.571	1.332	0.000	0.720	Yes	4	3.333	1.280	1.000	0.700
Acro45	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro46	Yes	4	3.571	1.332	0.000	0.720	Yes	1	1.000	0.000	0.000	0.000
Acro47	Yes	3	2.273	0.950	0.000	0.560	No	0	0.000	0.000	0.000	0.000
Acro48	Yes	4	4.000	1.386	0.000	0.750	Yes	1	1.000	0.000	0.000	0.000
Acro49	Yes	3	2.667	1.040	0.000	0.625	Yes	2	1.220	0.325	0.200	0.180
Acro50	Yes	3	2.632	1.030	0.200	0.620	Yes	2	1.471	0.500	0.400	0.320
Acro52	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.600	0.560
Acro54	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro55	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320

Acro57	Yes	3	2.273	0.950	0.000	0.560	Yes	4	3.846	1.366	0.600	0.740
Acro58	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro62	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro63	Yes	2	1.471	0.500	0.000	0.320	No	0	0.000	0.000	0.000	0.000
Acro64	Yes	3	2.667	1.040	0.000	0.625	No	0	0.000	0.000	0.000	0.000
Acro69	Yes	3	2.273	0.950	0.000	0.560	Yes	4	3.846	1.366	0.200	0.740
Acro71	Yes	3	2.273	0.950	0.000	0.560	Yes	3	2.273	0.950	0.000	0.560
Acro72	Yes	2	1.280	0.377	0.250	0.219	Yes	2	1.471	0.500	0.000	0.320
Acro73	Yes	4	1.923	0.940	0.400	0.480	Yes	2	1.471	0.500	0.000	0.320
Acro74	Yes	3	2.632	1.030	0.600	0.620	Yes	3	2.273	0.950	0.000	0.560
Acro76	Yes	3	2.381	0.943	0.200	0.580	Yes	1	1.000	0.000	0.000	0.000
Acro77	Yes	2	1.600	0.562	0.000	0.375	Yes	3	2.462	0.974	0.250	0.594
Acro80	Yes	3	2.133	0.900	0.250	0.531	Yes	1	1.000	0.000	0.000	0.000
Acro82	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro84	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro87	No	0	0.000	0.000	0.000	0.000	Yes	4	2.381	1.089	0.800	0.580
Acro88	Yes	3	2.273	0.950	0.000	0.560	Yes	3	2.632	1.030	0.400	0.620
Acro89	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro91	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.778	1.055	0.000	0.640
Acro92	Yes	2	1.471	0.500	0.000	0.320	Yes	6	5.556	1.748	0.600	0.820
Acro93	Yes	3	2.273	0.950	0.000	0.560	Yes	4	1.923	0.940	0.600	0.480
Acro94	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro97	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.400	0.320
Acro98	Yes	2	1.220	0.325	0.200	0.180	No	0	0.000	0.000	0.000	0.000
Acro99	Yes	4	3.571	1.332	0.000	0.720	Yes	3	2.778	1.055	0.000	0.640
Acro100	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro101	Yes	3	2.778	1.055	0.000	0.640	Yes	1	1.000	0.000	0.000	0.000
Acro102	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro103	Yes	3	1.515	0.639	0.200	0.340	Yes	3	2.273	0.950	0.000	0.560
Acro105	Yes	2	1.923	0.673	0.000	0.480	Yes	1	1.000	0.000	0.000	0.000
Acro106	Yes	3	2.273	0.950	0.400	0.560	No	0	0.000	0.000	0.000	0.000
Acro108	No	0	0.000	0.000	0.000	0.000	Yes	5	3.571	1.418	0.800	0.720
Acro109	Yes	2	1.471	0.500	0.000	0.320	Yes	4	2.941	1.221	0.800	0.660
Acro110	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro111	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro113	Yes	3	2.778	1.055	0.000	0.640	No	0	0.000	0.000	0.000	0.000
Acro116	Yes	3	2.778	1.055	0.000	0.640	No	0	0.000	0.000	0.000	0.000
Acro118	No	0	0.000	0.000	0.000	0.000	Yes	4	3.571	1.332	0.000	0.720
Acro121	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro122	Yes	4	2.381	1.089	0.200	0.580	Yes	1	1.000	0.000	0.000	0.000
Acro124	No	0	0.000	0.000	0.000	0.000	Yes	4	3.571	1.332	0.800	0.720
Acro125	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro126	Yes	3	2.273	0.950	0.000	0.560	Yes	3	2.381	0.943	0.400	0.580
Acro130	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro136	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro137	Yes	2	1.220	0.325	0.200	0.180	Yes	3	2.778	1.055	0.400	0.640
Acro138	Yes	3	2.778	1.055	0.000	0.640	Yes	1	1.000	0.000	0.000	0.000

