

# UC Davis

## UC Davis Electronic Theses and Dissertations

### Title

Physiological and Molecular Processes Driving Fruit Quality during Tomato and Pistachio Fruit Ripening

### Permalink

<https://escholarship.org/uc/item/7wm127v7>

### Author

Adaskaveg, Jaclyn Alexandra

### Publication Date

2023

### Supplemental Material

<https://escholarship.org/uc/item/7wm127v7#supplemental>

Peer reviewed|Thesis/dissertation

Physiological and Molecular Processes Driving Fruit Quality during  
Tomato and Pistachio Fruit Ripening

By

JACLYN ALEXANDRA ADASKAVEG  
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Plant Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

---

Barbara Blanco-Ulate, Chair

---

Georgia Drakakaki

---

Cai-Zhong Jiang

Committee in Charge

2023

*For Alex, and his love, support, and patience.*

# CONTENTS

|   |           |
|---|-----------|
| List of Figures . . . . .   | vii       |
| List of Tables . . . . .  | ix        |
| Abstract . . . . .  | x         |
| Acknowledgments . . . . .   | xiii      |
| <br>  |           |
| <b>1 Targeting ripening regulators to develop fruit with high quality and extended shelf life</b>   | <b>1</b>  |
| 1.1 Publication Statement . . . . .   | 1         |
| 1.2 Abstract . . . . .  | 1         |
| 1.3 Introduction . . . . .  | 2         |
| 1.4 Increasing fruit pigmentation for visual appeal . . . . .   | 3         |
| 1.5 Biofortifying fruit to provide a reliable dietary source of nutrients . . . . .   | 5         |
| 1.6 Enhancing flavor to satisfy consumer expectations . . . . .   | 8         |
| 1.7 Modifying fruit texture for longer shelf life . . . . .   | 9         |
| 1.8 Looking ahead: toward the commercial success of bioengineered fruit . . . . .   | 11        |
| <br>  |           |
| <b>2 Single and Double Mutations in Tomato Ripening Transcription Factors Have Distinct Effects on Fruit Development and Quality Traits</b> | <b>23</b> |
| 2.1 Publication Statement . . . . .   | 23        |
| 2.2 Abstract . . . . .  | 23        |
| 2.3 Introduction . . . . .  | 24        |
| 2.4 Materials and Methods . . . . .   | 27        |
| 2.4.1 Plant material . . . . .  | 27        |
| 2.4.2 Mutant Genotyping . . . . .   | 27        |
| 2.4.3 Fruit Trait Phenotyping . . . . .   | 29        |
| 2.4.4 RNA extraction . . . . .  | 30        |
| 2.4.5 cDNA Preparation and RT-qPCR . . . . .  | 30        |



|          |   |           |
|----------|---|-----------|
| 2.4.6    | cDNA Library Preparation, RNA Sequencing, and Sequencing Data Processing . . . . .                                  | 31        |
| 2.4.7    | Differential expression analysis, Functional Annotations, and Enrichment Analysis . . . . .                         | 32        |
| 2.4.8    | Measurement of phytohormones . . . . .  | 32        |
| 2.4.9    | Data availability . . . . .   | 33        |
| 2.5      | Results . . . . .   | 33        |
| 2.5.1    | Ripening Mutants Display Distinct Phenotypes and Transcriptional Profiles Throughout Fruit Development . . . . .    | 33        |
| 2.5.2    | <i>Cnr</i> Fruit Display Transcriptional Differences From Wild Type Before Ripening . . . . .                       | 36        |
| 2.5.3    | Transcriptional Misregulation in the Ripening Mutants Leads to Inhibition or Delay of Molecular Processes . . . . . | 38        |
| 2.5.4    | Ripening Mutants Present Alterations in Hormone Networks . . . . .  | 40        |
| 2.5.5    | Ripening Mutations Influence the Expression Dynamics of <i>CNR</i> , <i>NOR</i> , and <i>RIN</i> in Fruit . . . . . | 46        |
| 2.5.6    | Phenotypic Differences in Double Mutants Reveal Genetic Relationships . . . . .                                     | 46        |
| 2.5.7    | Double Mutant <i>Cnr/nor</i> Shows Gene Expression Unique From Both Parents . . . . .                               | 49        |
| 2.6      | Discussion . . . . .  | 52        |
| 2.6.1    | Delay or Inhibition of Ripening Events Vary in <i>nor</i> and <i>rin</i> . . . . .                                  | 52        |
| 2.6.2    | <i>Cnr</i> Is More Than a Ripening Mutant . . . . .   | 54        |
| 2.6.3    | The <i>Cnr</i> Mutant Produces Ethylene Beyond Basal Levels . . . . .   | 55        |
| 2.6.4    | ABA Biosynthesis and Accumulation Is Affected in <i>nor</i> and <i>rin</i> . . . . .                                | 56        |
| 2.7      | Conclusion . . . . .  | 58        |
| 2.8      | Supplemental Material . . . . .   | 66        |
| <b>3</b> | <b>Defining pistachio fruit ripening and its impacts on quality</b>   | <b>73</b> |
| 3.1      | Abstract . . . . .  | 73        |

|        |   |     |
|--------|---|-----|
| 3.2    | Introduction . . . . .  | 74  |
| 3.3    | Materials and Methods . . . . .   | 76  |
| 3.3.1  | Sample collection and preparation for genome sequencing . . . . .                                       | 76  |
| 3.3.2  | Genome size estimation . . . . .  | 77  |
| 3.3.3  | Genome assembly and chromosome construction . . . . .   | 77  |
| 3.3.4  | Genome annotation . . . . .   | 78  |
| 3.3.5  | Genome synteny analysis . . . . .   | 79  |
| 3.3.6  | Sample collection for physiological data and RNAseq experiments   | 79  |
| 3.3.7  | Physiological measurements . . . . .  | 80  |
| 3.3.8  | Statistical analysis of physiological parameters . . . . .  | 81  |
| 3.3.9  | RNA extraction of fruit tissues . . . . .   | 82  |
| 3.3.10 | cDNA library preparation, RNA sequencing, and sequencing data<br>processing . . . . .                   | 82  |
| 3.3.11 | Functional annotation and enrichments . . . . .   | 83  |
| 3.3.12 | Coexpression Analysis . . . . .   | 83  |
| 3.3.13 | Volatile measurements . . . . .   | 84  |
| 3.3.14 | Fatty acid profiles . . . . .   | 85  |
| 3.4    | Results . . . . .   | 85  |
| 3.4.1  | A complete pistachio genome assembly as a valuable resource for<br>fruit development research . . . . . | 85  |
| 3.4.2  | Pistachio fruits develop in four distinct stages . . . . .  | 90  |
| 3.4.3  | Pistachio developmental stages have distinct patterns of gene ex-<br>pression . . . . .                 | 91  |
| 3.4.4  | Hormone-related gene expression changes in a tissue- and temporal-<br>specific manner . . . . .         | 93  |
| 3.4.5  | Peak in volatiles production preceding ripening . . . . .   | 95  |
| 3.4.6  | Fruit color changes are associated with carotenoid biosynthesis gene<br>expression . . . . .            | 97  |
| 3.4.7  | Expression of pectin degrading enzymes increase during hull softening                                   | 100 |

|       |   |     |
|-------|---|-----|
| 3.4.8 | Shell hardening corresponds to an increase in expression of the lignin biosynthesis pathway . . . . . | 102 |
| 3.4.9 | Kernel fatty acids accumulate at a maximum during ripening . . .                                      | 104 |
| 3.5   | Discussion . . . . .  | 105 |
| 3.5.1 | The most complete pistachio genome to date . . . . .  | 105 |
| 3.5.2 | Pistachios exhibit a unique asynchronous development of fruit and seed . . . . .                      | 108 |
| 3.5.3 | Pistachio fruit exhibit a non-climacteric ripening pattern . . . . .                                  | 110 |
| 3.5.4 | Volatiles produced at the onset of ripening offer insights to fruit-disperser relationships . . . . . | 111 |
| 3.5.5 | Hulls undergo physiological and gene expression changes consistent with fruit ripening . . . . .      | 112 |
| 3.5.6 | Seed maturity coincides with fruit ripening . . . . .   | 114 |
| 3.6   | Conclusions . . . . .   | 115 |
| 3.7   | Supplemental Material . . . . .   | 125 |

## LIST OF FIGURES

|   |     |
|---|-----|
| 1.1 Regulators of fruit ripening. . . . .   | 4   |
| 1.2 Factors driving fruit quality traits and examples of ways to improve them. . . . .  | 6   |
| 1.3 Summary of fruit quality and the tools to improve them. . . . .   | 12  |
| 2.1 Fruit traits of the tomato single ripening mutants. . . . .   | 34  |
| 2.2 Transcriptional misregulation in fruit from the single ripening mutants<br><i>Cnr</i> , <i>nor</i> , and <i>rin</i> . . . . . | 37  |
| 2.3 Plant hormone networks altered in the single ripening mutants <i>Cnr</i> , <i>nor</i> ,<br>and <i>rin</i> . . . . .           | 43  |
| 2.4 Impact of ripening mutations on the CNR, NOR, and RIN transcription<br>factors. . . . .                                       | 47  |
| 2.5 Fruit traits of the homozygous double ripening mutants. . . . .   | 48  |
| 2.6 Comparison of <i>Cnr/nor</i> double mutant gene expression to parents. . . . .  | 51  |
| 2.S1 Color of mature green fruit. . . . .   | 66  |
| 2.S2 Total expression of early development in mutants . . . . .   | 67  |
| 2.S3 Functional enrichments of mutation-related DEGs . . . . .  | 68  |
| 2.S4 Functional enrichments of ripening-related DEGs . . . . .  | 70  |
| 2.S5 Double mutant fruit . . . . .  | 71  |
| 3.1 Chromosome-scale genome assembly of <i>Pistacia vera</i> cv. Kerman. . . . .  | 86  |
| 3.2 Defining pistachio development stages with physiological measurements. . . . .  | 92  |
| 3.3 Coexpressed gene expression module patterns and their correlations to traits. . . . .   | 94  |
| 3.4 Transcriptional induction of hormone pathways. . . . .  | 96  |
| 3.5 Hull volatile biosynthesis during fruit development. . . . .  | 98  |
| 3.6 Expression of pectin degrading enzymes is correlated with hull softening. . . . .   | 101 |
| 3.7 Lignin biosynthesis pathway . . . . .   | 103 |
| 3.8 Kernel fatty acid accumulation. . . . .   | 106 |
| 3.S1 WGCNA modules . . . . .  | 126 |

|  |     |
|--|-----|
| 3.S2 WGCNA modules-trait relationships . . . . . | 127 |
|--|-----|

## LIST OF TABLES

|     |  |    |
|-----|--|----|
| 2.1 | Differential expression of key genes associated with tomato fruit traits in the single ripening mutants . . . . .                        | 41 |
| 3.1 | Statistics of PacBio HiFi sequencing and assembly of Pistacia vera cv. Kerman. . . . .   | 87 |
| 3.2 | The summary of BUSCO (Benchmarking Universal Single-Copy Orthologs) assessment of Kerman primary contig assembly and 15 chromosomes. . . | 88 |
| 3.3 | Summary statistics of assembly repeat content. . . . .   | 89 |
| 3.4 | Statistics of the number of transcripts in different tissue types. . . . .   | 90 |

## ABSTRACT

### **Physiological and Molecular Processes Driving Fruit Quality during Tomato and Pistachio Fruit Ripening**

Fruit ripening is a complex developmental process in which fruit undergo physiological and biochemical changes, transforming fruit into nutritious and flavorful foods. Fruit quality traits, such as flavor and aroma, texture, and color, are essential for product marketability and increase during ripening. The characteristic changes in fruit physiology and biochemical composition during ripening result from extensive transcriptional reprogramming. While some regulatory mechanisms governing ripening have been identified, such as, transcription factors (TFs), plant hormones, and post-transcriptional and epigenetic factors, the genetic factors regulating fruit quality are not fully understood.

The overarching goal of this Ph.D. dissertation was to explain the physiological and genetic events determining fruit quality traits during ripening. First, I assessed the current knowledge of ripening regulators of fruit quality traits across diverse fruit species (Chapter 1). Five main categories emerged—color, flavor, nutrition, texture, and shelf life. Transcription factor regulators can act in a pathway-specific manner or as “master” regulators of multiple pathways. Pathway-specific regulators are promising targets for improving fruit quality because they cause fewer unwanted effects on other traits (Chapter 1). In contrast, manipulations to master regulators impact multiple quality attributes causing pleiotropic effects (Chapter 1, Chapter 2). Plant hormones act with transcription factors and are often non-specific to quality traits (Chapter 1).

I employed multiple approaches to provide a system-wide understanding of ripening regulation and events associated with fruit quality in two economically important crops for California, tomato (*Solanum lycopersicum*) and pistachio (*Pistacia vera*). In both studies (Chapters 2 and 3), fruit physiological parameters were measured in field conditions to provide evidence of physical changes occurring in the fruit through ripening. Biochemical measurements provided further evidence of the specific compounds being altered in each

crop. Then, the information was integrated with genome-wide expression analysis to identify underlying molecular mechanisms for each study.

