UNIVERSIDADE DE SÃO PAULO
Faculdade de Ciências Farmacêuticas
Programa de Pós-Graduação em Ciência dos Alimentos
Área de Bromatologia

Biologia de sistemas aplicada ao estudo do amadurecimento de mamões: influência do etileno

Caroline Giacomelli Soares

Dissertação para obtenção do Título
de Mestre
Orientador: Prof. Dr. João Paulo Fabi

São Paulo
2021

UNIVERSIDADE DE SÃO PAULO
Faculdade de Ciências Farmacêuticas
Programa de Pós-Graduação em Ciência dos Alimentos
Área de Bromatologia

Systems biology applied to the study of papaya fruit ripening: the influence of ethylene

Caroline Giacomelli Soares

Versão corrigida

Dissertação para obtenção do Título
de Mestre
Orientador: Prof. Dr. João Paulo Fabi

São Paulo
2021

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:
Marlene Aparecida Vieira - CRB - 8/5562

S676s

Soares, Caroline Giacomelli
Systems biology applied to the study of papaya fruit ripening: the influence of ethylene / Caroline Giacomelli Soares. - São Paulo, 2021.
107 p.

Dissertação (mestrado) - Faculdade de Ciências Farmacêuticas da Universidade de São Paulo. Departamento de Alimentos e Nutrição Experimental.
Orientador: Fabi, João Paulo

1. Amadurecimento de frutos. 2. Etileno. 3. Mamão papaia. 4. Transcriptomica. 5. Fisiologia pós-colheita. I. T. II. Fabi, João Paulo, orientador.

Caroline Giacomelli Soares

Systems biology applied to the study of papaya fruit ripening: the influence of ethylene

Comissão Julgadora
da
Dissertação para obtenção do Título de Mestre

Prof. Dr. João Paulo Fabi
orientador/presidente

_____
1o. Examinador

_____
2o. Examinador

_____
3o. Examinador

São Paulo, _____ de _____ de 2021.

“The greatest scientific discovery was the discovery of ignorance.”

(HARARI, 2016, p.126)

ACKOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Professor Dr. João Paulo Fabi for providing me the opportunity of a graduation at such a renowned University and the continuous guidance, motivation, and patience throughout the development of my research. His insights and encouragement were essential to me successfully finish this step of my academic career.

A special thanks to Professor Dr. Sónia Andrade from the Department of Genetics and Evolutionary Biology, Biosciences Institute, USP, who introduced me to the world of Bioinformatics and helped me processing and analyzing the RNA-seq data. Getting to know a little about her laboratory routine and joining the Journal Club allowed me to spread my wings about Molecular Biology.

Many thanks to Professor Dr. Eduardo Purgatto and his students Isabel and Eric, who kindly allowed me to follow their practice of metabolomics and proteomics protocols. Those were hectic but worthwhile, fun weeks.

I am very grateful to Lab Technicians of the Laboratory of Food Chemistry, Biochemistry and Molecular Biology, for guiding me in lab practice and organization and teaching me proper usage of equipment. Samira and Victor were especially helpful to me when I had just joined the Department. I also would like to thank my labmates. Even though I spent a short time in the lab, I keep good memories.

I would like to acknowledge the assistance of the National Council for Scientific and Technological Development (CNPq) for granting the Master fellowship (#168378/2018-3) and a research project (#441789/2014-5). I extend my appreciation to the Food Research Center (FoRC) for providing cutting-edge lab equipment and infrastructure, and also to all the people who work in the Department of Food Science.

Finally, I am deeply indebted to my parents and grandfather, who have provided me the privilege to devote myself to my studies. I could not finish this list of acknowledgments without mentioning this special little creature who came to teach me that even when everything seems dark, there is always a bright side. From now on, it is all about you, Alice.

RESUMO

SOARES, C. G. **Biologia de sistemas aplicada ao estudo do amadurecimento de mamões: a influência do etileno**. 2021. 107f. Dissertação (Mestrado) – Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, 2021

O amadurecimento de frutos é um processo bioquímico que resulta em sabor, odor, textura e cor adequados para o consumo humano, além de propiciar o acesso a nutrientes importantes. Apesar do amadurecimento promover incrementos sensoriais e nutricionais nos frutos, ocorre também um aumento da suscetibilidade a danos físicos, como é o caso do mamão. Essas transformações ocorrem devido às alterações nos padrões de expressão gênica nos diferentes estádios de amadurecimento, cujo controle e coordenação decorrem da ação combinada de hormônios vegetais, principalmente do etileno. Como a ação deste hormônio na regulação da expressão gênica ainda é elusiva, a presente dissertação buscou abordar a análise global do transcriptoma em um amplo estudo dos processos moleculares envolvidos no amadurecimento de mamões tratados e não tratados com etileno. Os fatores de transcrição relacionados com a síntese e a sinalização do etileno tiveram sua atividade aumentada perante o tratamento exógeno com etileno. Consequentemente, as enzimas reguladas por esse hormônio tiveram seus genes de codificação expressos diferencialmente, como foi o caso de genes relacionados à síntese de carotenoides, linalool e vitaminas, que atuam no aumento da cor, aroma e atividade antioxidante, respectivamente. Vias metabólicas relacionadas com à síntese de açúcares foram reprimidas enquanto genes codificantes da enzima responsável pela síntese de sacarose mantiveram uma expressão basal, evidenciando que o acúmulo de açúcares ocorre antes do processo de amadurecimento. A firmeza da casca e da polpa dos frutos foram fortemente influenciadas pelo tratamento com etileno e pelo tempo de experimento, sofrendo ação de inúmeras enzimas relacionadas com a degradação da parede celular. A principal enzima responsável pelo amolecimento da polpa foi a poligalacturonase, em conjunto com a atividade de outras pectinases e celulases. Em contraste com a necessidade da ação pré-climatérica da pectato liase e da pectinesterase relatada em outras frutas carnosas, como tomates e morangos, o mamão não apresentou uma diferença significativa na expressão das mesmas. A meta-análise de diversos transcriptomas do amadurecimento do mamão reafirmaram o perfil de expressão observado no RNA-seq, além de prover enriquecimento estatístico às narrativas biológicas. Por fim, o presente estudo reuniu uma gama de informações robustas sobre a regulação gênica do processo de amadurecimento do mamão papaia, o que abrange a possibilidade para futuras abordagens de análise transcriptomica e valida o uso do mamão como modelo para tais estudos.

Palavras-chaves: *Carica papaya*; etileno; amadurecimento de frutos; fruto climatérico; transcriptoma; amolecimento da parede celular;

ABSTRACT

SOARES, C. G. **Systems biology applied to the study of papaya fruit ripening: the influence of ethylene.** 2021. 107p. Dissertation (Master) – Faculty of Pharmaceutical Sciences, University of São Paulo, São Paulo, 2021

Fruit ripening is a biochemical process that results in flavor, odor, texture, and color suitable for human consumption, in addition to providing access to important nutrients. Although ripening promotes sensory and nutritional increases in fruits, there is also an increased susceptibility to physical damage, as is the case with papaya. These transformations occur due to changes in gene expression patterns at different stages of maturity, whose control and coordination result from the combined action of plant hormones, especially ethylene. As the action of this hormone in the regulation of gene expression is still elusive, this dissertation sought to address the global analysis of the transcriptome in an overview study of molecular processes involved in the ripening of ethylene-treated and non-treated papaya. Transcription factors related to ethylene synthesis and signaling had increased activity towards exogenous-ethylene treatment. Consequently, ethylene-induced enzymes had their coding genes differentially expressed, like genes related to the synthesis of carotenoids, linalool, and vitamins, which increase color, aroma, and antioxidant activity, respectively. Metabolic pathways related to the synthesis of sugars were suppressed while genes encoding the enzyme responsible for sucrose synthesis maintained a basal expression, showing that the accumulation of sugars occurs before the ripening process. The firmness of the peel and pulp of the fruits were strongly influenced by the treatment with ethylene and by the time of the experiment, suffering the action of numerous enzymes related to the degradation of the cell wall. The main enzyme responsible for softening the pulp was polygalacturonase, together with the activity of other pectinases and cellulases. In contrast to the need for the pre-climacteric action of pectate lyase and pectinesterase reported in other fleshy fruits, such as tomatoes and strawberries, papaya did not show a significant difference in their expression. The meta-analysis of several papaya ripening transcriptomes confirmed the expression profile observed in the previous RNA-seq, besides providing statistical enrichment to the biological narratives. Finally, the present study gathered a range of robust information on the gene regulation of the papaya ripening process, which opens possibilities for future approaches to transcriptomic analysis and validates the use of papaya as a model for such studies.

Keywords: *Carica papaya*; ethylene; fruit ripening; climacteric fruit; transcriptome; cell wall disassembling; post-harvest.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

INTRODUCTION

The understanding of organisms at a molecular level is aimed since the inception of DNA's structure by Watson & Crick (1953). Many techniques were elaborated and improved, beginning with a seventy-seven-nucleotide tRNA sequencing (HOLLEY *et al.*, 1965), escalating to detect the sequence of entire genes (MAXAM, GILBERT, 1977), and finally, Sanger *et al.* (1977) successfully deciphered the whole genome of a bacteriophage, becoming the well-known Sanger sequencing method. The development of new techniques escalated quickly, and eventually, one of the scientific community goals was achieved: the human genome sequencing by the International Human Genome Sequencing Consortium (2004). Despite the moral and ethical implications of genetic engineering, the Human Genome Project was a major scientific breakthrough in the early 21st century. Still, as the available methods were costly and slow, there was an urge for new, low-cost methods. Not long after, the so-called Next-generation Sequencing emerged with the parallelization of the sequencing process, which has been on revolutionary progress with several platforms, methods, and outputs.

Together with sequencing platforms evolution, the volume of biological data generated increased exponentially, requiring agile and automated tools to compose the information cohesively and understandably, which makes bioinformatics an indispensable ally for the evolution of this type of research. The domain over joining programming and biology has also evolved, allowing efficient data integration in order to obtain a robust and accurate interpretation, as well as modeling *in silico* systems to enable the simulation of natural and modified systems (KARAHALIL, 2016). Systems biology is an interdisciplinary approach that consolidates biology and bioinformatics to reveal the molecular components and cellular processes within a biological system under varying conditions and temporal changes, considering specifics behavior that cannot be predicted when solely observed (KARAHALIL, 2016).

Nowadays, scientists of the most wide-ranging fields rely on Next-Generation Sequencing (NGS) technologies to investigate their hypotheses and explore their discoveries. In the medical field, a recent genome-wide association study detected a fifty-kilobases-segment inherited from Neanderthals which is related to a major risk of developing severe symptoms of COVID-19 (ZEBERG; PÄÄBO, 2020). Regarding the area of Ecology, CRANDALL *et al.* (2020) describe how -omics technologies can be

helpful to manage plant disease towards climate changes. Agriculture counts on NGS technologies to prospect plant-microbe symbiosis (AGRAHARI *et al.*, 2020), crops design (TIAN *et al.*, 2021), the use of biofertilizers aiming at sustainability (IGIEHON *et al.*, 2017), among others. Some prominent studies bring into focus food attributes, such as improving nutrient availability (RODRÍGUEZ *et al.*, 2020) and post-harvest quality of fruit (DER AGOPIAN *et al.*, 2020).

FAO (2011) estimates that 40% of the vegetable and fruit loss occurs during the post-harvest period due to poor handling and storage, causing physical damage and premature ripening. Despite fruit ripening increments sensory and nutritional quality, it also increases fruit susceptibility to physical damage as the pulp and skin soften (FABI *et al.*, 2009a). Therefore, the ripening process has been extensively studied. It consists of a series of biochemical transformations favoring the phenotype to attract consumer animals that promote seed dispersion (FABI *et al.*, 2010), besides reaching flavor, odor, texture, color, and nutritional quality suitable for consumption (TUCKER *et al.*, 2018). One of the main factors that coordinate the ripening process is the hormone ethylene, known to regulate gene expression on several stages of a fruit (LIU *et al.*, 2020).

Ethylene's production in plants is endogenous and autocatalytic, that is, there is stimulation by molecular signs of the transcription of the genes of the two main enzymes of synthesis (FABI; DO PRADO, 2019). Furthermore, exogenous treatment increases the transcription of the genes regulated by this hormone (FABI *et al.*, 2009b). In this way, an effective methodology to investigate the expression profile during ripening is transcriptomics, which identifies the complete set of a sample's mRNA content. An NGS technology that has been widely applied lately is RNA sequencing, outstanding for its accurate quantification of gene expression level, its capacity of distinguishing polymorphic transcripts corresponding to the same gene, and the possibility to scan specific stages of an organism (STARK *et al.*, 2019). In contrast to previous methods to quantify transcriptomes, RNA-seq does not require existing genomic sequences neither suffer interference from background signals, thus it can screen non-model organisms with high specificity (PRADHAN *et al.*, 2019).

Deeper knowledge about the genetic control of ripening is limited to a few fruits, such as tomatoes, which represent an important experimental model of a climacteric fruit and has already gotten its spatiotemporal global expression in several tissues elucidated (SHINOZAKI *et al.*, 2018). Similar studies have been conducted for strawberry (SÁNCHEZ-SEVILLA *et al.*, 2017), melon (CHEN *et al.*, 2021), mango

(DESHPANDE, *et al.*, 2017), peach (PEI *et al.*, 2019), banana (LI *et al.*, 2019), among others. Some tropical fruits of great commercial importance, such as papaya, are still restricted to a surface knowledge on global expression profile, with most of research focused on the study of the expression of specific enzymes (FABI, *et al.*, 2009; FU *et al.*, 2021; LIU *et al.*, 2017; ZHOU *et al.*, 2019).

Papaya (*Carica papaya* L.) is a climacteric fleshy fruit of great demand in Brazil's domestic and foreign markets. In 2019, Brazil ranked second only behind India as the largest papaya producer, with 1,161,808 tones cultivated in an area of 27,556 hectares, mostly in the states of Espírito Santo and Bahia (data from IBGE - Produção Agrícola Municipal). Exports yielded more than US$46 million on sales destined mainly to Europe and the United States, surpassed only by Mexico. Primarily, papaya is a fleshy fruit and a plentiful source of carotene and vitamins A and C, when ripe is usually consumed raw but also used to make jams, candies, and nectar (DAAGEMA *et al.*, 2020). In the industry, the extracted enzyme papain is commonly used in tenderizing meat, breweries, production of chewing gums, and cosmetics. There are also medicinal applications as the amount of antioxidant compounds improves the immune system, besides pectin extracted from a specific fruit stage could have anti-cancer properties (DO PRADO *et al.*, 2017). However, the production of this fruit is greatly affected by post-harvest losses due to excessive softening of the pulp, which favors microbial attack and mechanical damage.

Because papaya is a climacteric fruit, the pulp and peel softening occurs relatively quickly and management techniques are limited in terms of preventing and minimizing damage (FABI *et al.*, 2009). In this way, papaya is not only a fruit of commercial importance but also an excellent model for studying the ripening process and molecular and biochemical modifications that lead to changes in the firmness of the pulp. Thus, the deepen of knowledge about the ripening process of papaya can contribute to the understanding of the regulatory mechanisms involved and how it affects its final quality. Although papaya already has most of its genome sequenced (MING *et al.*, 2008), many studies focus solely on changes in the cell wall which culminates in pulp softening (MANRIQUE & LAJOLO, 2004; SHIGA *et al.*, 2009; DO PRADO *et al.*, 2016). SHEN *et al.*, 2017 studied the transcriptome of papaya treated with ethylene and its inhibitor (1-methylcyclopropene) to visualize differences in certain transcripts after the respective treatments. Regarding their study, analyzing more ripening-representative physiological parameters would add accuracy on data

analysis, thereby resulting in a more biologically significative outcome. Another proposal would be investigating the transcriptome of other cultivars and even integrating data from several studies with similar experimental design, suppressing possible bias due to thorough conduction of experiments thus enhancing statistical representation of the transcriptomic analysis.

Therefore, the present dissertation describes three approaches to the transcriptomic analysis of ethylene-treated and non-treated papayas. **Chapter I** sets up a descriptive overview of the global expression profile, determining the differentially expressed genes and the enriched and suppressed pathways. On the other hand, **Chapter II** focuses on genes related to the plant fruit wall disassembling and establishes a co-expression correlation network. Lastly, **Chapter III** portraits a meta-analysis of several papaya transcriptomes and compares to the data of the first Chapter. This study brings new perspectives on the complex physiological phenomenon of the ripening of fleshy fruits at a molecular level.

OBJECTIVES

General

To investigate the effects of the hormone ethylene on the ripening process of papaya fruit applying the Systems Biology approach on a temporal transcriptomics analysis.

Specific

(a) To provide an overview of the global expression profile, the differentially expressed genes, and the enrichment analysis of the ripening stages of papaya, both natural and exogenous ethylene-induced ripening.

(b) To investigate the genes involved in pulp softening by disassembling the fruit cell wall.

(c) To integrate transcriptomic data of papaya ripening studies available in the GEO database in order to generate corroborated results through meta-analysis.

Chapter I

Genome-wide transcriptome analysis of ethylene-treated and non-treated papaya during ripening

Caroline Giacomelli Soares[1,2], Samira Bernardino Ramos do Prado[1,2], Sónia Cristina da Silva Andrade[3], João Paulo Fabi[1,2,4]

[1] Department of Food Science and Experimental Nutrition, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.

[2] Food Research Center (FoRC), CEPID-FAPESP (Research, Innovation and Dissemination Centers, São Paulo Research Foundation), São Paulo, Brazil.

[3] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade São Paulo, Brazil.

[4] Food and Nutrition Research Center (NAPAN), University of São Paulo, São Paulo, Brazil.

Abstract

Ethylene induces many downstream modifications during climacteric ripening in a complex coordinated transcription level. Papaya is a fleshy fruit that undergoes heavy ethylene-induced modifications and turns into a palatable but fragile fruit due to pulp softening. RNA-seq analysis of ethylene-treated and non-treated papaya enabled a wide transcriptome overview to elucidate the expression profile during ripening. It was detected over seventeen thousand differentially expressed genes which were assigned to more than five thousand Gene Ontology terms. Ethylene induced and repressed pathways related to developmental process, response to stimulus, cellular component organization or biogenesis, and both primary and secondary catalytic and metabolic processes. Correlation among expression and physiological data led to clarification about ripening-induced modifications in the expression profile of the autocatalytic synthesis and signaling of ethylene, gene regulation by transcription factors, as well as regulation of changes in fruit color, firmness, aroma, and flavor. Hub genes involved in cellular signaling and synthesis of vitamins and antioxidants were detected by modular and correlation network analyses, portraying the several approaches a transcriptomics analyses result. This study provides an overview of the several genetic and molecular changes related to the ripening process and induced by exogenous ethylene. In addition, this transcriptomic analysis set precedents to unveiling candidate genes that may be useful to increase the post-harvest life of papaya fruit.

Keywords: ethylene; climacteric fruit; Papaya; transcription factors; ripening-related modifications; post-harvest; transcriptome.

Abbreviations: ACO, ACC oxidase; ACS, ACC synthase; ACC, 1-aminocyclopropane-carboxylic acid; SAM, s-adenosylmethionine; MTA, 5-methylthioadenosine; TFs, transcription factors; 1-MCP, 1-methylcyclopropene; ppm, particles per million; NCBI, National Center of Biological Information; GO, Gene Ontology; DEG, differentially expressed gene; GS, gene significance; MM, modular membership.

Introduction

Fleshy fruit are extremely popular for their high content of nutrients, sugar, and water when ripe, which makes them a delicious, health – and lucrative – snack for humans as well as an attractive for animals responsible for seed-dispersal. However, the waste rate of fleshy fruit is towering mostly due to post-harvest losses caused by fast ripening and consequent short shelf-life (FAO, 2011). Climacteric fleshy fruit ripening is coordinated by the mutual action of plant hormones, such as ethylene, which regulate several stages of the global gene expression of a fruit (TUCKER *et al.*, 2017). The autocatalytic production of ethylene occurs through the action of ACS and ACO enzymes on SAM and ACC molecules respectively, in which the final product is ethylene. Meanwhile, the Yang Cycle is fed with a side product of SAM, the MTA, which is converted to methionine, the general precursor of ethylene biosynthesis (CHEN *et al.*, 2018). Downstream regulation of climacteric ripening is then carried on by a set of transcription factors, specific for each fruit. A scheme established by LÜ *et al.* (2018) claims that tomatoes make use of MADS TFs and papayas make use of of NAC TFs, while bananas combine both of systems generating the ability to synthesize ethylene even towards its inhibitor, 1-MCP.

On the onset of ripening, there is a burst of ethylene that triggers a cascade of molecular and biochemical pathways which accelerates and induces the ripening process, promoting pulp softening, color changes, and enhancement on aroma, flavor, and nutrient content (FABI *et al.*, 2012). However, such changes also turn the fruit more susceptible to physical damage and microbial pathogens, which is the case of papaya. On account of the commercial interest in reverting the post-harvest losses, many studies on genetic regulation of ripening have been done in an attempt to understand the action of hormones on gene transcription and the detailed steps involved in this regulation, such as the isolation of primary and secondary transcription factors and specific receptors (LI *et al.*, 2013; FU *et al.*, 2021). In addition, physiological modifications stimulated by TFs and the enzymes involved in these processes are heavily investigated (FABI *et al.*, 2014; DO PRADO *et al.*, 2016; FABI, DO PRADO, 2019).

Since the modulation of gene expression is associated with the transcription level of DNA to mRNA, the study of the transcriptome might reveal a comprehensive, accurate estimate of the contribution of the genes to the physiological change of

interest. Nevertheless, in-depth knowledge about the full genetic control of papaya ripening is still elusive, thus the present study seeks to establish a general panorama of gene expression during ripening based on RNA-seq technology, availing of exogenous ethylene treatment to accelerate and intensify ripening effects.

Material and methods

Plant material and experimental design

Papaya fruit (*Carica papaya* L. cv. 'Golden') at the pre-climacteric stage were acquired from a commercial producer in the municipality of Linhares/ES, Brazil, (19°21'33.9"S 40°08'15.6"W) between one and two days after harvest, still with up to one-quarter of yellow peel and in three consecutive harvesting times (August, September, October 2017). Right after the fruit arrived in the laboratory and following sanitization with chlorine (100 ppm), some of them were immediately characterized according to FABI *et al.* (2007) and frozen to compose the 0-hour control group. The remaining fruit were separated into two groups: treated and control. Control groups were left to ripen with the parameters described in FABI *et al.* (2007) and fruit samples were taken after 12hs and 24hs. Treatment with ethylene at 0hs was done according to FABI *et al.* (2007) by exposing randomly selected fruit to a concentration of 100 ppm (100 μLL^{-1}) of ethylene, kept in constant flux for 17 minutes for gas saturation and 12 hours more in a closed system. After 12hs, fruit were removed from the ethylene chamber and exhaustively air-vented for one hour. Ethylene-treated fruit were frozen (12hs group) and other fruit were left to ripen as the control group for more twelve hours, totalizing 24hs after fruit have reached the laboratory (group 24hs treated). In all five groups (0hs, 12hs control, 24hs control, 12hs ethylene treated and 24hs ethylene treated) fruit were characterized for ripening parameters, peeled, their seeds were removed, and the pulp was cut in cubes, frozen in liquid nitrogen, and stored at -80 ºC (Figure S1). The experiment used three sets of harvested fruit collected at 3 different times of the year (biological triplicate) and each group was composed of at least five fruit (which were pooled before storage at -80 ºC).

Analysis of the ripening parameters

The ethylene produced by fruit were analyzed through flame ionization gas chromatography (GC-FID) as described in FABI *et al.* (2007). Besides visual inspection, a colorimeter (CR 410, Konica Minolta) was used to measure the color of the fruit peel. Following the CIE System, *a*, *b*, and L parameters were taken from six spots and then transformed in Hue angle, which is higher if the peel is green and lower if it is yellow (FABI *et al.,* 2007). The peel resistance was measured using a penetration probe of 9 mm diameter at three different spots of the fruit, aiming to represent the whole fruit's thickness. Then, fruit were cut longitudinally, and pulp resistance was measured using a penetration probe of 6 mm diameter at six different spots (FABI *et al.*, 2007).

Total RNA extraction

The total RNA of the fruit was isolated using Concert™ Plant RNA Reagent (Invitrogen®) and purified by treatment with the DNA-free™ kit (Invitrogen®), following the protocol described by the manufacturer. Nucleic acids were quantified spectrophotometrically using the Implen® N50 spectrophotometer; one microliter of the sample was used for quantification, performed in technical triplicate. Quality was assessed by the absorbance readings at 260 nm/280 nm and 260 nm/230 nm - relative to concentrations of nucleic acids concerning possible impurities such as proteins and secondary metabolites and/or components from the extraction process, respectively – and 1 % agarose gel electrophoresis. At the end of the entire process, the RNA samples were subjected to a second quantity, purity, and quality analysis using the Agilent 2100 Bioanalyzer apparatus (Agilent Technologies ©).

Illumina sequencing (RNA-Seq)

The cDNA libraries were sequenced at the Center of Functional Genomics, ESALQ-USP (Piracicaba/Brazil). A total of 15 cDNA libraries were prepared using the Illumina TruSeq Stranded mRNA LT Sample Prep Protocol and paired-end sequenced using HiSeq SBS Kit v4 on the HiSeq2500 system (Illumina) and were deposited under the access number of **GSE128577**. The raw RNA-seq data was initially checked for its

quality through FastQC software (ANDREWS, 2010), filtered by SeqyClean (ZHBANNIKOV *et al.*, 2017), and the alignment of the reads against the papaya reference genome, available at NCBI (Papaya1.0, GCF_000150535.2), was performed by the software STAR (DOBIN *et al.*, 2012) along with the count of the reads uniquely aligned. The BlastX tool was used to align the extracted transcript sequences to Viridiplantae (txid33090) nr database.

Differential expression analysis from RNA-seq

For the statistical analysis of the differentially expressed genes, the R (R Core Team, 2014) software and the DESeq2 package (LOVE; HUBER; ANDERS, 2014), specific for differential expression analysis, were used. The comparisons were done between all samples, thus developing a temporal overview of maturation and verifying the effect of the treatment. Genes with adjusted *p*-value ≤ 0.05 were considered differentially expressed.

Enrichment analysis

All genes were annotated and mapped to referent Gene Ontology terms using the Blast2GO tool (Biobam Bioinformatics S.L., Valencia, Spain). The identified differentially expressed genes with log2FC ≥ |1.5| and adjusted p-value ≤ 0.05 were submitted to GO enrichment analysis. Fisher's Exact Test was performed in two-tailed mode to detect both over and underrepresented GO Terms with a threshold of FDR ≤ 0.005, for the categories of biological process, molecular function, and cellular component. Gene Set Enrichment Analysis (GSEA – SUBRAMANIAN *et al.*, 2005) was carried for specific sets of genes.

Co-expression analyses

Co-expression network analysis was performed using the WGCNA (Weighted Gene Correlation Network Analysis) R package (LANGFELDER; HORVATH, 2008). The tool calculates the connectivity of genes and the topological overlap based on the Pearson correlation raised to a β constant and then detects and selects modules and eigengenes, a set of genes present in the first Principal Component that represent the

expression profile of the modules. Furthermore, it is made a second correlation between the modules and experimental traits data, which in the case of this study was the ethylene production measurements. Pearson's correlations are established for each gene significance to the treat of interest, and its respective module. In this respect, the result is a proportional and clustered correlation network concerning the degree of connectivity of the transcripts. The visualization of the correlation network was built using Cytoscape software (SHANNON *et al.*, 2003). The pipeline of the RNA-seq methodology used in this work is summarized in Figure 1.

Statistics

Experimental results are expressed as the mean ± standard deviation (SD) obtained from all fruit of each one of the five groups from the three biological replicates, with $p \leq 0.05$ considered to represent statistical significance. Data were analyzed using GraphPad Prism 6.0 software (GraphPad Software, San Diego, CA, USA). One-way analysis of variance (ANOVA) with Tukey's test (to assess differences among all groups) was used as a post hoc test. FastQC results were aggregated and analyzed through the MultiQC tool (EWELS *et al.*, 2016). Differential expression analysis followed the empirical Bayes approach to assign moderated estimation of dispersion and fold change values of pairwise comparison between all samples, while co-expression assay was based on Pearson correlation (LANGFELDER; HORVATH, 2008; LOVE; HUBER; ANDERS, 2014). Enrichment analysis was run through Fisher's exact test with a cutoff of adjusted FDR ≤ 0.005 (GÖTZ *et al.*, 2008). GSEA uses weighted Kolmogorov–Smirnov test to assign an Enrichment Score (ES) based on a ranked list of genes (SUBRAMANIAN *et al.*, 2005). Venn diagrams were drawn to observe differentially expressed genes between two or more comparisons, and also those exclusive to each sample. The same was done to enriched pathways from different sets of genes. Co-expression correlations greater than 0.05 were considered significant, with further filtering by the cut-offs of GS ≥ |0.7| and GS.p-value ≤ 0.05, as well as MM ≥ |0.09| and MM.p-value ≤ 0.05.