Acro142	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.400	0.320
Acro143	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro144	Yes	2	1.600	0.562	0.000	0.375	No	0	0.000	0.000	0.000	0.000
Acro145	Yes	4	3.333	1.280	0.800	0.700	Yes	5	3.846	1.471	1.000	0.740
Acro146	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro147	No	0	0.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.800	0.480
Acro148	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.471	0.500	0.000	0.320
Acro149	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro151	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.220	0.325	0.200	0.180
Acro152	No	0	0.000	0.000	0.000	0.000	Yes	3	2.632	1.030	0.600	0.620
Acro153	Yes	3	1.515	0.639	0.200	0.340	Yes	2	1.280	0.377	0.250	0.219
Acro155	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro156	Yes	1	1.000	0.000	0.000	0.000	Yes	4	2.941	1.221	0.200	0.660
Acro158	Yes	2	1.600	0.562	0.000	0.375	Yes	1	1.000	0.000	0.000	0.000
Acro159	No	0	0.000	0.000	0.000	0.000	Yes	3	1.515	0.639	0.400	0.340
Acro160	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro162	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro164	Yes	5	3.571	1.418	0.400	0.720	Yes	3	2.273	0.950	0.000	0.560
Acro165	Yes	2	1.923	0.673	0.000	0.480	Yes	1	1.000	0.000	0.000	0.000
Acro166	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro167	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro168	Yes	5	4.571	1.560	0.250	0.781	Yes	2	1.923	0.673	0.000	0.480
Acro170	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro171	Yes	4	3.571	1.332	0.000	0.720	Yes	3	1.852	0.802	0.200	0.460
Acro172	Yes	2	1.923	0.673	0.000	0.480	Yes	5	3.846	1.471	0.400	0.740
Acro174	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.923	0.673	0.000	0.480
Acro178	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.778	1.055	0.400	0.640
Acro180	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro182	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro183	Yes	4	2.381	1.089	0.200	0.580	Yes	2	1.923	0.673	0.000	0.480
Acro187	Yes	3	2.571	1.011	0.333	0.611	No	0	0.000	0.000	0.000	0.000
Acro188	Yes	6	5.556	1.748	0.600	0.820	Yes	3	2.778	1.055	0.400	0.640
Acro189	Yes	3	2.462	0.974	0.250	0.594	Yes	3	1.852	0.802	0.200	0.460
Acro190	Yes	3	2.381	0.943	1.000	0.580	Yes	2	1.220	0.325	0.200	0.180
Acro191	Yes	3	1.515	0.639	0.200	0.340	Yes	3	2.462	0.974	0.250	0.594
Acro192	Yes	4	3.571	1.332	0.000	0.720	Yes	1	1.000	0.000	0.000	0.000
Acro193	Yes	2	1.600	0.562	0.000	0.375	Yes	1	1.000	0.000	0.000	0.000
Acro195	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro196	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro197	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro198	Yes	2	1.923	0.673	0.400	0.480	No	0	0.000	0.000	0.000	0.000
Acro199	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro201	Yes	2	1.471	0.500	0.000	0.320	Yes	4	2.941	1.221	0.200	0.660
Acro203	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.000	0.560
Acro205	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.600	0.562	0.000	0.375
Acro206	Yes	2	1.600	0.562	0.000	0.375	No	0	0.000	0.000	0.000	0.000
Acro208	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.220	0.325	0.200	0.180

Acro209	Yes	3	2.273	0.950	0.000	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro210	Yes	3	1.852	0.802	0.200	0.460	Yes	5	4.167	1.505	0.800	0.760
Acro211	No	0	0.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro212	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro213	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro214	Yes	4	2.381	1.089	0.200	0.580	No	0	0.000	0.000	0.000	0.000
Acro215	No	0	0.000	0.000	0.000	0.000	Yes	4	2.941	1.221	0.600	0.660
Acro216	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro217	No	0	0.000	0.000	0.000	0.000	Yes	4	2.381	1.089	0.400	0.580
Acro218	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro220	Yes	3	2.462	0.974	0.750	0.594	Yes	4	2.941	1.194	1.000	0.660
Acro221	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.220	0.325	0.200	0.180
Acro223	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro225	Yes	2	1.800	0.637	0.000	0.444	Yes	2	1.923	0.673	0.000	0.480
Acro227	Yes	4	3.571	1.332	0.000	0.720	Yes	4	3.600	1.330	0.333	0.722
Acro228	Yes	3	2.778	1.055	0.000	0.640	Yes	3	2.632	1.030	1.000	0.620

	Juçara palm						African oil palm					
	Transferability	Na	Ne	I	Ho	He	Transferability	Na	Ne	I	Ho	He
Acro2	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.000	0.560
Acro3	Yes	1	1.000	0.000	0.000	0.000	Yes	4	3.000	1.242	0.667	0.667
Acro6	Yes	2	1.800	0.637	0.000	0.444	Yes	3	2.571	1.011	0.667	0.611
Acro10	No	0	0.000	0.000	0.000	0.000	Yes	5	2.500	1.228	0.400	0.600
Acro14	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro15	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro16	Yes	3	2.273	0.950	0.600	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro17	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro20	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro22	No	0	0.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro23	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.000	0.560
Acro24	Yes	2	1.220	0.325	0.200	0.180	No	0	0.000	0.000	0.000	0.000
Acro25	Yes	1	1.000	0.000	0.000	0.000	Yes	3	1.684	0.736	0.250	0.406
Acro27	Yes	3	2.273	0.950	0.400	0.560	Yes	3	1.515	0.639	0.200	0.340
Acro29	Yes	2	1.220	0.325	0.200	0.180	No	0	0.000	0.000	0.000	0.000
Acro30	Yes	2	1.220	0.325	0.200	0.180	Yes	2	1.923	0.673	0.000	0.480
Acro32	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro33	Yes	2	1.471	0.500	0.000	0.320	Yes	3	1.852	0.802	0.200	0.460
Acro34	Yes	3	2.381	0.943	1.000	0.580	Yes	1	1.000	0.000	0.000	0.000
Acro36	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.400	0.560
Acro37	Yes	4	2.941	1.221	0.800	0.660	Yes	2	1.471	0.500	0.000	0.320
Acro38	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.941	1.089	0.200	0.660
Acro39	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro41	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro42	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro44	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.174	0.898	0.200	0.540
Acro45	Yes	4	2.941	1.221	0.600	0.660	Yes	3	2.778	1.055	0.400	0.640
Acro46	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000