In Chapter 2, I studied the impact of mutations in ripening-related transcription factors (NOR, RIN, and CNR) on tomato fruit quality through trait phenotyping, transcriptional profiling, and hormone measurements over multiple field seasons. I hypothesized each mutant had distinct defects in fruit quality attributes. I determined that ripening in two mutants (*nor* and *rin*) is not entirely inhibited; instead, some ripening processes are delayed. I also demonstrate that the *Cnr* mutation affects fruit development long before the onset of ripening. We also generated homozygous double mutants of these genotypes for the first time to study the combined effect of the mutations on development and ripening. This study contributed new knowledge regarding the tomato ripening mutants, which have been employed to understand fruit ripening for at least the past two decades. Also, given the importance of both *rin* and *nor* in breeding, quality trait data from these mutants are of high value.

In Chapter 3, I applied the knowledge gained in the model system of tomato to investigate a nut crop, pistachio, in which ripening had not been previously described. With collaborators, we generated critical genomic resources to study pistachio ripening: a complete genome assembly for *P. vera* (c.v. Kerman) and a large transcriptomic study expanding 15 weeks of pistachio growth and development and three different fruit tissues (hull, shell, kernel). I also performed a multiyear multilocation analysis of pistachio phenology, allowing me to model multiple fruit physiological measurements across the growing season. The molecular and physiological data allowed me to determine four distinct stages of pistachio development, with Stage IV corresponding to fruit ripening and kernel maturation. I then hypothesize that transcriptional changes at Stage IV are critical for gaining and establishing nut quality traits. I identified hormones associated with fruit ripening and downstream genetic pathways that can explain the physiological changes observed in the nut. I gave particular attention to genes and pathways related to key quality attributes in pistachio, hull softening, shell hardening, and kernel fat composition. Overall, this work ascertained a complete characterization of pistachio development for



the first time, generated valuable genetic resources for future nut research, and provided novel insights into the genetic programs governing pistachio ripening.

The results of this dissertation improve the understanding of the molecular basis of fruit quality traits. My research provides new knowledge about transcriptional regulation required for tomato fruit quality and the implications for better utilizing ripening mutants in breeding for hybrids with extended shelf-life. The genetic resources created for pistachio development will be a basis for continued research in pistachio and other nut tree crops. Further, identifying the timing and mechanisms of ripening in pistachio involved in hull softening and color changes will help develop management strategies for harvest time.

## ACKNOWLEDGMENTS

First and foremost, I would like to thank Barbara Blanco-Ulate for her guidance, mentorship, and support. I thank her for believing in me and helping me believe in myself. I greatly appreciate her exceptional dedication to her students and her expertise, creativity, and intellect that shaped my project for the best. She has helped me grow as a scientist and as a person and has fostered a highly collaborative and caring lab environment that has made an inevitably stressful time enjoyable.

I would like to thank Ann Powell, whose lab I joined during my undergraduate degree, for supporting me and encouraging me to go to graduate school, and for catalyzing my fascination with ripening mutant tomatoes. Thank you to my committee members Georgia Drakakaki and Cai-Zhong Jiang for their advice and support. Thank you to the collaborators in the pistachio project including Giulia Marino, Grey Monroe, Chaehee Lee, and Selina Wang for all their support.

I thank each current and former Blanco Lab member for their efforts to help me in times of need no matter how many tomato or pistachio measurements I had to complete (often too many). My lab mates have provided support, joy, and laughter on the hardest days and have made my graduate school experience all the better. Thank you to my friend and lab mate Christian Silva, for our many conversations (scientific and not) that created many fond memories over the years and for always being willing to help me with bioinformatics. Thank you to Isabel Ortega Salazar, for your many years of friendship and encouragement in and out of the lab. Thank you to my lab mate Saskia Mesquida Pesci for her friendship and kindness and always being willing to help no matter the task. Thank you to Stefan Petrasch for always being available to talk about plants and genetics. Thank you to the Mann Lab Building friends for lending a hand when needed. And thank you to my cohort and the Plant Biology Graduate Group for their support.

I would also like to acknowledge all the undergraduate interns and visiting scholars who have been in the Blanco Lab over the past 6 years that have been instrumental to my research project. It was a joy to mentor my undergraduate assistants over the years. Yiduo Wei provided incredible support in the bioinformatics and analysis of pistachio

genetics. I also thank her for her patience and listening ear when analyzing and writing the Pistachio chapter in my final year. A big thank you to Peng Huang, Aaron Lee, Sarah Lee who helped with many tomato mutant projects. And thank you to Saskia Mesquida Pesci, Fangyi Wang, and Caio Cattai for their tireless efforts collecting data for the pistachio project.

A sincere thank you to my family and friends who have cheered me on from near and far. I especially want to thank my parents for their love, encouragement, and care. My father, Jim Adaskaveg, gave me my first research experience in his lab and shared his deep knowledge of plants with me and my mother Peggy Adaskaveg has been my biggest cheerleader and has shared her passion for plant research with me. Thank you to my husband Alex Cisneros, who has been by my side every step of this journey and continues to lift me up and encourage me to chase my dreams. Thank you to my brother Adam Adaskaveg for his listening ear and for making me laugh in times of stress. Thank you to Zully and Steve Crossway and family for their love and support on my journey. And finally thank you to my dog, Zuki who sat with me for many hours as I wrote this dissertation.

# Chapter 1

## Targeting ripening regulators to develop fruit with high quality and extended shelf life

### 1.1 Publication Statement

The content of this chapter was peer-reviewed and published as follows:

Adaskaveg, J.A. and Blanco-Ulate, B. (2023). Targeting ripening regulators to develop fruit with high quality and extended shelf life. *Current Opinions in Biotechnology*, 79:102872.

### 1.2 Abstract

Fruit quality directly impacts fruit marketability and consumer acceptance. Breeders have focused on fruit quality traits to extend shelf life, primarily through fruit texture, but, in some cases, have neglected other qualities such as flavor and nutrition. In recent years, integrative biotechnology and consumer-minded approaches have surfaced, aiding in the development of flavorful, long-lasting fruit. Here, we discussed how specific transcription factors and hormones involved in fruit ripening can be targeted to generate high-quality fruit through traditional breeding and bioengineering. We highlight regulators that can be used to generate novel-colored fruit or biofortify fresh produce with health-promoting nutrients, such as vitamin C. Overall, we argue that addressing grower and industry needs must be balanced with consumer-based traits.

## 1.3 Introduction

Fruit are rich in essential nutrients, yet most people do not consume the recommended amount of fresh produce to sustain healthy diets and reduce disease risks (US Department of Agriculture (USDA) ERS — Food Availability and Consumption, URL: <https://www.ers.usda.gov>). To promote consumption, expand access, and reduce waste, fruit quality and shelf life need to be increased through breeding and biotechnology, alongside adequate harvest practices, transportation logistics, and postharvest treatments. Commercial fruit crops are primarily bred for high yield and extended shelf life to meet the expectations for mass production and global markets, however, recently, there has been a shift of focus toward developing new crop varieties that meet consumer demands for better flavor and nutrition. Balancing shelf life with consumer-based quality traits is perhaps the biggest challenge breeders and researchers face in the quest for better quality fruit, mainly because these attributes appear to have negative genetic correlations in many crops (Farcuh et al., 2020; Petrasch et al., 2022).

Quality (i.e. color, nutritional value, flavor, and texture) peaks when fruit reach their optimum ripeness. Many studies on fruit ripening of various plant species have emerged in the past decade, helping to identify genetic pathways and molecular regulators that can be manipulated for crop improvement (Figure 1.1). Moreover, biotechnology advances have provided access to high-quality genomic resources and tools, supporting breeding strategies, genetic modification, and gene editing in traditional and nontraditional fruit crops. Here, we review current knowledge of the genetics of fruit traits and argue that manipulating transcription factors (TFs) is a promising approach to enhance fruit quality. We discuss how pleiotropic effects could potentially be avoided by targeting TFs that exclusively regulate specific pathways (i.e. function-specific regulators) instead of master regulators. However, ripening master regulators may remain useful if their effects on gene expression can be fine-tuned (e.g. creating alleles with different impact on gene expression). Similarly, the timing and coordination of regulators need to be considered (e.g. using tissue- and stage-specific promoters) to achieve desired effects on fruit traits (Smirnova and Kochetov, 2020). Finally, we consider the current climate surrounding

consumer acceptance of genetically modified (GM) and gene-edited fruit.

## 1.4 Increasing fruit pigmentation for visual appeal

Fruit color is a significant indicator of ripeness and nutritional quality. It is often used as a harvest maturity index and can predict shelf-life potential in certain crops (Nordey et al., 2019). Fruit color is dependent on pigments that accumulate in the exocarp (skin) and mesocarp (flesh). As fruit ripen, chlorophyll (green) is replaced by carotenoids (yellow, orange, and red), anthocyanins (orange, red, and blue), or betalains (yellow, violet) (Figure 1.2). Fruit gloss, influenced by the structure and composition of the cuticle layer, also affects color perception (Lara et al., 2019).

There have been multiple efforts to enhance pigmentation in fruit crops by modulating genes in pigment biosynthesis pathways (Meng et al., 2022; Zheng et al., 2020). Pink-flesh pineapples have been engineered through the introduction of a phytoene synthase gene (*CrPSY*) from tangerine in combination with suppression via RNAi of two endogenous lycopene genes (*AcLYC-B* and *AcLYC-E*) to inhibit the conversion of lycopene into other carotenoids (Firoozabady et al., 2013). This pineapple, known as Pinkglow™ is one of the few bioengineered fruits available in the United States. Yet, targeting or introducing single genes involved in pigment metabolism can be meticulous or have unintended consequences on other metabolites. For example, the overexpression of *SlLYC-B* or *SlZDS* in tomato fruit increased carotenoid content but also increased the production of abscisic acid (ABA), a derivative of this pathway, which affected the fruit's normal ripening progression (Diretto et al., 2020; McQuinn et al., 2020).

A more effective means to manipulate pigment content in fruit might be dissecting function-specific regulators, particularly those that directly control the biosynthesis pathway and act downstream of ripening regulatory networks. This is the case of R2R3-MYB TFs involved in regulating anthocyanin biosynthesis as first demonstrated in apple and grape (Azuma et al., 2008; Espley et al., 2007; Lin-Wang et al., 2010). Wild and cultivated tomatoes are red because of a mutation in the R2R3-MYB TF SlAN2-like that suppresses anthocyanin biosynthesis. However, overexpressing the functional gene in fruit turned on

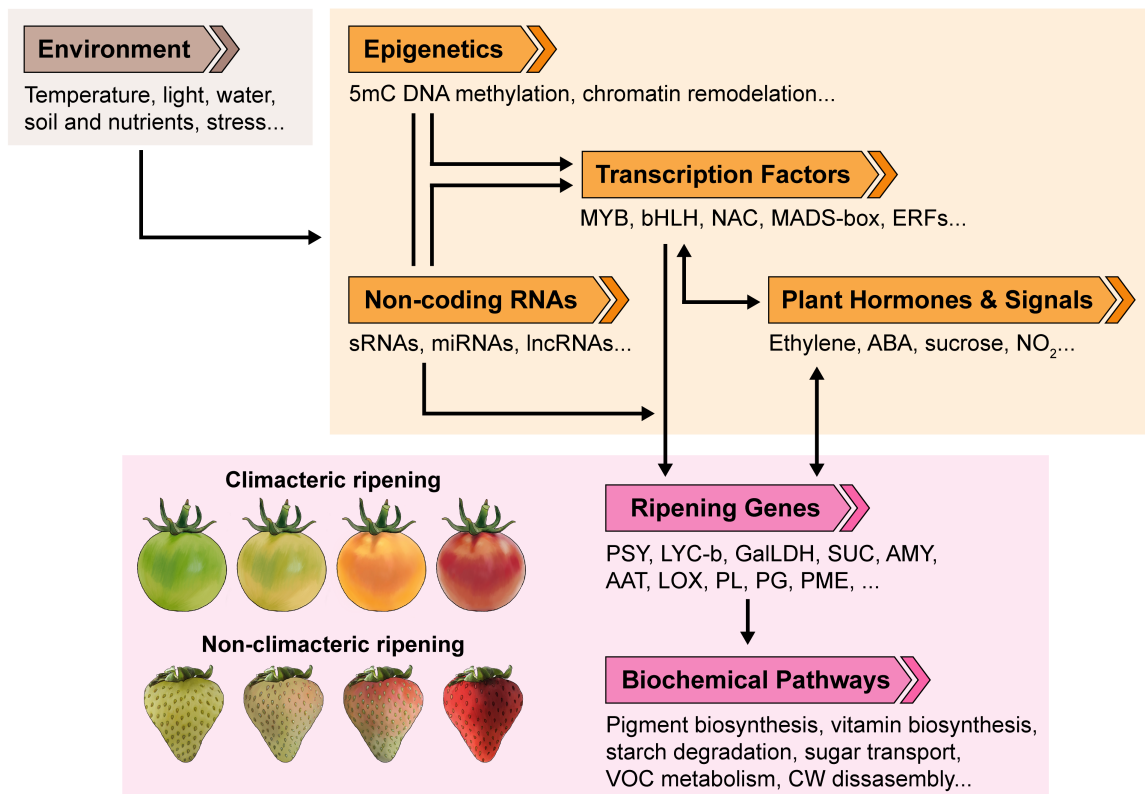


Figure 1.1: Regulators of fruit ripening. Fruit ripening is a complex developmental program that determines the quality and shelf life of fresh produce (Seymour and Granell, 2014). The activation and progression of ripening processes (pink box) are tightly controlled by multiple regulatory levels (orange box), which in turn are influenced by environmental stimuli (brown box). Specific ripening regulators, such as plant hormones and signals (e.g. ethylene, ABA, sucrose, or  $\text{NO}_2$ ), may have different roles depending on the fruit crop and whether the fruit is climacteric (e.g. tomato) or nonclimacteric (e.g. strawberry) (Jia et al., 2016; Liu et al., 2020). TFs are central in the control of ripening and generally act upstream of plant hormones (e.g. MYB; bHLH; NAM, ATAF and CUC, NAC; MCM1, AGAMOUS, DEFICIENS, SRF, MADS; ethylene response factors, ERF). Ripening regulatory mechanisms involving noncoding RNAs (small RNA, sRNA; microRNA, miRNA; long noncoding RNA, lncRNA) and epigenetic factors appear to be conserved across multiple plant species (Li et al., 2022b). Ripening genes generally encode enzymes in biochemical pathways that influence fruit quality traits.