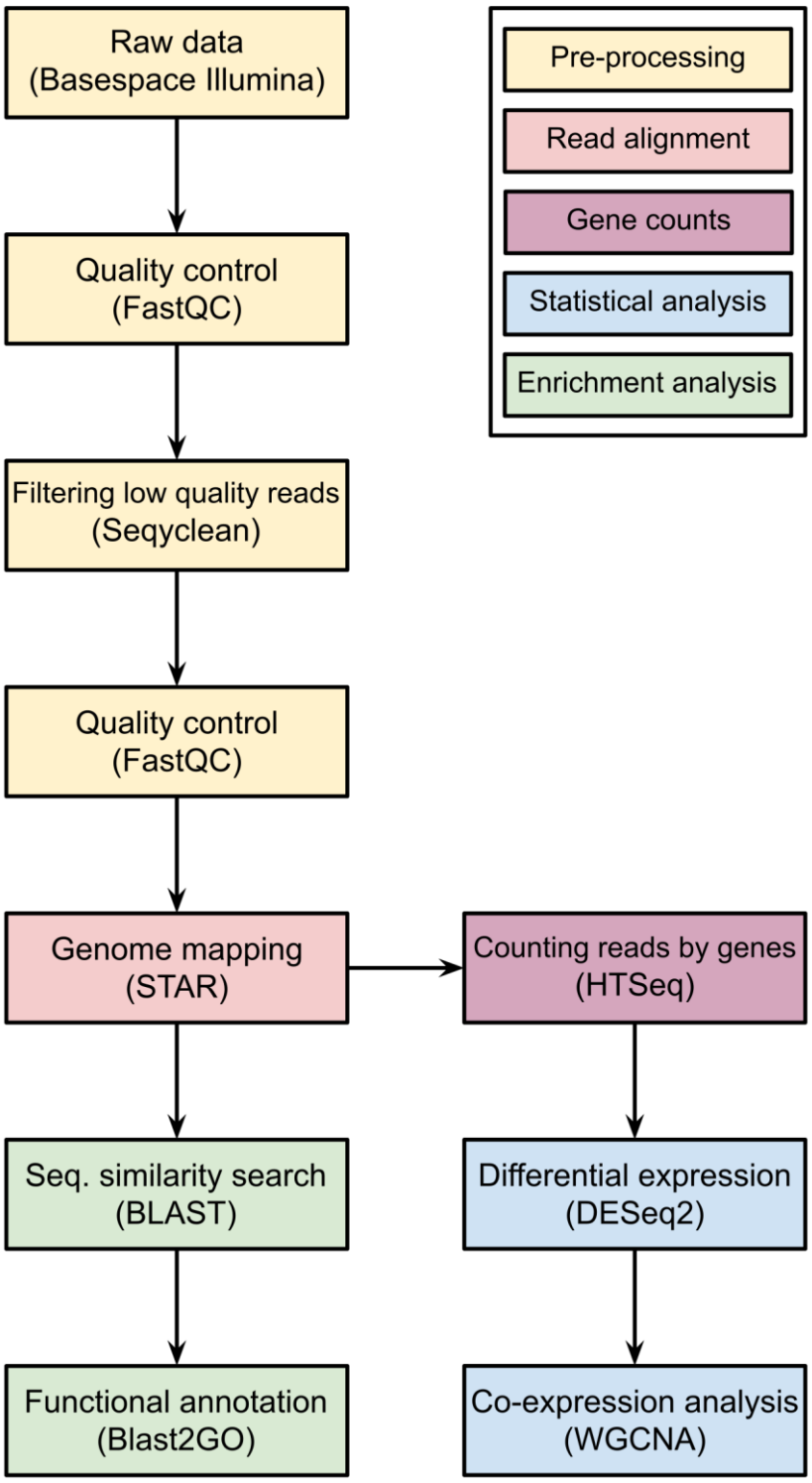


Figure 1. A pipeline of papaya pulp RNA-seq processing and treatment. Flowchart of activities performed during the processing and analysis of RNA-seq data. The colors represent the different stages of the pipeline.

Results and discussion

Ethylene treatment induced the papaya ripening

Papayas are climacteric fleshy fruit characterized by fast pulp softening in the post-harvest period (FABI *et al.*, 2007). As expected, ethylene treatment triggered the production of endogenous ethylene as it can be compared with control samples after 12h and 24h the treatment timeline (Figure 2). It is observed in Figure 2 that the release of ethylene showed an increasing pattern in treated fruit, while control ones maintained similar levels during the experiment. Texture measures from both whole fruit and fruit pulp showed decay in treated fruit with no changes in control ones. Skin color changes were only detected in one replicate (B), due to the short period in which the samplings were carried out. In general, the results obtained were consistent with what was expected after treatment with ethylene and simulating the post-harvest phase. From these results, it is possible to describe a maturity profile of the samples, which were made with fruit of the same origin and at the same time of the year, in addition to making it possible to collect plant material corresponding to the three biological replicates for RNA-seq analysis. The values obtained for all the analyzed parameters indicated a good homogeneity of all samples, a fact that confirms the similarity of the stage of ripening of the fruit when harvested and, mainly, unifies the ripening process (FABI *et al.*, 2007, 2014).

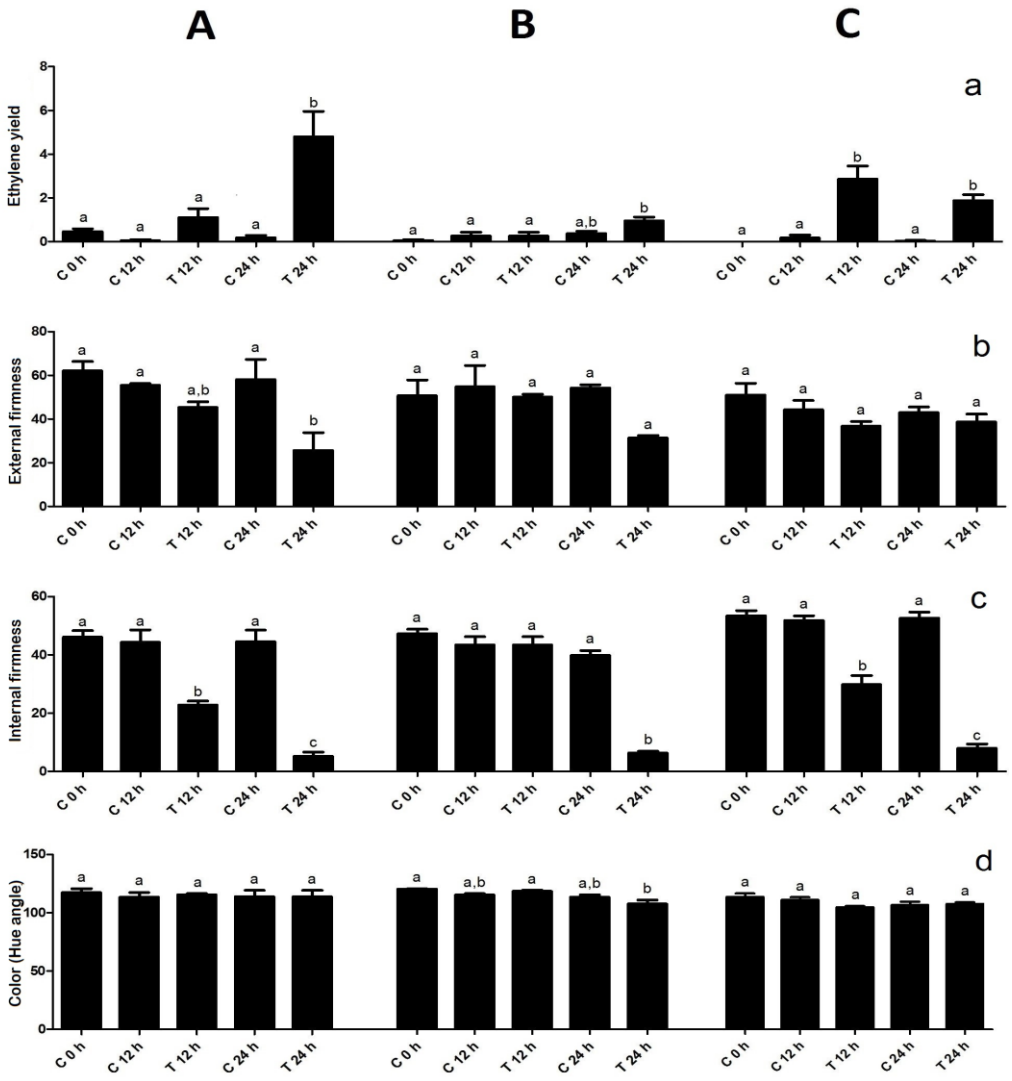


Figure 2. Ripening parameters measured for the three biological replicates for papaya treatment. Papayas for each time point and each biological replicate (A, B and C) were treated or not treated with ethylene and parameters such as ethylene production (µL.Kg-1.h-1 - a), external firmness (N - b), internal firmness (N - c), and pulp color (hue angle - d) were measured in each time point.

Transcriptomics

Approximately 167.9 million fragments were mapped against the reference genome (GCF_000150535.2, disposed of 370.5 million base pairs with a size of 372 megabases—MING *et al.*, 2008), corresponding to about 86 % of the fragments that were kept after the filtering process (Figure S2C), while just over 13 % were not mapped. Only 1,736 of a total of 23,332 genes on the reference genome had no fragment aligned. Although the number of fragments per sample varied, the proportion

of mapped fragments remained similar, and the base content was balanced after trimming (Figures S2A, S2B). Based on RNA-seq data, it was possible to elucidate a temporal line of both the control and ethylene-induced ripening process and metabolic changes of papaya fruit.

A BLASTx against the non-redundant data set from NCBI identified a total of 19,316 genes, a second search against *Arabidopsis thaliana* genome resulted in 18,378 homologous genes. After a manual screening, only 1,355 genes remained unidentified. Twelve thousand hundred fifty genes were assigned to 5,349 GO terms within the categories of Biological Process (3,020), Molecular Function (1,695), and Cellular Component (634). Functional pathways with the highest number of genes refer to cellular and metabolic processes, responses to chemical signals and stress, system development in general, binding, catalytic activity, transcription regulation, and intracellular components and organelles (Figure S3).

Changes on transcriptome profile during ripening

Pairwise differential expression analysis between all control and treated samples identified 8,936 and 8,726 significantly up- and down-regulated genes, respectively. In a refined perspective, the number of genes whose expression had at least 50% of variation is represented in Figure 3. Comparisons between control and treated samples showed a higher number of DEG while differences between control samples remained minimal, suggesting that exogenous ethylene enhanced gene transcription indeed. The similarity of the comparisons between Control 24h and both treated samples implies that the control sample has initiated endogenous-ethylene induced modifications only at the 24h point of the experiment, evidencing the acceleration of the ripening process by the exogenous ethylene treatment.

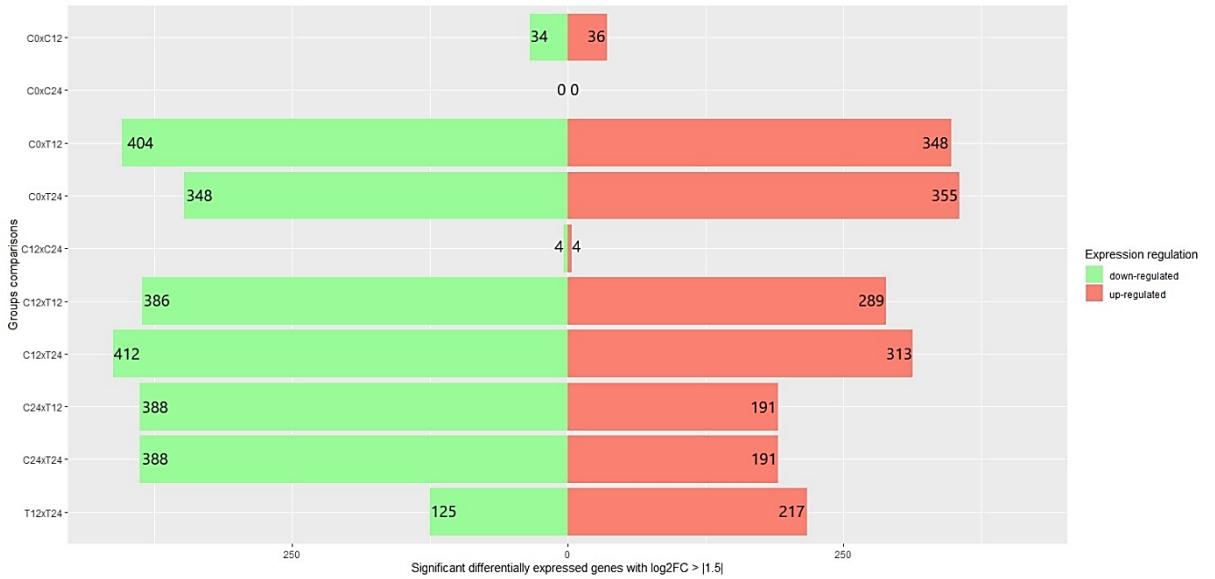


Figure 3. Differentially expressed genes. Total numbers of differentially expressed genes with p-adj ≤ 0.05 and log2FC ≥ |1.5| between pairwise comparisons. Green bars represent down-regulated while red represent up-regulated genes. Samples are identified as C (control) or T (treated) followed by the time point.

Among control groups, down-regulated genes are enriched in many pathways of response to extracellular stimulus and anatomical structures development, with active transmembrane transporter activity, while repressed GO terms are mostly involved in RNA metabolic processes. Up-regulated genes in control samples are strongly involved in response to stimulus and hormones, catalytic activity of many enzymes such as mannosidase and glucosidase, and transport of fluids, whereas repressed pathways are related to intracellular anatomical structures and translation. Considering comparisons between control and treated samples, up-regulated genes had a positive regulation for programmed cell death, camalexin biosynthetic process (pathogen defense), abscisic acid metabolic process, plant-type cell wall organization with the activity of xyloglucan:xyloglucosyl transferase and UDP-glucosyltransferase, root development, and leaf senescence. In contrast, repressed pathways of up-regulated genes are generally involved in translation. Ethylene-treated samples' expression underwent down-regulation in genes enriched to negative regulation of cell growth, inflorescence morphogenesis, and external stimulus, while translation and intracellular protein transport decreased. In order to clearly visualize changes, enriched pathways detected through GSEA of Control 0h versus Treated 24h are shown in Figure S4.

Correlation among expression and phenotypic data

After clustering the samples, control and treated samples got separated into two main branches (Figure 4), demonstrating that exogenous-ethylene treatment resulted in big, plain modifications in expression profile. Sample C24h from sampling B was an exception, although it still differs from treated samples on firmness values. Ethylene production rate increase is evident in the treated groups, while external and internal firmness displayed the opposite behavior. Peel color got slightly lighter, even though changes were minimal, probably because peel color is not uniform, and the plotted value is the mean of Hue angles of several points of the fruit.

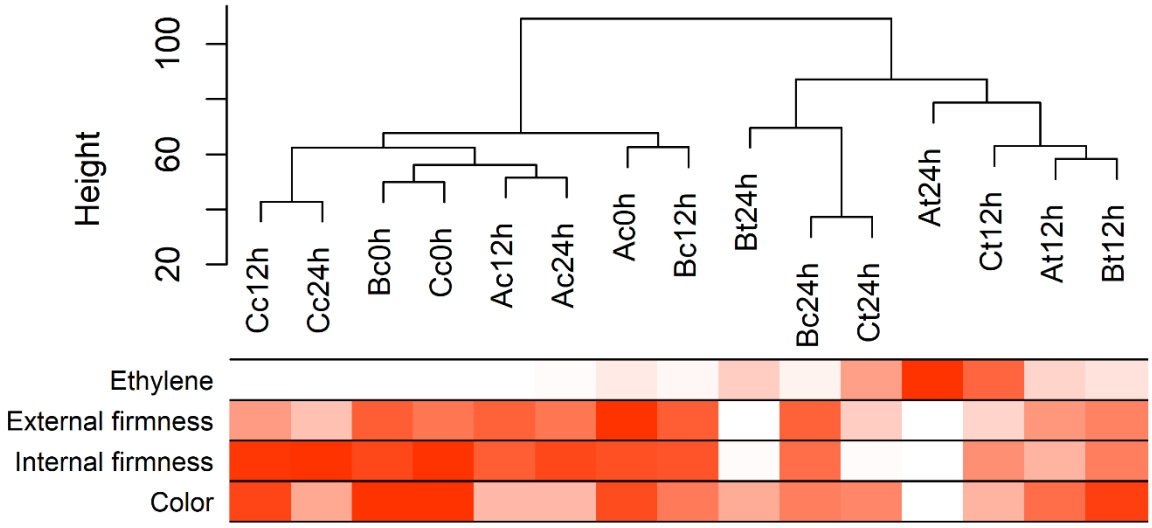


Figure 4. Sample dendrogram and phenotypic heatmap. Cluster and heatmap of the parameter's measures, where white is low and red is high relative value.

Modular analysis by expression profile correlation identified 29 modules with 40 (darkolivegreen) to 7,649 (brown) members. Further correlation among modules and phenotypic parameters followed the heatmap (Figure 4) pattern: modules with a high correlation to ethylene yield are inversely correlated to other phenotypes (Figure 5). Significant correlations (p-value ≤ 0.05) were detected in modules brown, orange, black, midnightblue, and lightgreen to ethylene yield, as well as inverse correlation to the other three parameters (except lightgreen). This opposite correlation pattern suggests that, in addition to the ethylene-related genes, members of these modules are also involved in maturation processes such as cell wall degradation and carotenoid synthesis.

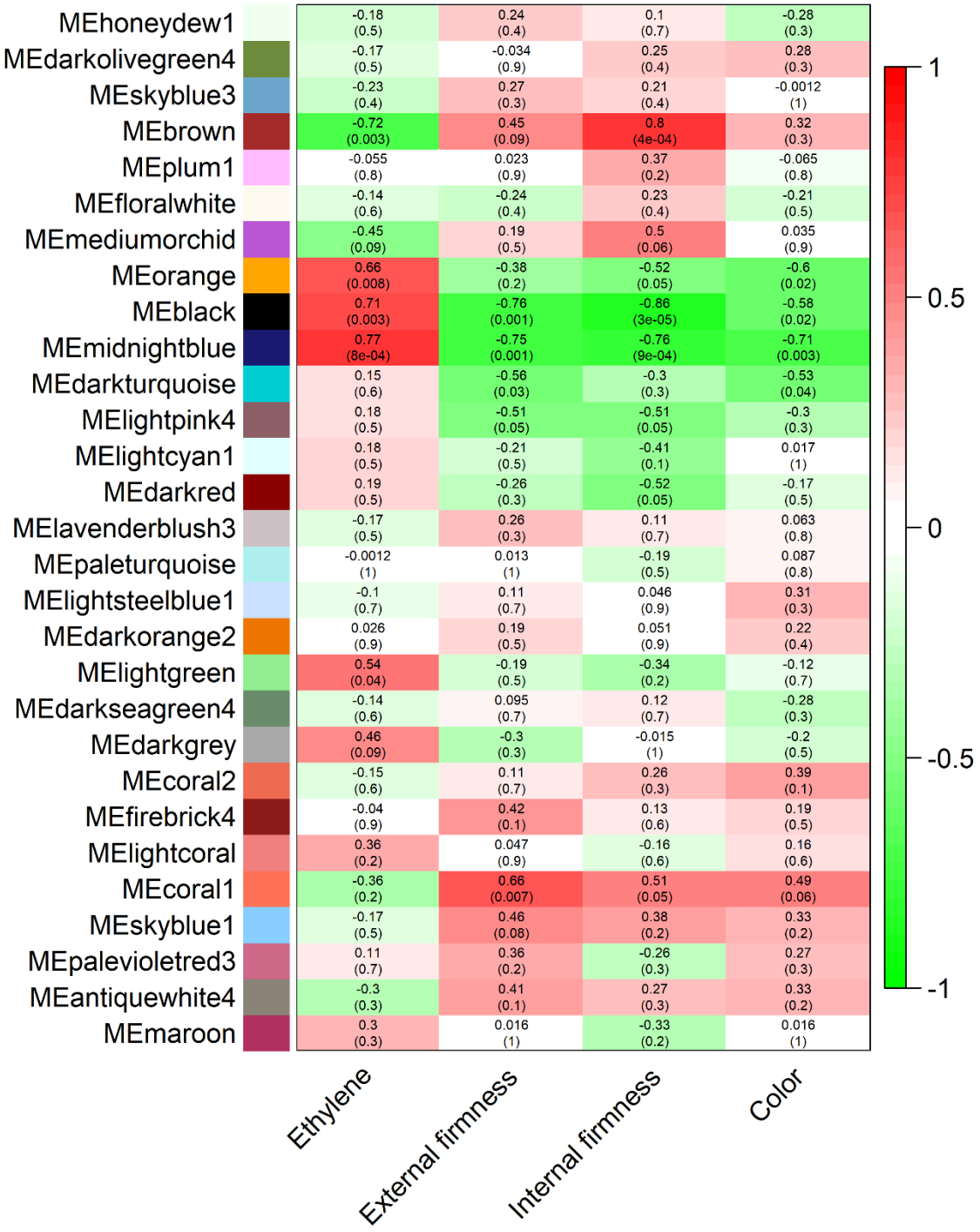


Figure 5. Module-phenotype correlation. Heatmap of correlation between modules and physiological measures. Each cell informs the correlation level and *p*-value.

Enriched pathways of genes in black module are associated with processes of uracil catabolism, and lipid and fatty acid metabolism, while the activity of cellular differentiation and development, and plant-type cell wall organization or biogenesis were repressed. Midnightblue members are assigned to high involvement in antioxidant activity, like vitamin E and fat-soluble vitamin biosynthetic process.

Lightgreen genes are pointed to positive participation in the primary metabolic process, and negative in transcription. Orange and brown modules had no significantly enriched pathways. Genes and functional pathways of the primary metabolism with positive regulation, such as fatty acid and lipid metabolism, suggest an increased activity related to the ripening process with a consequent increase in energy intake of the fruit (FABI *et al.*, 2012). Lipid content is also associated with carotenoids level, thus the negative correlation of black module with color measurements corroborate to an increase of major carotenoids and decrease of total chlorophyll in yellow-peel fruit (SHEN *et al.*, 2019). A negative correlation between black module and firmness, together with the repression of fruit cell wall organization and biogenesis, endorses the disassembling of the fruit cell wall during ripening (DO PRADO *et al.*, 2016).

As modules black and midnightblue are significantly correlated to all physiological parameters, they were further investigated through a correlation network in order to identify hub genes that might regulate the expression profile of other members. Many more hub genes were identified in black module due to its number of members, which is over three times the total of members of midnightblue (Figure 6). In midnightblue module, the nodes positively correlated to ethylene increase with more neighbors are tocopherol O-methyltransferase (XP_021899610.1), laccase (XP_021897387.1) and a MLO-like protein (XP_021906880.1), while RNA polymerases I and III subunit (XP_021905792.1) and 10 kDa chaperonin (XP_021907659.1) had an inverse, strong correlation. Hub genes of the black module either directly or inversely correlated to ethylene, are mostly involved in cellular signaling, such as actin-related protein (XP_021911702.1), F-box protein SKIP5 (XP_021886875.1), and a CCR4-NOT (XP_021898814.1). Hub genes are highly associated with the modules' functional pathways previously described (Table S1), indicating that this co-expression modular approach is a powerful method to tape significant results among such quantity of information that transcriptomics reveals.

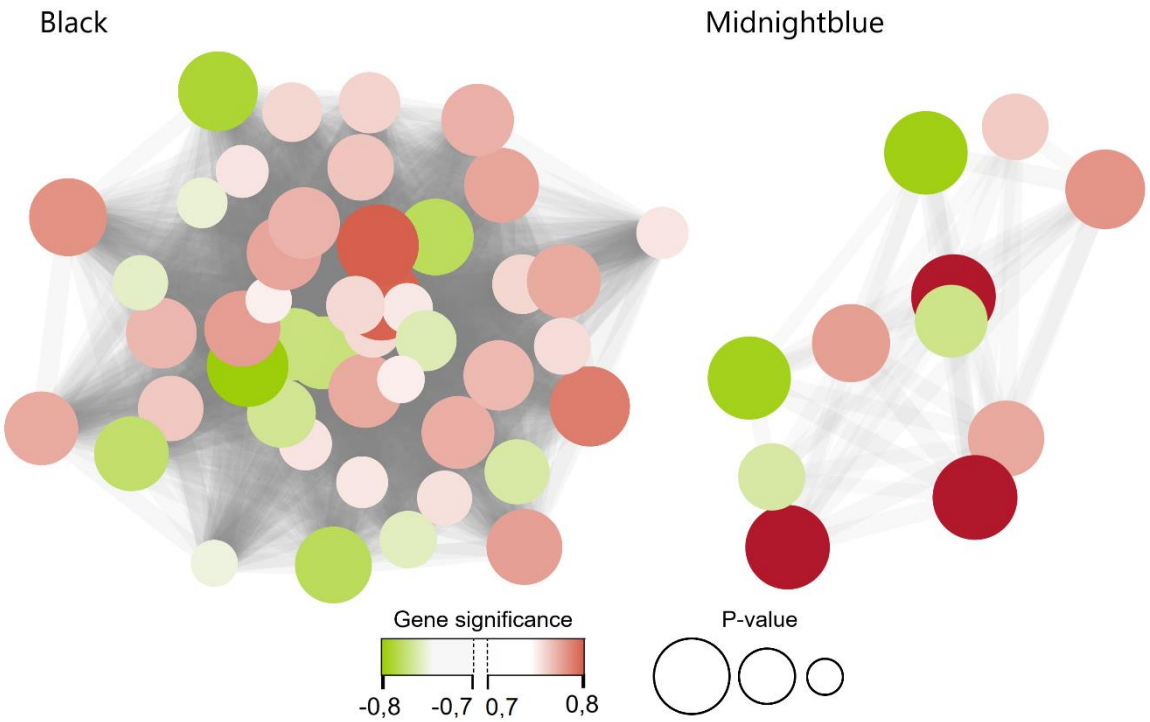


Figure 6. Co-expression correlation network of modules Black and Midnightblue hubgenes. Each circle is a gene whose color is negative (green) and positive (red) significance for the module correlation to ethylene yield measures. The size of the circle is proportional to the *p*-value of the significance, being the lower the *p*-value is, the bigger the circle is. The length and te color of the edges are related to the weight of the correlation, being the longer and lighter an edge is, the lower is the weight.

Transcription factors

Expression profile is predominantly regulated by several key proteins able to recognize and bind to specific DNA targets and either promote or repress their transcription in a coordinated and cooperative manner (SPITZ; FURLONG, 2012). Investigating the expression of these proteins, known as transcription factors, can elucidate a large part of the changes happening during the ripening process. Control fruit had an alteration in only three TF: CCR4-NOT complex, a negative regulator of transcription (LIANG *et al.*, 2009), was repressed; and two genes encoding WRKY, related to the response to abiotic and biotic stress (PAN; KIANG, 2014), were up-regulated. Treated fruit had no difference in the expression of CRR4-NOT, but WRKY was intensely suppressed, endorsing the many response pathways that were repressed in the functional analysis. The MYB family, associated with anthocyanin synthesis and cell development, had three genes with decreased expression even with the increase in the color intensity of the fruit pulp, possibly because papaya has a greater influence on carotenoids than anthocyanins (FABI *et al.*, 2007). Seven genes encoding for basic helix-loop-helix (bHLH), one of the largest families of transcription

factors involved in many basal biological processes such as plant growth, development, and defense, presented isoforms with different levels of expression in treated groups (YANG *et al.,* 2020). Heat shock proteins (HPSs) are numerous and involved in a wide range of processes, usually stress-mediated due to their chaperone-like function (WATERS; VIERLING, 2020). Their expression was opposite in the treated groups, being decreased in 12h treatment and increased in 24h treatment, indicating a possible increase in protein folding and consequently greater biological activity of proteins with the advancement of the fruit ripening process.

Hormones have a direct influence on TFs activity acting both synergistically and antagonistically among them. Auxin is mainly related to elongation and reproductive organs, besides participating with ethylene in the regulation of the ripening process (LIU *et al.*, 2015). It was detected 3 ARFs (auxin-response factor) up- and down-regulated in the treated groups, while auxin response repressor genes (IAA) were suppressed and those in response to excess of auxin (GH3) were induced. Thus, there was some activity related to auxin, but its TFs specific function is still not completely elucidated. TFs related to ethylene response were abundant and presented strongly different levels of expression (Figure 7). ACO, ACS, and SAM-Mtase were predominately up-regulated, indicating a high rate of ethylene biosynthesis. EIN is a mediator between ethylene synthesis and signaling, acting as a “gateway” when CTR1, a negative regulator, is reduced by the presence of ethylene. However, there were few EIN genes differentially expressed, and they were mostly repressed. Thus, ethylene signaling operates through other routes independently of CTR1 or EIN (BINDER, 2020).

AP2/ERF-encoding genes are involved in abiotic stress response and are very active during development and ripening processes (LI *et al.*, 2013). From AP2/ERF family, almost all RAPs and half of the ERFs were repressed while other ERFs and AP2 were up-regulated. The NAC family is one of the largest families of plant transcription factors whose genes regulate several processes of development and maintenance of the system, since embryogenesis until leaf senescence (NURUZZAMAN; SHARONI; KIKUCHI, 2013). NAC genes related to the regulation of the development of secondary cell wall fibers and several stress responses, and auxin signaling were up-regulated, while the ones involved in the response to cold and pathogens was repressed (ZHONG *et al.*, 2008; NURUZZAMAN; SHARONI; KIKUCHI, 2013). The MADS-box family is also a big TFs family involved in the

regulation of many routes and had its genes expression mostly increased. In contrast, LÜ *et al.* (2018) unveiled that the positive feedback to the autocatalytic cycle of ethylene in papayas is not regulated by MADS-box family, but by NAC. As these TFs families are too vast and the specific function of each gene is still not completely elucidated, the assumption would be that MADS-box DEGs are related to activities other than ethylene signaling regulation.