Acro47	Yes	1	1.000	0.000	0.000	0.000	Yes	5	3.846	1.471	0.200	0.740
Acro48	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.778	1.055	0.400	0.640
Acro49	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro50	Yes	4	2.381	1.089	0.400	0.580	Yes	1	1.000	0.000	0.000	0.000
Acro52	No	0	0.000	0.000	0.000	0.000	Yes	4	3.571	1.332	0.000	0.720
Acro54	Yes	2	1.923	0.673	0.000	0.480	Yes	2	1.923	0.673	0.000	0.480
Acro55	No	0	0.000	0.000	0.000	0.000	Yes	4	3.333	1.280	0.400	0.700
Acro57	Yes	5	4.545	1.557	0.200	0.780	Yes	2	1.923	0.673	0.000	0.480
Acro58	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro62	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro63	Yes	1	1.000	0.000	0.000	0.000	Yes	3	1.515	0.639	0.200	0.340
Acro64	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro69	No	0	0.000	0.000	0.000	0.000	Yes	2	2.000	0.693	0.000	0.500
Acro71	Yes	4	3.556	1.321	0.500	0.719	Yes	4	3.846	1.366	0.600	0.740
Acro72	Yes	2	1.923	0.673	0.400	0.480	Yes	3	1.684	0.736	0.500	0.406
Acro73	No	0	0.000	0.000	0.000	0.000	Yes	4	2.381	1.089	0.200	0.580
Acro74	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro76	Yes	3	2.174	0.898	0.200	0.540	Yes	2	1.800	0.637	0.000	0.444
Acro77	Yes	2	1.600	0.562	0.000	0.375	Yes	2	1.220	0.325	0.200	0.180
Acro80	Yes	5	3.846	1.471	0.800	0.740	Yes	1	1.000	0.000	0.000	0.000
Acro82	No	0	0.000	0.000	0.000	0.000	Yes	3	2.571	1.011	0.667	0.611
Acro84	Yes	3	2.778	1.055	0.000	0.640	Yes	3	2.778	1.055	0.000	0.640
Acro87	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.667	1.040	0.000	0.625
Acro88	Yes	4	3.333	1.280	0.800	0.700	Yes	3	2.941	1.089	0.200	0.660
Acro89	Yes	3	1.515	0.639	0.200	0.340	Yes	4	2.778	1.168	0.800	0.640
Acro91	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro92	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro93	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro94	Yes	1	1.000	0.000	0.000	0.000	Yes	4	2.778	1.168	0.400	0.640
Acro97	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro98	Yes	5	4.167	1.505	0.400	0.760	Yes	3	2.273	0.950	0.400	0.560
Acro99	Yes	3	1.852	0.802	0.200	0.460	Yes	3	2.273	0.950	0.000	0.560
Acro100	Yes	3	1.515	0.639	0.400	0.340	Yes	1	1.000	0.000	0.000	0.000
Acro101	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro102	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro103	Yes	1	1.000	0.000	0.000	0.000	Yes	3	1.515	0.639	0.400	0.340
Acro105	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro106	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro108	No	0	0.000	0.000	0.000	0.000	Yes	4	1.923	0.940	0.400	0.480
Acro109	Yes	2	2.000	0.693	1.000	0.500	No	0	0.000	0.000	0.000	0.000
Acro110	Yes	3	2.273	0.950	0.000	0.560	Yes	1	1.000	0.000	0.000	0.000
Acro111	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro113	No	0	0.000	0.000	0.000	0.000	Yes	3	2.667	1.040	1.000	0.625
Acro116	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.724	0.611	0.200	0.420
Acro118	Yes	3	2.273	0.950	0.000	0.560	Yes	3	2.381	0.943	0.400	0.580
Acro121	Yes	2	1.471	0.500	0.000	0.320	No	0	0.000	0.000	0.000	0.000
Acro122	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.632	1.030	0.200	0.620