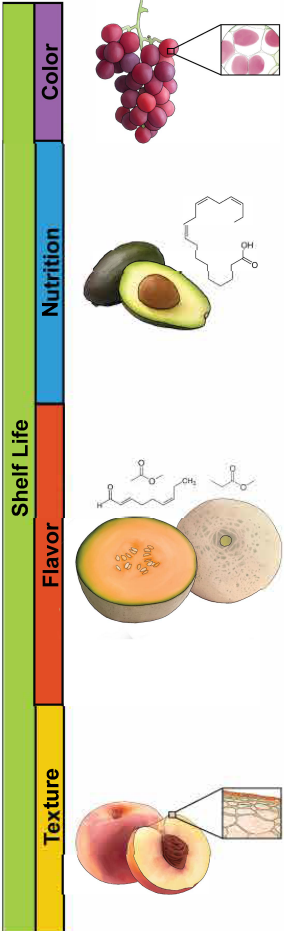
Figure 1.1: (Continued from previous page) For example, pigment biosynthesis genes (e.g. phytoene synthase, *PSY*; lycopene beta-cyclase, *LYC-b*), genes involved in the production of vitamins and antioxidants (e.g. vitamin-C biosynthesis gene L-galactono-1,4-lactone dehydrogenase, *GalLDH*), genes encoding CWDEs (e.g. *PL*, *PG*, *PME*), genes involved in starch degradation (e.g. amylase, *AMY*) and sugar transport (e.g. sucrose transport protein, *SUC*), and genes related with the production of aromatic VOC (e.g. alcohol acetyl transferase, *AAT*; lipoxygenase, *LOX*).

the pathway and produced anthocyanin-rich purple-fleshed tomatoes (Sun et al., 2020). Purple tomatoes with high anthocyanin content have also been generated by introducing two TFs from snapdragon, the R2R3-MYB TF *AmRos1* and the basic helix loop helix (bHLH) TF *AmDel* under a fruit-specific promoter (Butelli et al., 2008). R2R3-MYBs also appear to be critical for the control of betalain biosynthesis, uncovered through the recent assembly of the pitaya genome (ye Chen et al., 2021). A pitaya R2R3-MYB TF (*HuMYB1*) involved in regulation was recently identified and followed a very similar activation mechanism as with anthocyanins (Xie et al., 2021). Because R2R3-MYB TFs exist in diverse plant species, this class of TFs could be easily edited with clustered regularly interspaced short palindromic repeat (CRISPR) technologies to engineer new fruit colors, for example, novel ‘white’ fruit such as the wild strawberry species *Fragaria nilgerrensis*, carrying a mutation in the *FvMYB10* TF (Zhang et al., 2020b). Pigmentation can also be controlled at higher-level regulatory points, but this usually comes with consequences to other quality traits. For instance, exogenous applications of the hormone ABA induce anthocyanin biosynthesis in table grapes, positively influencing color, but negatively affected berry texture (Koyama et al., 2018).

## 1.5 Biofortifying fruit to provide a reliable dietary source of nutrients

Fruit are used as food source for macronutrients (e.g. fiber, sugars, and lipids) and micronutrients (e.g. vitamins, minerals, and antioxidants) (Figure 1.2). Macronutrients are essential to provide energy and maintain the body’s structure and functions. Vitamins are





|                       | <b>Attributes</b>   | <b>Metabolites &amp; Structures</b>  | <b>Examples of Loci</b>  |  |   |
|-----------------------|---|--|--|--|---|
|                       |   |  | <b>Gene</b>  | <b>Fruit Crop</b>  | <b>Reference</b>  |
| <b>Color</b>          | <b>Pigmentation</b>   | Chlorophyll, carotenoids, anthocyanins, betalains  | <i>LYC-B</i><br><i>SIAN2-like</i><br><i>HuMYB1</i>                   | Pineapple<br>Tomato<br>Pitaya                            | Firoozabady 2013<br>Sun 2020<br>Xie 2021  |
|                       | <b>Gloss</b>  | Cuticle (lipids, fatty acids, cutin)   | <i>CgRuby</i>  | Blood orange   | Huang 2019  |
| <b>Nutrition</b>      | <b>Macronutrients</b>   | Carbohydrates (starch)<br>Fiber (cellulose)<br>Fats (fatty acids)<br>Protein   | <i>MabHLH6</i>   | Banana   | Xiao 2018   |
|                       | <b>Micronutrients</b>   | Antioxidants (i.e., carotenoids, phenolics, monoterpenoids, Vitamin C, Vitamin E)<br>Vitamins (i.e., Vitamin D, Vitamin B) | <i>S17-DR2</i><br><i>AcGGP</i><br><i>SINFYA10</i><br><i>SibHLH59</i> | Tomato<br>Kiwifruit<br>Tomato<br>Tomato                  | Li 2022a<br>Laing 2015<br>Ye 2019a<br>Ye 2019b                                  |
| <b>Flavor</b>         | <b>Taste</b>  |  |  |  |   |
|                       | Sweet   | Soluble sugars (i.e., sucrose, fructose, glucose)  | <i>FvebZIPs1.1</i><br><i>SITomLoxC</i>                               | Strawberry<br>Tomato                                     | Xing 2020<br>Tieman 2017  |
|                       | Sour  | Organic acids (i.e., citric acid, malic acid)  | <i>Ma10</i>  | Apple  | Ma 2019   |
|                       | Bitter  | Alkaloids, phenolics, terpenoids   |  |  |   |
|                       | Umami   | Amino acids, ribonucleotides   |  |  |   |
| <b>Aroma</b>          | Volatile organic compounds (i.e., esters, aldehydes, ketones) | <i>AcNAC4</i><br><i>PpNAC1</i><br><i>MdNAC5</i>  | Kiwifruit<br>Peach<br>Apple  | Wang 2022;<br>Nieuwenhuizen 2015<br>Cao 2021<br>Cao 2021 |   |
| <b>Texture</b>        | <b>Firmness</b>   |  |  |  |   |
|                       | Cell wall disassembly   | CW polymers (cellulose, hemicelluloses, pectins)   | <i>SINOR</i><br><i>SILOB1</i><br><i>SIPL</i>                         | Tomato<br>Tomato<br>Tomato                               | Adaskaveg 2021; Kumar 2018;<br>Wang 2019b; Rose 2009<br>Shi 2021;<br>Wang 2019a |
|                       | Water loss  | Cuticle (lipids, fatty acids, cutin)   | <i>FvPL</i><br><i>FvPME</i>  | Strawberry<br>Strawberry                                 | Xue 2020<br>Bruening 2000   |
| <b>Juiciness</b>      |   |  |  |  |   |
| Cell-to-cell adhesion | Middle lamella pectins  | <i>PpCYP82A3</i>   | Peach  | Lurie 2021; Ma 2019.                                     |   |
| Turgor pressure       | Water, solutes  |  |  |  |   |

Figure 1.2: Factors driving fruit quality traits and examples of ways to improve them. Fruit color, nutrition, flavor, and texture determine the overall quality and shelf life of fruit. This table breaks up these categories of fruit quality (left colored blocks) into specific attributes. For example, color perception is comprised of pigmentation and gloss of a fruit. The attributes are further defined by characteristic structures or metabolites. Examples of loci (genes and QTLs) that can act as potential breeding or bioengineering targets are included. These targets are listed with their associated fruit crop and reference.

Figure 1.2: (Continued from previous page) The orange font represents genes that are TF regulators, while the pink font represents ripening genes part of biochemical pathways. (Firoozabady et al., 2013), (Sun et al., 2020), (Xie et al., 2021), (Huang et al., 2019), (Xiao et al., 2018), (Li et al., 2022a), (Laing et al., 2015), (Ye et al., 2019a), (Ye et al., 2019b), (Xing et al., 2020), (Tieman et al., 2017), (Ma et al., 2019), (Wang et al., 2022), (Nieuwenhuizen et al., 2015), (Cao et al., 2021), (Adaskaveg et al., 2021), (Kumar et al., 2018), (Wang et al., 2019b), (Rose and Isaacson, 2009), (Shi et al., 2021), (Wang et al., 2019a), (Xue et al., 2020), (Bruening et al., 2000), (Lurie, 2021),

required for various bodily functions and can only be found in food sources such as fruit. Antioxidants inhibit cell damage caused by oxidative agents. Carotenoids (pro-vitamin A), and phenolics (anthocyanins and betalains), reviewed above, are all antioxidants, as well as vitamin C (ascorbic acid).

Vitamin-D deficiency is a global health problem due to few dietary sources of this vitamin. Biofortification of vitamin D in tomatoes has recently become possible by engineering its biosynthesis from a pre-existing pathway (Li et al., 2022a). Owing to partial duplication of the pathway, a single enzyme could be knocked out with CRISPR-Cas9 to convert the precursor into vitamin D without an expense to other metabolites. This discovery has further implications for other *Solanaceae* plants.

Ascorbic acid, an important antioxidant and nutrient for immune health and wound healing, has proven to be less easily biofortified into fruit because increasing biosynthesis also leads to activation of catabolic and recycling pathways (Mellidou et al., 2021). Post-transcriptional regulation of an ascorbic acid biosynthesis enzyme from kiwifruit has been demonstrated using tobacco leaves (Laing et al., 2015). Removing the upstream open-reading frame (uORF) that repressed translation increased ascorbic acid concentration in the leaves. Function-specific TFs can increase ascorbic acid in fruit without negative impact on quality as demonstrated in tomato. *SlHZ24*, a bHLH TF, regulates ascorbic acid biosynthesis and catabolism genes, and its transient overexpression has been reported to increase the accumulation of this vitamin (Hu et al., 2016). Other TFs, *SlNL33* and *SlNFYA10*, have been found to regulate the pathway negatively, and silencing them also

increased ascorbic acid level (Chen et al., 2020; Ye et al., 2019b). Understanding regulatory pathways governing nutrients can also facilitate traditional breeding programs. For example, a genome-wide association study (GWAS) led to the discovery and validation that alleles in *SibHLLH59* determine ascorbic acid content in tomato cultivars (Ye et al., 2019a).

A less specific effect on nutrient accumulation can be achieved with hormone applications. For example, ethylene application in kiwifruit, a climacteric fruit, increased ascorbic acid and other antioxidants during ripening (Choi et al., 2022). Ascorbic acid also increased after nitric oxide (NO<sub>2</sub>) application in sweet pepper, a nonclimacteric fruit (Rodríguez-Ruiz et al., 2017).

## 1.6 Enhancing flavor to satisfy consumer expectations

Fruit flavor is a critical quality trait for consumer acceptance. Flavor includes all sensations experienced when eating, consisting of taste, aroma, and texture (Figure 1.2). In fruit, taste is mainly defined by a balance between sweetness and acidity but can include bitterness and umami. Fruit aroma comes from specific classes and combinations of volatile organic compounds (VOCs). For example, the unique kiwifruit flavor is associated with esters, mainly ethyl butanoate and methyl butanoate (Wang et al., 2022).

Modulation of function-specific transcriptional and post-transcriptional regulators offers an effective solution for flavor improvement. A bHLH TF in banana (*MabHLLH6*) activates 11 starch-degrading genes expressed during fruit ripening and is a likely candidate for increasing sugar and sweetness (Xiao et al., 2018). In strawberry, editing the uORF of a bZIP TF that controls sucrose biosynthesis led to its translational activation and higher sugar content in fruit (Xing et al., 2020).