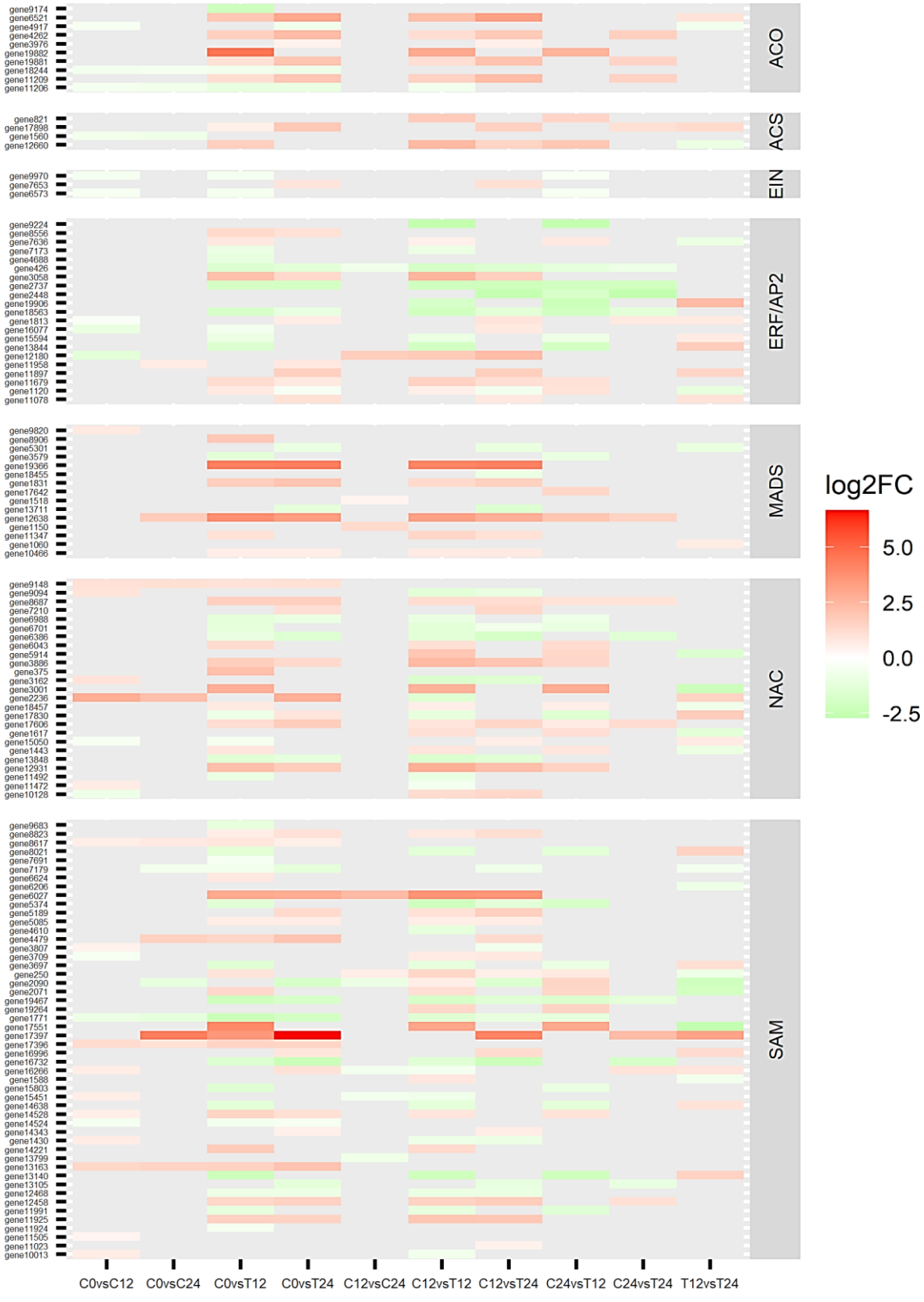


Figure 7. Ethylene-related genes expression. ACO: 1-aminocyclopropane-1-carboxylate oxidase; ACS: 1-aminocyclopropane-1-carboxylate synthase; EIN: ethylene-insensitive protein; ERF/AP2: ethylene-responsive transcription factor; MADS: agamous-like MADS-box protein; NAC: NAC domain-containing protein; SAM: S-adenosyl-L-methionine-dependent methyltransferases superfamily protein.

Discussion

RNA-seq analysis provided thousands of differentially expressed genes between ethylene-treated and non-treated papaya pulp, which together with the evident contrast of physiological parameters among samples, represent distinct stages of the fruit ripening. Major changes previously described as caused by ethylene-induced enzymes were observed both in physiological measures and in the expression profile of the genes that encode to the enzymes. Several fleshy fruits undergo a quick softening of pulp resulted from the cell wall polysaccharides disassembling promoted by many enzymes, such as strawberries and tomatoes (ZHANG *et al.*, 2015; SHINOZAKI *et al.*, 2018). Several genes involved in this process were DEGs, such as polygalacturonase, pectinesterase, cellulase, among others, and they were also enriched to the transport of fluids and water and repressed to fruit cell wall organization functional pathways. In conclusion, the action of these enzymes dissolute fruit cell wall polysaccharides, decreasing cell turgor and promoting pulp softening (DO PRADO; FABI, 2019). Carotenoid-accumulation is the major factor in color changing in papaya fruit and acts as an antioxidant, highly associated with ethylene levels which synthesis is related to lipid, fatty acids, and pyruvate metabolism, as well as lycopene biosynthetic process (SAINI; ZAMANY; KEUM, 2017). Ethylene-treated samples showed an enhancement of pulp color together with the over expression of many carotenoid-related genes, whose functional pathways were enriched. In addition, there were two carotenoid-related genes acting as hub genes in the black module.

Aroma enhancement is significantly represented by the enrichment of alcohol dehydrogenase and linalool synthase activity, which synthesizes the main volatile compound of papaya and is directly induced by ethylene (PINO, 2014). Although papaya is popular for its sweet taste, the sucrose metabolic process was suppressed in treated samples while many functional pathways of sucrose and glucose were enriched for down-regulated genes. Still, there were few up-regulated genes encoding sucrose synthase as described in GOMEZ; LAJOLO; CORDENUNSI (2002). Although total soluble sugars content does not change much after harvesting, sucrose synthase keeps its activity by consuming carbon from the degalactosylation of cell wall polysaccharides. Another factor that makes papaya a highly consumed fruit is its antioxidant and nutrient properties. Even though the maturity of papaya leads to a decrease in flavonoids and phenolic content (SANCHO; YAHIA; GONZÁLEZ-

AGUILAR, 2011), it was still detected genes and functional pathways induced in the treated 12h sample. Vitamins A, C, and E gene-related also had their pathways enriched and were previously related to increasing during ripening (FABI *et al.*, 2009).

Importantly, ethylene synthesis and signaling routes were greatly enriched with many genes being differently expressed among non-treated and treated fruit, pointing to the success of exogenous ethylene treatment to accelerate the ripening process and accentuate ripening-related modifications, making papaya a promising candidate to post-harvest researches as a model fruit for fast pulp softening and overall ripening other than tomato. Whilst both fruits share many primary and secondary pathways, there are still inconsistencies between their molecular functioning, such as the regulation of ethylene-related transcription factors. Moreover, the smaller size of the papaya genome related to species commonly used in ripening studies, such as tomato, is a facilitator regarding gene function prediction (PAULL *et al.*, 2008).

Conclusion

Transcriptomics is a robust approach to study the spatiotemporal profile of the expression, generating a lot of important data which should be minutely investigated to fully elucidate the pathways that led and suffered modifications. In this work, it was presented the most significant changes and their role in the ripening process of papaya. Differences between control and treated samples confirm the efficiency of exogenous ethylene treatment to induce ripening. Genes related to ethylene metabolism, carotenoids, fruit cell wall modifications, response to several stimuli, and transcription activity have been identified and superficially described. Co-expression correlation modules facilitate the assignment of functional pathways, spreading insights on their members general role. Their submission to network analysis enabled the identification of hub genes, which is an important approach to design a complete directed pathway. Further investigation of subsets of genes is required to fully understand the complex coordinated processes that a climacteric fleshy fruit undergoes during ripening.

References

ANDREWS, S. FastQC: a quality control tool for high throughput sequence data. 2010; Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

BINDER, B. M. Ethylene signaling in plants, **The Journal of biological chemistry**, v. 295, n. 22, p. 7710–7725, 2020. DOI: 10.1074/jbc.REV120.010854.

CHEN, Y., GRIMPLET, J., DAVID, K., *et al.* Ethylene receptors and related proteins in climacteric and non-climacteric fruits, **Plant Science**, v. 276, p. 63–72, 2018. DOI: 10.1016/j.plantsci.2018.07.012.

DO PRADO, S. B. R., MELFI, P. R., CASTRO-ALVES, V. C., *et al.* Physiological degradation of pectin in papaya cell walls: Release of long chains galacturonans derived from insoluble fractions during postharvest fruit ripening, **Frontiers in Plant Science**, v. 7, n. JULY2016, p. 1–11, 2016. DOI: 10.3389/fpls.2016.01120. .

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., *et al.* STAR: ultrafast universal RNA-seq aligner, **Bioinformatics**, v. 29, n. 1, p. 15–21, 1 jan. 2013. DOI: 10.1093/bioinformatics/bts635.

EWELS, P., MAGNUSSON, M., LUNDIN, S., *et al.* MultiQC: summarize analysis results for multiple tools and samples in a single report, **Bioinformatics**, v. 32, n. 19, p. 3047–3048, 1 out. 2016. DOI: 10.1093/bioinformatics/btw354.

FABI, J. P., CORDENUNSI, B. R., BARRETO, G. P. D. M., *et al.* Papaya fruit ripening: Response to ethylene and 1-methylcyclopropene (1-MCP), **Journal of Agricultural and Food Chemistry**, v. 55, n. 15, p. 6118–6123, 2007. DOI: 10.1021/jf070903c.

FABI, J. P., CORDENUNSI, B. R., SEYMOUR, G. B., *et al.* Molecular cloning and characterization of a ripening-induced polygalacturonase related to papaya fruit softening, **Plant Physiology and Biochemistry**, v. 47, n. 11, p. 1075–1081, 2009. DOI: https://doi.org/10.1016/j.plaphy.2009.08.002.

FABI, J. P., SEYMOUR, G. B., GRAHAM, N. S., *et al.* Analysis of ripening-related gene expression in papaya using an Arabidopsis-based microarray, **BMC Plant Biology**, v. 12, n. 1, p. 242, 2012. DOI: 10.1186/1471-2229-12-242.

FAO. Global food losses and food waste -Extent, causes and prevention. Rome, 2011. Available at http://www.fao.org/docrep/014/mb060e/mb060e00.pdf

FU, C.-C., CHEN, H.-J., GAO, H.-Y., *et al.* Papaya CpMADS4 and CpNAC3 co-operatively regulate ethylene signal genes CpERF9 and CpEIL5 during fruit ripening, **Postharvest Biology and Technology**, v. 175, p. 111485, 2021. DOI: https://doi.org/10.1016/j.postharvbio.2021.111485.

GOMEZ, M., LAJOLO, F., CORDENUNSI, B. Evolution of Soluble Sugars During Ripening of Papaya Fruit and its Relation to Sweet Taste, **Journal of Food Science**, v. 67, n. 1, p. 442–447, 1 jan. 2002. DOI: https://doi.org/10.1111/j.1365-2621.2002.tb11426.x.

GÖTZ, S., GARCÍA-GÓMEZ, J. M., TEROL, J., *et al.* High-throughput functional annotation and data mining with the Blast2GO suite, **Nucleic acids research**, v. 36, n. 10, p. 3420–3435, jun. 2008. DOI: 10.1093/nar/gkn176.

LANGFELDER, P., HORVATH, S. WGCNA: an R package for weighted correlation network analysis, **BMC Bioinformatics**, v. 9, n. 1, p. 559, 2008. DOI: 10.1186/1471-2105-9-559.

LI, X., ZHU, X., MAO, J., *et al.* Isolation and characterization of ethylene response factor family genes during development, ethylene regulation and stress treatments in papaya fruit, **Plant Physiology and Biochemistry**, v. 70, p. 81–92, 2013. DOI: https://doi.org/10.1016/j.plaphy.2013.05.020.

LIANG, W., LI, C., LIU, F., *et al.* The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses, **Cell Research**, v. 19, n. 3, p. 307–316, 2009. DOI: 10.1038/cr.2008.317.

LIU, K., YUAN, C., LI, H., *et al.* Genome-wide identification and characterization of auxin response factor (ARF) family genes related to flower and fruit development in papaya (*Carica papaya* L.), **BMC Genomics**, v. 16, n. 1, p. 901, 2015. DOI: 10.1186/s12864-015-2182-0.

LOVE, M. I., HUBER, W., ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, **Genome Biology**, v. 15, n. 12, p. 550, 2014. DOI: 10.1186/s13059-014-0550-8.

LÜ, P., YU, S., ZHU, N., *et al.* Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening, **Nature Plants**, v. 4, n. 10, p. 784–791, 2018. DOI: 10.1038/s41477-018-0249-z.

MING, R., HOU, S., FENG, Y., *et al.* The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus), **Nature**, v. 452, n. 7190, p. 991–996, 2008. DOI: 10.1038/nature06856.

NURUZZAMAN, M., SHARONI, A. M., KIKUCHI, S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. **Frontiers in Microbiology**. 2013. DOI10.3389/fmicb.2013.00248.

PAN, L., JIANG, L. Identification and expression of the WRKY transcription factors of *Carica papaya* in response to abiotic and biotic stresses, **Molecular Biology Reports**, v. 41, n. 3, p. 1215–1225, 2014. DOI: 10.1007/s11033-013-2966-8.

PAULL, R., IRIKURA, B., WU, P., *et al.* Fruit Development, Ripening and Quality Related Genes in the Papaya Genome, **Tropical Plant Biology**, v. 1, p. 246–277, 1 dez. 2008. DOI: 10.1007/s12042-008-9021-2.

PINO, J. A. Odour-active compounds in papaya fruit cv. Red Maradol, **Food Chemistry**, v. 146, p. 120–126, 2014. DOI: 10.1016/j.foodchem.2013.09.031.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2014) URL http://www.R-project.org/

SAINI, R. K., ZAMANY, A. J., KEUM, Y.-S. Ripening improves the content of carotenoid, α-tocopherol, and polyunsaturated fatty acids in tomato (*Solanum lycopersicum* L.) fruits, **3 Biotech**, v. 7, n. 1, p. 43, 2017. DOI: 10.1007/s13205-017-0666-0.

SANCHO, L. E. G.-G., YAHIA, E. M., GONZÁLEZ-AGUILAR, G. A. Identification and quantification of phenols, carotenoids, and vitamin C from papaya (*Carica papaya* L., cv. Maradol) fruit determined by HPLC-DAD-MS/MS-ESI, **Food Research International**, v. 44, n. 5, p. 1284–1291, 2011. DOI: https://doi.org/10.1016/j.foodres.2010.12.001.

SHANNON, P., MARKIEL, A., OZIER, O., *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks, **Genome research**, v. 13, n. 11, p. 2498–2504, nov. 2003. DOI: 10.1101/gr.1239303.

SHEN, Y. H., YANG, F. Y., LU, B. G., *et al.* Exploring the differential mechanisms of carotenoid biosynthesis in the yellow peel and red flesh of papaya, **BMC Genomics**, v. 20, n. 1, p. 49, 2019. DOI: 10.1186/s12864-018-5388-0.

SHINOZAKI, Y., NICOLAS, P., FERNANDEZ-POZO, N., *et al.* High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening, **Nature Communications**, v. 9, n. 1, p. 364, 2018. DOI: 10.1038/s41467-017-02782-9.

SPITZ, F., FURLONG, E. E. M. Transcription factors: from enhancer binding to developmental control, **Nature Reviews Genetics**, v. 13, n. 9, p. 613–626, 2012. DOI: 10.1038/nrg3207.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, **Proceedings of the National Academy of Sciences**, v. 102, n. 43, p. 15545 LP – 15550, 25 out. 2005. DOI: 10.1073/pnas.0506580102.

TUCKER, G., YIN, X., ZHANG, A., *et al.* Ethylene, **Food Quality and Safety**, v. 1, n. 4, p. 253–267, 2017. DOI: 10.1093/fqsafe/fyx024.

WATERS, E. R., VIERLING, E. Plant small heat shock proteins – evolutionary and functional diversity, **New Phytologist**, v. 227, n. 1, p. 24–37, 1 jul. 2020. DOI:10.1111/nph.16536.

YANG, M., ZHOU, C., YANG, H., *et al.* Genome-wide analysis of basic helix-loop-helix transcription factors in papaya (*Carica papaya* L.), **PeerJ**, v. 8, p. e9319, 2020. DOI: 10.7717/peerj.9319.

ZHANG, Y., LI, W., DOU, Y., *et al.* Transcript Quantification by RNA-Seq Reveals Differentially Expressed Genes in the Red and Yellow Fruits of *Fragaria vesca*, **PLOS ONE**, v. 10, n. 12, p. e0144356, 4 dez. 2015. DOI: 10.1371/journal.pone.0144356.

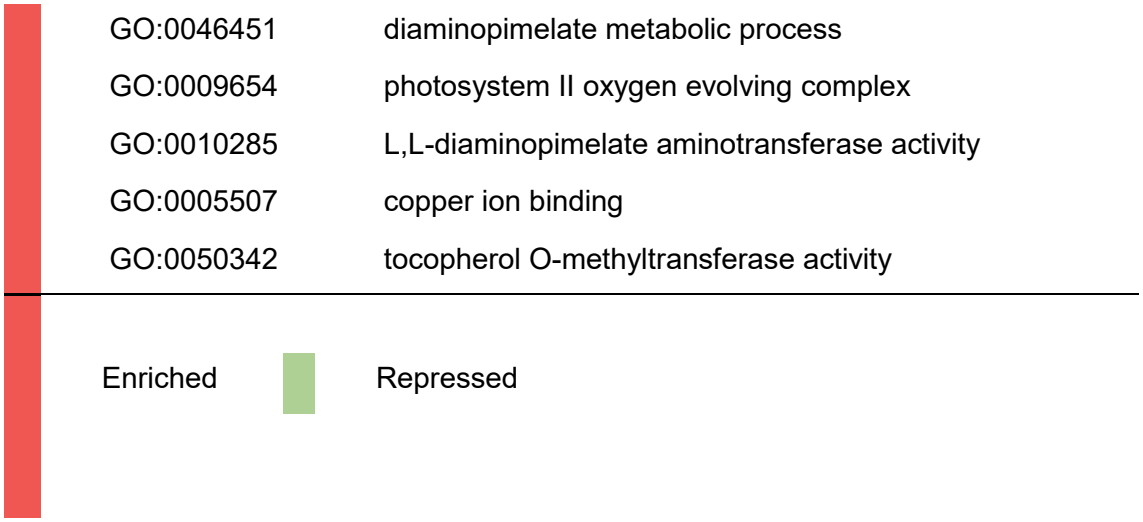
ZHBANNIKOV, I. Y., HUNTER, S. S., FOSTER, J. A., *et al.* SeqyClean: A Pipeline for High-Throughput Sequence Data Preprocessing. 2017. **Anais** [...] New York, NY, USA, Association for Computing Machinery, 2017. p. 407–416. DOI: 10.1145/3107411.3107446.

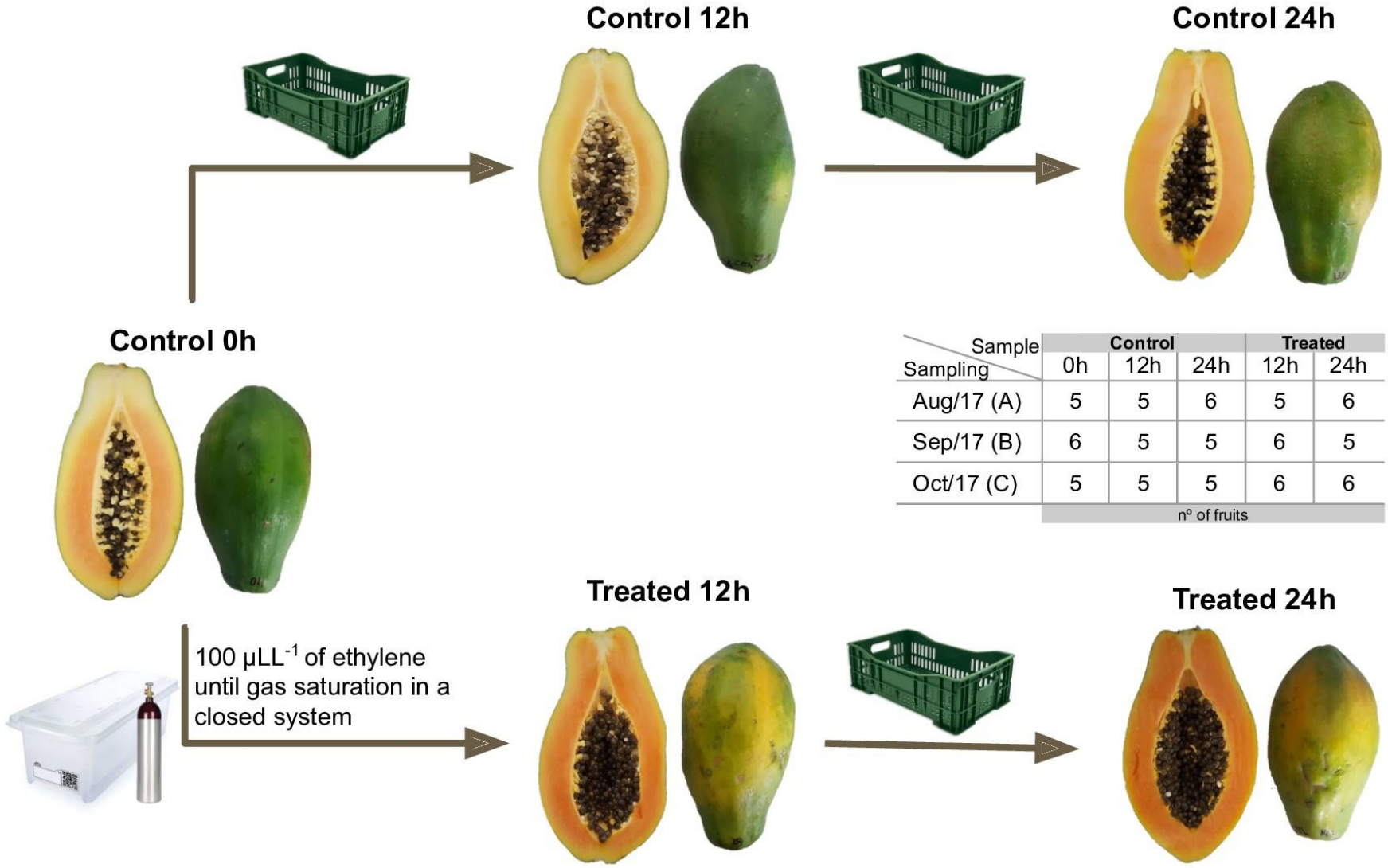
ZHONG, R., LEE, C., ZHOU, J., *et al.* A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in *Arabidopsis*, **The Plant Cell**, v. 20, n. 10, p. 2763–2782, 1 out. 2008. DOI: 10.1105/tpc.108.061325.

Supplementary Material

Table S1. Enriched and repressed function pathways of the genes in the network correlation.

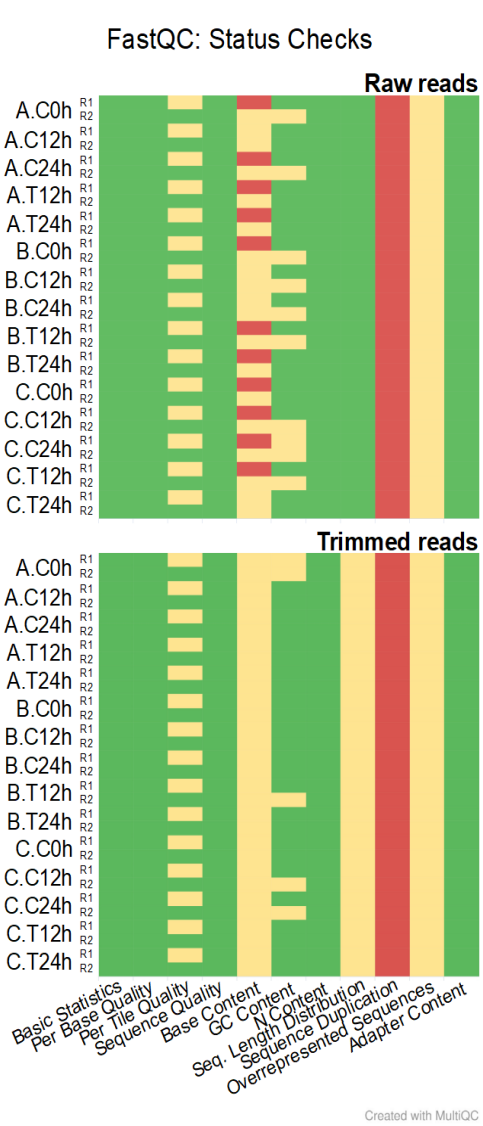
GO ID	GO Name
Black	
GO:0032993	protein-DNA complex
GO:0035035	histone acetyltransferase binding
GO:0044815	DNA packaging complex
GO:0044764	multi-organism cellular process
GO:0005771	multivesicular body
GO:0009294	DNA mediated transformation
GO:0009292	genetic transfer
GO:0010353	response to trehalose
GO:0046068	cGMP metabolic process
GO:0000786	nucleosome
GO:0009190	cyclic nucleotide biosynthetic process
GO:0009187	cyclic nucleotide metabolic process
GO:0052652	cyclic purine nucleotide metabolic process
GO:0015629	actin cytoskeleton
GO:0008154	actin polymerization or depolymerization
GO:0030906	retromer, cargo-selective complex
GO:0051117	ATPase binding
GO:0006182	cGMP biosynthetic process
Midnightblue	
GO:0006775	fat-soluble vitamin metabolic process
GO:0042360	vitamin E metabolic process
GO:0042362	fat-soluble vitamin biosynthetic process
GO:0006553	lysine metabolic process
GO:0009085	lysine biosynthetic process
GO:0009089	lysine biosynthetic process via diaminopimelate
GO:0010189	vitamin E biosynthetic process


GO:0046451 diaminopimelate metabolic process
GO:0009654 photosystem II oxygen evolving complex
GO:0010285 L,L-diaminopimelate aminotransferase activity
GO:0005507 copper ion binding
GO:0050342 tocopherol O-methyltransferase activity
Enriched
Repressed



Sample / Sampling	Control			Treated	
	0h	12h	24h	12h	24h
Aug/17 (A)	5	5	6	5	6
Sep/17 (B)	6	5	5	6	5
Oct/17 (C)	5	5	5	6	6
	nº of fruits				

Figure S1. Sampling scheme. Scheme for dividing groups and sampling points. Each group consists of at least five fruits. This process was repeated three times, resulting in a biological triplicate.



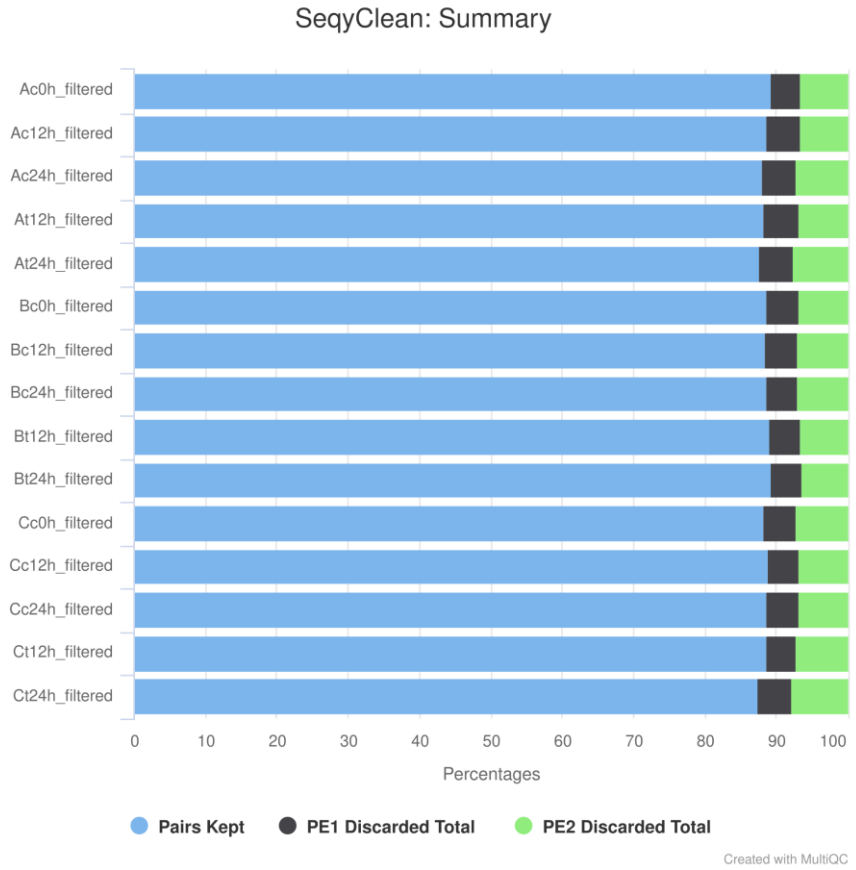


Figure S2. Pre-processing quality assay. (A) FastQC assessment of raw reads. (Green: entirely normal, orange: slightly abnormal, red: very unusual). (B) FastQC assessment after trimming. Base and GC contents were balanced; sequence length distribution raised due to differences in discarded pairs on trimming, overrepresented sequences kept red probably due to highly biologically significance. (C) Trimming proportions. A mean of 86 % of the reads was kept.