Acro124	No	0	0.000	0.000	0.000	0.000	Yes	2	1.724	0.611	0.200	0.420
Acro125	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro126	Yes	3	1.852	0.802	0.600	0.460	Yes	2	1.724	0.611	0.200	0.420
Acro130	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro136	Yes	2	1.471	0.500	0.000	0.320	Yes	4	3.333	1.280	0.200	0.700
Acro137	Yes	2	1.220	0.325	0.200	0.180	Yes	4	3.333	1.280	0.600	0.700
Acro138	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro142	No	0	0.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro143	Yes	4	2.941	1.221	0.400	0.660	Yes	2	1.724	0.611	0.600	0.420
Acro144	No	0	0.000	0.000	0.000	0.000	Yes	2	1.280	0.377	0.250	0.219
Acro145	Yes	5	3.846	1.471	1.000	0.740	No	0	0.000	0.000	0.000	0.000
Acro146	No	0	0.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro147	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro148	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.778	1.055	0.400	0.640
Acro149	No	0	0.000	0.000	0.000	0.000	Yes	4	3.333	1.280	0.200	0.700
Acro151	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro152	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.667	1.040	0.000	0.625
Acro153	Yes	1	1.000	0.000	0.000	0.000	Yes	5	4.545	1.557	0.600	0.780
Acro155	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro156	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.381	0.943	0.200	0.580
Acro158	Yes	2	1.471	0.500	0.000	0.320	Yes	2	1.471	0.500	0.000	0.320
Acro159	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.381	0.943	0.200	0.580
Acro160	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro162	Yes	2	1.471	0.500	0.000	0.320	No	0	0.000	0.000	0.000	0.000
Acro164	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro165	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro166	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro167	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro168	Yes	2	1.923	0.673	0.000	0.480	Yes	4	2.941	1.221	0.200	0.660
Acro170	Yes	2	1.220	0.325	0.200	0.180	Yes	3	2.381	0.943	0.200	0.580
Acro171	No	0	0.000	0.000	0.000	0.000	Yes	4	3.000	1.242	0.667	0.667
Acro172	Yes	2	1.471	0.500	0.000	0.320	Yes	5	4.545	1.557	0.800	0.780
Acro174	Yes	2	1.923	0.673	0.000	0.480	Yes	3	2.273	0.950	0.000	0.560
Acro178	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro180	No	0	0.000	0.000	0.000	0.000	Yes	3	2.273	0.950	0.400	0.560
Acro182	Yes	4	3.333	1.280	0.200	0.700	Yes	1	1.000	0.000	0.000	0.000
Acro183	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro187	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro188	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro189	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.923	0.673	0.000	0.480
Acro190	Yes	5	3.571	1.418	0.600	0.720	Yes	3	2.273	0.950	0.400	0.560
Acro191	Yes	2	1.471	0.500	0.000	0.320	Yes	5	3.571	1.418	1.000	0.720
Acro192	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro193	Yes	1	1.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro195	Yes	2	1.471	0.500	0.000	0.320	Yes	1	1.000	0.000	0.000	0.000
Acro196	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro197	Yes	1	1.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000

Acro198	Yes	1	1.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro199	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro201	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro203	No	0	0.000	0.000	0.000	0.000	Yes	2	1.600	0.562	0.000	0.375
Acro205	Yes	4	3.571	1.332	0.000	0.720	Yes	2	1.471	0.500	0.400	0.320
Acro206	Yes	2	1.600	0.562	0.000	0.375	Yes	1	1.000	0.000	0.000	0.000
Acro208	Yes	2	1.471	0.500	0.000	0.320	No	0	0.000	0.000	0.000	0.000
Acro209	Yes	6	6.000	1.792	1.000	0.833	Yes	1	1.000	0.000	0.000	0.000
Acro210	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro211	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro212	No	0	0.000	0.000	0.000	0.000	Yes	3	1.852	0.802	0.200	0.460
Acro213	No	0	0.000	0.000	0.000	0.000	Yes	2	1.471	0.500	0.000	0.320
Acro214	Yes	1	1.000	0.000	0.000	0.000	Yes	4	3.333	1.280	0.200	0.700
Acro215	Yes	2	1.882	0.662	0.250	0.469	Yes	3	2.909	1.082	0.750	0.656
Acro216	Yes	5	3.846	1.471	0.800	0.740	Yes	2	1.220	0.325	0.200	0.180
Acro217	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro218	No	0	0.000	0.000	0.000	0.000	Yes	4	3.571	1.332	0.000	0.720
Acro220	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.632	1.030	0.400	0.620
Acro221	No	0	0.000	0.000	0.000	0.000	Yes	1	1.000	0.000	0.000	0.000
Acro223	No	0	0.000	0.000	0.000	0.000	No	0	0.000	0.000	0.000	0.000
Acro225	Yes	2	1.471	0.500	0.000	0.320	Yes	3	2.632	1.030	0.200	0.620
Acro227	Yes	3	2.273	0.950	0.000	0.560	Yes	2	1.923	0.673	0.000	0.480
Acro228	Yes	1	1.000	0.000	0.000	0.000	Yes	3	2.381	0.943	0.800	0.580

4. GENERAL DISCUSSION

This exploratory study brings important results related to the macaúba palm, a native palm tree which is considered a novel bioenergy crop to oil production since it has high oil content (comparing with African oil palm), and it is observed and adapted in many environments. In the first moment, we developed De Novo comprehensive transcriptome to describe macaúba palm transcripts, the tissue-specific transcripts and elucidate gene family expansion. Furthermore, we identify a set of EST-SSR markers from reference assembly approach, which they were validated in the macaúba palm and transferred in *Acrocomia intumescens, Acrocomia totai,* African oil palm (*Elaeis guineensis*), Hat palm (*Sabal causiarum*), Juçara palm (*Euterpe edulis*), Peach palm (*Bactris gasipaes*).

In the first chapter, we assembled 34.293 transcripts in all libraries joined in a single reference. 5062 transcripts were annotated by PANNZER and 25.246 transcripts (73.61%) were annotated by CDD. 4.129 transcripts were associated with at least one GO term, which 12 transcripts were classified in unique GO terms. Also, 48.545 terms were assigned to a molecular function category, 81.511 terms in the biological process category, and 20.421 terms in the cellular component category.