Combining datasets generated through genomic, transcriptomic, metabolomic, and consumer panel studies has proven to be an effective strategy for identifying flavor-related genes and TFs. In tomato, metabolite data, associated loci, and consumer panels were analyzed across hundreds of varieties to determine key genes contributing to flavor (Tie-man et al., 2017). Coupling metabolites relevant to flavor with transcriptomic analyses

can produce gene networks and identify regulatory TFs involved in flavor pathways. This approach was used in kiwifruit, where the *AcNAC4* TF regulating a key gene in ester biosynthesis (*AcAAT10*) was validated (Nieuwenhuizen et al., 2015; Wang et al., 2022). Homologs of these NAC TFs have been implicated in ester formation in peach (*PpNAC1*) and apple (*MdNAC5*), suggesting that their functions are conserved across diverse families of climacteric fruit (Cao et al., 2021).

Flavor is impacted by postharvest handling. Chilling during transportation and storage can alter fruit flavor. Epigenetic factors (i.e. DNA methylation) have been shown to regulate the suppression of VOC biosynthesis in chilled tomatoes (Zhang et al., 2016). Most climacteric fruit are picked unripe (but mature) and later treated with ethylene to induce ripening. However, this practice has been associated with poor flavor development. A study in off-vine ripened tomatoes confirmed that the fruit presented reduced VOC emission and a low sugar/acid ratio, due to alterations in gene expression and decreased methylation of their promoters (Zhang et al., 2020a).

Genomic resources can help anticipate consumer preference and assist in breeding to select fruit with enhanced flavor. A population genomic study revealed that distinct consumer preferences between eastern and western countries drove selection for peach cultivars with different acidity (Yu et al., 2021). Fruit VOC profile data can help predict consumer liking before performing sensory panels, allowing for a more efficient selection of high-flavor fruit, as seen in tomato and blueberry (Colantonio et al., 2022). Moreover, the tomato pan-genome helped identify a rare favorable allele selected against during domestication that could be incorporated back into new cultivars to improve fruit flavor (Gao et al., 2019).

## 1.7 Modifying fruit texture for longer shelf life

Fruit texture is associated with freshness, flavor, and shelf-life potential. Texture involves many attributes, such as firmness, juiciness, crispiness, and meltiness (Figure 1.2). Loss of firmness (i.e. fruit softening) is a hallmark of fruit ripening and negatively correlates with shelf life. Fruit softening is mainly attributed to the remodeling and degradation of

the polysaccharides in the primary cell walls (CWs). The cuticle layer, deposited on the CWs of epidermal tissue, also contributes to fruit firmness by preventing water loss and maintaining cell turgor pressure (Lara et al., 2019).

Traditional breeding has focused on creating firmer fruit that withstand transportation and have longer shelf life. This has been accomplished in tomato by developing hybrid lines between elite varieties and nonripening mutants such as *ripening-inhibitor* (*rin*) and *non-ripening* (*nor*) (Osei et al., 2017). These mutants have defects in TFs considered master regulators of many ripening processes, including the induction of cell wall-degrading enzymes (CWDEs) and changes in cuticle composition (Adaskaveg et al., 2021; Kosma et al., 2010). Other efforts to improve firmness and shelf life in climacteric fruit have taken advantage of mutations affecting ethylene biosynthesis and perception, as this hormone is also known to regulate genes encoding CWDEs, among others (Huang et al., 2022; Tucker et al., 2017). However, as already discussed, modulating master regulators or hormone pathways has numerous drawbacks to other quality traits, such as color and flavor. Thus, downstream TFs controlling specific CW enzymes, such as LOB TF (*SILOB1*) in tomato, may prove to be better targets (Shi et al., 2021). Another possibility is to leverage the availability of natural or induced allelic variants in ripening master regulators to produce a range of fruit phenotypes. For example, spontaneous (i.e. *alcobaca* (*alc*) or equivalent *delayed fruit deterioration* (*dfd*)) and CRISPR–Cas9-generated mutations in the ripening regulator NOR can extend shelf life with minimal impact on other fruit attributes (Kumar et al., 2018; Rose and Isaacson, 2009; Wang et al., 2019b), compared with the canonical mutant *nor* (a gain-of-function mutation) (Adaskaveg et al., 2021; Wang et al., 2020).

RNAi and gene editing approaches have been used to target ripening-specific CWDEs that influence fruit firmness, such as polygalacturonase (PG), pectate lyase (PL), and pectin methyl esterase (PME) (Wang et al., 2018). The first GM fruit product, the FLAVR SAVR™ tomato, was engineered with antisense RNA against SIPG, however, it did not show a phenotype for fruit firmness (Kramer and Redenbaugh, 1994). In contrast, the CRISPR–Cas9 *SIPL* knockout in tomato significantly improved fruit firmness and shelf life (Wang et al., 2019a). Similar observations were previously reported in strawberry

using RNAi knockdowns of *FvPL* and *FvPME* (Xue et al., 2020; Yang et al., 2017), 61]. In addition to firmer fruit, tomato and strawberry *SIPL* mutants have reduced fruit susceptibility to fungal disease (Silva et al., 2021; Yang et al., 2017).

Breeding strategies and molecular studies have also focused on ameliorating textural defects such as fruit mealiness caused by cold storage. Mealiness, considered the opposite of juiciness, occurs when neighboring cells lose adhesion and detach while remaining intact (Li et al., 2020). In peach, quantitative trait loci (QTLs) associated with cold-tolerant varieties have been identified to support breeding for less mealy fruit (Lurie, 2021). Beyond QTLs, understanding the genetic mechanisms behind the trait provides avenues for targeting breeding and genetic modifications. For instance, peach mealiness appears to be associated with increased DNA methylation, leading to the downregulation and hypermethylation of mealy-associated genes such as *PpCYP82A3* (Rothkegel et al., 2021). Finding molecular approaches to avoid the deposition of methyl groups in the promoters of key ripening genes in response to cold storage can serve as a potential solution to mealiness.

## 1.8 Looking ahead: toward the commercial success of bioengineered fruit

The FLAVR SAVR™ tomato hit the market in the early 90s promising a product with longer shelf life. However, this GM fruit had high production costs and was not well accepted by consumers, which led to its removal from the marketplace (Bruening et al., 2000). Since then, other bioengineered crops with improved plant disease resistance or production-related traits have become available worldwide (Lobato-Gómez et al., 2021). These products have not sparked much enthusiasm mainly because they were not generated considering consumer-based traits or due to public fear of GM organisms.

Recently, two fruits bioengineered for quality attributes were approved by both US and Canadian regulatory entities and are available to consumers. These are the Pinkglow™ pineapple and the Arctic™ apple. The latter was bioengineering to reduce oxidative browning in the cut fruit (Lobato-Gómez et al., 2021). These fruit are considered novelty

items because they are less available in the marketplace and significantly more expensive than traditional cultivars. Both fruits are primarily sold online and offered in limited supplies. The Pinkglow™ pineapple costs nearly ten times more than a common yellow pineapple. Despite their limitations, these fruits were developed with the consumer in mind, which may entice the public more than previous GM products.

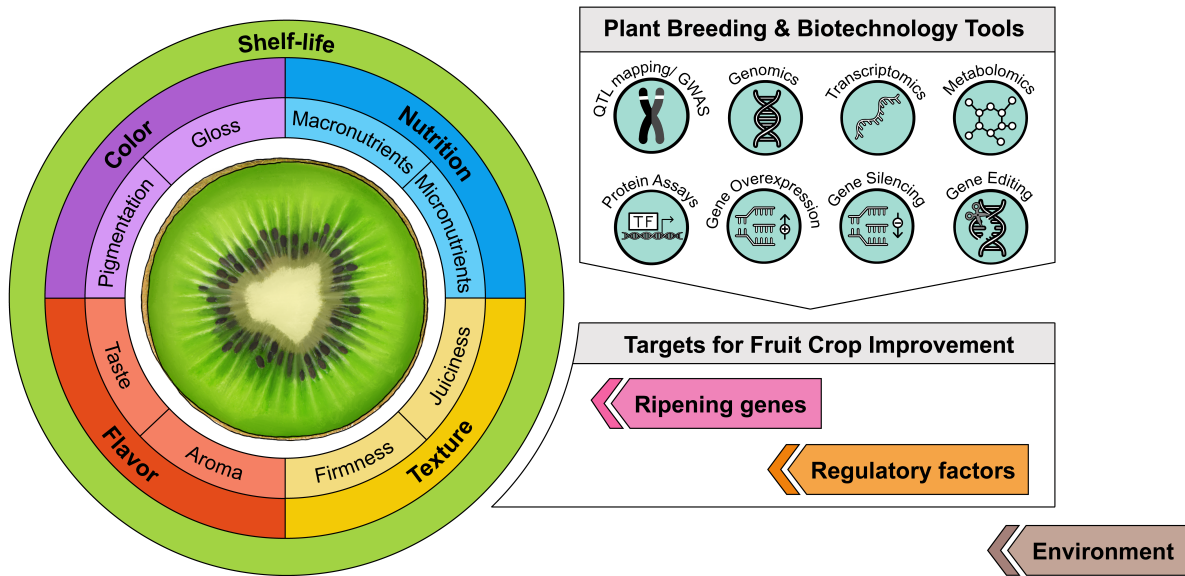


Figure 1.3: Summary of fruit quality and the tools to improve them. Fruit color, nutrition, flavor, and texture determine the overall quality and shelf life of fruit. Examples of specific attributes of each quality trait are in the inner circle. Targets for improving quality discussed in the review are represented by the right arrows. Finally, the methods that can be used as tools to study and improve these traits are represented above in blue circles.

In 2022, the USDA deregulated the purple tomato developed with the expression of two snapdragon TFs (discussed above, (Butelli et al., 2008)). Novel fruit colors and potential higher nutrition may draw consumers to a new emerging category of bioengineered produce. These fruit will also need to meet high consumer expectations of flavor, affordability, and food safety to ensure their success. Gene editing techniques such as CRISPR–Cas9 enable the fine-tuning of quality traits in a variety of fruit crops and may be more well received than previous GM products. For example, Japan started selling

the first Cas9-edited fruit in the world in 2021, a health-promoting  $\gamma$ -aminobutyric acid (GABA)-enriched tomato (Waltz, 2022). Gene-edited fruit without foreign DNA have a more straightforward regulatory path in the United States than GM products, increasing the speed to market and reducing costs associated with authorization. However, this is not the case in other parts of the world, such as the European Union.

In conclusion, researchers are armed with knowledge on fruit ripening and tools to improve fruit quality and generate greater access to fresh, flavorful, and nutritious food. Demands for better-tasting, more sustainable fruit, are in reach (Figure 1.3). We sit on the edge of an era where gene-edited and bioengineered commodities can become a new category in the market, if the consumer will allow it.

## References

- Adaskaveg, J. A., Silva, C. J., Huang, P., and Blanco-Ulate, B. (2021). Single and Double Mutations in Tomato Ripening Transcription Factors Have Distinct Effects on Fruit Development and Quality Traits. *Frontiers in Plant Science*, 12:672.
- Azuma, A., Kobayashi, S., Mitani, N., Shiraishi, M., Yamada, M., Ueno, T., Kono, A., Yakushiji, H., and Koshita, Y. (2008). Genomic and genetic analysis of Myb-related genes that regulate anthocyanin biosynthesis in grape berry skin. *Theoretical and Applied Genetics*, 117(6):1009–1019.
- Bruening, G., Lyons, J., et al. (2000). The case of the flavr savr tomato. *California Agriculture*, 54(4):6–7.
- Butelli, E., Titta, L., Giorgio, M., Mock, H. P., Matros, A., Peterek, S., Schijlen, E. G., Hall, R. D., Bovy, A. G., Luo, J., and Martin, C. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology*, 26(11):1301–1308.
- Cao, X., Wei, C., Duan, W., Gao, Y., Kuang, J., Liu, M., Chen, K., Klee, H., and Zhang, B. (2021). Transcriptional and epigenetic analysis reveals that NAC transcription factors regulate fruit flavor ester biosynthesis. *Plant Journal*, 106(3):785–800.