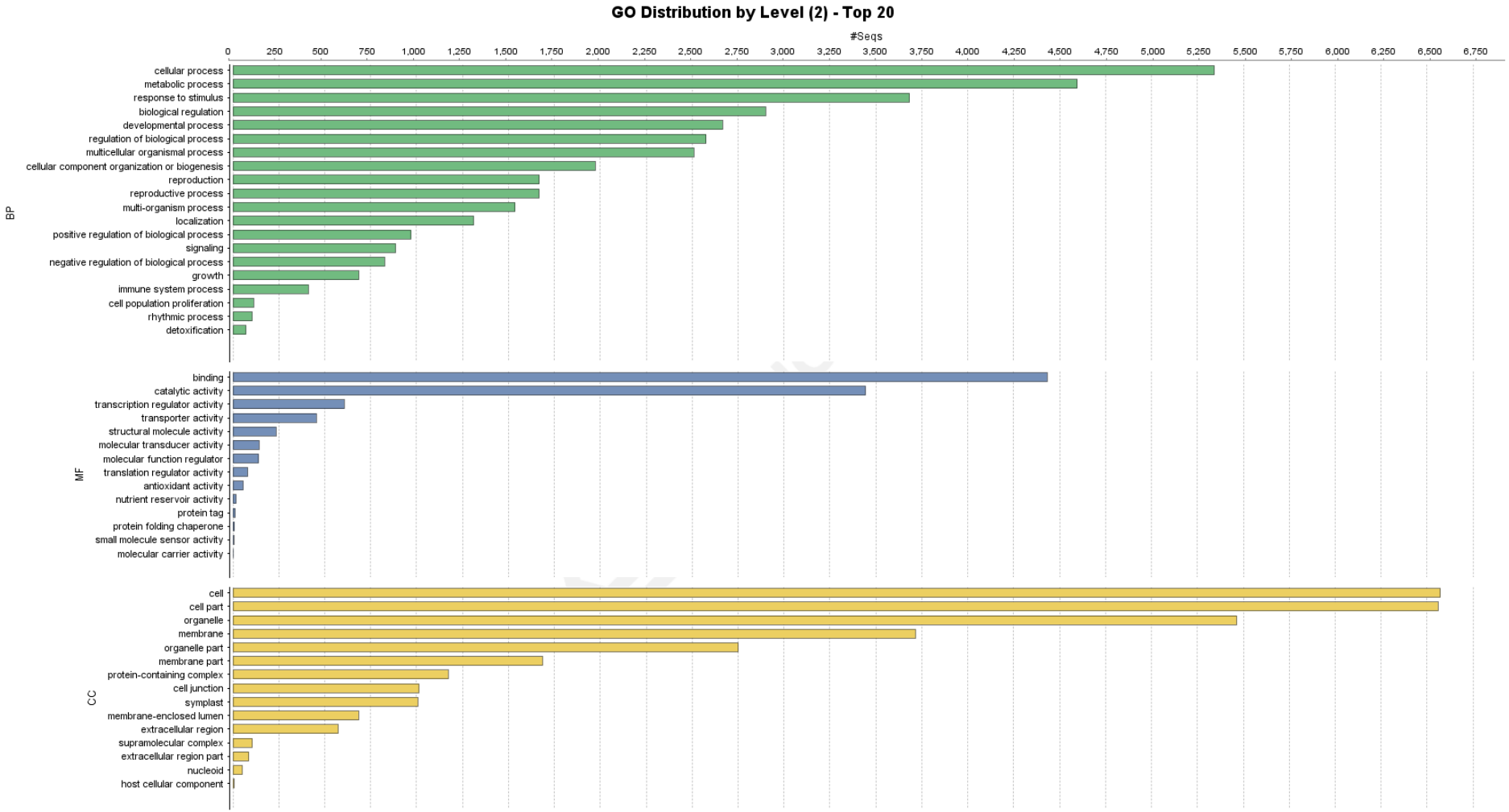


Figure S3. GO terms distribution per category. For each category, the 20 GO terms with the highest number of genes are represented. BP: Biological Process, MF: Molecular Function, CC: Cellular Component.

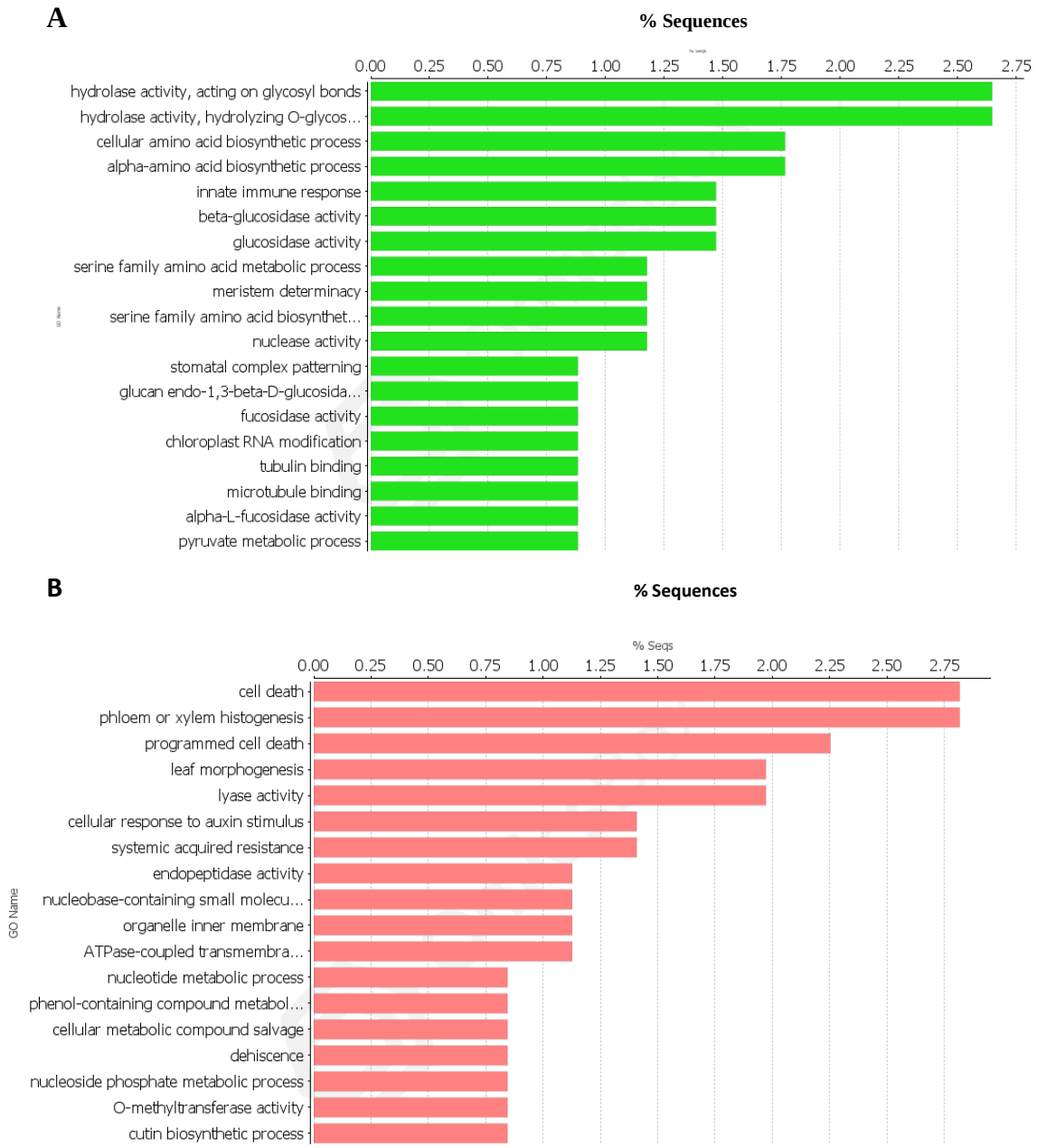


Figure S4. GSEA between Control 0h and Treated 24h. GO Terms assigned to down-regulated (A) and up-regulated (B) DEGs among the extreme time points of the experimental design, ordered by Enrichment Score.

Chapter II

Coordinated expression of ethylene-triggered pectinases from papaya pulp linked to fast pulp softening

Caroline Giacomelli Soares[1,2], Samira Bernardino Ramos do Prado[1,2], Sónia Cristina da Silva Andrade[3], João Paulo Fabi[1,2,4]

[1] Department of Food Science and Experimental Nutrition, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.

[2] Food Research Center (FoRC), CEPID-FAPESP (Research, Innovation and Dissemination Centers, São Paulo Research Foundation), São Paulo, Brazil.

[3] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade São Paulo, Brazil.

[4] Food and Nutrition Research Center (NAPAN), University of São Paulo, São Paulo, Brazil.

Abstract

Pulp softening is the main factor that could limit fruit sensorial quality and post-harvest life. Sometimes it occurs slowly, and sometimes it occurs very quickly depending on fruit climacteric pattern, precluding any effort for postponing shelf-life storage. Papaya cell wall disassembly could be considered a model of fast pulp softening since ethylene triggering increases the expression of pectinases in high amounts that degrade pectin very fast. The most studied pectinases in fruit are polygalacturonases, pectate lyases, and pectinesterases. Their coordinated expression is very important for controlling pulp softening and it occurs in different ways throughout fruit. In this way, this work describes the massive sequencing of papaya transcripts after ethylene treatment and the correlation of pectinases expressions to fast pulp softening compared to strawberries and tomatoes, which need a previous action of pectate lyases and pectinesterases in order to polygalacturonases hydrolyze pectin. Papaya fruit does not need this previous action of pectate lyases, being polygalacturonases the main enzymes responsible for fruit cell wall disassembling. Results presented herein connect the concept of the main set of cell wall-degrading genes that are differentially expressed during papaya ripening and the implication of these coordinated actions in the sensorial quality of fleshy fruit.

Keywords: papaya pulp softening; ethylene; pectinases; cell wall; climacteric fruit; non-climacteric fruit.

Abbreviations: PCW, plant cell wall; GalA, galacturonic acid; Rha, rhamnose; RG-I, Rhamnogalactunoronan type I; PG, polygalactunorase; PL, pectate lyase; PME, pectinesterase; GO, Gene Ontology; AGAL, α-galactosidase; BGAL, β-galactosidase; XTH, xyloglucan endo-transglycosylase/hydrolase; 1-MCP, 1-methylcyclopropene; HG, homogalacturonan.

Introduction

It is very important to elucidate the structure, functions, and gene expression regulation of the plant cell wall during fruit ripening as commercial interests in postponing the shelf-life storage are crucial to maintaining the sensorial fruit quality (FABI; DO PRADO, 2019). The final structure of the fruit cell wall will determine the fruit quality since a fast pulp softening could lead to fruit injuries and postharvest losses (HUSSEIN; FAWOLE; OPARA, 2020). The primary flesh fruit cell wall is mainly formed by crosslinked polysaccharides such as hemicelluloses and pectins, which confer rigidity to the fruit cell walls (TUCKER et al., 2017). Pectic polysaccharides could be found as water-soluble and chelate-soluble fractions in primary cell walls and middle lamella (VOINICIUC; PAULY; USADEL, 2018). Pectin regulates essentials properties in wall functioning such as porosity, hydration, and wall swelling (MAJDA; ROBERT, 2018). Pectin is comprised of several structures, but the main ones in fleshy fruit are homogalacturonans and rhamnogalacturonans type I, but not limited to. Homogalacturonan (HG) accounts for the biggest portion of pectin, composed of a linear polymer of galacturonic acid (GalA) that could be esterified with methyl groups (WANG, D. et al., 2018). Rhamnogalactunoronan type I (RG-I) is composed of intercalated GalA (1,4-α-D-GalpA) and rhamnose (Rha) (1,2-α-L-Rhap) with branched sugars in the O-4 position of the Rha, formed by arabinose (arabinan), galactose (galactan), or both (arabinogalactan) (TUCKER et al., 2017). The pectic network is a continuous backbone of HG with intercalated RG-I domains and this is determinant in maintaining the fleshy fruit structure (POSÉ et al. 2019).

Papaya fruit could be considered a model of fast pulp softening (DO PRADO et al., 2016). Papaya pulp pectin is hydrolyzed by pectinases expressed during ripening and the fruit softening occurs because of fruit cell wall disassembling (FABI et al., 2014). The most intense structural modifications of papaya cell walls occur on pectins induced by pectinases enzymes such as polygalactunorases (PG) (FABI et al., 2014; TUCKER et al., 2017). In other fleshy fruit such as tomatoes (climacteric) and strawberries (non-climacteric), other pectinases act together with PG during ripening such as pectate lyases (PL) and pectinesterases

(PME). Since only a few ethylene-dependent pectinases transcripts were studied in papaya pulp metabolism and studies of those transcripts are still elusive, this work describes the massive sequencing of papaya transcripts after ethylene treatment and the correlation of papaya pectinases expressions to pulp softening. The results were compared to pectinases transcripts of strawberries and tomatoes and the study of the pectinases actions during ripening showed significant differences between those fruits. The results presented herein associate the main set of cell wall-degrading genes that are differentially expressed during papaya ripening and the implication of this model of coordinated pectinases actions in sensorial quality of fleshy fruit that exhibit fast pulp softening.

Material and methods

Sampling, phenotipic analysis, RNA sequencing, bioinformatics analyses, differential expression analysis, and statistics are the ones described on **Chapter I**.

Identifying papaya cell wall-related genes

As Kubicek, Starr e Glass (2014) have listed, some of the cell wall-degrading enzymes families are polygalacturonase, pectate lyase, rhamnogalacturonate lyase, pectinesterase, cellulase, glycosidase, xylanases, xyloglucanases, α-Galactosidases, β-Mannanase, α-Arabinosidases, β-Galactosidases, β-Glucuronidases and, xyloglucan endotransglucosylase/hydrolase. Based on Blast and Gene Ontology analysis, genes related to these enzymes were selected from previous differential analysis through programming tools using the cutoff of P-value ≤ 0.05 to be considered differentially expressed.

Enrichment analysis

Blast2GO tool (Biobam Bioinformatics S.L., Valencia, Spain) was used to perform GO enrichment analysis with the test set being the selected PCW genes,

for the categories of biological process, molecular function, and cellular component, with a cutoff of FDR ≤ 0.005.

Co-expression analyses of papaya cell wall-related genes

Co-expression network analysis was performed using the WGCNA (Weighted Gene Correlation Network Analysis) R package (LANGFELDER; HORVATH, 2008), using the pulp firmness measurements as the experimental trait data to be correlated to modules. From the resulting correlation, a text file from the the package was provided, and it was selected only interactions between the previously selected cell wall-related genes using programming tools. The visualization of the correlation network was built using Cytoscape software (SHANNON *et al.*, 2003). The network visualization on Cytoscape was designed by an algorithm called “Edge-weighted Spring-Embedded Layout”, which treats the weight of correlations as an inverse physical force, resulting in smaller edge lengths to higher correlations between nodes (SHANNON *et al.*, 2003). The pipeline of the RNAseq methodology used in this work is summarized in Figure 1.

Differential expression analysis through Real-Time PCR

The experiments were done rigorously according to what was previously described and by using the ripening samples accordingly to do Prado *et al.* (2016). The genes used were as followed, with primers sequences, GenBank access and curve efficiency listed in Tables S1, S2, S3: Gene9059 - PG1; gene9058 - PG2; gene5336 - PG3; gene9513 - PL1; gene171 - PL2; gene14164 - AGAL1; gene1360 - AGAL3; gene7019 - BGAL1; gene2838 - BGAL3; gene13517 - PME1; gene3087 - PME2; gene15057 – PME3.

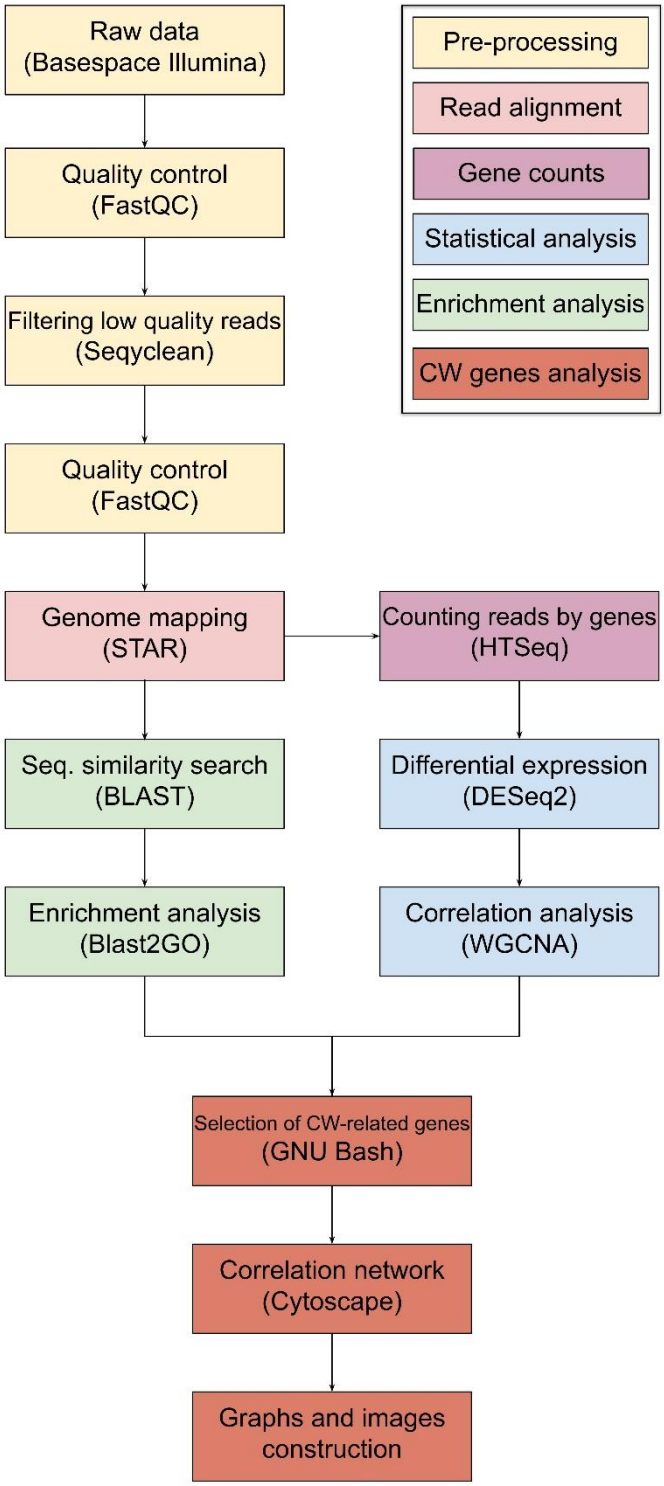


Figure 1. A pipeline of papaya pulp RNA-seq processing and treatment followed by the filtering of PCW-related genes.

Results

Transcriptomic analysis reveals a diverse expression of papaya cell wall-related genes

Papaya pulp transcriptome showed 157 expressed genes related to cell wall metabolism according to the blasted transcripts (Table S4), including some up- and down-regulated ones, as well as some genes known to be ethylene-induced (Figure 2).

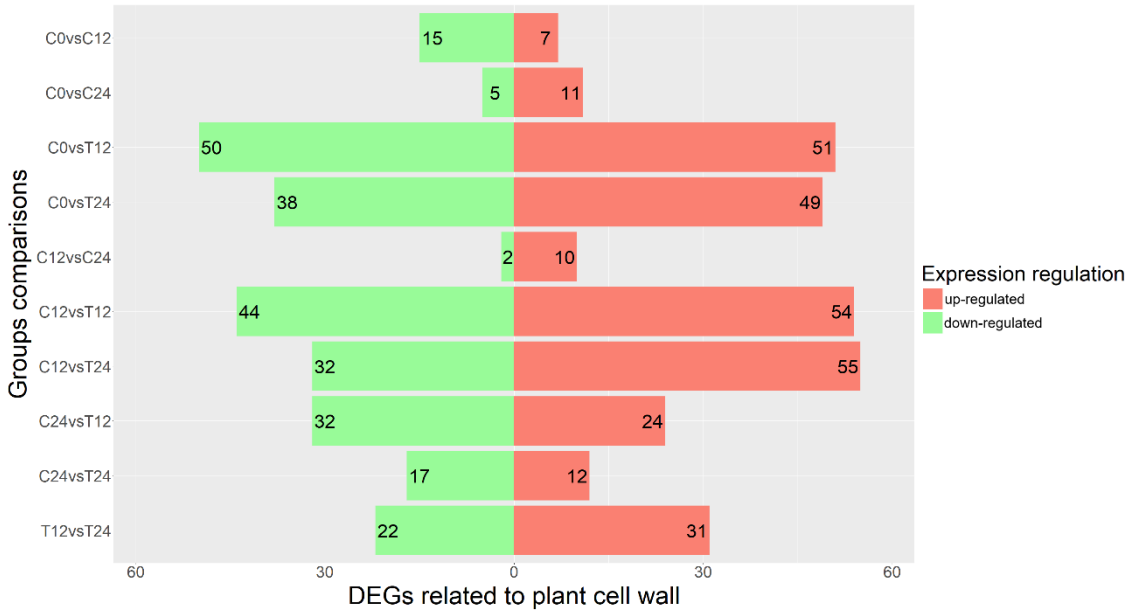


Figure 2. Differentially expressed genes related to fruit cell wall metabolism in papayas. Green bars represent down-regulated differentially expressed genes while red bars represent up-regulated ones. Numbers inside bars are the total of genes represented. Groups are classified by control (C) or treated (T) followed by the respective point (0h, 12h, or 24h).

In general, genes related to cell wall disassembly did not change their expression at the first 12hs and 24hs of the ripening process in control fruit (Figure S1). This corroborates that changes in ethylene production and pulp texture were not observed (**Chapter I**, Figure 2), with no apparent fruit cell wall alteration in the first and second days after harvesting (DO PRADO *et al.*, 2016). In ethylene-treated fruit (both 12hs and 24hs groups), a huge number of cell wall-related genes were observed as being differentially expressed (Figure 2). A considerable set of genes from PG, rhamnogalacturonate lyases, galactanases, and XTH were up-regulated, while most of PL and PME genes did not change their expression or were down-regulated, bias also observed on GO analysis (Figure 3).

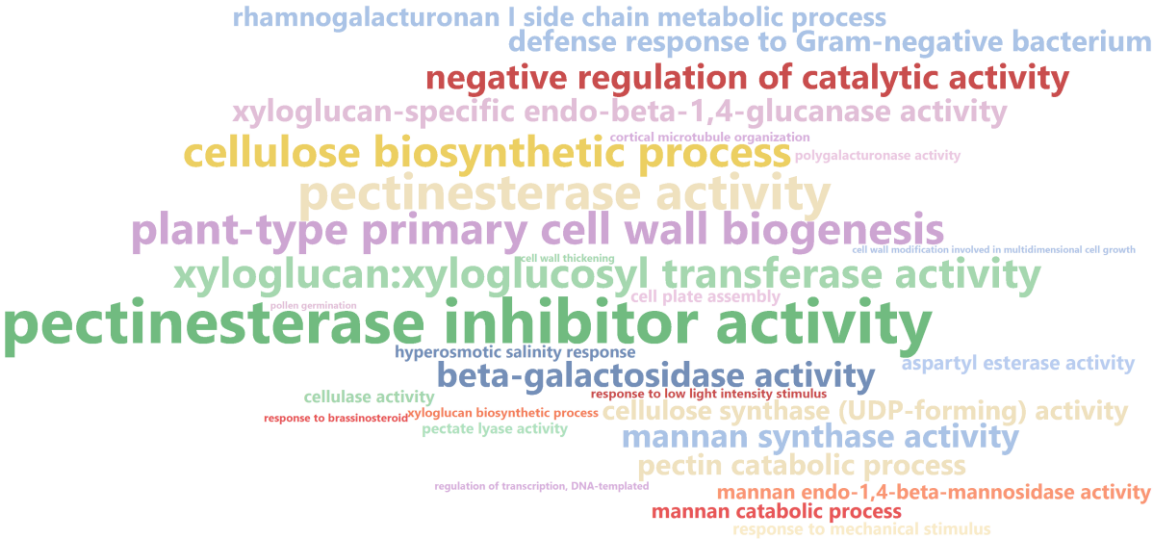


Figure 3. Enriched pathways of papaya cell wall genes. Word cloud of the thirty most significant GO terms of the enriched and repressed pathways assigned with p-value ≤ 0.005 to DEG related to the cell wall of the fruit, with the word size inversely proportional to the p-value.

Genes related to the cell wall were assigned to 73 GO terms referring to the functional pathways in which the genes participate, being 30 repressed and 43 enriched (Table S5). Repressed pathways are related to the regulation of stimulus (GO:0048583, GO:0009755, GO:0009737) and cellular, nuclear, DNA, and gene expression processes (GO:0048523, GO:0000166, GO:0043565, GO:0006355). Enriched biological processes involve the development and maintenance of the general biological system (GO:0009833, GO:0050829, GO:0000919, GO:0043622), biosynthesis of pectin, cellulose, and xyloglucan (GO:0048358, GO:0030244, GO:0009969), response to fungi (GO:0009620), pectin and mannan catabolic process (GO:0045490, GO:0046355) and several cell wall modifications (GO:0052386, GO:0042547, GO:0009830), among others. Regarding the molecular function category, enriched pathways are related to the activity of several enzymes, such as PME (GO:0046910, GO:0030599), PG (GO:0004650), PL (GO:0030570), xyloglucan transferase (GO:0016762, GO:0033946), beta-galactosidase (GO:0004565), and mannanase (GO:0051753, GO:0016985).

In general, genes related to cell wall degradation did not alter their expression in control samples, such as XTH and PG genes (FABI *et al.*, 2012). However, ethylene-treated groups showed several upregulated cell wall-related genes (XTH and PG) and some genes such as PME and cellulose synthase were down-regulated, though. Other relevant results were the up-regulation of PME inhibitor genes while PG inhibitor was down-regulated (Table S2). An interesting result

was the up-regulation of rhamnogalacturonate lyase genes, enzymes responsible for hydrolyzing RG-I portions of pectin through the β-elimination mechanism. Despite it was the first time these genes were identified in papayas, they have already been identified in tomatoes and potatoes (OCHOA-JIMÉNEZ *et al.*, 2018). In order to help the visualization of the differentially expressed genes in the RNAseq experiments, a heat-map figure was done (Figure S1).

Network files generated by WGCNA were visualized through Cytoscape towards the application of the “Edge-weighted Spring-Embedded Layout” algorithm (Figure 4). To better understanding, some information was turned into graphical parameters. Node sizes are set by their connectivity degree representing the total number of correlations they participate in, whereas their color indicates the enzyme to which the node (gene) is related. Edge lengths are inversely proportional to weight values of correlations whilst their width is directly proportional, meaning the smaller and the thicker an edge is, the higher is the correlation between the nodes.

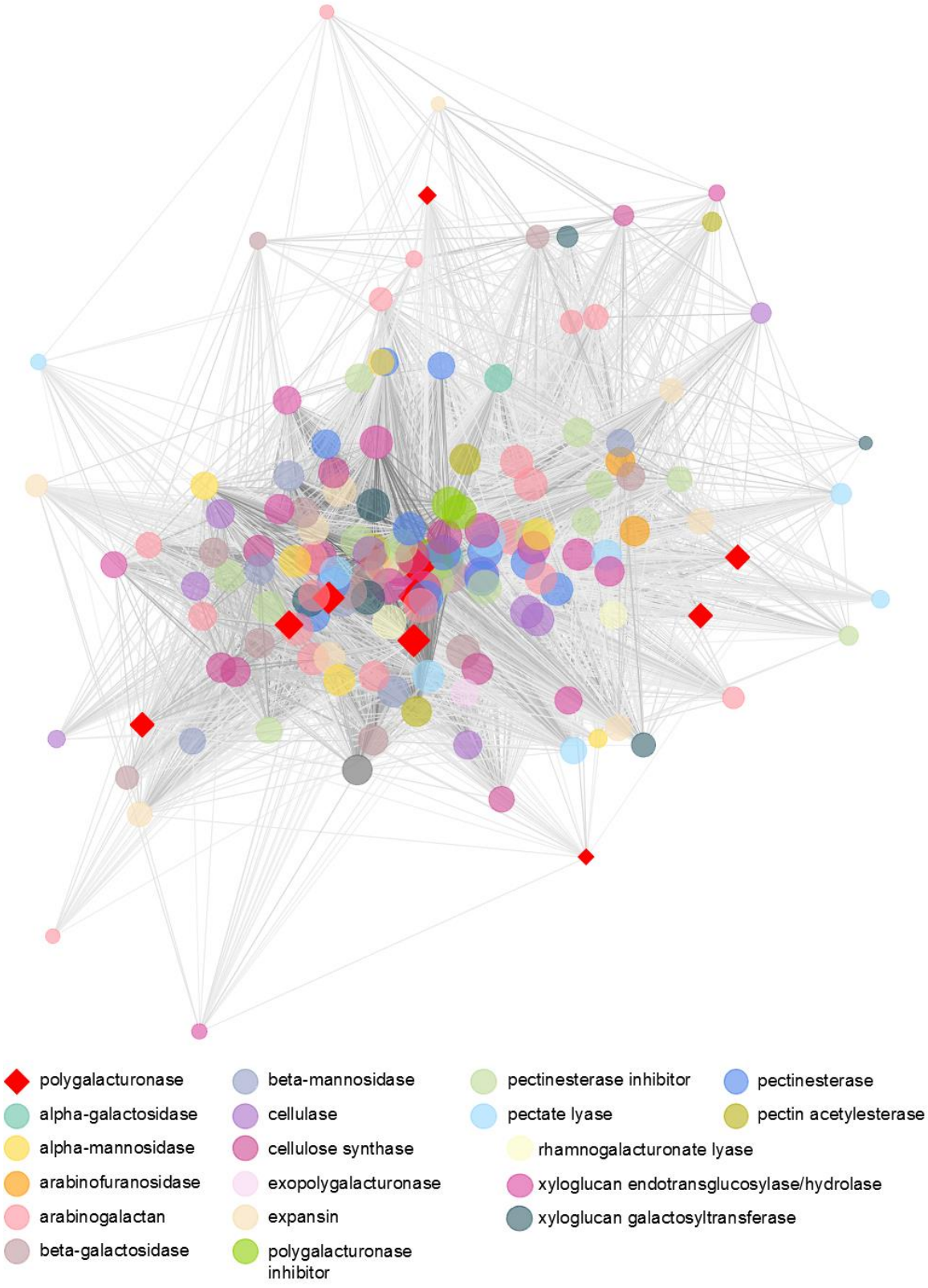


Figure 4. Co-expression correlation network of cell wall genes. Nodes size represents their connectivity degree. Edges vary in length and width meaning the smaller and the thicker an edge is, the higher is the correlation between the nodes.