Transcript abundance (TPM values) analysis showed that TRINITY_DN32851_c0_g1, annotated as TCTP were highly expressed in leaf sheaves, bulb, root, female flower and fruit tissues. TCTP is ubiquitously found in all eukaryotes and it is regulated in response to diverse functions, such as cell growth (TAO et al., 2015) and proliferation (TOSCANO-MORALES et al., 2015), as well as protection against various cellular stresses (WANG et al., 2015) and other immunological functions.

Considering the SPM value ≥ 0.95, eight transcripts expressed only in fruit tissues, two in leaf tissue, eight in bulb tissue, five in female flowers, 14 in male flowers, 8053 in root tissues and none in leaf sheave tissues. Applying a TPM cut-off of ≥ 20 we observed, two transcripts highly expressed and fruit-specific.

Enoyl-ACP reductases (TRINITY_DN64786_c0_g1 – FabL and TRINITY_DN25411_c0_g1 - Trans-2-enoyl-CoA reductase) were highly expressed and fruit-specific. Enoyl-[acyl-carrier-protein (ACP)] reductase is involved in the elongation step of fatty acid biosynthesis. The reaction is mediated by four enzymes and the enoyl catalyze the last step of the cycle, in which the trans-2-enoyl-ACP is reduced to an acyl-ACP. Depending on the plant source, the reductase may be specific for NADH or may be able to utilize NADPH as well (VICK et al., 2015). Quite a lot is known about the gene (or genes) coding for the enoyl reductase, as well as the reaction mechanism which involves a compulsory-ordered ternary complex Trans-2-Enoyl-CoA Reductase of *Euglena gracilis* was reported to be useful for increasing lipid content in plants based on overexpression (CIRPUS et al., 2005).

In root, we identified 341 transcripts high expressed. Important classes of proteins were annotated, as serine/threonine kinases (136 transcripts observed), Cytochrome P450 (80 transcripts), and 14 pathogenesis-related proteins (PR), suggesting that macaúba palm has a rusticity and responses to abiotic and biotic stress, since these protein families play a major role in plant defense against various biotic and abiotic stresses. Also, in male flower, we detected two transcripts annotated as Non-specific lipid-transfer protein (nsLTP), which are expressed in diverse organs, as germinating and maturing seeds, roots, stems, and pollens. Since we observed a male flower specific nsLTP, we can suppose a possible protein related to pollen formation, anther development or reproductive functions.

Moreover, orthologue analysis between *Elaeis guineensis, Musa acuminata, Phoenix dactylifera, Oryza sativa*, and *Acrocomia aculeata* detected 15.283 families, which 12.664 families with at least one representant of macaúba palm. BadiRate analysis assessed *family expansion* between macaúba and oil palm. A total of 24 families were identified being expanded in the macaúba palm (comparing with oil palm). Four families were annotated as PPR proteins, two families were like heat-shock proteins and three families were like protein kinase; six families were exclusive to macaúba and 11 of them had at least one TS transcripts, being all root-specific transcripts. When we observed expression level, only the OG0000038 family had overexpressed transcripts like ubiquitin, a regulatory protein that plays function in the most different biochemical processes. we can propose a molecular adaptation of macaúba to environmental adaptation.

Interestingly, transcripts related to drought stress were grouped in 114 families containing at least one macaúba transcript. It was possible to analyze and suggest that the plant adjusts expressing regulatory proteins. Including the TS families, proteins from the 14-3-3 family and CAM had high expression. In addition, transcripts of macaúba samples were expressed in a high level in 13 families of stress-related transcription factors and 11 families of kinases. Among all the data obtained with the multi-tissue transcriptome, it was possible to observe that the macaúba palm overexpressed genes involved in the signal transduction to stress tolerance and probably economical and efficient regulation of biological processes to deal with environmental adaptation. Although it is not possible to correlate with plant evolution and plant adaptation to the different habitats, our data may suggest molecular adaptations of the plant.

In the second chapter, we identify and develop a set of EST-simple sequence repeat (SSR) markers by using transcriptome sequencing. Further, we apply in previous genetic diversity analysis and cross transferability with other palm trees with environmental and economic importance. In this study, we identified 85.014 redundant EST-SSRs from transcript sequences of all samples, with a frequency of 1/1 Mbp, to prevent the validation of more than one locus in linkage equilibrium, according to assembly in the African oil palm genome. The SSR density inside EST depends on the search tool, criteria used to identify SSRs in the database or redundancy in the set of SSR, which can lead to the identification of multiple markers at the same locus (VARSHNEY; GRANER; SORRELLS, 2005).

We observed 418 non-redundant EST-SSRs as representative of each locus, and we selected 232 EST-SSRs, with trinucleotide repeats being the most frequent motif, representing 380 (90.9%). This suggests a result of selection and evolution because of tri- and hexa-SSRs do not change the coding frame of the gene regions. Other repetitions change the frame and induce negative mutation when there is variation in SSR length (QIU et al., 2010). The most frequent motifs were GAG/CAG, CCT/CCG and GGA/GGC, which encode glutamic acid, proline, and glycine, respectively. Among the proteins containing 10 or more repetitions of single amino acid, glutamine, alanine, glycine, glutamic acid, and serine repeats were more frequent than other amino acids (ASP et al., 2007; KATTI et al., 2000). KATTI et al. (2001) observed that codon repeats corresponding to small hydrophilic amino acids are possibly more easily tolerated, and selection pressure possibly eliminates codon repeats encoding hydrophobic and basic amino acids. Consequently, the high level of occurrence of these motifs is substantial because amino acids produced by them are observed to a high extent in proteins.