- Chen, W., Hu, T., Ye, J., Wang, B., Liu, G., Wang, Y., Yuan, L., Li, J., Li, F., Ye, Z., and Zhang, Y. (2020). A CCAAT-binding factor, SINFYA10, negatively regulates ascorbate accumulation by modulating the d-mannose/l-galactose pathway in tomato. *Horticulture Research*, 7(1).
- Choi, H. R., Baek, M. W., Cheol, L. H., Jeong, C. S., and Tilahun, S. (2022). Changes in metabolites and antioxidant activities of green ‘Hayward’ and gold ‘Haegeum’ kiwifruits during ripening with ethylene treatment. *Food Chemistry*, 384:132490.
- Colantonio, V., Ferrão, L. F. V., Tieman, D. M., Bliznyuk, N., Sims, C., Klee, H. J., Munoz, P., and Resende, M. F. (2022). Metabolomic selection for enhanced fruit flavor. *Proceedings of the National Academy of Sciences of the United States of America*, 119(7).
- Diretto, G., Frusciante, S., Fabbri, C., Schauer, N., Busta, L., Wang, Z., Matas, A. J., Fiore, A., K.C. Rose, J., Fernie, A. R., Jetter, R., Mattei, B., Giovannoni, J., and Giuliano, G. (2020). Manipulation of  $\beta$ -carotene levels in tomato fruits results in increased ABA content and extended shelf life. *Plant Biotechnology Journal*, 18(5):1185–1199.
- Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., and Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant Journal*, 49(3):414–427.
- Farcuh, M., Copes, B., Le-Navenec, G., Marroquin, J., Cantu, D., Bradford, K. J., Guinard, J. X., and Van Deynze, A. (2020). Sensory, physicochemical and volatile compound analysis of short and long shelf-life melon (*Cucumis melo* L.) genotypes at harvest and after postharvest storage. *Food Chemistry: X*, 8:100107.
- Firoozabady, E., Young, T. R., and Del Monte Fresh Produce Company (2013). "Pineapple plant named Rosé (EF2-114)".
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-Chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez,

- M. J., Blanca, J., Canizares, J., Xu, Y., van der Knaap, E., Huang, S., Klee, H. J., Giovannoni, J. J., and Fei, Z. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics*, 51(6):1044–1051.
- Hu, T., Ye, J., Tao, P., Li, H., Zhang, J., Zhang, Y., and Ye, Z. (2016). The tomato HD-Zip i transcription factor SlHZ24 modulates ascorbate accumulation through positive regulation of the d -mannose/ l -galactose pathway. *Plant Journal*, 85(1):16–29.
- Huang, D., Yuan, Y., Tang, Z., Huang, Y., Kang, C., Deng, X., and Xu, Q. (2019). Retrotransposon promoter of Ruby1 controls both light- and cold-induced accumulation of anthocyanins in blood orange. *Plant Cell and Environment*, 42(11):3092–3104.
- Huang, W., Hu, N., Xiao, Z., Qiu, Y., Yang, Y., Yang, J., Mao, X., Wang, Y., Li, Z., and Guo, H. (2022). A molecular framework of ethylene-mediated fruit growth and ripening processes in tomato. *The Plant Cell*, 34(9):3280–3300.
- Jia, H., Jiu, S., Zhang, C., Wang, C., Tariq, P., Liu, Z., Wang, B., Cui, L., and Fang, J. (2016). Abscisic acid and sucrose regulate tomato and strawberry fruit ripening through the abscisic acid-stress-ripening transcription factor. *Plant Biotechnology Journal*, 14(10):2045–2065.
- Kosma, D. K., Parsons, E. P., Isaacson, T., Lü, S., Rose, J. K., and Jenks, M. A. (2010). Fruit cuticle lipid composition during development in tomato ripening mutants. *Physiologia Plantarum*, 139(1):107–117.
- Koyama, R., Roberto, S. R., de Souza, R. T., Borges, W. F., Anderson, M., Waterhouse, A. L., Cantu, D., Fidelibus, M. W., and Blanco-Ulate, B. (2018). Exogenous abscisic acid promotes anthocyanin biosynthesis and increased expression of flavonoid synthesis genes in *vitis vinifera* × *vitis labrusca* table grapes in a subtropical region. *Frontiers in Plant Science*, 9:323.
- Kramer, M. G. and Redenbaugh, K. (1994). Commercialization of a tomato with an anti-sense polygalacturonase gene: The FLAVR SAVR™ tomato story. *Euphytica*, 79(3):293–297.

- Kumar, R., Tamboli, V., Sharma, R., and Sreelakshmi, Y. (2018). NAC-NOR mutations in tomato Penjar accessions attenuate multiple metabolic processes and prolong the fruit shelf life. *Food Chemistry*, 259(April):234–244.
- Laing, W. A., Martínez-Sánchez, M., Wright, M. A., Bulley, S. M., Brewster, D., Dare, A. P., Rassam, M., Wang, D., Storey, R., Macknight, R. C., and Hellens, R. P. (2015). An upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in arabidopsis. *Plant Cell*, 27(3):772–786.
- Lara, I., Heredia, A., and Domínguez, E. (2019). Shelf life potential and the fruit cuticle: The unexpected player. *Frontiers in Plant Science*, 10:770.
- Li, J., Scarano, A., Gonzalez, N. M., D’Orso, F., Yue, Y., Nemeth, K., Saalbach, G., Hill, L., de Oliveira Martins, C., Moran, R., Santino, A., and Martin, C. (2022a). Biofortified tomatoes provide a new route to vitamin D sufficiency. *Nature Plants*, 8(6):611–616.
- Li, Q., Xu, R., Fang, Q., Yuan, Y., Cao, J., and Jiang, W. (2020). Analyses of microstructure and cell wall polysaccharides of flesh tissues provide insights into cultivar difference in mealy patterns developed in apple fruit. *Food Chemistry*, 321:126707.
- Li, X., Wang, X., Zhang, Y., Zhang, A., and You, C. X. (2022b). Regulation of fleshy fruit ripening: from transcription factors to epigenetic modifications. *Horticulture Research*, 9.
- Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T. K., Espley, R. V., Hellens, R. P., and Allan, A. C. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biology*, 10(1):1–17.
- Liu, Y., Tang, M., Liu, M., Su, D., Chen, J., Gao, Y., Bouzayen, M., and Li, Z. (2020). The Molecular Regulation of Ethylene in Fruit Ripening. *Small Methods*, 4(8):1900485.
- Lobato-Gómez, M., Hewitt, S., Capell, T., Christou, P., Dhingra, A., and Girón-Calva,

- P. S. (2021). Transgenic and genome-edited fruits: background, constraints, benefits, and commercial opportunities. *Horticulture Research*, 8(1):166.
- Lurie, S. (2021). Genomic and transcriptomic studies on chilling injury in peach and nectarine. *Postharvest Biology and Technology*, 174:111444.
- Ma, B., Liao, L., Fang, T., Peng, Q., Ogutu, C., Zhou, H., Ma, F., and Han, Y. (2019). A Ma10 gene encoding P-type ATPase is involved in fruit organic acid accumulation in apple. *Plant Biotechnology Journal*, 17(3):674–686.
- McQuinn, R. P., Gapper, N. E., Gray, A. G., Zhong, S., Tohge, T., Fei, Z., Fernie, A. R., and Giovannoni, J. J. (2020). Manipulation of ZDS in tomato exposes carotenoid- and ABA-specific effects on fruit development and ripening. *Plant Biotechnology Journal*, 18(11):2210–2224.
- Mellidou, I., Koukounaras, A., Kostas, S., Patelou, E., and Kanellis, A. K. (2021). Regulation of vitamin c accumulation for improved tomato fruit quality and alleviation of abiotic stress. *Genes*, 12(5).
- Meng, F., Li, Y., Li, S., Chen, H., Shao, Z., Jian, Y., Mao, Y., Liu, L., and Wang, Q. (2022). Carotenoid biofortification in tomato products along whole agro-food chain from field to fork. *Trends in Food Science Technology*, 124:296–308.
- Nieuwenhuizen, N. J., Chen, X., Wang, M. Y., Matich, A. J., Perez, R. L., Allan, A. C., Green, S. A., and Atkinson, R. G. (2015). Natural variation in monoterpene synthesis in kiwifruit: Transcriptional regulation of terpene synthases by NAC and ETHYLENE-INSENSITIVE3-like transcription factors. *Plant Physiology*, 167(4):1243–1258.
- Nordey, T., Davrieux, F., and Léchaudel, M. (2019). Predictions of fruit shelf life and quality after ripening: Are quality traits measured at harvest reliable indicators? *Postharvest Biology and Technology*, 153:52–60.
- Osei, M. K., Danquah, A., E.T., B., Danquah, E., and Adu-Dapaah (2017). An overview

- of tomato fruit-ripening mutants and their use in increasing shelf life of tomato fruits. *African Journal of Agricultural Research*, 12(51):3520–3528.
- Petrasch, S., Mesquida-Pesci, S. D., Pincot, D. D., Feldmann, M. J., López, C. M., Famula, R., Hardigan, M. A., Cole, G. S., Knapp, S. J., and Blanco-Ulate, B. (2022). Genomic prediction of strawberry resistance to postharvest fruit decay caused by the fungal pathogen *Botrytis cinerea*. *G3: Genes, Genomes, Genetics*, 12(1).
- Rodríguez-Ruiz, M., Mateos, R. M., Codesido, V., Corpas, F. J., and Palma, J. M. (2017). Characterization of the galactono-1,4-lactone dehydrogenase from pepper fruits and its modulation in the ascorbate biosynthesis. Role of nitric oxide. *Redox Biology*, 12:171–181.
- Rose, J. and Isaacson, T. (2009). Delayed fruit deterioration allele in plants and methods of detection.
- Rothkegel, K., Espinoza, A., Sanhueza, D., Lillo-Carmona, V., Riveros, A., Campos-Vargas, R., and Meneses, C. (2021). Identification of DNA Methylation and Transcriptional Profiles Associated With Fruit Mealiness in *Prunus persica* (L.) Batsch. *Frontiers in Plant Science*, 12.
- Seymour, G. B. and Granell, A. (2014). Fruit development and ripening. 65(16):4489–4490.
- Shi, Y., Vrebalov, J., Zheng, H., Xu, Y., Yin, X., Liu, W., Liu, Z., Sorensen, I., Su, G., Ma, Q., Evanich, D., Rose, J. K., Fei, Z., van Eck, J., Thannhauser, T., Chen, K., and Giovannoni, J. J. (2021). A tomato LATERAL ORGAN BOUNDARIES transcription factor, SILOB1, predominantly regulates cell wall and softening components of ripening. *Proceedings of the National Academy of Sciences of the United States of America*, 118(33):e2102486118.
- Silva, C. J., Van Den Abeele, C., Ortega-Salazar, I., Papin, V., Adaskaveg, J. A., Wang, D., Casteel, C. L., Seymour, G. B., and Blanco-Ulate, B. (2021). Host susceptibility

- factors render ripe tomato fruit vulnerable to fungal disease despite active immune responses. *Journal of Experimental Botany*, 72(7):2696–2709.
- Smirnova, O. G. and Kochetov, A. V. (2020). Choice of the Promoter for Tissue and Developmental Stage-Specific Gene Expression. *Methods in Molecular Biology*, 2124:69–106.
- Sun, C., Deng, L., Du, M., Zhao, J., Chen, Q., Huang, T., Jiang, H., Li, C. B., and Li, C. (2020). A Transcriptional Network Promotes Anthocyanin Biosynthesis in Tomato Flesh. *Molecular Plant*, 13(1):42–58.
- Tieman, D., Zhu, G., Resende, M. F., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., and Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.)*, 355(6323):391–394.
- Tucker, G., Yin, X., Zhang, A., Wang, M., Zhu, Q., Liu, X., Xie, X., Chen, K., and Grierson, D. (2017). Ethylene and Softening. 1(4):253–267.
- Waltz, E. (2022). GABA-enriched tomato is first CRISPR-edited food to enter market. *Nature Biotechnology*, 40(1):9–11.
- Wang, D., Samsulrizal, N. H., Yan, C., Allcock, N. S., Craigon, J., Blanco-Ulate, B., Ortega-Salazar, I., Marcus, S. E., Bagheri, H. M., Perez-Fons, L., Fraser, P. D., Foster, T., Fray, R., Paul Knox, J., and Seymour, G. B. (2019a). Characterization of CRISPR mutants targeting genes modulating pectin degradation in ripening tomato 1. *Plant Physiology*, 179(2):544–557.
- Wang, D., Yeats, T. H., Uluisik, S., Rose, J. K., and Seymour, G. B. (2018). Fruit Softening: Revisiting the Role of Pectin. *Trends in Plant Science*, 23(4):302–310.
- Wang, R., Lammers, M., Tikunov, Y., Bovy, A. G., Angenent, G. C., and de Maagd, R. A.