It was detected 8,294 interactions with a weight ≥ 0.05 between genes related to cell wall-modifying enzymes (Figure 4). The PGs, highlighted in red in the shape of a diamond, interact with almost all genes of the network except for some arabinogalactan-protein, expansins, and PLs, among others. It is possible to observe the centrality of the deeply connected PG right in the center of the network while other PGs are distributed all over around, indicating a heavy influence on the behavior of other members. The only exo-PG exhibits an interesting template: it is intensely correlated to the nodes in the center of the network but not to the peripheral ones. PMEs are highly connected to the network, showing no correlation with only four genes of the 157 genes related to cell-wall enzymes. However, PME inhibitors also followed the same pattern. PLs, although present in a much smaller number, also interact with almost all members of the network. Galactosidases, xylanases, and cellulases are fully interconnected with all members of the network. The rhamnogalacturonate lyases showed intense correlation mainly with genes encoding xylanases, PG, PME, PG and PME inhibitors, expansins, cellulose synthases β-galactosidases, and arabinogalactans.

Some of these genes had their expression confirmed by qPCR (Figure 5) using the control versus 12hs ethylene treatment, while also using a regular ripening curve, as previously published (DO PRADO *et al.*, 2016). These results corroborated what was observed in the RNA-seq and in the correlation analyses, in which papayas are indeed climacteric fruit with fast softening that is achieved exclusively by the massive action of the ethylene-induced PG (DO PRADO *et al.*, 2016, 2017; FABI *et al.*, 2014).

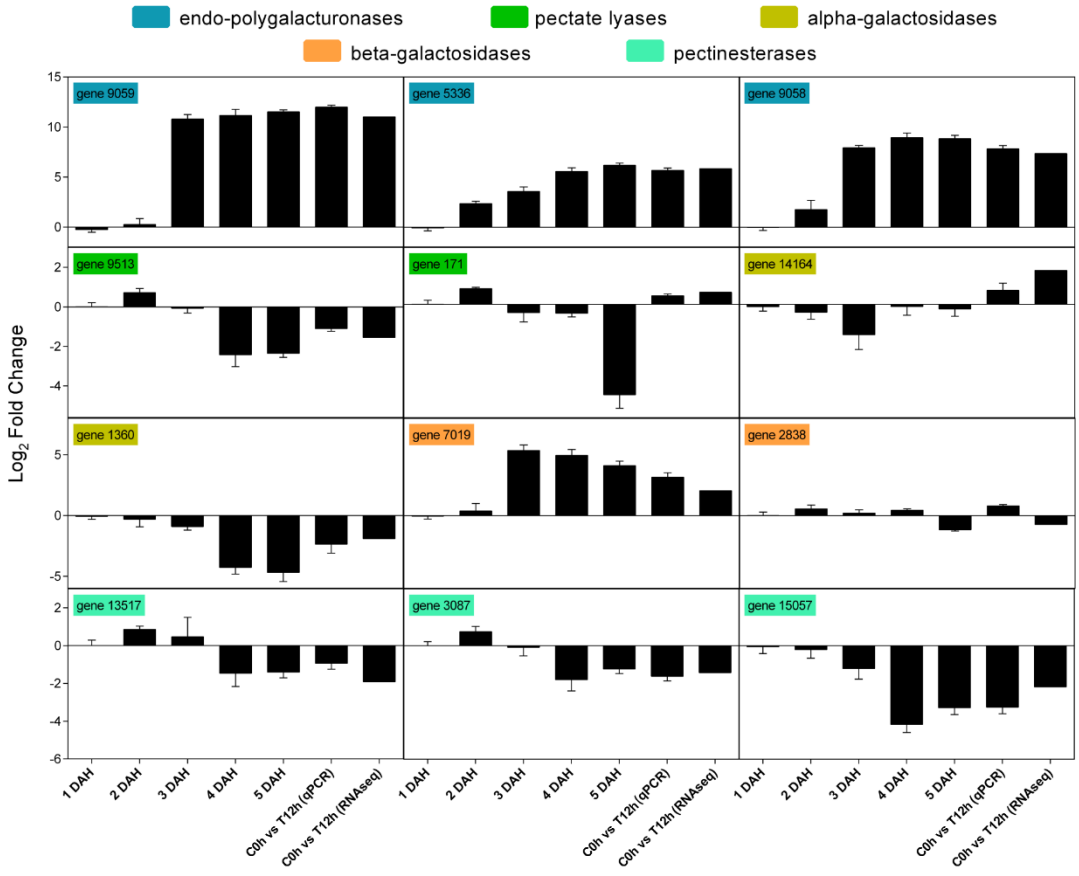


Figure 5. Confirmative gene expression experiment of cell-wall related genes that were differentially expressed after ethylene-treatment of papaya fruit. The expression of differentially expressed genes between the control group (0h) vs. ethylene-treated fruit (12h) was confirmed in the same sample (C0h X T12h) and in a ripening curve that were previously published with quantitative Real Time-PCR being done according to DO PRADO *et al.* (2016). Gene ID can be verified in Table S4.

Discussion

Papaya pulp softening: different mechanisms compared to models from climacteric fruit (tomato) and non-climacteric fruit (strawberry)

The plant hormone ethylene is the main factor responsible for the fruit ripening phenomenon (TUCKER *et al.*, 2017). Transcriptomic data reveals that climacteric fruit distributes higher amounts of ethylene receptors genes; on the other hand, non-climacteric fruit expresses these genes right on the onset of ripening. Therefore, it is reasonable to establish that this classification must be made upon fruit's sensitivity to ethylene (CHEN *et al.*, 2018). Ethylene is a major regulatory factor in climacteric fruit, not only in the ripening process but also in

growth and development; it triggers a cascade of metabolic events that eventually result in cell wall disassembly (FABI; DO PRADO, 2019).

Many cell wall modifying enzymes in fleshy fruit are ethylene-responsive (FABI *et al.*, 2012). The most explored pectin-modifying enzymes are PG, PL, and PME, although their individual actions are not enough to deeply modify cell wall structure (DO PRADO *et al.*, 2016). Earlier studies using tomatoes have shown that PG activities are three times increased in ripe fruit compared to unripe fruit, following a similar pattern of ethylene production (KE *et al.*, 2018). This behavior is also noticed when an ethylene biosynthesis gene is suppressed, indicating that PG expression happens even at a very low ethylene level, besides being induced by other signaling pathways (TUCKER *et al.*, 2017). Another research study with mangoes revealed that PL is active during ripening as well as PG, and remained with significative expression after the climacteric peak, but their expression was completely abolished after 1-MCP treatment (CHOURASIA; SANE; NATH *et al.,* 2006; DAUTT-CASTRO *et al.,* 2015). Controversially, strawberry cell wall-related genes showed a considerable increase in expression after 1-MCP treatment (PME, β-xylosidase, endoglucanase, xyloglucan endotransglycosidase/hydrolase, arabinofuranosidase, and cellulase genes), while they were down-regulated after ethylene treatment (VILLARREAL *et al.,* 2016). In fact, non-climacteric fruit show a relevant activity of PL together with PG, although the optimum pH of PL is significantly higher than that of PG. These enzymes have different mechanisms of action using the same substrate (de-esterified HG), suggesting that their activities are synergetic (POSÉ *et al.,* 2015). Studies conducted in which scientists have silenced PG and PL expression in strawberries indicate that PL has a more specific activity on covalent bonds between pectin and cell walls, while PG affects polyuronides both from the primary cell wall and middle lamella (POSÉ *et al.,* 2015).

Cell wall loosening in fleshy fruit and the consequent pulp softening are mainly achieved by the migration of pectin from insoluble to more soluble water fractions (DO PRADO *et al.,* 2016). However, the molecular and biochemical pathways to eventual cell wall disassembly are greatly different regarding enzyme expression and timing of the complete ripening process among distinct species (LÜ *et al.,* 2018). Slow pulp softening of non-climacteric fruit such as strawberries

is marked by a previous activity of PME and PL with a concomitant action of PG in ethylene-independent ways (PANIAGUA *et al.*, 2017; POSÉ *et al.*, 2015). Slow pulp softening of climacteric fruit such as some melons and peach varieties demonstrate depolymerization in the later stage of pulp softening after PG-independent solubilization of pectins, probably through the cleavage of galactose residues from galactan side chains by galactosidases (GOULAO; OLIVEIRA, 2008). Regarding peaches, if they are hard-pulp varieties, there is an only exo-PG activity with little depolymerization and solubilization of pectin (BRUMMELL *et al.*, 2004). Otherwise, if they are melting-pulp varieties, fast softening triggered by ethylene occurs through both exo- and endo-PG activities leading to time-dependent massive pectin depolymerization and solubilization, also with increased PME and galactosidases activities (BRUMMELL *et al.*, 2004).

Climacteric fruits that show mid-fast pulp softening, such as tomatoes, showed previous activities of PME and PL – just like strawberries – with pectin depolymerization at a mid-softening stage in a PG-dependent manner with galactanases activities appearing only after the onset of pulp softening (GOULAO; OLIVEIRA, 2008; WANG, D. *et al.*, 2018). Climacteric fruit with fast pulp softening, such as papayas, present a basal expression of PME and decreased expression of PL but with a huge increment in endo-PG activity, leading to massive depolymerization and solubilization of pectins (DO PRADO *et al.*, 2016, 2017; FABI *et al.*, 2014). The differences in cell wall–related genes among strawberries, tomatoes, and papayas regarding gene expression assays are depicted in Figure 6.

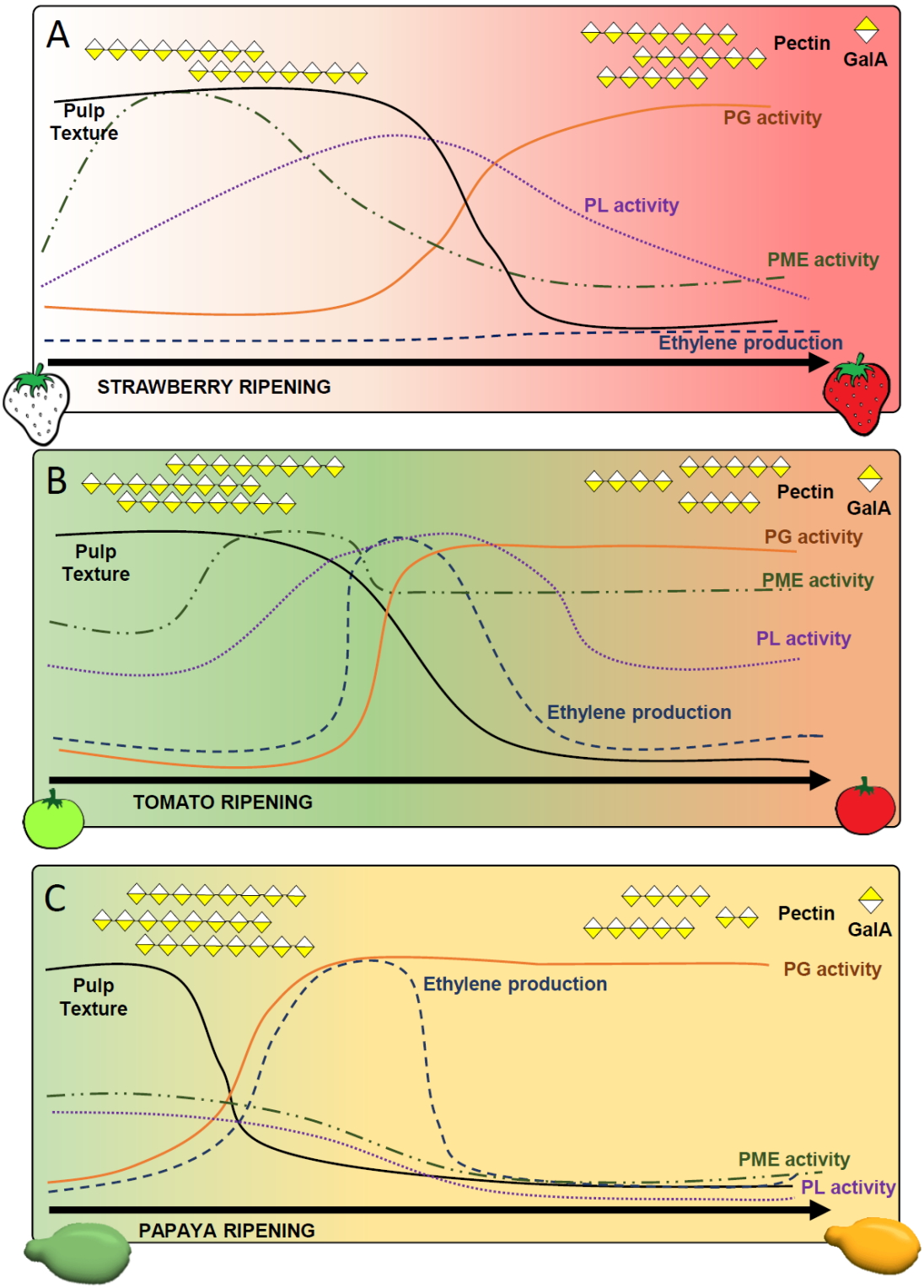


Figure 6. General representation of strawberry, tomato, and papaya fruit ripening against pectinases expression. (A) Non-climacteric strawberry ripening is marked by low levels of ethylene production and a low pulp softening. It follows an increment in PG expression and a decrease in PL and PME expressions that prepare the pectin to be accessed by PG. (B) Tomato ripening is marked by a climacteric ethylene peak production accompanied by a mid-fast pulp softening and the increased PG expression while PL and PME expressions decrease. PL and PME pre-climacteric expressions make the pectin more vulnerable to PG action. (C) Papaya ripening is very fast and is marked by a fast pulp softening due to a massive pectin mobilization driven by PG higher expression. The low PME and PL expressions make this process distinct from the A and B processes. GalA: galacturonic acid.

Polygalacturonases are intensely ripening-induced enzymes that hydrolyze the α-1,4-galacturonosyl linkages in the de-esterified HG backbone (endo-activity), while the exo-PG remove residues of galacturonosyl (FABI *et al.*, 2009; GOULAO; OLIVEIRA, 2008; KOZIOŁ *et al.*, 2017). Endo-PG has a key role in fruit softening since they act inside the pectin molecules, resulting in pectic depolymerization (lowering the average molecular weight), leading to water solubilization of pectic fractions. Pectate lyases (PL) activity is similar to PG in also depolymerizing HG but with a lower enzymatic rate. Besides, PL acts through a different mechanism by using β-elimination with the presence of calcium ions as a requisite (KOZIOŁ *et al.*, 2017; TUCKER *et al.*, 2017). In another way, pectinesterases (PME) catalyze the GalA de-esterification of the methyl groups from methylated GalA. This would enable a calcium ion to bind to the remaining negatively charged carboxylic acid, forming the egg-box as the outcome. However, PME expression is basal in papayas after harvesting as well as during cell-wall loosening (DO PRADO *et al.*, 2017; KIRTIL *et al.*, 2014; MAJDA; ROBERT, 2018). In summary, the PME action prepares the pectin to be hydrolyzed by the PG and PL (HOCQ; PELLOUX; LEFEBVRE, 2017). Although some other enzymes act in fleshy fruit cell wall architecture, such as galactanases, arabinofuranosidases, xylanases, xyloglucan endotransglicosylases/hydrolases (XTH), cellulases, expansins, and rhamnogalacturonases, the expression of PG, PL, and PME genes is the main factor that leads to pulp softening. While three different fruit (strawberries, tomatoes, and papayas) exhibit different rates of pulp softening and plant cell wall gene expressions, some studies were analyzed herein to corroborate those differentially expressed profiles throughout ripening, specifically focusing on the expression of PG, PL, and PME.

Strawberries are a non-climacteric fleshy fruit that undergoes pulp softening after postharvest ripening. Although the pulp softening is still elusive, it might rely on solubilization of primary cell wall pectins due to RG-I neutral sugars side-chains cleavage and depolymerization, and subsequently solubilization of pectins through the action of PG and PL (PANIAGUA *et al.*, 2017). Pectin degradation is proven to be the major reason for strawberries softening since PG and/or PL inhibitions lead to greater pulp firmness at a ripe stage (POSÉ *et al.*, 2015). Cell

wall disassembling starts with the PME activity with a concomitant PL action, which reduces the HG sizes. Afterward, the pectin is depolymerized by PG (GOULAO; OLIVEIRA, 2008; POSÉ *et al.*, 2015). Paniagua *et al.* (2017) relate a considerable decrease in arabinose concentration as well as in pectin soluble fractions during ripening, while hemicellulosic fractions showed just a little decrease. Cellulase activities were detected through XTH activity, while ethylene-induction decreased β-xylosidase and endoglucanase activities, contributing to hemicellulose network degradation (VILLARREAL *et al.*, 2016). Expansins are thought to play an important role in strawberry pulp softening as well (GOULAO; OLIVEIRA, 2008). All these presented data were confirmed through the analyses of the RNA-seq experiments during strawberry ripening (access number: SRX1294640— WANG, Q. *et al.*, 2017; ZHANG *et al.*, 2015). Seeking PG, PL, and PME transcripts, one could indicate that the strawberry is a slow pulp-softening non-climacteric fruit with a previous activity of PME and PL to “prepare” the pectin molecules to be degraded by PG (Figure 6A).

Tomatoes are an extensively explored fruit regarding ripening-related biochemical pathways and gene expression due to their importance to agriculture as well as their mid-fast pulp softening, which makes them a great fleshy fruit model (WANG, D. *et al.*, 2018). The onset of tomato softening is marked by a significant decrease in xyloglucan fraction due to depolymerization of hemicelluloses by XTH action and a high rate of loss of neutral sugars by depolymerization and solubilization of polyuronides through the concomitant action of PG and PME at the mid-softening stage (SEYMOUR *et al.*, 2013). There are detectable increasing activities of both endo- and exo-PG during ripening but little or none of galactosidases or xylosidases (GOULAO; OLIVEIRA, 2008). Expansins demonstrate a strong correlation with ripening as their inhibition results in increased fruit firmness, in contrast to their over-expression, which seems to enhance softening because of hydrogen-bonding cleavage of xyloglucan fractions (WANG, D. *et al.*, 2018). However, the suppression of several cell wall-related genes had no significant effects on tomato softening (WANG, D. *et al.*, 2018). One exception would be the silencing of a PL gene that had dramatically decreased cell wall disassembling during ripening, thus maintaining the pulp firmness (ULUISIK *et al.*, 2016). This would be due to a lack

of “preparation” of pectins to be degraded by PG since PL might have not cleaved the crosslinked HG before the climacteric stage, not enabling further PG action. Furthermore, Yang *et al.* (2017) conducted PL-silencing experiments in tomatoes, concluding that PL acts mainly on HG of the middle lamella and tricellular junction zones, in addition to saving a higher concentration of cellulose and hemicellulose content in cell walls. These data corroborated the outcome of firmer fruit with an extended shelf-life observed by Uluisik *et al.* (2016) with the desired attribute of lower susceptibility to pathogen attacks. All of these PME, PL, and PG combined actions in tomato pulp softening were confirmed through the analyses of some RNA-seq experiments (access number: SRP109982 — SHINOZAKI *et al.*, 2018; SRP076745 — ZHAO *et al.*, 2018). Seeking the expression patterns of PME, PL, and PG, it can be speculated that tomatoes are climacteric fruit with mid-fast pulp softening marked by a previous activity of PME and PL to “prepare” the pectin molecules to be degraded by PG, as occurs with strawberries (Figure 6B).

The mechanisms behind papaya pulp softening are different from other climacteric and non-climacteric fruit such as tomatoes and strawberries, and the possible silencing of papaya PL might not result in harder fruit, as occurred with these other two species. Further studies must be done to completely elucidate the molecular and biochemical profiles of ripening-induced modifications of climacteric and non-climacteric fruit for both economical and biotechnological purposes. It should be noted that this is a difficult task since the studies should merge complex networks involving thousands of genes, transcripts, proteins, enzymes, and other molecules that interact among themselves and with the environment — the so-called system biology approach — to eventually result in desired (or not) sensorial characteristics of the fruit. The tomato seems to be a good research model, but some other fruits are also interesting to explore due to individual specificities, as is the case of papaya pulp softening. Perhaps a promising pathway is to keep looking for full ethylene regulation to obtain the ability to maintain its desired effects and changing undesired ones, like delaying fruit ripening or enhancing both the pulp texture and the defense from pathogens, in order to increase the sensorial fruit quality.

Conclusion

It is clear that pulp softening derives from a chain of complex processes coordinated by many ethylene-dependent, partially dependent, and non-dependent enzymes. Some enzymes have minor but significant activity, such as galactosidases and XTH, whilst others have shown great changes in the expression of coding genes, like pectinases and cellulases. Expression profile suggests PG is induced by exogenous ethylene and reiterate it is the main enzyme responsible for changes in pectin structure during papaya ripening. In contrast to other fleshy fruit like tomato and strawberry, papaya does not seem to need a pre-climacteric action of PL and PME to begin the disassembling process of the pulp cell wall.

References

BRUMMELL, D. A., DAL CIN, V., CRISOSTO, C. H., *et al.* Cell wall metabolism during maturation, ripening and senescence of peach fruit, **Journal of Experimental Botany**, v. 55, n. 405, p. 2029–2039, 1 set. 2004. DOI: 10.1093/jxb/erh227.

CHEN, Y., GRIMPLET, J., DAVID, K., *et al.* Ethylene receptors and related proteins in climacteric and non-climacteric fruits, **Plant Science**, v. 276, p. 63–72, 2018. DOI: 10.1016/j.plantsci.2018.07.012.

CHOURASIA, A., SANE, V. A., NATH, P. Differential expression of pectate lyase during ethylene-induced postharvest softening of mango *(Mangifera indica* var. Dashehari), **Physiologia Plantarum**, v. 128, n. 3, p. 546–555, 1 nov. 2006. DOI: 10.1111/j.1399-3054.2006.00752.x.

DAUTT-CASTRO, M., OCHOA-LEYVA, A., CONTRERAS-VERGARA, C. A., *et al.* Mango (*Mangifera indica* L.) cv. Kent fruit mesocarp de novo transcriptome assembly identifies gene families important for ripening, **Frontiers in Plant Science**, v. 6, 2015. DOI: 10.3389/fpls.2015.00062

DO PRADO, S. B. R., MELFI, P. R., CASTRO-ALVES, V. C., *et al.* Physiological degradation of pectin in papaya cell walls: Release of long chains galacturonans derived from insoluble fractions during postharvest fruit ripening, **Frontiers in Plant Science**, v. 7, p. 1–11, 2016. DOI: 10.3389/fpls.2016.01120.

DO PRADO, S. B. R., FERREIRA, G. F., HARAZONO, Y., *et al.* Ripening-induced chemical modifications of papaya pectin inhibit cancer cell proliferation, **Scientific Reports**, v. 7, n. 1, p. 16564, 2017. DOI: 10.1038/s41598-017-16709-3.

FABI, J. P., BROETTO, S. G., DA SILVA, S. L. G. L., *et al.* Analysis of papaya cell wall-related genes during fruit ripening indicates a central role of polygalacturonases during pulp softening, **PLoS ONE**, v. 9, n. 8, 2014. DOI: 10.1371/journal.pone.0105685.

FABI, J. P., CORDENUNSI, B. R., SEYMOUR, G. B., *et al.* Molecular cloning and characterization of a ripening-induced polygalacturonase related to papaya fruit softening, **Plant Physiology and Biochemistry**, v. 47, n. 11, p. 1075–1081, 2009. DOI: 10.1016/j.plaphy.2009.08.002.

FABI, J. P., DO PRADO, S. B. R. Fast and furious: Ethylene-triggered changes in the metabolism of papaya fruit during ripening, **Frontiers in Plant Science**, v. 10, 2019. DOI: 10.3389/fpls.2019.00535.

FABI, J. P., SEYMOUR, G. B., GRAHAM, N. S., *et al.* Analysis of ripening-related gene expression in papaya using an Arabidopsis-based microarray, **BMC Plant Biology**, v. 12, n. 1, p. 242, 2012. DOI: 10.1186/1471-2229-12-242.

GOULAO, L. F., OLIVEIRA, C. M. Cell wall modifications during fruit ripening: when a fruit is not the fruit, **Trends in Food Science & Technology**, v. 19, n. 1, p. 4–25, 2008. DOI: 10.1016/j.tifs.2007.07.002.

HOCQ, L., PELLOUX, J., LEFEBVRE, V. Connecting Homogalacturonan-Type Pectin Remodeling to Acid Growth, **Trends in Plant Science**, v. 22, n. 1, p. 20–29, 2017. DOI: 10.1016/j.tplants.2016.10.009.

HUSSEIN, Z., FAWOLE, O. A., OPARA, U. L. Harvest and Postharvest Factors Affecting Bruise Damage of Fresh Fruits, **Horticultural Plant Journal**, v. 6, n. 1, p. 1–13, 2020. DOI: 10.1016/j.hpj.2019.07.006.

KE, X., WANG, H., LI, Y., *et al.* Genome-Wide Identification and Analysis of Polygalacturonase Genes in *Solanum lycopersicum*, **International journal of molecular sciences**, v. 19, n. 8, p. 2290, 4 ago. 2018. DOI: 10.3390/ijms19082290.

KIRTIL, E., OZTOP, M. H., SIRIJARIYAWAT, A., *et al.* Effect of pectin methyl esterase (PME) and CaCl2 infusion on the cell integrity of fresh-cut and frozen-thawed mangoes: An NMR relaxometry study, **Food Research International**, v. 66, p. 409–416, 2014. DOI: 10.1016/j.foodres.2014.10.006.

KOZIOŁ, A., CYBULSKA, J., PIECZYWEK, P. M., *et al.* Changes of pectin nanostructure and cell wall stiffness induced in vitro by pectinase, **Carbohydrate Polymers**, v. 161, p. 197–207, 2017. DOI: 10.1016/j.carbpol.2017.01.014.

KUBICEK, C. P., STARR, T. L., GLASS, N. L. Plant Cell Wall–Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi, **Annual Review of Phytopathology**, v. 52, n. 1, p. 427–451, 4 ago. 2014. DOI: 10.1146/annurev-phyto-102313-045831.

LANGFELDER, P., HORVATH, S. WGCNA: an R package for weighted correlation network analysis, **BMC Bioinformatics**, v. 9, n. 1, p. 559, 2008. DOI: 10.1186/1471-2105-9-559.

LÜ, P., YU, S., ZHU, N., *et al.* Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening, **Nature Plants**, v. 4, n. 10, p. 784–791, 2018. DOI: 10.1038/s41477-018-0249-z.

MAJDA, M., ROBERT, S. The Role of Auxin in Cell Wall Expansion, **International journal of molecular sciences**, v. 19, n. 4, p. 951, 22 mar. 2018. DOI: 10.3390/ijms19040951.

OCHOA-JIMÉNEZ, V., BERUMEN-VARELA, G., BURGARA-ESTRELLA, A., *et al.* Functional analysis of tomato rhamnogalacturonan lyase gene Solyc11g011300 during fruit development and ripening, **Journal of Plant Physiology**, v. 231, 1 set. 2018. DOI: 10.1016/j.jplph.2018.09.001.

PANIAGUA, C., SANTIAGO-DOMÉNECH, N., KIRBY, A. R., *et al.* Structural changes in cell wall pectins during strawberry fruit development, **Plant physiology and biochemistry**, v. 118, p. 55–63, 2017. DOI: 10.1016/j.plaphy.2017.06.001.

POSÉ, S., KIRBY, A. R., PANIAGUA, C., *et al.* The nanostructural characterization of strawberry pectins in pectate lyase or polygalacturonase silenced fruits elucidates their role in softening, **Carbohydrate Polymers**, v. 132, p. 134–145, 2015. DOI: 10.1016/j.carbpol.2015.06.018.

POSÉ, S., PANIAGUA, C., MATAS, A. J., *et al.* A nanostructural view of the cell wall disassembly process during fruit ripening and postharvest storage by atomic force microscopy, **Trends in Food Science and Technology**, v. 87, p. 47–58, 2019. DOI: 10.1016/j.tifs.2018.02.011.

SEYMOUR, G. B., CHAPMAN, N. H., CHEW, B. L., *et al.* Regulation of ripening and opportunities for control in tomato and other fruits, **Plant Biotechnology Journal**, v. 11, n. 3, p. 269–278, 1 abr. 2013. DOI: 10.1111/j.1467-7652.2012.00738.x.

SHANNON, P., MARKIEL, A., OZIER, O., *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks, **Genome research**, v. 13, n. 11, p. 2498–2504, nov. 2003. DOI: 10.1101/gr.1239303.

SHINOZAKI, Y., NICOLAS, P., FERNANDEZ-POZO, N., *et al.* High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening, **Nature Communications**, v. 9, n. 1, p. 364, 2018. DOI: 10.1038/s41467-017-02782-9.

TUCKER, G., YIN, X., ZHANG, A., *et al.* Ethylene, **Food Quality and Safety**, v. 1, n. 4, p. 253–267, 2017. DOI: 10.1093/fqsafe/fyx024.