145 EST-SSRs (62.5%) were validated for consistent amplification in 17 macaúba palm samples, and 100 were determined to be polymorphic with PIC values ranging from 0.25 to 0.77 (EST-SSR Acro205/ EST-SSR Acro201). Among these, 39 exhibited high PIC values, ranging from 0.77 (Acro205) to 0.50 (Acro220); 61 exhibited medium PIC value, ranging from 0.49 (Acro125) to 0.25 (Acro213); 31 with low PIC value, ranging from 0.24 (Acro33) to 0.05 (Acro124); and 14 with null PIC value (monomorphic markers), according to BOTSTEIN et al. (1980). The PIC values of our markers are higher than the values of the markers published by NUCCI et al. (2008) for the macaúba palm and are more valuable for marker-assisted selection and other applications.

Genetic diversity analysis was performed with the 20 most informative EST-SSR markers, based on PIC values and probability of identity (PI) for all EST-SSRs, to verify efficiency in the genetic analysis of the samples. Principal coordinate analysis (PCoA) was performed on the genotype data of 17 samples of macaúba palm based on the Nei distance (NEI, 1972). The PCoA results revealed distinct genetic grouped by their geographic location, containing samples from Itapira and Jaguariúna city from São Paulo State/Brazil; Rifaina from São Paulo State/Brazil, Serra da Canastra, and Capitólio from Minas Gerais State/Brazil; and Luz (Minas Gerais State/Brazil).

Also, these 145 markers were transferred in six other palm species resulting in transferability rates of 99% (144) in *Acrocomia intumescens*, 98% (143) in *Acrocomia totai*, 80.7% (117 EST-EST) in African oil palm (*Elaeis guineensis*) and peach palm (*Bactris gasipaes*) samples, 70% (102) in the juçara palm (*Euterpe edulis*) and 71.7% (104) in the hat palm (*Sabal causiarum*). The transferable nature of EST-SSR markers within related species or genera extends their usefulness in plant breeding and genetic studies, being limited in those species that have no available EST dataset. We have achieved a high cross-species transferability of EST-SSRs found in the macaúba palm to other palm genera. It is advantageous to save time and cost for developing SSR markers for species that have not been largely studied.

5. CONCLUSION

This was the first study to generate large-scale NGS data and multi-tissue transcriptome characterization in macaúba palm species, providing valuable genome resources for studying the evolution, gene expression, and genomic studies. The transcriptome harbor many novel transcripts and it was possible to detect expanded and exclusives families in comparison with African oil palm. Additionally, reference-based assembly was an effective strategy of obtaining efficient EST-SSRs for discriminate individuals inside and outside the species. The EST-markers identified in our study proved to be a valuable resource and provide a genomic tool for genetic mapping and diversity analysis of macaúba palm and related species;

Taken together, our data provide a source of macaúba palm transcripts and molecular markers. Consequently, we believe the data generated will be of great support for the in-progress research in Acrocomia genera, particularly.

6. REFERENCES

ABREU, A. G. et al. The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). **Genetics and Molecular Biology**, v. 35, n. 1, p. 116–121, 2012.

ABREU, I. S. et al. First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae) a promising plant for biodiesel production. **Australian Journal of Botany**, v. 59, n. 2, p. 149–155, 2011.

APROBIO. **IMPACTOS SOCIOECONÔMICOS DA INDÚSTRIA DE BIODIESEL NO BRASIL**. [s.l: s.n.].

ASP, T. et al. Frequency, type, and distribution of EST-SSRs from three genotypes of *Lolium perenne*, and their conservation across orthologous sequences of Festuca arundinacea, Brachypodium distachyon, and Oryza sativa. **BMC Plant Biology**, v. 7, p. 1–12, 2007.

BELTRÃO, N. E. M.; OLIVEIRA, M. I. P. **Oleaginosas potenciais do Nordeste para a produção de biodieselDocumentos**. [s.l: s.n.].

BERTON, L. H. C. **Avaliação de populações naturais , estimativas de parâmetros genéticos e seleção de genótipos elite de macaúba (*Acrocomia aculeata*)**. [s.l: s.n.].

BOTSTEIN, D. et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American journal of human genetics**, v. 32, n. 3, p. 314–31, 1980.

CICONINI, G. et al. Biometry and oil contents of *Acrocomia aculeata* fruits from the Cerrados and Pantanal biomes in Mato Grosso do Sul, Brazil. **Industrial Crops and Products**, v. 45, n. March 2016, p. 208–214, 2013.

CIRPUS, P. et al. **Trans-2-enoyl-coa reductase gene of euglena gracilis**, 2005.

COLOMBO, C. A. et al. Macauba: a promising tropical palm for the production of vegetable oil. **Oilseeds & Fats Crops and Lipids**, v. 25, n. 1, p. 9, 2018.

EVARISTO, A. B. et al. Actual and putative potentials of macauba palm as feedstock for solid biofuel production from residues. **Biomass and Bioenergy**, v. 85, p. 18–24, 2016.

IRENA. Renewable Energy and Jobs: Annual Review 2018. p. 28, 2018.

KATTI, M.; RANJEKAR, P.; GUPTA, V. Differential Distribution of Simple Sequence Repeats in Eukaryotic Genome Sequences. **Mol. Biol. Evol.**, v. 18, n. 7, p. 1161–1167, 2001.