- (2020). The rin, nor and Cnr spontaneous mutations inhibit tomato fruit ripening in additive and epistatic manners. *Plant Science*, 294:110436.
- Wang, R., Shu, P., Zhang, C., Zhang, J., Chen, Y., Zhang, Y., Du, K., Xie, Y., Li, M., Ma, T., Zhang, Y., Li, Z., Grierson, D., Pirrello, J., Chen, K., Bouzayen, M., Zhang, B., and Liu, M. (2022). Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytologist*, 233(1):373–389.
- Wang, R., Tavano, E. C. d. R., Lammers, M., Martinelli, A. P., Angenent, G. C., and de Maagd, R. A. (2019b). Re-evaluation of transcription factor function in tomato fruit development and ripening with CRISPR/Cas9-mutagenesis. *Scientific Reports*, 9(1):1696.
- Xiao, Y. Y., Kuang, J. F., Qi, X. N., Ye, Y. J., Wu, Z. X., Chen, J. Y., and Lu, W. J. (2018). A comprehensive investigation of starch degradation process and identification of a transcriptional activator MabHLH6 during banana fruit ripening. *Plant Biotechnology Journal*, 16(1):151–164.
- Xie, F., Hua, Q., Chen, C., Zhang, Z., Zhang, R., Zhao, J., Hu, G., Chen, J., and Qin, Y. (2021). Genome-wide characterization of R2R3-MYB transcription factors in pitaya reveals a R2R3-MYB repressor HuMYb1 involved in fruit ripening through regulation of betalain biosynthesis by repressing betalain biosynthesis-related genes. *Cells*, 10(8):1949.
- Xing, S., Chen, K., Zhu, H., Zhang, R., Zhang, H., Li, B., and Gao, C. (2020). Fine-tuning sugar content in strawberry. *Genome Biology*, 21(1):1–14.
- Xue, C., Guan, S. C., Chen, J. Q., Wen, C. J., Cai, J. F., and Chen, X. (2020). Genome wide identification and functional characterization of strawberry pectin methylesterases related to fruit softening. *BMC Plant Biology*, 20(1):1–17.
- Yang, L., Huang, W., Xiong, F., Xian, Z., Su, D., Ren, M., and Li, Z. (2017). Silencing of SIPL, which encodes a pectate lyase in tomato, confers enhanced fruit firmness, pro-

- longed shelf-life and reduced susceptibility to grey mould. *Plant Biotechnology Journal*, 15(12):1544–1555.
- Ye, J., Li, W., Ai, G., Li, C., Liu, G., Chen, W., Wang, B., Wang, W., Lu, Y., Zhang, J., Li, H., Ouyang, B., Zhang, H., Fei, Z., Giovannoni, J. J., Ye, Z., and Zhang, Y. (2019a). Genome-wide association analysis identifies a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genetics*, 15(5):e1008149.
- Ye, J., Liu, G., Chen, W., Zhang, F., Li, H., Ye, Z., and Zhang, Y. (2019b). Knockdown of SINL33 accumulates ascorbate, enhances disease and oxidative stress tolerance in tomato (*Solanum lycopersicum*). *Plant Growth Regulation*, 89(1):49–58.
- ye Chen, J., fang Xie, F., ze Cui, Y., bin Chen, C., jin Lu, W., di Hu, X., zhu Hua, Q., Zhao, J., jiang Wu, Z., Gao, D., ke Zhang, Z., kai Jiang, W., ming Sun, Q., bing Hu, G., and hua Qin, Y. (2021). A chromosome-scale genome sequence of pitaya (*Hylocereus undatus*) provides novel insights into the genome evolution and regulation of betalain biosynthesis. *Horticulture Research*, 8(1):164.
- Yu, Y., Guan, J., Xu, Y., Ren, F., Zhang, Z., Yan, J., Fu, J., Guo, J., Shen, Z., Zhao, J., Jiang, Q., Wei, J., and Xie, H. (2021). Population-scale peach genome analyses unravel selection patterns and biochemical basis underlying fruit flavor. *Nature Communications*, 12(1):1–13.
- Zhang, B., Tieman, D. M., Jiao, C., Xu, Y., Chen, K., Fe, Z., Giovannoni, J. J., and Klee, H. J. (2016). Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44):12580–12585.
- Zhang, C., Duan, W., Chen, K., and Zhang, B. (2020a). Transcriptome and methylome analysis reveals effects of ripening on and off the vine on flavor quality of tomato fruit. *Postharvest Biology and Technology*, 162.



- Zhang, J., Lei, Y., Wang, B., Li, S., Yu, S., Wang, Y., Li, H., Liu, Y., Ma, Y., Dai, H., Wang, J., and Zhang, Z. (2020b). The high-quality genome of diploid strawberry (*Fragaria nilgerrensis*) provides new insights into anthocyanin accumulation. *Plant Biotechnology Journal*, 18(9):1908–1924.
- Zheng, X., Giuliano, G., and Al-Babili, S. (2020). Carotenoid biofortification in crop plants: citius, altius, fortius. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1865(11):158664.

## Chapter 2

# Single and Double Mutations in Tomato Ripening Transcription Factors Have Distinct Effects on Fruit Development and Quality Traits

### 2.1 Publication Statement

The content of this chapter was peer-reviewed and published as follows:

Adaskaveg, J. A., Silva, C. J., Huang P., and Blanco-Ulate, B. (2021). Single and Double Mutations in Tomato Ripening Transcription Factors Have Distinct Effects on Fruit Development and Quality Traits. *Frontiers in Plant Sciences*, 12:647035.

For this publication, I wrote the paper and was credited as first author. Additionally, I performed the field experiments, measured all phenotypic traits, and performed all genotyping, qPCR, and statistical analyses.

### 2.2 Abstract

Spontaneous mutations associated with the tomato transcription factors COLORLESS NON-RIPENING (SPL-CNR), NON-RIPENING (NAC-NOR), and RIPENING-INHIBITOR (MADS-RIN) result in fruit that do not undergo the normal hallmarks of ripening but are phenotypically distinguishable. Here, we expanded knowledge of the physiological, molecular, and genetic impacts of the ripening mutations on fruit development beyond ripening. We demonstrated through phenotypic and transcriptome analyses that *Cnr* fruit exhibit a broad range of developmental defects before the onset of fruit ripening, but fruit still undergo some ripening changes similar to wild type. Thus, *Cnr* should be considered as

a fruit developmental mutant and not just a ripening mutant. Additionally, we showed that some ripening processes occur during senescence in the *nor* and *rin* mutant fruit, indicating that while some ripening processes are inhibited in these mutants, others are merely delayed. Through gene expression analysis and direct measurement of hormones, we found that *Cnr*, *nor*, and *rin* have alterations in the metabolism and signaling of plant hormones. *Cnr* mutants produce more than basal levels of ethylene, while *nor* and *rin* accumulate high concentrations of abscisic acid. To determine genetic interactions between the mutations, we created for the first time homozygous double mutants. Phenotypic analyses of the double ripening mutants revealed that *Cnr* has a strong influence on fruit traits and that combining *nor* and *rin* leads to an intermediate ripening mutant phenotype. However, we found that the genetic interactions between the mutations are more complex than anticipated, as the *Cnr/nor* double mutant fruit has a *Cnr* phenotype but displayed inhibition of ripening-related gene expression just like *nor* fruit. Our reevaluation of the *Cnr*, *nor*, and *rin* mutants provides new insights into the utilization of the mutants for studying fruit development and their implications in breeding for tomato fruit quality.

## 2.3 Introduction

Fleshy fruit gain most of their quality traits, such as color, texture, flavor, and nutritional value, as a result of physiological and biochemical changes associated with ripening. Fruit ripening has been studied for decades, yet there are still many unanswered questions about the timing and coordination of the biological processes related to this developmental program. Much of this research has been done in the model for fleshy fruit ripening, tomato (*Solanum lycopersicum*), and has utilized the spontaneous single ripening mutants *Cnr* (*Colorless non-ripening*), *nor* (*non-ripening*), and *rin* (*ripening inhibitor*) (Giovannoni et al., 2004; Manning et al., 2006; Robinson and Tomes, 1968; Thompson et al., 1999; Tigchelaar et al., 1973). Each of these mutations produces pleiotropic defects to ripening and occur in or near genes encoding the transcription factors (TFs) SPL-CNR, NAC-NOR, and MADS-RIN, belonging to the SQUAMOSA promoter binding protein-

like (SPL), NAM, ATAF1/2, CUC2 (NAC) and, MCM1, AG, DEF, SRF (MADS) TF families, respectively. Each TF family functions in diverse developmental processes and have distinct spatiotemporal expression patterns (Karlova et al., 2014; Shinozaki et al., 2018).

These mutants were used to study ripening under the assumption that the mutations cause a complete loss of function to the corresponding protein. Recently, it has been discovered that the *nor* and *rin* mutations produce proteins that are still functional and gain the ability to negatively regulate their targets (Gao et al., 2020, 2019; Ito et al., 2017; Li et al., 2018, 2019b; Wang et al., 2019). In *nor*, the two base pair deletion truncates the protein but still produces a functional DNA-binding and dimerizing NAC domain (Gao et al., 2020). In *rin*, a large deletion creates a chimeric protein with the neighboring gene MACROCALYX (MC), producing a functional protein with suppression activity (Ito et al., 2017). The *Cnr* mutation is also thought to be a gain of function mutation, although the mechanism has yet to be understood (Gao et al., 2019). The *Cnr* mutation results from hypermethylation upstream of the gene near the promoter and has been shown to inhibit the genome-wide demethylation cascade associated with normal tomato ripening (Zhong et al., 2013). Previously, these TFs were regarded as master regulators of ripening; however, given the new information about the nature of the mutations in *Cnr*, *nor*, and *rin*, it is less clear the precise roles the TFs are playing in ripening (Giovannoni et al., 2017; Wang et al., 2020a).

The *nor* and *rin* mutants have been utilized in breeding for developing tomato hybrids with extended shelf life or extended field harvest depending on their purpose for the fresh market and processing tomato industries (Garg et al., 2008; Kitagawa et al., 2005; Kopeliovitch et al., 1979; Osei et al., 2017). Hybrids between elite varieties and the ripening mutants have a delayed ripening progression, but with the tradeoff of decreased fruit quality attributes, such as color, taste, and aroma (Kitagawa et al., 2005; Tieman et al., 2017). Although there are some publications dedicated to evaluating the physiological characteristics of mutant or hybrid fruit (Agar et al., 1994; Garg et al., 2008; Tigchelaar et al., 1978), up to this point, much of what we know about the ripening mu-

tations is based on controlled greenhouse experiments with limited fruit and few ripening stages examined. A complete dataset of phenotypic data produced from large-scale field trials evaluating fruit ripening and senescence is lacking to provide information relevant to breeding, particularly in the new context of the molecular mechanisms behind the *nor* and *rin* mutations.

The *Cnr* mutant provides a unique opportunity to study the role of epigenetics in fruit ripening but is not used in breeding because the mutant phenotype is dominant. *Cnr* has been regarded as a ripening mutant due to its unique colorless phenotype and additional ripening defects (Thompson et al., 1999). It has been suggested that *Cnr* fruit undergo normal growth and development (Lai et al., 2020); however, fruit appear different from wild type (WT) even before ripening, with a smaller size, alterations in cell wall enzyme expression, and earlier chlorophyll degradation (Eriksson et al., 2004; Wang et al., 2020b). To better utilize *Cnr* as a tool for studying fruit development and ripening, a broader understanding of the physiological and transcriptomic alterations in this mutant is necessary.

These spontaneous single mutants need to be reevaluated as tools to understand the wide-ranging biological processes regulated by each TF. Previous literature has generally assumed that the mutations block ripening, resulting in similar processes affected (Giovannoni et al., 2017; Giovannoni, 2007; Karlova et al., 2014; Osorio et al., 2020). This study demonstrates that each mutant has a unique ripening phenotype, resulting from a combination of inhibited and delayed developmental processes. We integrated phenotypic data with gene expression data and hormone measurements in the *Cnr*, *nor*, and *rin* mutants across ripening and senescence to characterize the extent and timing of the ripening defects. Tomatoes grown under field conditions were assessed for fruit traits over multiple seasons. We then performed a transcriptomic analysis to gain more definition of the timing in which mutant fruit deviated from WT in their development and to determine specific molecular functions altered in each mutant. Due to their pivotal role in regulating ripening, we focused on defects in hormone networks, including biosynthesis and accumulation. We analyzed the influence of each mutation on the expression of the other TF

throughout ripening and senescence. Finally, to better understand the combined genetic effects of the mutants on fruit ripening, we generated homozygous double mutants of *Cnr*, *nor*, and *rin* and used phenotyping and transcriptional data to evaluate the relationships between the mutants.

## 2.4 Materials and Methods

### 2.4.1 Plant material

Tomato plants (*Solanum lycopersicum*) of c.v. ‘Alisa Craig’ and the isogenic ripening mutants *Cnr*, *nor*, and *rin* were grown in randomized plots under standard field conditions in Davis, CA, United States, during the 2016, 2017, 2018, and 2020 seasons. Fruit tagged at 10 days post-anthesis (dpa), which corresponds to 7 mm in fruit diameter, were harvested at stages equivalent to the WT fruit. Fruit were sampled at the mature green (MG), turning (T), red ripe (RR), and overripe (OR) stages, corresponding to 37, 45, 50, and 57 dpa, respectively. The term “RR” is used throughout the manuscript to refer to the 50 dpa stage of all genotypes, even when the mutant fruit do not turn red. Fruit stages for each of the mutants were further validated by external color analysis (see details on fruit trait phenotyping).