ULUISIK, S., CHAPMAN, N. H., SMITH, R., *et al.* Genetic improvement of tomato by targeted control of fruit softening, **Nature Biotechnology**, v. 34, n. 9, p. 950–952, 2016. DOI: 10.1038/nbt.3602.

VILLARREAL, N. M., MARINA, M., NARDI, C. F., *et al.* Novel insights of ethylene role in strawberry cell wall metabolism, **Plant Science**, v. 252, p. 1–11, 2016. DOI: 10.1016/j.plantsci.2016.06.018.

VOINICIUC, C., PAULY, M., USADEL, B. Monitoring polysaccharide dynamics in the plant cell wall, **Plant Physiology**, v. 176, n. 4, p. 2590–2600, 2018. DOI: 10.1104/pp.17.01776.

WANG, D., YEATS, T. H., ULUISIK, S., *et al.* Fruit Softening: Revisiting the Role of Pectin, **Trends in Plant Science**, v. 23, n. 4, p. 302–310, 1 abr. 2018. DOI: 10.1016/j.tplants.2018.01.006.

WANG, Q.-H., ZHAO, C., ZHANG, M., *et al.* Transcriptome analysis around the onset of strawberry fruit ripening uncovers an important role of oxidative phosphorylation in ripening, **Scientific Reports**, v. 7, n. 1, p. 41477, 2017. DOI: 10.1038/srep41477.

YANG, L., HUANG, W., XIONG, F., *et al.* Silencing of SlPL, which encodes a pectate lyase in tomato, confers enhanced fruit firmness, prolonged shelf-life and reduced susceptibility to grey mould, **Plant biotechnology journal**, v. 15, n. 12, p. 1544–1555, dez. 2017. DOI: 10.1111/pbi.12737.

ZHANG, Y., LI, W., DOU, Y., *et al.* Transcript Quantification by RNA-Seq Reveals Differentially Expressed Genes in the Red and Yellow Fruits of *Fragaria vesca*, **PLOS ONE**, v. 10, n. 12, p. e0144356, 4 dez. 2015. DOI: 10.1371/journal.pone.0144356.

ZHAO, X., YUAN, X., CHEN, S., *et al.* Role of the tomato TAGL1 gene in regulating fruit metabolites elucidated using RNA sequence and metabolomics analyses, **PLOS ONE**, v. 13, n. 6, p. e0199083, 12 jun. 2018. DOI: 10.1371/journal.pone.0199083.

ZHBANNIKOV, I. Y., HUNTER, S. S., FOSTER, J. A., *et al.* Seqyclean: A pipeline for high-throughput sequence data preprocessing, **ACM-BCB 2017**, p. 407–416, 2017. DOI: 10.1145/3107411.3107446.

Supplementary material

Table S1. Nucleotide sequences used in qPCR.

Gene	Sequence (5'→3')
PG1	TGG TGG TGC GTA TAG ATG GA
	ACA AAA CCC AGT ACC CAC CA
PG2	TCC TGA AGC TCA CCC TTC AT
	CCT CAA TGC CTT TGA AGC TC
PG3	TTG GAG GGC AGC TTG TTT AG
	CAC CCA AGC CTT TAT TGT TCC
PL1	TTC CCT GTG GGC TTA CAA TC
	AGG GAG GTC TGC CAT TAC AT
PL2	CAT GTT CTT GTC CTG CGT GT
	CCA TCC ACG GCT CTA ATT TC
AGAL1	GTA TAG GCG GAA GGT GGA T
	AAG GCC ACC TCT CTG GAT
AGAL3	TGT CCG TTG ATC CTT CAG TG
	AGT CCG GAA GAA TGC TGA TG
BGAL1	GTG CTT GCA ACT ATG CTG GA
	ATA GGT TCG CAG TTG GGT TG
BGAL3	CCA AAG TGG GGA CAT TTG AG
	ACA CCC AGA CTT CGA CTT G
PME1	TAT CTT GGT AGG CCC TGG A
	AGG CCA GTG TTT CGG TAC T
PME2	GTG GTT TGT TCC TCA GCA CA
	TGG AAC GTA ACT GCA AGT GG
PME3	GCA AGC TTT AGG GGT GTT GA
	AGG CCT GCA GAG CTT ATT GA
ACT	CGT GAC CTT ACT GAT CAC TTG
	GTC AAG GGC AAT GTA AGA CAG
EF1	GTT AAG AAC GTT GCC GTG AAG
	ATG TGA AGT TGG CTG CTT CCT
UBQ	ACT CAC CGG CAA GAC CAT
	GTG GAG AGT CGA TTC CTT TTG

Table S2. GenBank Annotations

PG1	FJ007644
PG2	GQ479791
PG3	GQ479794
PL1	DQ660903
PL2	ABIM01001816
AGAL1	ABIM01008846
AGAL3	ABIM01016598
BGAL1	AF064786
BGAL3	ABIM01026480
PME1	GR486204
PME2	ABIM01018702
PME3	ABIM01014785

Table S3. Calibration curves for relative gene expression.

Name of the gene	Efficiency (10-slope)	y=ax+b	R^2
PG1	1.96	-3.4408x + 22.295	1
PG2	1.94	-3.4841x + 27.584	0.99
PG3	1.93	-3.4997x + 28.439	1
PL1	1.98	-3.3722x + 23.508	1
PL2	1.82	-3.8546x + 35.543	0.99
AGAL1	2.2	-2.9239x + 25.249	0.99
AGAL3	2.17	-2.9673x + 25.612	0.97
BGAL1	2.05	-3.2145x + 24.034	0.99
BGAL3	1.93	-3.4933x + 25.775	1
PME1	1.95	-3.4408x + 22.295	1
PME2	1.94	-3.4841x + 27.584	0.99
PME3	1.93	-3.4997x + 28.439	1
ARF	2	-3.3321x + 27.523	0.99
XYL	1.94	-3.4742x + 23.389	1
CELL	1.86	-3.7239x + 24.695	0.99
XTH	1.74	-4.1503x + 28.827	0.99
ACT	2.03	-3.2446x + 21.922	0.99
EF1	1.92	-3.5886x + 25.540	0.98
UBQ	1.89	-3.4841x + 22.049	0.99

Table S4. Expressed genes related to cell wall metabolism during papaya ripening.

Gene	Gene ID (NCBI)	Description
gene10120	LOC110814542	probable pectate lyase 13
gene1032	LOC110813559	lysine-rich arabinogalactan protein 18
gene10338	LOC110814908	expansin-A15-like
gene10377	LOC110814932	arabinogalactan peptide 13-like
gene10529	LOC110815026	exopolygalacturonase-like
gene10533	LOC110815072	probable pectinesterase/pectinesterase inhibitor 58
gene10588	LOC110815157	endoglucanase 25-like
gene10632	LOC110815166	cellulose synthase A catalytic subunit 2 [UDP-forming]-like
gene10728	LOC110815331	mannosylglycoprotein endo-beta-mannosidase
gene10769	LOC110815463	pectinesterase (Protein of unknown function, DUF538)
gene10774	LOC110815485	pectinesterase inhibitor-like
gene10836	LOC110815417	probable pectinesterase 29
gene10889	LOC110815466	pectin acetylesterase 12-like isoform X1
gene11115	LOC110815766	cellulose synthase A catalytic subunit 7 [UDP-forming]-like
gene11392	LOC110816092	classical arabinogalactan protein 9-like
gene11393	LOC110816167	classical arabinogalactan protein 1
gene11454	LOC110816184	probable pectinesterase/pectinesterase inhibitor 51
gene11696	LOC110816435	beta-galactosidase 10
gene11704	LOC110816515	beta-galactosidase 6 isoform X2
gene1193	LOC110815318	pectin acetylesterase 8 isoform X1
gene11930	LOC110816665	probable cellulose synthase A catalytic subunit 5 [UDP-forming] isoform X2
gene12026	LOC110816610	classical arabinogalactan protein 10
gene12117	LOC110817030	xyloglucan endotransglucosylase/hydrolase protein 9
gene12118	LOC110816923	fasciclin-like arabinogalactan protein 21
gene12216	LOC110816958	mannan endo-1,4-beta-mannosidase 2-like
gene12218	LOC110816900	alpha-L-arabinofuranosidase 1-like
gene12309	LOC110817105	Cellulase (glycosyl hydrolase family 5) protein
gene12310	LOC110817068	Cellulase (glycosyl hydrolase family 5) protein
gene12311	LOC110817069	Cellulase (glycosyl hydrolase family 5) protein
gene12343	LOC110817139	fasciclin-like arabinogalactan protein 17

gene12359	LOC110817077	probable pectinesterase/pectinesterase inhibitor 46
gene12413	LOC110817099	alpha-L-arabinofuranosidase 1-like isoform X2
gene12584	LOC110817416	polygalacturonase QRT3-like isoform X1
gene12707	LOC110817457	pectinesterase inhibitor 10
gene12708	LOC110817609	pectinesterase inhibitor 9-like isoform X3
gene12726	LOC110817606	probable xyloglucan galactosyltransferase GT19
gene12775	LOC110817680	beta-galactosidase 15 isoform X1
gene12777	LOC110817723	probable polygalacturonase
gene12802	LOC110817677	beta-galactosidase 15 isoform X1
gene12891	LOC110817819	non-classical arabinogalactan protein 31-like
gene13023	LOC110817778	pectinesterase inhibitor 5-like
gene13050	LOC110817890	cellulose synthase-like protein E6
gene13051	LOC110817796	cellulose synthase-like protein E6 isoform X1
gene13333	LOC110818133	expansin-A11
gene13357	LOC110818178	putative pectinesterase 11
gene13406	LOC110818407	beta-galactosidase 1 isoform X1
gene13517	LOC110818404	pectinesterase 1
gene1360	LOC110817097	alpha-galactosidase 1-like
gene13783	LOC110818785	endoglucanase 3
gene14044	LOC110818908	expansin-like A2
gene14153	LOC110819370	expansin-A8
gene14164	LOC110819226	alpha-galactosidase 3
gene14281	LOC110819182	Fasciclin-like arabinogalactan family protein
gene14581	LOC110819632	endoglucanase 25-like
gene147	LOC110817465	cellulose synthase A catalytic subunit 6 [UDP-forming]-like isoform X2
gene14718	LOC110819800	pectate lyase-like
gene14736	LOC110819677	pectinesterase 1
gene14785	LOC110819881	polygalacturonase inhibitor protein, partial
gene14835	LOC110819981	classical arabinogalactan protein 27
gene14914	LOC110820068	cellulose synthase-like protein G3 isoform X1
gene15057	LOC110820172	pectinesterase 31
gene15557	LOC110820823	fasciclin-like arabinogalactan protein 1
gene15943	LOC110821202	probable polygalacturonase At1g80170 isoform X2
gene15998	LOC110821276	pectinesterase inhibitor-like
gene16066	LOC110821401	probable polygalacturonase

gene16191	LOC110821415	pectinesterase (Protein of unknown function, DUF538)
gene16192	LOC110821414	pectinesterase (Protein of unknown function, DUF538)
gene16436	LOC110821790	putative xyloglucan endotransglucosylase/hydrolase protein 1
gene16477	LOC110821594	cellulose synthase, putative (DUF1644)
gene16482	LOC110821787	probable pectate lyase 8 isoform X2
gene1695	LOC110820693	uncharacterized protein LOC110820693
gene16952	LOC110822316	mannan endo-1,4-beta-mannosidase 7
gene16987	LOC110822129	beta-galactosidase 3
gene17040	LOC110822284	endoglucanase 24-like
gene171	LOC110819499	probable pectate lyase 4, partial
gene17454	LOC110822705	probable pectate lyase 3
gene17545	LOC110823044	probable pectate lyase 5
gene17689	LOC110823327	putative pectinesterase 63
gene17706	LOC110823186	arabinogalactan peptide 22-like
gene17738	LOC110823232	probable pectinesterase/pectinesterase inhibitor 34
gene17892	LOC110823425	xyloglucan galactosyltransferase XLT2
gene17935	LOC110823499	pectinesterase inhibitor 4
gene17936	LOC110823071	pectinesterase inhibitor 11
gene17939	LOC110823299	probable pectinesterase/pectinesterase inhibitor 59
gene1797	LOC110821988	probable polygalacturonase
gene18197	LOC110823748	probable xyloglucan endotransglucosylase/hydrolase protein 28
gene18257	LOC110823831	endoglucanase 8
gene18313	LOC110823677	expansin-A8-like, partial
gene18427	LOC110824296	xyloglucan endotransglucosylase/hydrolase protein 31
gene18615	LOC110824087	lysine-rich arabinogalactan protein 19
gene18763	LOC110824164	expansin-A4
gene19062	LOC110824624	cellulose synthase-like D3
gene19582	LOC110825189	polygalacturonase inhibitor (DUF639)
gene19599	LOC110825207	expansin-like B1
gene19834	LOC110825467	polygalacturonase inhibitor (DUF639)
gene19957	LOC110825603	cellulose synthase A catalytic subunit 3 [UDP-forming]-like
gene19959	LOC110825605	probable rhamnogalacturonate lyase B, partial

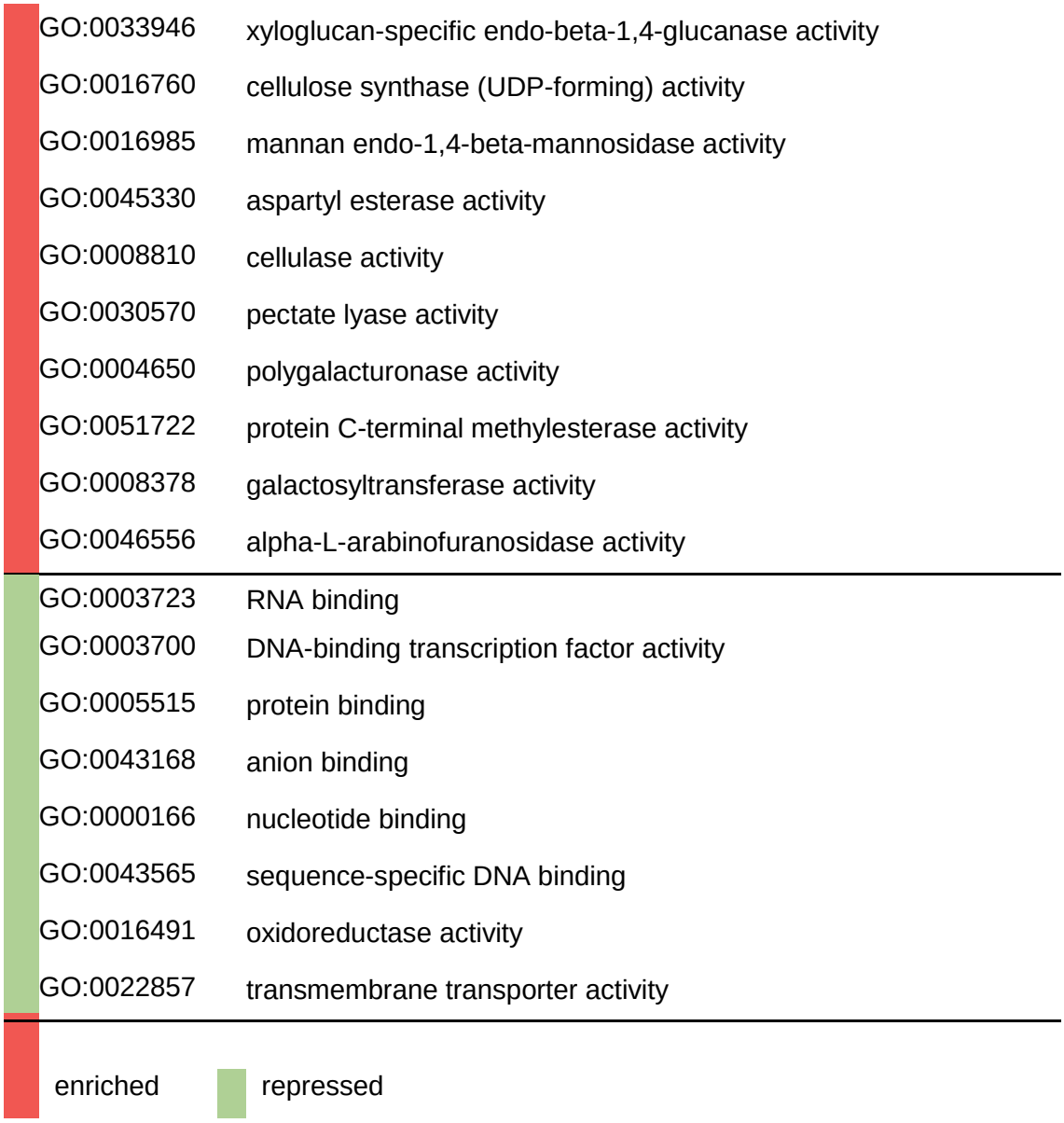
gene19960	LOC110825606	probable rhamnogalacturonate lyase B, partial
gene20068	LOC110825722	endoglucanase 6-like
gene2480	LOC110826208	expansin-A6-like
gene2536	LOC110826274	polygalacturonase-like
gene2660	LOC110826398	probable xyloglucan endotransglucosylase/hydrolase protein 25
gene2663	LOC110826405	xyloglucan endotransglucosylase/hydrolase protein 24-like
gene2664	LOC110826410	probable xyloglucan endotransglucosylase/hydrolase protein 23
gene2775	LOC110806382	fasciclin-like arabinogalactan protein 17
gene2830	LOC110806455	classical arabinogalactan protein 7-like
gene2838	LOC110806458	beta-galactosidase 8
gene2950	LOC110806596	probable xyloglucan endotransglucosylase/hydrolase protein 30
gene3084	LOC110806756	probable xyloglucan endotransglucosylase/hydrolase protein 8
gene3087	LOC110806763	pectinesterase
gene3486	LOC110807205	probable polygalacturonase
gene3555	LOC110807298	expansin-A13
gene3582	LOC110807303	probable rhamnogalacturonate lyase B isoform X1
gene38	LOC110807182	polygalacturonase inhibitor (DUF639)
gene3830	LOC110807600	mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS3, partial
gene3877	LOC110807696	fasciclin-like arabinogalactan protein 13
gene4133	LOC110807966	probable xyloglucan galactosyltransferase GT14
gene4134	LOC110807938	probable xyloglucan galactosyltransferase GT14
gene4246	LOC110808054	putative pectate lyase 2, partial
gene4292	LOC110808164	pectinesterase-like
gene4328	LOC110808110	endoglucanase 25-like
gene4520	LOC110808377	probable pectinesterase/pectinesterase inhibitor 36
gene4522	LOC110808380	pectinesterase inhibitor 9
gene4527	LOC110808384	probable pectinesterase 53
gene4550	LOC110808409	pectin acetylesterase 7-like
gene4805	LOC110808736	fasciclin-like arabinogalactan protein 2
gene5145	LOC110809080	probable xyloglucan endotransglucosylase/hydrolase protein 6
gene5336	LOC110809302	polygalacturonase QRT2-like

gene5562	LOC110809511	arabinogalactan peptide 13-like
gene5787	LOC110809801	cellulose synthase A catalytic subunit 1 [UDP-forming]
gene6084	LOC110810147	alpha-mannosidase I MNS5 isoform X1
gene6671	LOC110810818	beta-galactosidase 5
gene6889	LOC110811093	mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS1-like isoform X1
gene6942	LOC110811082	probable xyloglucan endotransglucosylase/hydrolase protein 23
gene699	LOC110810171	probable pectinesterase/pectinesterase inhibitor 12
gene7019	LOC110811182	beta-galactosidase-like
gene7049	LOC110811267	polygalacturonase-like isoform X1
gene7317	LOC110811520	xyloglucan galactosyltransferase MUR3, partial
gene7425	LOC110811647	mannan endo-1,4-beta-mannosidase 6
gene7693	LOC110811947	beta-galactosidase 9-like
gene7908	LOC110812108	probable xyloglucan galactosyltransferase GT11
gene803	LOC110811190	fasciclin-like arabinogalactan protein 4 isoform X1
gene8551	LOC110812917	cellulose synthase A catalytic subunit 2 [UDP-forming]
gene8626	LOC110813037	probable xyloglucan endotransglucosylase/hydrolase protein 5
gene879	LOC110811995	mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS1-like
gene8799	LOC110813200	mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS2-like, partial
gene9058	LOC110813488	polygalacturonase-like
gene9059	LOC110813489	polygalacturonase-like
gene9153	LOC110813600	beta-galactosidase 16 isoform X1
gene9155	LOC110813537	beta-galactosidase 16-like isoform X1
gene940	LOC110812631	mannan endo-1,4-beta-mannosidase 1, partial
gene9513	LOC110813995	probable pectate lyase 1
gene9560	LOC110814068	alpha-mannosidase I MNS4
gene957	LOC110812813	cellulose synthase-like protein D2, partial
gene9649	LOC110814146	probable polygalacturonase
gene9661	LOC110814143	pectin acetylesterase 6-like isoform X1
gene9675	LOC110814153	mannan endo-1,4-beta-mannosidase 1, partial

Table S5. Functional pathways associated with genes involved in papaya cell wall disassembling.

GO ID	GO Name
BIOLOGICAL PROCESS	
GO:0009833	plant-type primary cell wall biogenesis
GO:0030244	cellulose biosynthetic process
GO:0043086	negative regulation of catalytic activity
GO:0050829	defense response to Gram-negative bacterium
GO:0045490	pectin catabolic process
GO:0010400	rhamnogalacturonan I side chain metabolic process
GO:0046355	mannan catabolic process
GO:0000919	cell plate assembly
GO:0009612	response to mechanical stimulus
GO:0042538	hyperosmotic salinity response
GO:0009969	xyloglucan biosynthetic process
GO:0043622	cortical microtubule organization
GO:0009645	response to low light intensity stimulus
GO:0009741	response to brassinosteroid
GO:0052386	cell wall thickening
GO:0009846	pollen germination
GO:0042547	cell wall modification involved in multidimensional cell growth
GO:0009624	response to nematode
GO:0048358	mucilage pectin biosynthetic process
GO:0009834	plant-type secondary cell wall biogenesis
GO:0009409	response to cold
GO:0048359	mucilage metabolic process involved in seed coat development
GO:0009620	response to fungus
GO:0048767	root hair elongation
GO:0010047	fruit dehiscence
GO:0006949	syncytium formation

GO:0009826	unidimensional cell growth
GO:0009830	cell wall modification involved in abscission
GO:0006355	regulation of transcription, DNA-templated
GO:0065008	regulation of biological quality
GO:0048583	regulation of response to stimulus
GO:0048523	negative regulation of cellular process
GO:0006796	phosphate-containing compound metabolic process
GO:0071702	organic substance transport
GO:0031325	positive regulation of cellular metabolic process
GO:0055114	oxidation-reduction process
GO:0055085	transmembrane transport
GO:0006464	cellular protein modification process
GO:0010604	positive regulation of macromolecule metabolic process
GO:0009755	hormone-mediated signaling pathway
GO:0009737	response to abscisic acid
GO:0019752	carboxylic acid metabolic process
GO:0051641	cellular localization
GO:0006629	lipid metabolic process
GO:0051173	positive regulation of nitrogen compound metabolic process
GO:0033036	macromolecule localization
GO:0006396	RNA processing
GO:0006811	ion transport
GO:0010605	negative regulation of macromolecule metabolic process
GO:0071705	nitrogen compound transport
MOLECULAR FUNCTION	
GO:0046910	pectinesterase inhibitor activity
GO:0030599	pectinesterase activity
GO:0016762	xyloglucan:xyloglucosyl transferase activity
GO:0004565	beta-galactosidase activity
GO:0051753	mannan synthase activity


GO:0033946 xyloglucan-specific endo-beta-1,4-glucanase activity
GO:0016760 cellulose synthase (UDP-forming) activity
GO:0016985 mannan endo-1,4-beta-mannosidase activity
GO:0045330 aspartyl esterase activity
GO:0008810 cellulase activity
GO:0030570 pectate lyase activity
GO:0004650 polygalacturonase activity
GO:0051722 protein C-terminal methylesterase activity
GO:0008378 galactosyltransferase activity
GO:0046556 alpha-L-arabinofuranosidase activity
GO:0003723 RNA binding
GO:0003700 DNA-binding transcription factor activity
GO:0005515 protein binding
GO:0043168 anion binding
GO:0000166 nucleotide binding
GO:0043565 sequence-specific DNA binding
GO:0016491 oxidoreductase activity
GO:0022857 transmembrane transporter activity
enriched
repressed

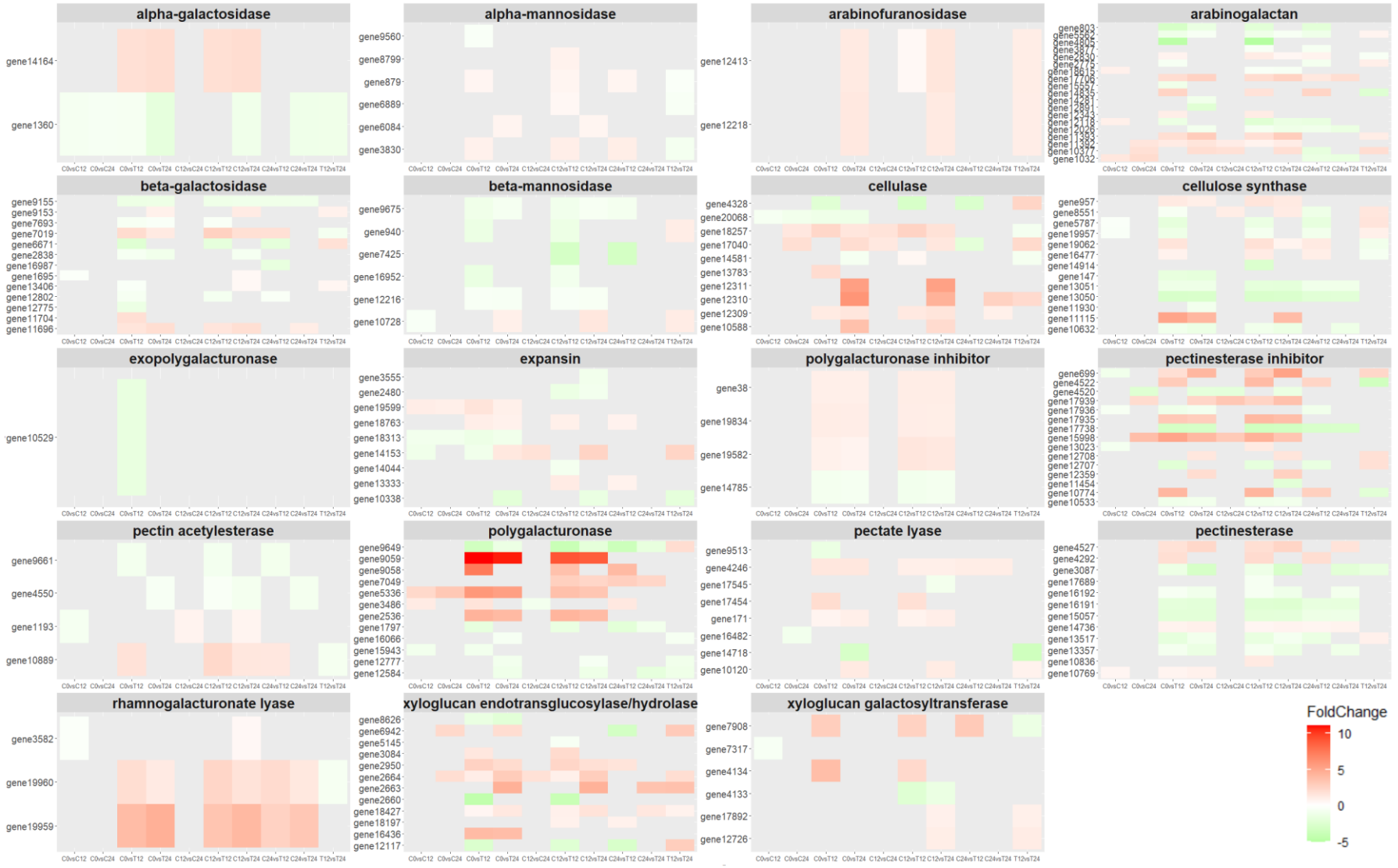


Figure S1. Heatmap of cell wall genes differentially expressed. Heatmaps of the expression levels of the genes coding for each enzyme related to cell wall disassembling.

Chapter III

Transcriptomics meta-analysis on the effect of ethylene in papaya ripening

Caroline Giacomelli Soares[1,2], João Paulo Fabi[1,2,3]

[1] Department of Food Science and Experimental Nutrition, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.

[2] Food Research Center (FoRC), CEPID-FAPESP (Research, Innovation and Dissemination Centers, São Paulo Research Foundation), São Paulo, Brazil.

[3] Food and Nutrition Research Center (NAPAN), University of São Paulo, São Paulo, Brazil.

Abstract

The availability of robust NGS data in public repositories allows *in silico* studies through bioinformatics tools. Regarding the integration of data, it is indispensable to carry a reliable statistical adjustment in order to remove batch effect and technical bias. Transcriptomics is a valuable technology to explore changes in the global expression profile towards different biological conditions. Post-harvest studies on the ripening process have been benefited with such strategies due to the high rate of fruit losses, either because of poor management and processing conditions of sensitive fruit, or because of short shelf-life. Papaya is a climacteric fruit that undergoes fast modification during the ripening process, widely regulated by the hormone ethylene. Even though there is extensive literature on the ripening of climacteric fleshy fruit, papaya remains with huge gaps in the understanding of the complex coordination of the molecular changes induced by ethylene. Thus, the present work gathered online available transcriptomes of papaya ripening and deduced functional pathways induced and repressed on immature and ripe fruits. DEG analysis identified 47 candidate genes on the major regulation of downstream enzymatic chain that leads to pulp softening and enhancement of color, flavor, aroma and nutrient content. Deeper investigation on the activity of these genes is needed, yet the filtering in this significant subset of genes with statistical power and reproducibility is a great first step to the full elucidation of ripening expression profile.