KATTI, M. V et al. Amino acid repeat patterns in 1. Katti M V, Sami-Subbu R, Ranjekar PK, Gupta VS.

Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. Protein Sci. 2000;9:1203–9. protein sequences: their diversity and st. **Protein science : a publication of the Protein Society**, v. 9, n. 6, p. 1203–1209, 2000.

LANES, É. C. M. et al. Molecular characterization and population structure of the Macaw Palm, *Acrocomia aculeata* (Arecaceae), Ex Situ Germplasm collection using microsatellites markers. **Journal of Heredity**, v. 106, n. 1, p. 102–112, 2015.

LANES, É. C. M. et al. Mating system and genetic composition of the macaw palm (*Acrocomia aculeata*): Implications for breeding and genetic conservation programs. **Journal of Heredity**, v. 107, n. 6, p. 527–536, 2016.

LANGEVELD, J. W. A. et al. Analyzing the effect of biofuel expansion on land use in major producing countries: evidence of increased multiple cropping. **Biofuels, Bioproducts and Biorefining**, v. 8, n. 1, p. 49–58, 2013.

LI, Y. C. et al. Microsatellites within genes: Structure, function, and evolution. **Molecular Biology and Evolution**, v. 21, n. 6, p. 991–1007, 2004.

LORENZI H, NOBLICK L, K. F. AND F. E. **Arecaceae (Palms). Plantarum, Nova Odessa**. Nova Odessa: Instituto Plantarum de Estudos da Flora, 2010.

MINISTRY OF MINES AND ENERGY. **Brazil at COP21/CMP1121st CONFERENCE OF THE PARTIES TO THE UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE AND 11th CONFERENCE OF THE PARTIES SERVING AS MEETING OF THE PARTIES TO THE KYOTO PROTOCOL**, 2015.

MINISTRY OF MINES AND ENERGY. **Oilseeds and Products Annual Planted Area to Hit Record for Ninth Consecutive Year**. [s.l: s.n.].

NEI, M. Genetic Distance between Populations. **The American Naturalist**, v. 106, n. 949, p. 283–292, 1 maio 1972.

NUCCI, S. M. et al. Development and characterization of microsatellites markers from the macaw. **Molecular ecology resources**, v. 8, n. 1, p. 224–226, jan. 2008.

PALLIYARAKKAL, M. K.; RAMASWAMY, M.; VADIVEL, A. Microsatellites in palm (Arecaceae) sequences. **Bioinformation**, v. 7, n. 7, p. 347–51, 2011.

QIU, L. et al. Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). **BMC plant biology**, v. 10, n. 1, p. 278, 2010.

SALLET, C. L.; ALVIM, A. M. Biocombustíveis: uma análise da evolução do biodiesel no Brasil. **Economia & Tecnologia**, v. 25, n. 7, p. 13, 2011.

SCARIOT, A.; LLERAS, E.; HAY, J. D. Flowering and Fruiting Phenologies of the Palm *Acrocomia aculeata*: Patterns and Consequences. **Biotropica**, v. 27, n. 2, p. 168–173, 1995.

STEINFELD, H. et al. Livestock's long shadow - environmental issues and options. **Food and Agriculture Organization of the United Nations**, v. 3, n. 1, p. 1–377, 2006.

TAO, J.-J. et al. Tobacco Translationally Controlled Tumor Protein Interacts with Ethylene Receptor Tobacco Histidine Kinase1 and Enhances Plant Growth through Promotion of Cell Proliferation. **Plant Physiology**, v. 169, n. 1, p. 96–114, 2015.

TOSCANO-MORALES, R. et al. AtTCTP2, an *Arabidopsis thaliana* homolog of Translationally Controlled Tumor Protein, enhances in vitro plant regeneration. **Frontiers in Plant Science**, v. 6, n. July, p. 1–17, 2015.

USDA. **United States Department of Agriculture**. [s.l: s.n.].

VARSHNEY, R. K. et al. Next-generation sequencing technologies and their implications for crop genetics and breeding. **Trends in Biotechnology**, v. 27, n. 9, p. 522–530, set. 2009.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: Features and applications. **Trends in Biotechnology**, v. 23, n. 1, p. 48–55, 2005.

VICK, J. E. et al. *Escherichia coli* enoyl-acyl carrier protein reductase (FabI) supports efficient operation of a functional reversal of the β-oxidation cycle. **Applied and Environmental Microbiology**, v. 81, n. 4, p. 1406–1416, 2015.

WANG, Z. Q. et al. OsTCTP, encoding a translationally controlled tumor protein, plays an important role in mercury tolerance in rice. **BMC Plant Biology**, v. 15, n. 1, p. 1–15, 2015.