Double mutant fruit were generated through reciprocal crosses: *Cnr* × *nor*, *nor* × *Cnr*, *Cnr* × *rin*, *rin* × *Cnr*, *nor* × *rin*, and *rin* × *nor*. Fruit were selfed after the initial cross to generate an F2 segregating generation. The double mutants were initially selected in the F2 generation through genotyping and phenotyping. At least two additional generations after F2 were obtained through selfing to ensure the stability of the double mutations and to perform the experiments in this study. Three seasons of data were collected for the *Cnr/nor* fruit (2016, 2017, and 2020) while only one season of data was collected for the *rin/nor* and *Cnr/rin* crosses.

### 2.4.2 Mutant Genotyping

The mutant lines were genotyped for their respective mutations. For *nor*, the Phire Plant Direct PCR Kit (Thermo Fisher Scientific, United States) was used to extract DNA and amplify the region of the gene containing the 2 bp mutation using the primers listed in

Supplementary Table 1 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). The PCRs were run on a SimpliAmp Thermal Cycler (Applied Biosystems, United States) with the following conditions denaturation: 99°C for 5 min; 35 cycles of 98°C for 5 s, 56°C for 25 s, and 72°C for 25 s; with a final extension of 72°C for 1 min. The PCR products were purified using Wizard SV Gel and PCR Clean-Up System (Promega, United States) and then sequenced with Sanger technology to confirm the absence of the two (AA) nucleotides. For rin, the Phire Plant Direct PCR Kit (Thermo Fisher Scientific, United States) was used to extract DNA and perform end-point PCRs using primers specific for the mutant and WT alleles (see Supplementary Table 1 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). The following PCR conditions were used for the WT allele primers: denaturation 99°C for 5 min; 35 cycles of 98°C for 5 s, 55°C for 25 s, and 72°C for 25 s; with a final extension of 72°C for 1 min. The PCR conditions for the mutant allele primers were: denaturation 98°C for 5 min; 40 cycles of 98°C for 5 s, 58°C for 25 s, and 72°C for 25 s; with a final extension of 72°C for 1 min. The PCR products were visualized as bands using a 1

The *Cnr* epimutation was genotyped by bisulfite sequencing. Extracted DNA was treated with the Zymo Gold bisulfite kit (Zymo Research, United States). Bisulfite treated-DNA was PCR amplified for the *CNR* promoter region containing the methylation changes (Manning et al., 2006) using the primers listed in Supplementary Table 1 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). The following PCR conditions were used: 94°C for 2 min; 40 cycles of 94°C for 30 s, 54°C for 30 s, and 60°C for 45 s, and a final extension of 60°C for 10 min. The PCR products were then Sanger sequenced and compared to the same region amplified in untreated controls with primers (see Supplementary Table 1 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). The following conditions were used to amplify the untreated DNA: 95°C for 2 min; 35 cycles of 95°C for 30 s, 56°C for 30 s and 72°C for 1 min, and a final extension of 72°C for 10 min. To ensure mutants were homozygous for the locus, we confirmed the double mutants by allowing the plants to self for at least two additional generations and checking that the progeny were not segregating for any fruit

phenotypes.

### 2.4.3 Fruit Trait Phenotyping

Fruit trait data were collected across four field seasons (2016, 2017, 2018, and 2020). The genotypes, developmental stages, number of biological replicates, and number of field seasons used for fruit trait phenotyping can be found in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). One season of phenotyping was performed for *Cnr/rin* and *rin/nor* double mutant fruits for color, firmness, and ethylene. Three seasons of data were collected for the *Cnr/nor* double mutant fruit for ethylene and two seasons of data for color and firmness. Fruit were collected from multiple plots or harvests to capture environmental variability. Fruit trait measurements were taken on the same day of harvest for all samples unless noted. Intact and halved fruit were imaged using the VideometerLab 3 (Videometer, Denmark) facilitated by Aginnovation LLC1. External color measurements were obtained from individual fruit with the CR-410 Chroma Meter (Konica Minolta Inc, Japan) and recorded in the  $L^*a^*b^*$  color space, where  $L^*$  quantifies lightness,  $a^*$  quantifies green/red color, and  $b^*$  quantifies blue/yellow color. Principal component analysis (PCA) of the color parameters was performed with the FactoMineR package and graphed with the FactoExtra package in R (Kassambara et al., 2017; Lê et al., 2008). Non-destructive firmness measurements were taken on the TA.XT2i Texture Analyzer (Texture Technologies, United States) using a TA-11 acrylic compression probe, a trigger force of 0.035 kg, and a test speed of 2.00 mm/sec with Exponent software (Texture Technologies Corporation, United States). Firmness values are reported as kilograms (kg) force. The size was measured by taking the largest diameter (mm) of the fruit with a handheld caliper.

Tomato juice was produced by pressing the fruit tissues with a juicer and filtering with cheesecloth to measure total soluble solids (TSS) and titratable acidity (TA). At least five biological replications of tomato juice were obtained from independent pools of 10–12 fruit from distinct plots in the field or at different harvest dates within the field season. TSS were measured as percent Brix with a Reichert AR6 Series automatic bench refractometer (Reichert Inc., United States) from the prepared juice with three technical



replicates. TA was measured using the tomato juice with the TitraLab TIM850 Titration Manager (Radiometer Analytics, Germany). Four grams of juice were diluted with water in 20 mL of deionized water to measure TA based on citric acid equivalents. Significant differences in fruit traits across genotypes and ripening stages were determined in R (R foundation for Statistical Computing) using Type I analysis of variance (ANOVA) tests, followed by a post hoc test (Tukey Honest Significant Differences, HSD) using the R package agricolae (De et al., 2017).

#### **2.4.4 RNA extraction**

On the day of harvest, the fruit pericarp tissues were dissected and flash-frozen in liquid nitrogen. Frozen tissues were then ground to a fine powder with the Retsch Mixer Mill MM 400 (Verder Scientific, Netherlands). One gram of ground tissue was used for RNA extractions as described in Blanco-Ulate et al. (Blanco-Ulate et al., 2013). RNA concentrations were quantified with Nanodrop One Spectrophotometer (Thermo Scientific, United States) and Qubit 3 (Invitrogen, United States). RNA integrity was then assessed on an agarose gel. Six biological replicates composed of 8–10 independent fruit were extracted per genotype and ripening stage from the 2016 and 2018 seasons.

#### **2.4.5 cDNA Preparation and RT-qPCR**

cDNA was prepared from 1 µg of RNA of all samples using M-MLV Reverse Transcriptase (Promega, United States) in the SimpliAmp Thermal Cycler (Applied Biosystems, United States). RT-qPCRs were performed using PowerSYBR Green PCR Master Mix (Applied Biosystems, United States) in the QuantStudio3 (Applied Biosystems, United States) following the preset qPCR conditions for the ‘Comparative CT method.’ The tomato *SlUBQ* (Solyc12g04474) was used as the reference gene for all relative expression analyses. Primers for the genes of interest were designed using Primer-BLAST (Ye et al., 2012) or obtained from previous studies (see Supplementary Table 1 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). For all new qPCR primer sets, efficiency was confirmed to be higher than 90 % using fourfold DNA or cDNA dilutions (0, 1:1, 1:4, 1:16, 1:64, and 1:256) in triplicate. Then, specificity was checked by analyzing

the melting curves at temperatures ranging from 60 to 95°C. Relative gene expression was calculated using the formula  $2^{(\text{reference gene Ct} - \text{gene of interest Ct})}$ .

#### **2.4.6 cDNA Library Preparation, RNA Sequencing, and Sequencing Data Processing**

Four biological replicates each of Cnr/nor MG and RR fruit RNA were used to prepare cDNA libraries. cDNA libraries were prepared with Illumina TruSeq RNA Sample Preparation Kit v.2 (Illumina, United States) from the extracted RNA. The quality of the barcoded cDNA libraries was assessed with the High Sensitivity DNA Analysis Kit in the Agilent 2100 Bioanalyzer (Agilent Technologies, United States) and then sequenced (50 bp single-end reads) on the Illumina HiSeq 4000 platform by the DNA Technologies Core at UC Davis Genome Center.

Raw RNAseq data from WT, Cnr, nor, and rin at MG and RR were obtained from a published dataset by our group (Silva et al., 2021), GEO accession GSE148217), while raw RNAseq data from the immature stages of the ripening mutants were extracted from (Lü et al., 2018) (GEO accession GSE116581). The RNAseq datasets for the Cnr/nor double mutant were generated in this study. The raw sequencing reads from the different datasets were analyzed de novo following the bioinformatics pipeline described below. Raw reads were trimmed for quality and adapter sequences using Trimmomatic v0.39 (Bolger et al., 2014) with the following parameters: maximum seed mismatches = 2, palindrome clip threshold = 30, simple clip threshold = 10, minimum leading quality = 3, minimum trailing quality = 3, window size = 4, required quality = 15, and minimum length = 36. Trimmed reads were then mapped using Bowtie2 (Langmead and Salzberg, 2013) to the tomato transcriptome (SL4.0 release2). Count matrices were made from the Bowtie2 results using sam2counts.py v0.913. A summary of all read mapping results can be found in Supplementary Table 3 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

### 2.4.7 Differential expression analysis, Functional Annotations, and Enrichment Analysis

The Bioconductor package DESeq2 (Love et al., 2014) in R was used to normalize read counts and perform PCAs and differential expression analyses for various comparisons (see Supplementary Tables 4, 5 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Differentially expressed genes (DEGs) for each comparison had an adjusted P-value of less than or equal to 0.05. Gene functional annotations were retrieved from the Kyoto Encyclopedia of Genes and Genomes (KEGG) using the KEGG Automatic Annotation Server (Moriya et al., 2007). Enrichment analysis for all functional annotations was performed using a Fisher test. The P-values obtained from the Fisher test were adjusted with the Benjamini and Hochberg method (Benjamini and Hochberg, 1995). Shared and unique DEGs among the comparisons were determined using the R package UpSetR (Conway et al., 2017).

### 2.4.8 Measurement of phytohormones

Ethylene production measurements were taken from MG, RR, and OR fruit on the day of harvest. At least five biological replicates of 5–7 fruit were used for the measurements. The genotypes, developmental stages, and number of biological replicates used for ethylene analysis in each field season can be found in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Fruit were weighed and placed in 1 L airtight glass jars. Headspace gas (3 ml) was extracted from the sealed containers after 60 min and was injected into a Shimadzu CG-8A gas chromatograph (Shimadzu Scientific Instruments, Japan). Sample peaks were measured against an ethylene standard. The rate of ethylene production ( $\text{nL kg}^{-1}$  fresh weight  $\text{h}^{-1}$ ) was calculated from the peak, fruit mass, and incubation time.

Frozen ground tissue prepared from the tomato fruit pericarp was lyophilized, weighed, and extracted in isopropanol:H<sub>2</sub>O:HCL1MOL(2:1:0.005) with 100 l of internal standard solution (1000 pg) as described in (Casteel et al., 2015). Abscisic acid (ABA) and 1-aminocyclopropane-1-carboxylate (ACC) were measured using liquid chromatography coupled to tandem mass spectrometry and internal standards as described in (Casteel

et al., 2015). The hormone concentrations were expressed as ng/g of dry weight. Four to six biological replicates composed of 8–10 fruit were used for these measurements for the 2017 season. Significant differences in hormone accumulation across genotypes and ripening stages were determined using Type I ANOVA in R, followed by an HSD test using the R package agricolae (De et al., 2017). In some cases, pairwise comparisons in hormone accumulation were also conducted by Student’s t-test in R.

#### **2.4.9 Data availability**

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/geo/>, GSE163745.