Keywords: ethylene; ripening-related modifications; Papaya; transcriptome; meta-analysis.

Abbreviations: NGS, Next Generations Sequencing; 1-MCP, 1-methylcyclopropene; PE, paired-end; GSEA, Gene Set Enrichment Analysis; ORA, Over Representation Analysis; GO, Gene Ontology; DEG, differentially expressed gene(s); ERF, ethylene-responsive fator; IAA, Indole-3-acetic acid; PME, pectinesterase.

Introduction

The advance in technology of research equipment increases exponentially with the availability of detailed data on all kinds of assays, from the genetic level to the phenotype. As long as genome sequencing uncovers a single cell's genetic material, variation in the expression profile engenders many combinations of transcripts depending on intra- and extracellular stimuli. In this regard, a suitable approach to binary "case-control" experiments is the inspection of the transcriptome, which enables the disclosure of the expression coordination towards distinct biological circumstances. Scientific community encourages the deposition of high throughput data in public repositories, "enabling researchers to plan experiments and to analyze existing data" (BLAXTER et al., 2016). In this way arose the so-called *in silico* science, feasible by bioinformatics tools and mathematics, entailing a migration from reductionist to a systems biology approach (PALSSON, 2000). Data integration should be carefully done since many technical variables can cause a misstatement in statistical reliability. Meta-analysis strategy consists of firstly correct background effects and normalize data, then proceed with further analyses for each study independently, and finally combine the results by their p-values into a z-score (EVANGELOU; IOANNIDIS, 2013).

Papaya is a climacteric fruit with great commercial importance but also a high rate of post-harvest losses due to pulp softening. Ethylene is an autocatalytic hormone that induces a coordinated and complex regulation of ripening on a transcription level, which is inhibited in the presence of its antagonist 1-MCP (FABI *et al*., 2007). Climacteric fruit ripening is strongly ethylene-induced, and although it provides nutrients, flavor, color, and aroma to the pulp, it also increases cell wall swelling by the degradation of pectin and hemicellulose polysaccharides promoted by ethylene-induced enzymes. Even though many studies have been developed on papaya transcriptome to investigate the ripening regulation, there are still many gaps in the understanding of molecular and metabolic pathways of the ripening process. In this manner, this study analyzed transcriptomic data available on a public repository in order to further investigate the expression profile changes related to ripening metabolism.

Material and methods

Data mining

The Sequence Read Archive (SRA) and Genome Expression Omnibus (GEO) repositories from NCBI were investigated on February 17, 2020, using the search terms “*Carica papaya* ripening”, “*Carica papaya* transcriptome”, “*Carica papaya* transcriptomics” and “*Carica papaya* RNA-seq”. In total, there were 342 search results. The filtering was carried using the criteria 1: no pulp samples, 2: not ripening-related, and 3: no raw data available. In addition, the RNA-seq experiment reported in **Chapter I** was also included. Selected studies are described in Table 1. Raw sequence data were downloaded using SRA ToolKit (https://github.com/ncbi/sra-tools). Workflow is schemed in Figure 1.

Processing of raw RNA-seq data

Reads of all samples were submitted to a quality assay in FastQC (ANDREWS, 2010) and trimmed by SeqyClean (ZHBANNIKOV *et al.*, 2017) and results were joined in MultiQC (EWELS *et al.,* 2016). Reads were mapped against the *Carica papaya* genome Papaya1.0 (GCF_000150535.2 – NCBI) with STAR (DOBIN *et al.*, 2013), followed by the counting of the reads by htseq-count (ANDERS *et al.*, 2015).

Statistical data integration

Considering the differences in RNA manipulation and sequencing platforms among experiments, raw reads of all samples were adjusted for batch effect through negative binomial regression in Combat-seq R package (ZHANG; PARMIGIANI; JOHNSON, 2020), with further normalization by variance stabilizing transformation (VST) using the DESeq2 package (LOVE; HUBER; ANDERS, 2014), both in R software (R Core Team, 2014). Principal component analysis (PCA) was applied in the three stages of data adjustment to verify the

Table 1: Data information. Bioproject and Run are accession numbers from https://trace.ncbi.nlm.nih.gov/Traces/sra/. PRJNA528193 is not available online yet.

Bioproject	Run	Sample	Tag	Sequencing	Reference
PRJNA352643	SRR5019672	MCP-treated papayas	MCP	Illumina HiSeq 2500	SHEN *et al*., 2017
	SRR5019673	Non-treated papayas	Ctrl		
	SRR5019674	ETH-treated papayas	ETH		
PRJNA449965	SRR6996528	flesh of green stage papaya fruit	GS	Illumina HiSeq 2500	SHEN et al., 2019
	SRR6996529	flesh of color break stage papaya fruit	CBS		
PRJNA381300	SRR5405093	Papaya (cv Tainong1) fruit flesh at stage4	Stage4	HiSeq X Ten	LÜ et al., 2018
	SRR5405094	Papaya (cv Tainong1) fruit flesh at stage3	Stage3		
	SRR5405095	Papaya (cv Tainong1) fruit flesh at stage2	Stage2		
	SRR5405096	Papaya (cv Tainong1) fruit flesh at stage1	Stage1		
	SRR6320462	Papaya (cv Golden) fruit flesh at stage1 (immature – 30DPA)	immature	Illumina HiSeq 4000	
	SRR6320463	Papaya (cv Golden) fruit flesh at stage4 (ripe – 150DPA)	ripe		
PRJNA528193	Private	Control 0h	C0h	Illumina HiSeq 2500	
		Control 12h	C12h		
		Control 24h	C24h		
		ETH-treated 12h	T12h		
		ETH-treated 24	T24h		

modifications. Samples were then clustered and submitted to a Pearson correlation test in order to check their similarity.

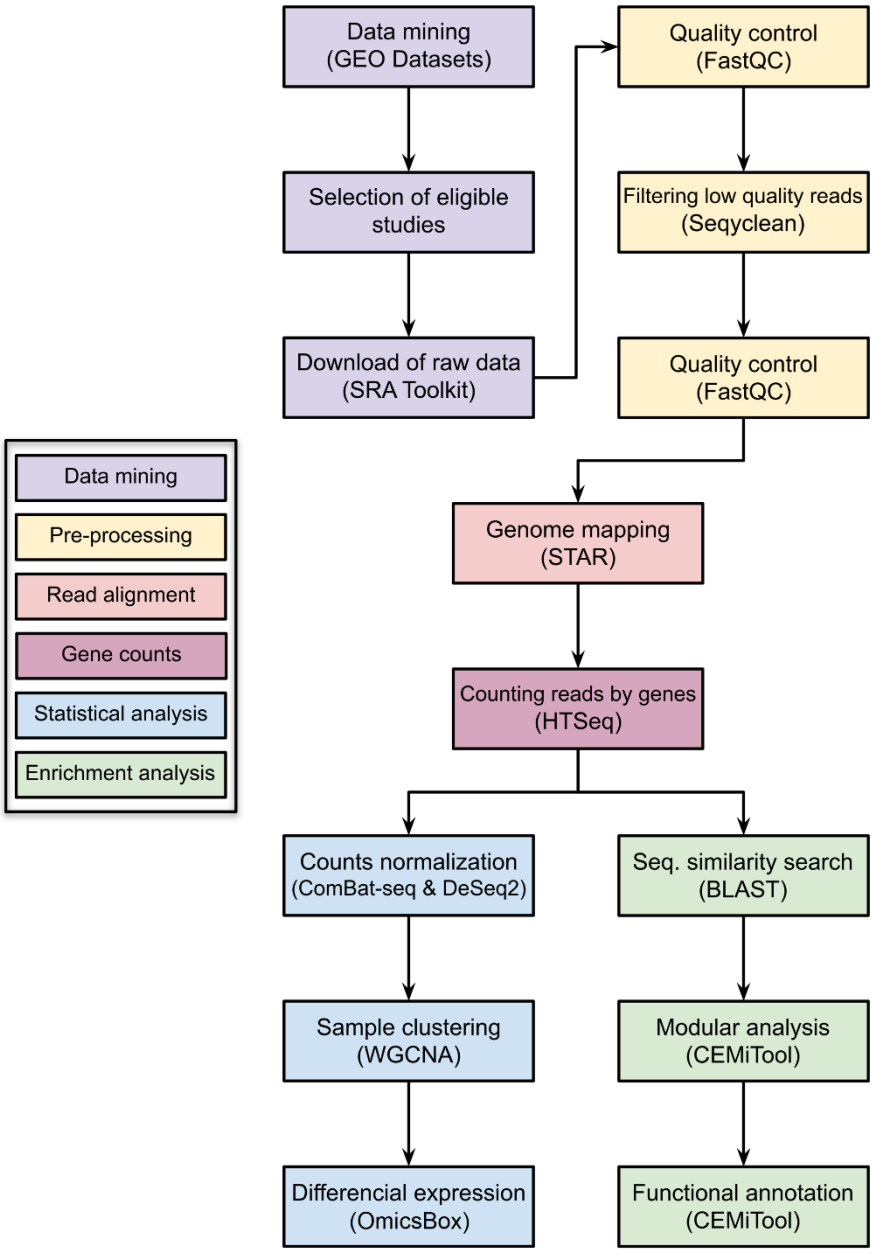


Figure 1. **Workflow of the development of this study.** Colors represent the different steps of the bioinformatics analysis.

Biological analyses

Differential expression analysis was run into OmicsBoxs, which is based on NOISeq package (TARAZONA et al., 2011). NOISeq simulated five technical replicates with variability of ±0.02 for each sample in order to calculate fold-change differences (M) and absolute expression differences (D), generating an overall

probability for each gene of being up- or down-regulated. Comparisons between the two most extreme samples of each study were carried independently. Genes with probability > 0.07 were considered significantly deferentially expressed. Gene Set Enrichment Analysis (GSEA) and Functional Enrichment Analysis were carried on webCEMiTool (https://cemitool.sysbio.tools/) and in Blast2GO (Biobam Bioinformatics S.L., Valencia, Spain).

Results

Collection and processing of raw data

Data were downloaded from de Sequence Read Archive (NCBI – accession number in column “Run”, Table 1) using the SRA TooKit command *fastq-dump.exe* with the argument *–split-files* to generate two files related to PE 1 and 2, as all the samples were sequenced in paired-end. Sequences were submitted to a quality assay, filtered from vectors, Poly A/T, and adapter remaining from the sequencing, and evaluated again. Results of FastQC (ANDREWS, 2010) assay were attached in MultiQC (EWELS *et al.,* 2016) to easier visualization (Figure 1S). Attributes that had no improvement are common bias in RNA-seq data: per sequence base content is affected by hexamers used in the preparation of cDNA libraries, and duplicated sequences are usually because there are enriched genes indeed. After trimming, it was kept around 80% of the total reads. As the quality of reads were satisfactory, they were mapped to the reference genome using STAR (DOBIN *et al.*, 2013) with --outFilterMultimapNmax 5 to limit to five the maximum number of multi-mapping reads.

Batch adjustment and normalization

After reads counting, raw counts were joined in a single expression table and were submitted to batch adjustment through Combat-Seq, then normalization by *getVarianceStabilizedData()* function of DESeq2 (LOVE; HUBER; ANDERS, 2014). A PCA plot was made to visualize the changes in all stages of data adjustment (Figure

2). Raw data was divided by studies and the batch effect correction grouped all of them into a single ellipse, but still demonstrating a sub-group within it. Variance transformation based on dispersion and size factors successfully normalized all the samples into a proportional and cohesive expression data.

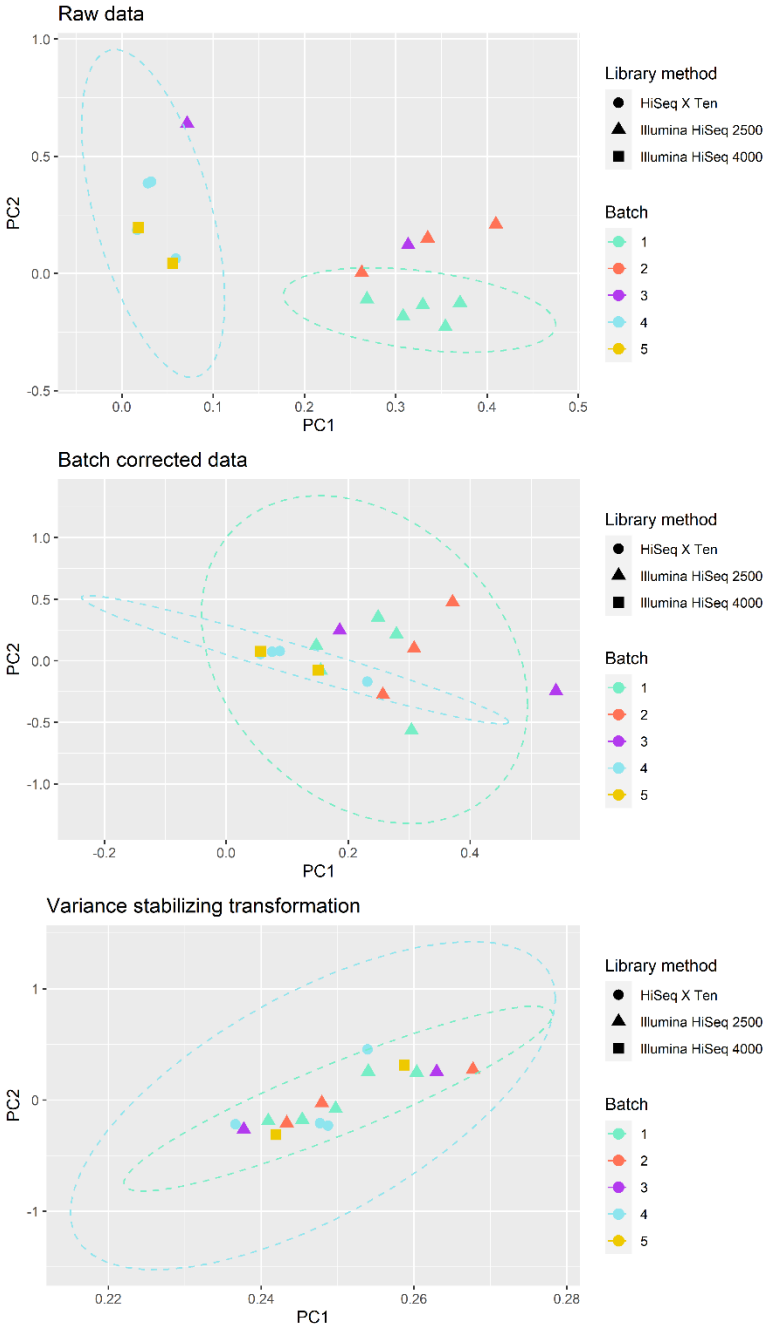


Figure 2. PCA plots of the steps to normalize the data. Shapes indicate the platform of sequencing, and colors represent the studies. First plot shows the raw data with samples randomly placed; second plot shows data after statistical correction for batch (studies) effect; last plot shows data after normalization, with all samples grouped into a single ellipse.

As a first overview of the expression profile, samples were clustered by similarity and Pearson correlation was calculated (Figure 3). The cluster dendrogram is divided into two main branches clearly grouping the conditions of interest: immature and ripe fruits. Samples are highly correlated to each other, but it is possible to see smaller correlation between most of the ripe and immature samples. High correlation suggests their expression counts are similar, which is a fair point when considering the whole transcriptome containing several basal and maintenance genes.

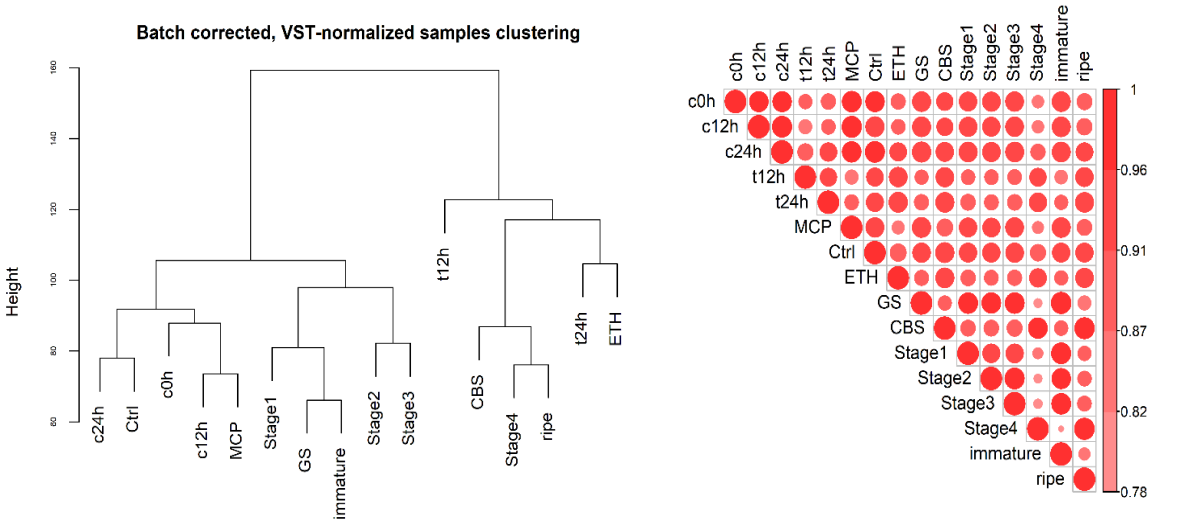


Figure 3. Cluster dendrogram and Pearson correlation of samples. The dendrogram shows the clustering of normalized data by expression level similarity. Pearson correlation between all samples is represented by the size of the circle and the intensity of the color. It is evident they have specific characteristics which differentiate them, although their expression behavior is similar.

Co-expression module analysis

CEMiTool (RUSSO *et al.*, 2018) detected 11 modules of genes with similar expression profile, and together with Gene Set Enrichment Analysis (GSEA – Figure 4) and Over Representation Analysis (ORA - Figure S2), it is possible to understand the major functional pathways the genes are involved into. Previously, it is important to elucidate the difference between this enrichment analysis: while GSEA assigns an enrichment score (ES, normalized in the plot of Figure 4) to each module towards a sample, ORA reveals the enriched pathways in the modules, represented in GO terms. M1 is enriched to mostly control samples, mainly GS and immature samples, in contrast to the suppression on Stage4, T12h and T24h. Genes within M1 are related to defense response, response to sucrose and cellular developmental process. M3 has no pattern in relation to samples, enriched and suppressed on both immature and ripe

samples. It is enriched to GO terms of response to chemical stimulus, cellular metabolic process and brassinosteroid activity. Genes on module 5 were strongly enriched to GS and immature samples but suppressed in other controls. They are mostly involved in response to the abiotic stimulus of red and blue light, as well as in biotic stress and cellular development process. M8 is involved in photosynthesis regulation by chloroplast allocation. Samples in earlier stages are suppressed and those in latter stages are enriched to M9, which acts in the negative regulation of ethylene signaling and positive regulation of auxin signaling. Finally, M10 in enriched to volatile biosynthetic process, such as linoleic acid. Non-cited modules had no significant enrichment.

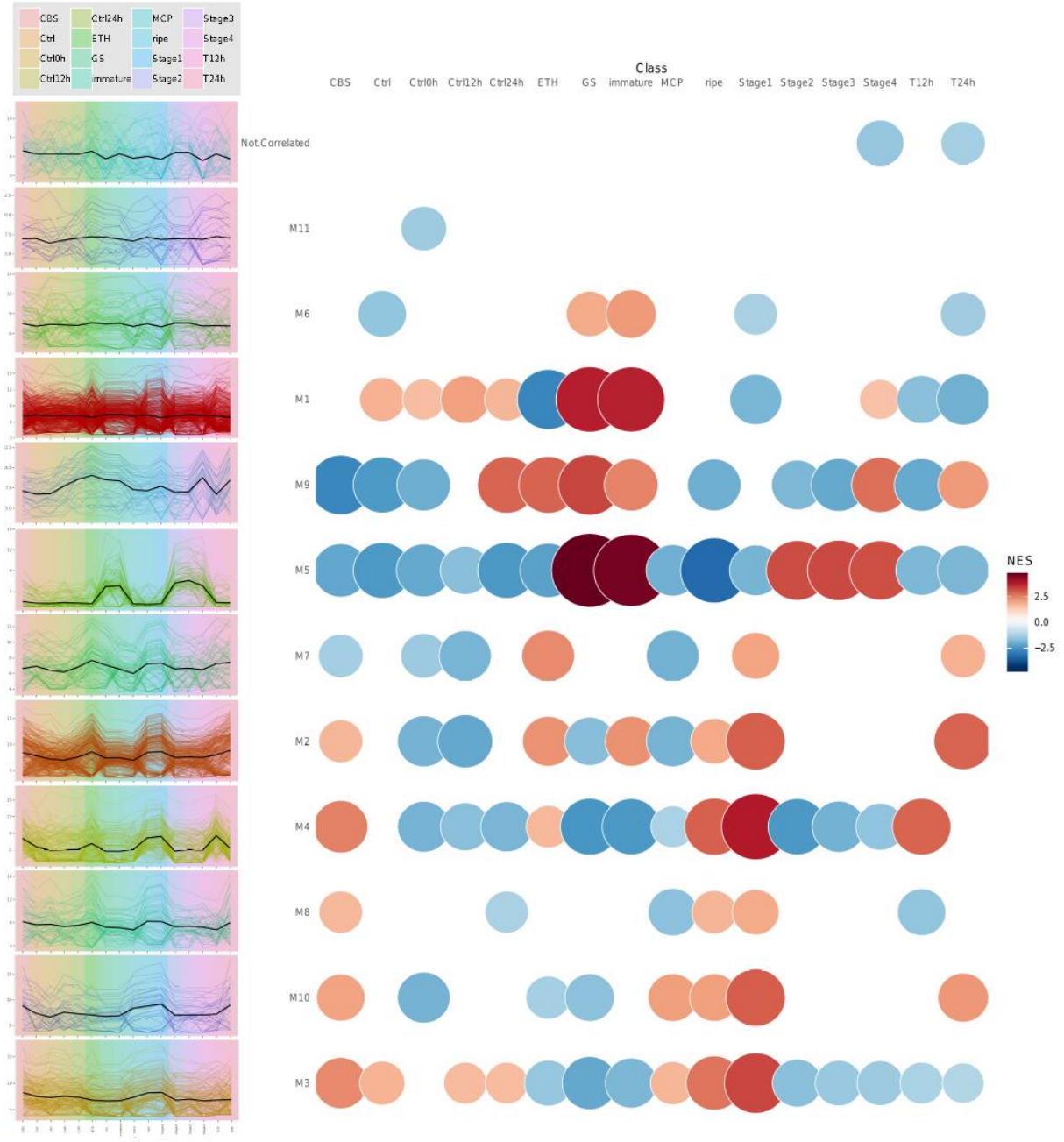


Figure 4. Expression profile and GSEA for modules detected in co-expression analysis. Line graphs in the left show the expression profile of modules in each sample (represented by the background color). GSEA plot shows the NES (Normalized Enrichment Score) of modules in each sample. Red is enriched while blue is suppressed; The size of the circles is proportional to the respective NES

Differential Expression Analysis

Independent analysis of each study detected 245, 198, 339, 278, and 182 DEGs in comparisons ETH x MCP, CBS x GS, Stage4 x Stage1, ripe x immature, and T24h x C0h, respectively. Among intercalation of all DEGs, it was detected 47 genes that

were differentially expressed in all comparisons (Figure S3). Many of the DEGs are key genes of ripening, such as transcription factors ERF (LOC110814497), SCARECROW (LOC110807644), IAA (LOC110808074), and MADS-box (LOC110824954); genes related to the plant cell wall, like PME (LOC110818405) and β–galactosidase (LOC110811182); hormones, ABA (LOC110815190); and vitamin E (LOC110815941), among others (Figure 5). These genes were enriched to functional pathways of response to hormones, transmembrane transporter activity, intracellular anatomical structures, and macromolecules and lipid metabolism (Figure S4).

Discussion

The integration of the transcriptomes presented a fair behavior regarding the biological condition of the samples. The maturity stages were clearly grouped within their similarity yet divided by treatment. The overall functions assigned to the significative modules are relatable to the ripening process, such as the enrichment of sucrose response on control samples, which happens until the onset of ripening (GOMEZ; LAJOLO; CORDENUNSI, 2002); organism maintenance-related functional pathways, such as response to stimulus and brassinosteroid activity (a hormone involved in the response to several cases of stress – ANWAR et al., 2018), were enriched and repressed for both groups. Module 9 presented significant functions in ripening, representing the antagonist activity of ethylene and auxin (MUDAY; RAHMAN; BINDER, 2012). The enrichment volatile biosynthetic process and linoleic acid suggest an enhance in the aroma of the fruits (FABI; DO PRADO, 2019). In an overview, it is possible to assert that the GSEA results corroborate to what is described in the literature regarding the general activities of fruit, showing biologically that the statistical integration of data was successfully normalized.

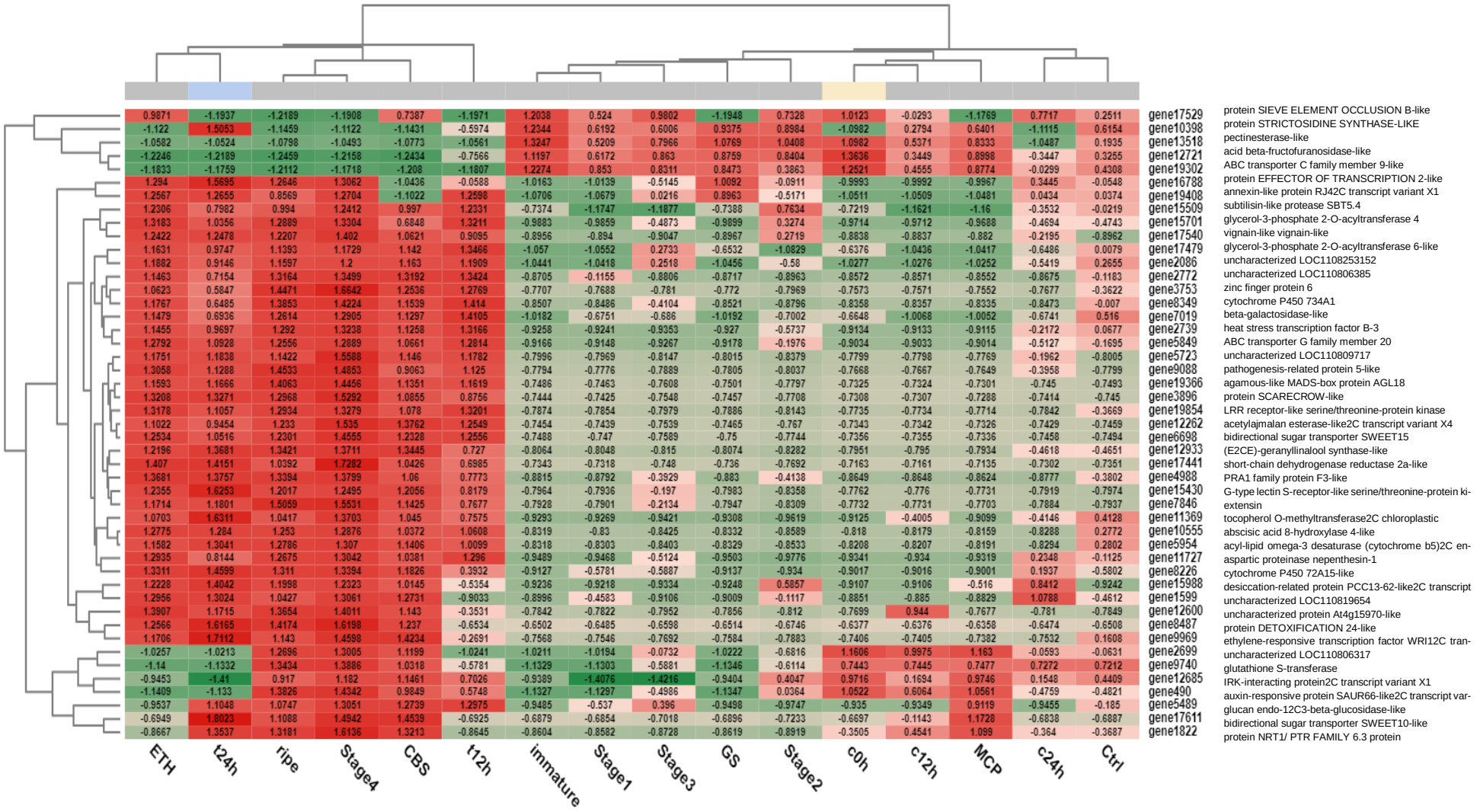


Figure 5. Heatmap of the 47 mutual DEGs. Red represents up-regulation, and green under-regulation. Numbers inside cells are the log2FoldChange. Lines are grouped by gene expression similarity while columns are grouped by sample similarity.

The 47 mutual differentially expressed genes were enriched to pathways highly related to the ripening process, such as response to fatty acids and lipid biosynthetic process, involved in carotenoids biosynthesis pathway (SHEN *et al.,* 2019). The up-regulation of a tocopherol-related gene also represents the ripening process of fruit (SAINI; ZAMANY; KEUM, 2017). The identification of well-known ripening-related genes, such as transcription factors and hormone-related genes as DEGs reinforces their indispensable activity on the coordination of the ripening process.

Conclusion

The meta-analysis of papaya ripening transcriptomes proved to be an effective approach to understand the coordination of gene expression. The variability of data generated by distinct studies enriches statistical power by representing several conditions that fruit can suffer during such a complex process, even towards the same experimental design. The detection of 47 mutual differentially expressed genes on papaya treated with ethylene provides promising candidate genes to act as key factors on ripening regulation. Further studies are needed to deepen these results, but the specificity and cohesion of the relation to the proposal of this work is a favorable signal that this approach is able to elucidate some gaps regarding the expression profile of papaya ripening.