APPENDIX

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

20/07/2018 870180062831
14:37
29409161711182795

Pedido nacional de Inveno, Modelo de Utilidade, Certificado de Adio de Inveno e entrada na fase nacional do PCT

Número do Processo: BR 10 2018 014889 3

Dados do Depositante (71)

Depositante 1 de 1

Nome ou Razão Social: UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP
Tipo de Pessoa: Pessoa Jurídica
CPF/CNPJ: 46068425000133
Nacionalidade: Brasileira
Qualificação Jurídica: Órgão Público
Endereço: Cidade Universitária Zeferino Vaz
Cidade: Campinas
Estado: SP
CEP: 13084-971
País: Brasil
Telefone:
Fax:
Email: patentes@inova.unicamp.br

Dados do Pedido

Natureza Patente:	10 - Patente de Invenção (PI)
Título da Invenção ou Modelo de Utilidade (54):	MÉTODO E KIT PARA A IDENTIFICAÇÃO E SEPARAÇÃO DE PLANTAS DE MACAÚBA (Acrocomia aculeata) E ESPÉCIES DO GÊNRO Acronomia sp.
Resumo:	Esta invenção descreve um método para a identificação e separação de espécies da planta macaúba (Acrocomia aculeata), bem como a diferenciação de indivíduos dentro e entre espécies do gênero Acrocomia. O referido método utiliza sequências flanqueadoras ou primers, conforme definidos nas SEQ. IDs. 1 a 290, as quais surgem como ferramentas úteis para traçar o perfil de variabilidade genética de uma população, desenvolver o mapa genético da espécie, identificar genótipos elite em populações naturais para melhoramento genético e caracterizar o banco de germoplasma. A invenção ainda prevê um kit compreendendo os referidos primers.
Figura a publicar:	1

Dados do Procurador

Procurador:

Nome ou Razão Social:	Ciro de la Cerda
Numero OAB:	
Numero API:	
CPF/CNPJ:	04584518866
Endereço:	Rua Roxo Moreira 1831, Caixa Postal 6131 - INOVA UNICAMP
Cidade:	Campinas
Estado:	SP
CEP:	13083-970
Telefone:	
Fax:	
Email:	patentes@inova.unicamp.br

Dados do Inventor (72)

Inventor 1 de 5

Nome: BARBARA REGINA BAZZO
CPF: 35444343851
Nacionalidade: Brasileira
Qualificação Física: Estudante de Pós Graduação
Endereço: Rua Emílio Henking, 343, Vila Rossi Borghi e Siqueira
Cidade: Campinas
Estado: SP
CEP: 13070-261
País: BRASIL
Telefone: (19) 352 16237
Fax:
Email: barbarareginabazzo@gmail.com

Inventor 2 de 5

Nome: LUCAS MIGUEL DE CARVALHO
CPF: 36198337880
Nacionalidade: Brasileira
Qualificação Física: Estudante de Pós Graduação
Endereço: Rua Écio Rossi, 233, Jd. São Lourenço
Cidade: Bragança Paulista
Estado: SP
CEP: 12908-611
País: BRASIL
Telefone: (19) 352 16651
Fax:
Email: lucasmigueel@gmail.com

Inventor 3 de 5

Nome: MARCELO FALSARELLA CARAZZOLLE

CPF: 28464839812

Nacionalidade: Brasileira

Qualificação Física: Pesquisador

Endereço: Rua Luis Bissoto, 240, Jd. Santa Rosa

Cidade: Valinhos

Estado: SP

CEP: 13275-110

País: BRASIL

Telefone: (19) 352 16651

Fax:

Email: mcarazzo@lge.ibi.unicamp.br

Inventor 4 de 5

Nome: GONÇALO AMARANTE GUIMARÃES PEREIRA

CPF: 28987039587

Nacionalidade: Brasileira

Qualificação Física: Professor do ensino superior

Endereço: Rua Dr. Lauro Pimentel, 323, Cidade Universitária - Parte II, Distrito de Barão Geraldo

Cidade: Campinas

Estado: SP

CEP: 13083-250

País: BRASIL

Telefone: (19) 352 16650

Fax:

Email: goncalo@unicamp.br

Inventor 5 de 5

Nome: CARLOS AUGUSTO COLOMBO
CPF: 01122383860
Nacionalidade: Brasileira
Qualificação Física: Pesquisador
Endereço: Rua Emerson José Moreira, 1667, Casa 04, Chácaras Primavera
Cidade: Campinas
Estado: SP
CEP: 13087-441
País: BRASIL
Telefone: (19) 320 21792
Fax:
Email: iac.colombo@gmail.com

Documentos anexados

Tipo Anexo	Nome
Relatório Descritivo	RELATORIO_DESCRITIVO_200718.pdf
Reivindicação	REIVINDICACOES_200718.pdf
Desenho	FIGURAS_200718.pdf
Resumo	RESUMO_200718.pdf
CADASTRO NO SISGEN	Cadastro no SISGEN_200718.pdf
Comprovante de pagamento de GRU 200	GRU_200718.pdf
Procuração	PROCURACAO_200818.pdf

Sequências Biológicas

☑ Declaro que a informação contida na 'Listagem de Sequências' apresentada em formato eletrônico está limitada ao conteúdo da matéria revelada pelas sequências de aminoácidos e/ou de nucleotídeos divulgadas no pedido de patente, conforme depositado

Tipos de Sequências Biológicas	Nome
Listagem de Sequências Biológicas em formato TXT	LISTAGEM_SEQUENCIA_DEFINITIVA_200718 .TXT

Acesso ao Patrimônio Genético

☑ Declaração Positiva de Acesso - Declaro que o objeto do presente pedido de patente de invenção foi obtido em decorrência de acesso à amostra de componente do Patrimônio Genético Brasileiro, realizado a partir de 30 de junho de 2000, e que foram cumpridas as determinações da Lei 13.123 de 20 de maio de 2015, informando ainda:

Número da Autorização de Acesso:	A9D3C49
Data da Autorização de Acesso:	19/07/2018

Declaração de veracidade

☑ Declaro, sob as penas da lei, que todas as informações acima prestadas são completas e verdadeiras.