References

ANDERS, S., PYL, P. T., HUBER, W. "HTSeq--a Python framework to work with high-throughput sequencing data", **Bioinformatics (Oxford, England)**, v. 31, n. 2, p. 166–169, 15 jan. 2015. DOI: 10.1093/bioinformatics/btu638.

ANDREWS, S. FastQC: a quality control tool for high throughput sequence data. 2010; Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

ANWAR, A., LIU, Y., DONG, R., *et al.* "The physiological and molecular mechanism of brassinosteroid in response to stress: a review", **Biological Research**, v. 51, n. 1, p. 46, 2018. DOI: 10.1186/s40659-018-0195-2

BLAXTER, M., DANCHIN, A., SAVAKIS, B., *et al.* "Reminder to deposit DNA sequences", **Science**, v. 352, n. 6287, p. 780 LP – 780, 13 maio 2016. DOI: 10.1126/science.aaf7672

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., *et al.* "STAR: ultrafast universal RNA-seq aligner", **Bioinformatics**, v. 29, n. 1, p. 15–21, 1 jan. 2013. DOI: 10.1093/bioinformatics/bts635.

EVANGELOU, E., IOANNIDIS, J. P. A. "Meta-analysis methods for genome-wide association studies and beyond", **Nature Reviews Genetics**, v. 14, n. 6, p. 379–389, 2013. DOI: 10.1038/nrg3472

EWELS, P., MAGNUSSON, M., LUNDIN, S., *et al.* "MultiQC: summarize analysis results for multiple tools and samples in a single report", **Bioinformatics**, v. 32, n. 19, p. 3047–3048, 1 out. 2016. DOI: 10.1093/bioinformatics/btw354

FABI, J. P., CORDENUNSI, B. R., BARRETO, G. P. D. M., *et al.* "Papaya fruit ripening: Response to ethylene and 1-methylcyclopropene (1-MCP)", **Journal of Agricultural and Food Chemistry**, v. 55, n. 15, p. 6118–6123, 2007. DOI: 10.1021/jf070903c.

FABI, J. P., DO PRADO, S. B. R. "Fast and furious: Ethylene-triggered changes in the metabolism of papaya fruit during ripening", **Frontiers in Plant Science**, v. 10, n. April, 2019. DOI: 10.3389/fpls.2019.00535.

GOMEZ, M., LAJOLO, F., CORDENUNSI, B. "Evolution of Soluble Sugars During Ripening of Papaya Fruit and its Relation to Sweet Taste", **Journal of Food Science**, v. 67, n. 1, p. 442–447, 1 jan. 2002. DOI: https://doi.org/10.1111/j.1365-2621.2002.tb11426.x

LÜ, P., YU, S., ZHU, N., *et al.* "Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening", **Nature Plants**, v. 4, n. 10, p. 784–791, 2018. DOI: 10.1038/s41477-018-0249-z.

MUDAY, G. K., RAHMAN, A., BINDER, B. M. "Auxin and ethylene: collaborators or competitors?", **Trends in Plant Science**, v. 17, n. 4, p. 181–195, 1 abr. 2012. DOI: 10.1016/j.tplants.2012.02.001

PALSSON, B. "The challenges of in silico biology", **Nature Biotechnology**, v. 18, n. 11, p. 1147–1150, 2000. DOI: 10.1038/81125.

RUSSO, P. S. T., FERREIRA, G. R., CARDOZO, L. E., *et al.* "CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses", **BMC Bioinformatics**, v. 19, n. 1, p. 56, 2018. DOI: 10.1186/s12859-018-2053-1

SAINI, R. K., ZAMANY, A. J., KEUM, Y.-S. "Ripening improves the content of carotenoid, α-tocopherol, and polyunsaturated fatty acids in tomato (*Solanum lycopersicum* L.) fruits", **3 Biotech**, v. 7, n. 1, p. 43, maio 2017. DOI: 10.1007/s13205-017-0666-0.

SHEN, Y. H., LU, B. G., FENG, L., *et al.* "Isolation of ripening-related genes from ethylene/1-MCP treated papaya through RNA-seq", **BMC Genomics**, v. 18, n. 1, p. 671, 2017. DOI: 10.1186/s12864-017-4072-0

SHEN, Y. H., YANG, F. Y., LU, B. G., *et al.* "Exploring the differential mechanisms of carotenoid biosynthesis in the yellow peel and red flesh of papaya", **BMC Genomics**, v. 20, n. 1, p. 49, 2019. DOI: 10.1186/s12864-018-5388-0

TARAZONA, S; GARCÍA-ALCALDE, F; DOPAZO, J; FERRER, A; CONESA, A. Differential expression in RNA-seq: a matter of depth. Genome Res. Dec;21, v. 12, p. 2213-23, 2011. doi: 10.1101/gr.124321.111.

ZHANG, Y., PARMIGIANI, G., JOHNSON, W. E. "ComBat-seq: batch effect adjustment for RNA-seq count data", **NAR Genomics and Bioinformatics**, v. 2, n. 3, 1 set. 2020. DOI: 10.1093/nargab/lqaa078.

ZHBANNIKOV, I. Y., HUNTER, S. S., FOSTER, J. A., *et al.* "SeqyClean: A Pipeline for High-Throughput Sequence Data Preprocessing". 2017. **Anais** [...] New York, NY, USA, Association for Computing Machinery, 2017. p. 407–416. DOI: 10.1145/3107411.3107446.

Supplementary material

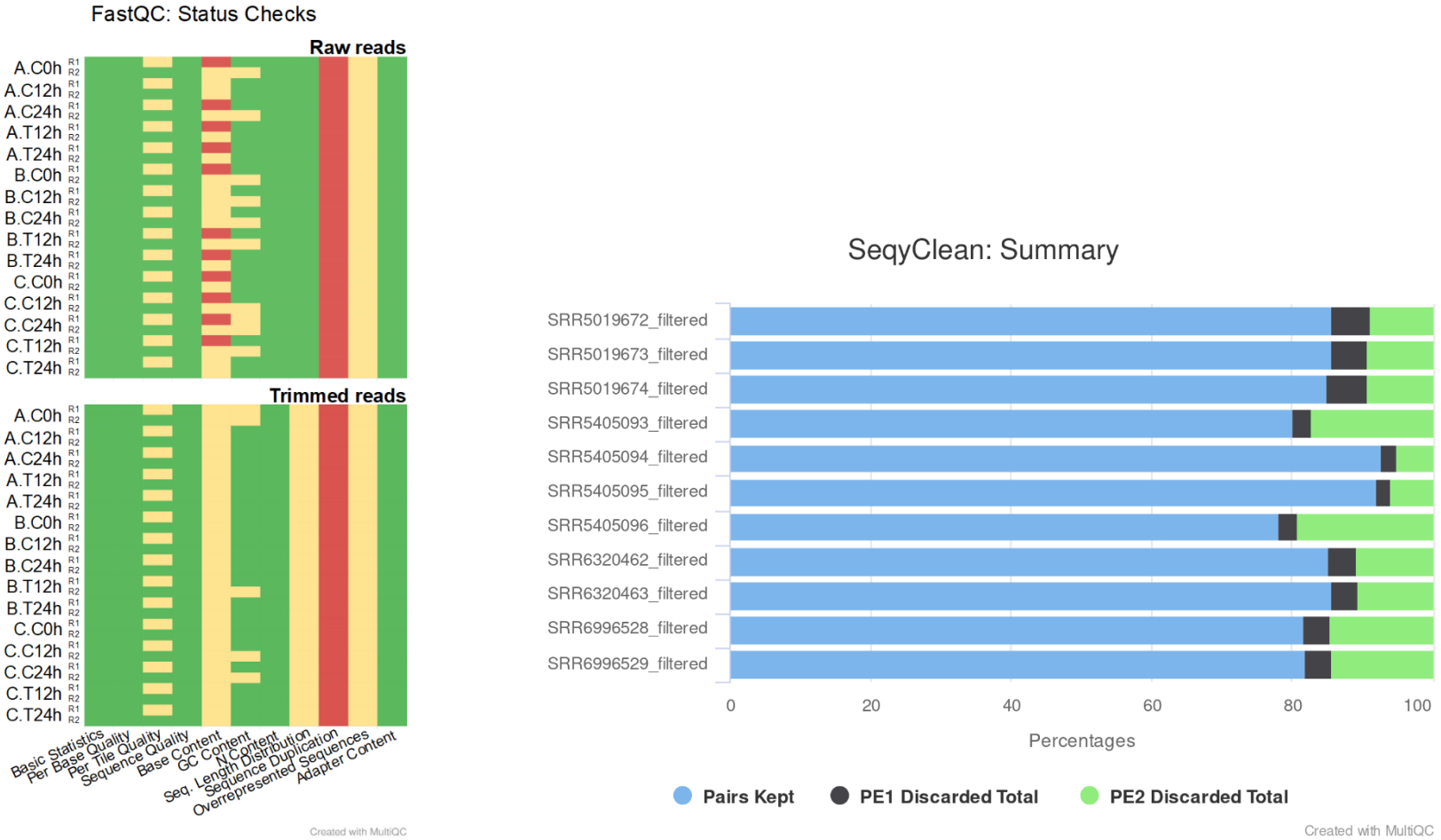


Figure S1. Quality assays of the pre-processing of raw data. (A) represents sequence attributes before and after filtering. (Green: entirely normal, orange: slightly abnormal, red: very unusual). (B) Proportion of base pairs throughout the filtering process.

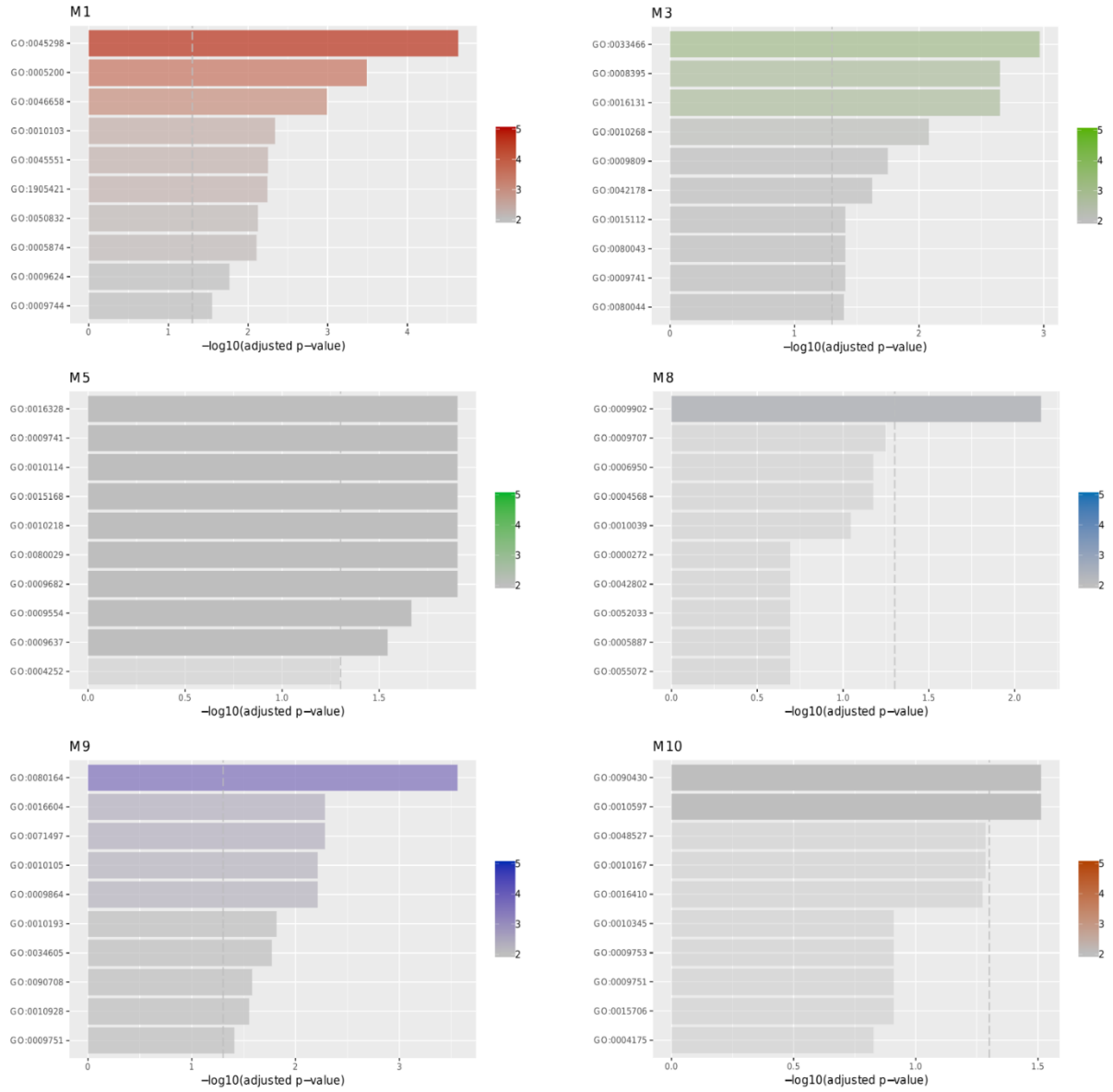


Figure S2. Significant ORA analysis. Modules not shown had no significant enriched pathways. Scale of color is for NES; bar size is for the p-value. Dashed lines mark the minimum value of significance.

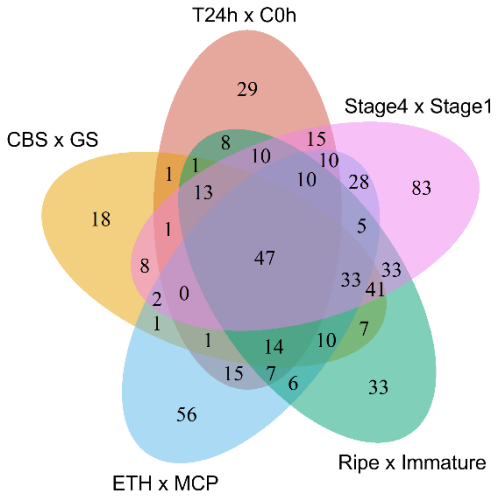


Figure S3. Venn diagram of DEGs of comparisons between each study independently.

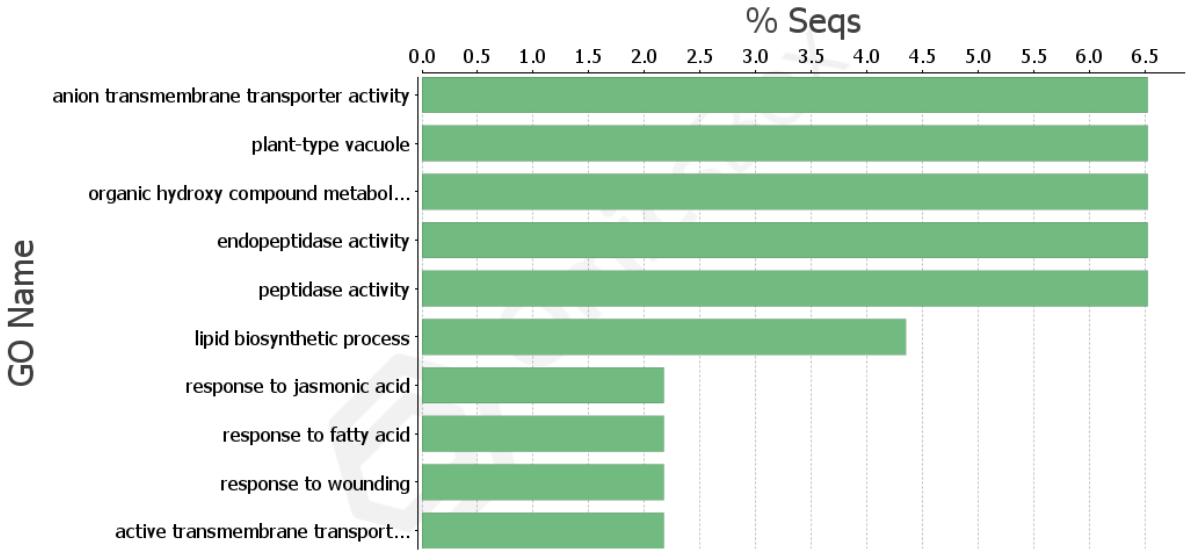


Figure S4. GSEA of the 47 mutual DEGs. The most enriched biological pathways for the 47 mutually, differentially expressed genes on all pairwise comparisons between each study of the meta-analysis.

FINAL CONSIDERATIONS

The fruit reaches its physiological maturation when cellular expansion is no longer happening and the fruit is at its maximum size. From this moment on, fruit can be detached from the mother plant with no harm to the prosecution of the physiological activity. Another maturation indicative is the full development of seeds, that were protected by the pericarp and are fully developed to be germinated. The fruit, then, enhance the sensorial properties of the pericarp in order to attract predators that will disperse seeds, thus proceeding with plant reproduction. Beyond the ecological duty, the pericarp comprises the edible portion in fleshy fruits, which are also benefited from the sensorial enhancement. Once the physiological maturation is achieved, the fruit starts the ripening process which will improve the aroma, flavor, color, and nutrient content of the fruit, whilst the pericarp is degraded to eventually expose the seeds. Regarding the human consumption of fleshy fruits, ripening is a two-sided factor: it turns the fruit palatable but also makes it susceptible to physical damages and pathogen attack.

Ripening process is the result of an intense biochemical activity that is coordinated by intra- and extracellular signaling. Climacteric fruits ripening undergoes a great influence of hormones, mainly ethylene. The autocatalytic production of ethylene makes the ripening of climacteric fruit a fast event, which leads to a short shelf-life hence a large-scale post-harvest loss. In this regard, ripening is exhaustively studied to achieve control of this process and reduce post-harvest losses. One of the fruits with the higher rates of loss is papaya, and because of its wide consumption and commercial importance, the pursuit for a solution to post-harvest losses is of great importance. Ripening is regulated at a transcription level, thus investigating the transcriptome is a reliable method to identify key factors which lead to molecular and biochemical modifications. The study of the transcriptomes of exogenous ethylene treated and non-treated papaya enabled elucidation on the global expression profile changes.

The treatment with exogenous ethylene accelerated and induced physiological changes in pulp firmness and color, as well as increase the activity of many ripening and ethylene-related enzymes. Autocatalytic production of ethylene was evinced by the high expression differences between treated and

control fruits. Genes encoding ethylene synthesis and signaling enzymes were highly up-regulated and were enriched to many ripening-related functional pathways. Furthermore, it was possible to detect many ethylene-induced enzymes with downstream activity, such as changes in the fruit cell wall. Pulp softening is promoted by the coordinated action of pectinases and hemicellulases which act directly on the disassembling of the pectic portion of the plant cell wall, causing pulp swelling and consequent softening. The enzyme with major activity on papaya fruit cell wall disassembling is the polygalacturonase, with no need of previous action of pectate lyase and pectinesterase as described to other fleshy fruits. Many other significant functional pathways and differentially expressed genes were observed, corroborating with the complex network that coordinates the climacteric ethylene-induced ripening process.

Transcriptomics is a good approach to the study of the expression profile on distinct conditions because transcription is what settles the connection between gene expression and phenotype. Moreover, the integration of transcriptomes from several experiments on the same subject allows an enriched, robust understanding of the molecular changes. The meta-analysis of papaya ripening transcriptomes led to the detection of 47 genes that were mutually different expressed. Further investigation is needed, but there are strong candidates to represent key factors in the genetic regulation of papaya ripening. In this way, papaya is a great model for the understanding of ripening regulation and effects, as its ripening process is fast and promotes brutal modifications. Transcriptomics is a robust approach to the global study of gene expression, and the possibility to integrate data of higher molecular-levels high throughput analysis, such as proteomics and metabolomics, might improve significance and specificity on the identification of key ripening regulators.

REFERENCES

AGRAHARI, R. K., SINGH, P., KOYAMA, H., & PANDA, S. K. Plant-microbe Interactions for Sustainable Agriculture in the Post-genomic Era. **Current Genomics**, n. *21, v.* 3, p. 168–178. 2020. DOI:10.2174/1389202921999200505082116

CHEN, S., LI, Y., ZHAO, Y., *et al.* iTRAQ and RNA-Seq analyses revealed the effects of grafting on fruit development and ripening of oriental melon (*Cucumis melo* L. var. makuwa), **Gene**. v. 766, p. 145142, 2021. DOI:0.1016/j.gene.2020.145142.

CONSORTIUM, I. H. G. S. Finishing the euchromatic sequence of the human genome, **Nature**. v. 431, n. 7011, p. 931–945, 2004. DOI:10.1038/nature03001.

CRANDALL, S. G., GOLD, K. M., JIMÉNEZ-GASCO, M. del M., *et al.* A multi-omics approach to solving problems in plant disease ecology, **PLOS ONE**. v. 15, n. 9, p. e0237975, 22 Sep. 2020. DOI:10.1371/journal.pone.0237975.

DER AGOPIAN, R. G., FABI, J. P., CORDENUNSI-LYSENKO, B. R. Metabolome and proteome of ethylene-treated papayas reveal different pathways to volatile compounds biosynthesis, **Food Research International**. v. 131, p. 108975, 2020. DOI:10.1016/j.foodres.2019.108975.

DESHPANDE, A. B., ANAMIKA, K., JHA, V., *et al.* Transcriptional transitions in Alphonso mango (*Mangifera indica* L.) during fruit development and ripening explain its distinct aroma and shelf life characteristics, **Scientific Reports**. v. 7, n. 1, p. 8711, 2017. DOI:10.1038/s41598-017-08499-5.

DO PRADO, S. B. R., FERREIRA, G. F., HARAZONO, Y., *et al.* Ripening-induced chemical modifications of papaya pectin inhibit cancer cell proliferation, **Scientific Reports**. v. 7, n. 1, p. 16564, 2017. DOI:10.1038/s41598-017-16709-3.

DO PRADO, S. B. R., MELFI, P. R., CASTRO-ALVES, V. C., *et al.* Physiological degradation of pectin in papaya cell walls: Release of long chains galacturonans derived from insoluble fractions during postharvest fruit ripening, **Frontiers in Plant Science**. v. 7, n. JULY2016, p. 1–11, 2016. DOI:10.3389/fpls.2016.01120.

FABI, J., FHG, P., GOMEZ, M. L. Papaya, Mango and Guava Fruit Metabolism during Ripening: Postharvest Changes Affecting Tropical Fruit Nutritional Content and Quality, **Fresh Produce**. 1 Mar. 2010.

FABI, J. P., CORDENUNSI, B. R., SEYMOUR, G. B., *et al.* Molecular cloning and characterization of a ripening-induced polygalacturonase related to papaya fruit softening, **Plant Physiology and Biochemistry**. v. 47, n. 11–12, p. 1075–1081, 2009. DOI:10.1016/j.plaphy.2009.08.002.

FABI, J. P., DO PRADO, S. B. R. Fast and furious: Ethylene-triggered changes in the metabolism of papaya fruit during ripening, **Frontiers in Plant Science**. v. 10, n. April, 2019. DOI:10.3389/fpls.2019.00535.

FABI, P., LAJOLO, F. M. Scientia Horticulturae Cloning and characterization of transcripts differentially expressed in the pulp of ripening papaya, **Sci. Hort**. v. 121, p. 159–165, 2009. DOI:10.1016/j.scienta.2009.01.036.

FU, C.-C., CHEN, H.-J., GAO, H.-Y., *et al.* Papaya CpMADS4 and CpNAC3 co-operatively regulate ethylene signal genes CpERF9 and CpEIL5 during fruit ripening, **Postharvest Biology and Technology**. v. 175, p. 111485, 2021. DOI:10.1016/j.postharvbio.2021.111485.

HOLLEY, R. W., APGAR, J., EVERETT, G. A., *et al.* Structure of a Ribonucleic Acid, **Science**. v. 147, n. 3664, p. 1462–1465, 1965. DOI:10.1126/science.147.3664.1462.

IGIEHON, N. O., BABALOLA, O. O. Biofertilizers and sustainable agriculture: exploring arbuscular mycorrhizal fungi, **Applied Microbiology and Biotechnology**. v. 101, n. 12, p. 4871–4881, 2017. DOI:10.1007/s00253-017-8344-z.

KARAHALIL, B. Overview of Systems Biology and Omics Technologies. **Current Medicinal Chemistry**. v. 23, n. 37, p. 4221-4230, 2016. DOI:10.2174/0929867323666160926150617

LI, T., YUN, Z., WU, Q., *et al.* Combination of Transcriptomic, Proteomic, and Metabolomic Analysis Reveals the Ripening Mechanism of Banana Pulp. **Biomolecules**. v. 9, n. 10 , 2019. DOI:10.3390/biom9100523

LIU, K., YUAN, C., FENG, S., *et al.* Genome-wide analysis and characterization of Aux/IAA family genes related to fruit ripening in papaya (*Carica papaya* L.), **BMC Genomics**. v. 18, n. 1, p. 351, 2017. DOI:10.1186/s12864-017-3722-6.

LIU, Y., TANG, M., LIU, M., *et al.* The Molecular Regulation of Ethylene in Fruit Ripening, **Small Methods**. v. 4, n. 1900485, 2020. DOI:10.1002/smtd.201900485.

MANRIQUE, G. D., LAJOLO, F. M. Cell-wall polysaccharide modifications during postharvest ripening of papaya fruit (*Carica papaya*), **Postharvest Biology and Technology**. v. 33, n. 1, p. 11–26, 2004. DOI:10.1016/j.postharvbio.2004.01.007.

MAXAM, A. M., GILBERT, W. A new method for sequencing DNA, **Proceedings of the National Academy of Sciences of the United States of America**. v. 74, n. 2, p. 560–564, Feb. 1977. DOI:10.1073/pnas.74.2.560.

MING, R., HOU, S., FENG, Y., *et al.* The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus), **Nature**. v. 452, n. 7190, p. 991–996, 2008. DOI:10.1038/nature06856.

PEI, M., GU, C., ZHANG, S. Genome-wide identification and expression analysis of genes associated with peach (*Prunus persica*) fruit ripening, **Scientia Horticulturae**. v. 246, p. 317–327, 2019. DOI:10.1016/j.scienta.2018.10.065.

PRADHAN, D., KUMAR, A., SINGH, H., *et al.*, Chapter 4 - High-throughput sequencing. **Data Processing Handbook for Complex Biological Data Sources**. p. 39–52, Academic Press, 2019. DOI:10.1016/B978-0-12-816548-5.00004-6.

RODRÍGUEZ, J. P., RAHMAN, H., THUSHAR, S., *et al.* Healthy and Resilient Cereals and Pseudo-Cereals for Marginal Agriculture: Molecular Advances for Improving Nutrient Bioavailability, **Frontiers in genetics**. v. 11, p. 49, 27 Feb. 2020. DOI:10.3389/fgene.2020.00049.

SÁNCHEZ-SEVILLA, J. F., VALLARINO, J. G., OSORIO, S., *et al.* Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (*Fragaria × ananassa*), **Scientific Reports**. v. 7, n. 1, p. 13737, 2017. DOI:10.1038/s41598-017-14239-6.

SANGER, F., AIR, G. M., BARRELL, B. G., *et al.* Nucleotide sequence of bacteriophage φX174 DNA, **Nature**. v. 265, n. 5596, p. 687–695, 1977. DOI:10.1038/265687a0.

SHEN, Y. H., LU, B. G., FENG, L., *et al.* Isolation of ripening-related genes from ethylene/1-MCP treated papaya through RNA-seq, **BMC Genomics**. v. 18, n. 1, p. 671, 2017. DOI:10.1186/s12864-017-4072-0.

SHIGA, T. M., FABI, J. P., DO NASCIMENTO, J. R. O., *et al.* Changes in Cell Wall Composition Associated to the Softening of Ripening Papaya: Evidence of Extensive Solubilization of Large Molecular Mass Galactouronides, **Journal of Agricultural and Food Chemistry**. v. 57, n. 15, p. 7064–7071, 12 Aug. 2009. DOI:10.1021/jf900073b.

SHINOZAKI, Y., NICOLAS, P., FERNANDEZ-POZO, N., *et al.* High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening, **Nature Communications**. v. 9, n. 1, p. 364, 2018. DOI:10.1038/s41467-017-02782-9.

STARK, R., GRZELAK, M., HADFIELD, J. RNA sequencing: the teenage years, **Nature Reviews Genetics**. v. 20, n. 11, p. 631–656, 2019. DOI:10.1038/s41576-019-0150-2.

TIAN, Z., WANG, J.-W., LI, J., *et al.* Designing future crops: challenges and strategies for sustainable agriculture, **The Plant Journal**. v. 105, n. 5, p. 1165–1178, 1 Mar. 2021. DOI:10.1111/tpj.15107.

TUCKER, G., YIN, X., ZHANG, A., *et al.* Ethylene † and fruit softening, **Food Quality and Safety**. v. 1, n. 4, p. 253–267, December 2017. DOI:10.1093/fqsafe/fyx024.

WATSON, J. D., CRICK, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid, **Nature**. v. 171, n. 4356, p. 737–738, 1953. DOI:10.1038/171737a0.

ZEBERG, H., PÄÄBO, S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals, **Nature**. v. 587, n. 7835, p. 610–612, 2020. DOI:10.1038/s41586-020-2818-3.

ZHOU, D., SHEN, Y., ZHOU, P., *et al.* Papaya CpbHLH1/2 regulate carotenoid biosynthesis-related genes during papaya fruit ripening, **Horticulture Research**. v. 6, n. 1, p. 80, 2019. DOI:10.1038/s41438-019-0162-2.