## **2.5 Results**

### **2.5.1 Ripening Mutants Display Distinct Phenotypes and Transcriptional Profiles Throughout Fruit Development**

Fruit from the *Cnr*, *nor*, and *rin* mutants fail to acquire most ripening-associated traits that make them appealing for consumption. Yet, each mutant can be distinguished by their unique phenotypes (Figure 2.1). To determine the impact of *Cnr*, *nor*, and *rin* mutations on the key fruit traits, we measured external color, firmness, total soluble solids (TSS), and titratable acidity (TA) at multiple ripening stages. Fruit from the isogenic mutants *Cnr*, *nor*, and *rin*, were harvested alongside WT from an experimental field at selected ripening stages, mature green (MG; 37 dpa), turning (T; 45 dpa), red ripe (RR; 50 dpa), and overripe (OR; 57 dpa) (Figure 2.1A). We captured field variability through large sample sizes and validated across two to four independent field seasons. A summary of all seasons is displayed in Figure 2.1 while a breakdown of the data by field season can be found in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

As expected, between the MG and RR stages WT fruit turned red internally and externally, reduced firmness, accumulated TSS, and became less acidic during ripening. *Cnr* fruit showed visual differences compared to all of the genotypes at the MG stage and

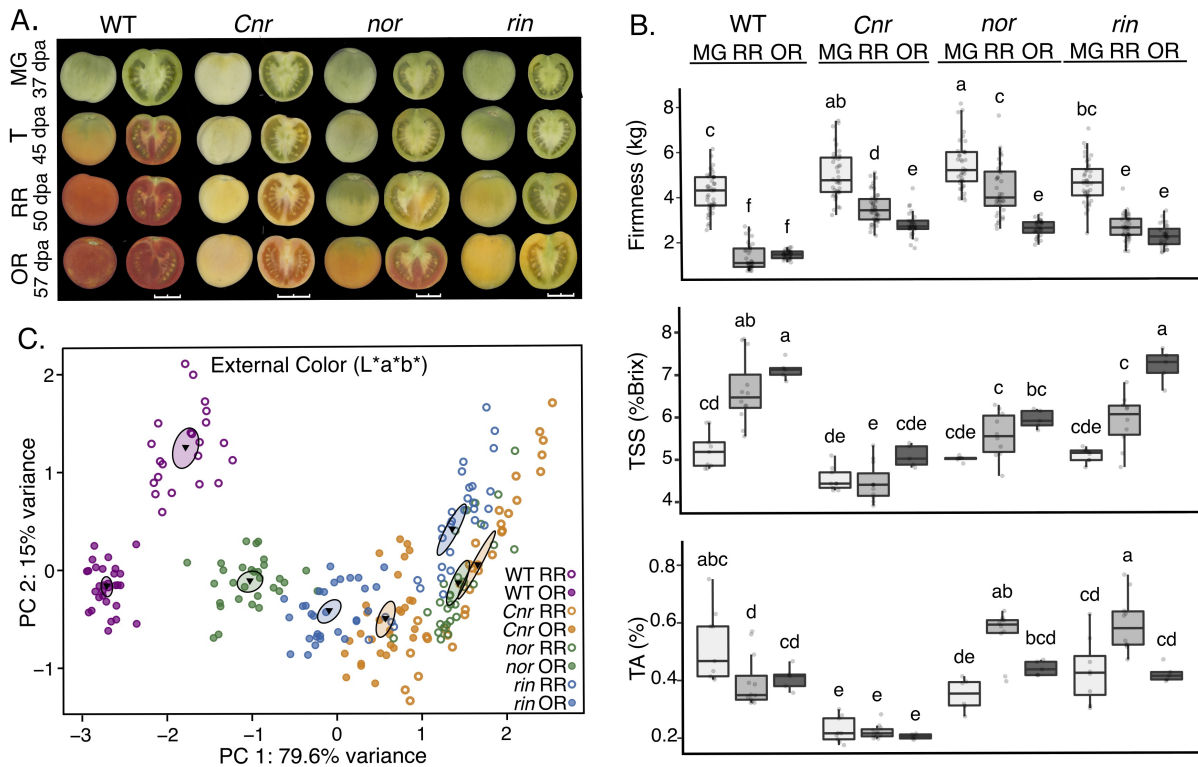


Figure 2.1: (A) Ripening progression of wild type (WT), c.v. ‘Ailsa Craig,’ compared to the isogenic mutants *Cnr*, *nor*, and *rin* across four developmental stages: mature green (MG, 37 days post anthesis (dpa)), turning (T, 45 dpa), red ripe (RR, 50 dpa), and overripe (OR, 57 dpa) shown at left whole and at right in longitudinal sections. Images were captured and processed with the VideometerLab instrument. Bars correspond to 2 cm. (B) Measurements of fruit firmness ( $n = 28-44$ ), total soluble solids (TSS) ( $n = 5-12$ ), and titratable acidity (TA) ( $n = 5-12$ ) for each MG, RR, and OR stages are presented. Error bars represent standard error between biological replicates of each sample. Letters indicate significant differences among genotypes and stages calculated by ANOVA and Tukey HSD ( $P \leq 0.05$ ). Principal component analysis of external color measured on the  $L^*a^*b^*$  color scale of each genotype at the RR ( $n = 22-34$ ) and OR stage ( $n = 28-40$ ). The center of gravity is represented by a triangle with surrounding ellipses indicating 95% confidence interval.

continuing through subsequent stages, including significantly smaller size and its characteristic colorless flesh, marked by an opaque yellow coloration of the pericarp (Figure 2.1 and Supplementary Figure 2.S1). Statistical analyses performed for color and size confirmed *Cnr* exhibited significant differences ( $P \leq 0.05$ ) consistently across each field season, as reported in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Fruit of *nor* and *rin* displayed a distinct absence of any red coloration compared to WT at the RR stage; instead, these fruit began to turn yellow externally. Fruit of all ripening mutants were significantly firmer across all stages than the WT, though this difference was especially pronounced in *Cnr* (Figure 2.1B). Overall, *Cnr* was consistently different from the other genotypes before and during ripening, while *nor* and *rin* remained similar to WT MG fruit.

The OR stage was selected to investigate if the ripening mutants displayed phenotypic changes at later time points that could be associated with a delay in fruit development. At this stage, *nor* fruit started to turn orange-red externally and red internally, similar to WT fruit. The *nor* OR fruit resembled WT fruit between the T and RR stages. We performed a PCA of the color data ( $L^*a^*b$  measurements) to compare the genotypes at the RR and OR stage, and found *nor* OR measured closely with WT RR in external coloration (Figure 2.1C). A summary of the color data and statistical analyses performed can be found in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). While *nor* OR fruit visually looked most similar to WT RR fruit, *rin* OR fruit consistently measured similarly to WT fruit at the RR and OR stages in the taste-related traits of TSS and TA. These phenotypes were especially noticeable in the OR stage, suggesting that *rin* exhibits a delay in these traits. In contrast, *Cnr* remained distinct from WT and the other mutants at the OR stage in all measurements (Figure 2.1B). Thus, in the OR stage, *nor* and *rin* behaved more similar to WT, suggesting they display more ripening phenotypes after the RR stage.

The distinct phenotypic differences observed between the ripening mutants indicate that each mutation has a unique impact on fruit molecular processes at specific developmental stages. We performed an RNAseq study of WT, *Cnr*, *nor*, and *rin* fruit at the

MG and RR stages to gain insights into the observed phenotypes. A principal component analysis (PCA) was performed using mapped normalized reads to the tomato predicted transcriptome (34,075 genes; SL4.0 release) from WT and mutant samples at MG and RR stages (Figure 2.2A). The PCA revealed that the genotypes were mainly separated by ripening stage (PC1, 60% variance) and that *Cnr* was distinct from WT and the other mutants (PC2, 23% variance). Remarkably, *Cnr* displayed the most similar pattern to WT across PC1 than any other mutant. Like their phenotypes suggested, *nor* and *rin* transcriptomic profiles showed little change between the MG and RR stage and clustered with the WT MG fruit. The separation driven by PC2 supported our observations that *Cnr* fruit was phenotypically different from other genotypes.

### 2.5.2 *Cnr* Fruit Display Transcriptional Differences From Wild Type Before Ripening

Because *Cnr* showed deviation from WT at the MG stage in both phenotype and transcriptional profiles, we hypothesized that gene expression across the genome was affected prior to the MG stage. To determine when the transcriptional profile of *Cnr* began to diverge from WT and other mutants, we obtained and reanalyzed raw RNAseq data from all genotypes at four early stages of fruit growth and development (7, 17, 27, and 37 dpa) (Lü et al., 2018). We performed a PCA for each developmental stage and found that *Cnr* was separated from other genotypes as early as 7 dpa in fruit development, while *nor* and *rin* were similar to WT throughout early development (Supplementary Figure 2.S2). When evaluating differentially expressed genes (DEGs,  $\text{textit{P}_{adj} \leq 0.05}$ ) between *Cnr* and WT fruit at 7 dpa, we detected 1,320 mutation-related DEGs while *nor* and *rin* had only 173 and 392, respectively (see Supplementary Table 5 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). These results suggest that *Cnr* fruit have different gene expression profiles from WT throughout fruit growth and maturation, even before ripening begins.

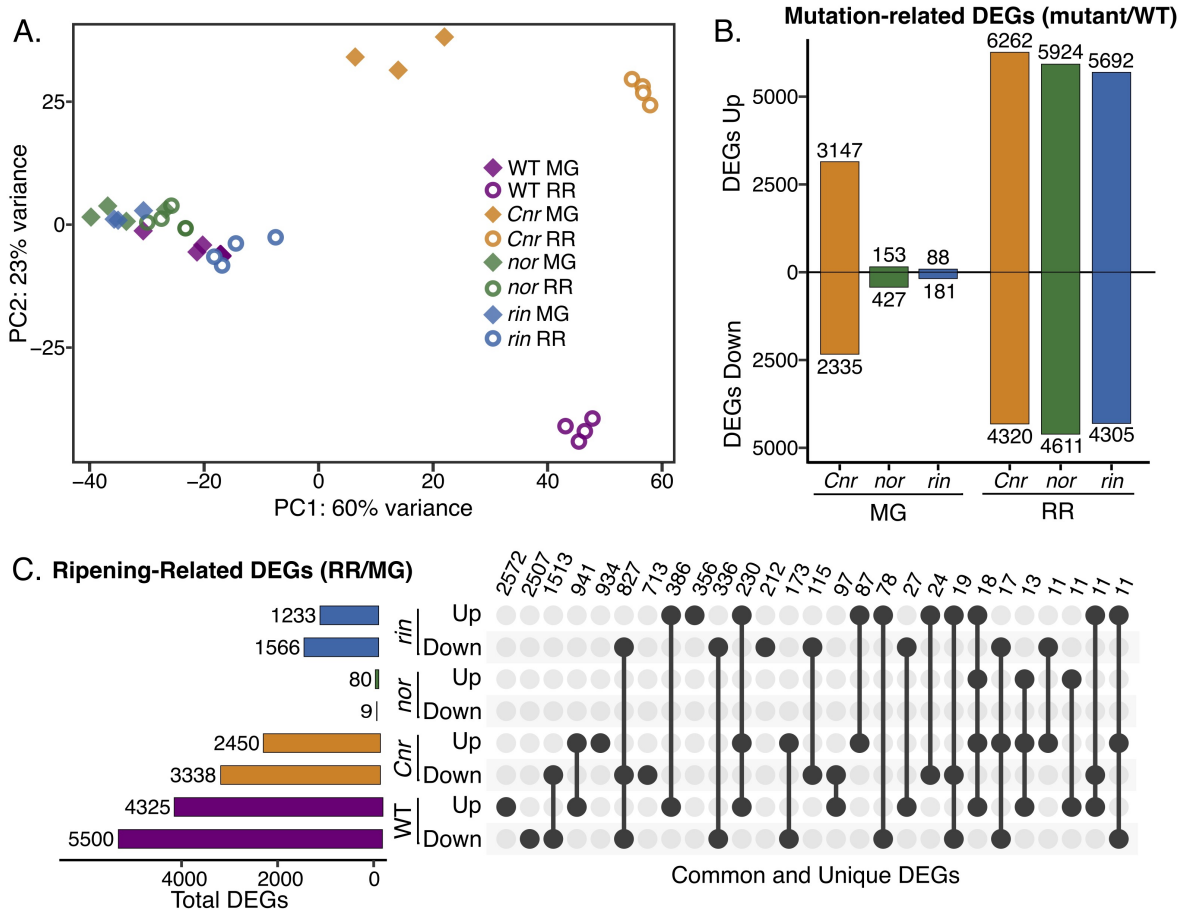


Figure 2.2: (A) Principal component analysis of total mapped RNAseq reads for *Cnr*, *nor*, *rin* and wild type (WT) fruit at the mature green (MG) and red ripe (RR) stages. (B) Number of differentially expressed genes (DEGs;  $P_{adj} \leq 0.05$ ) up- or down-regulated for each mutant compared to WT. (C) Total, unique, and intersecting up- and down-regulated DEGs across ripening (RR/MG) for each genotype visualized using UpSetR. Dots connected by lines indicate common DEGs between categories and single dots indicate unique DEGs, with the number of genes in each category listed above.



### 2.5.3 Transcriptional Misregulation in the Ripening Mutants Leads to Inhibition or Delay of Molecular Processes

We determined DEGs ( $P_{adj} \leq 0.05$ ) from the MG and RR stages to identify specific molecular functions altered in *Cnr*, *nor*, and *rin* fruit. First, we compared the ripening mutants to the WT at each stage and obtained a total of 16,085 mutation-related DEGs across all comparisons (Figure 2.2B). Like the PCA suggested (Figure 2.1A), *Cnr* MG fruit presented the largest amount of mutation-related DEGs (5,482), while *nor* and *rin* MG had considerably fewer DEGs when compared to the WT counterpart (580 and 269 DEGs, respectively). At the RR stage, large differences between each mutant and WT were observed, with *Cnr* RR fruit displaying once again the largest differences in the amount of mutation-related DEGs (10,582, Figure 2.2B). The large number of mutation-related DEGs shown by *Cnr* fruit further supports our hypothesis that the *Cnr* mutation more broadly affects fruit development and that *nor* and *rin* appear to be more ripening-specific mutations.

We examined molecular functions based on KEGG annotations that were significantly ( $P_{adj} \leq 0.05$ ) enriched among the mutation-related DEGs for each *Cnr*, *nor*, and *rin* fruit at MG and RR (Supplementary Figure 2.S3). Large differences in enriched functions were detected in the *Cnr* MG fruit, which mainly corresponded to alterations in carbohydrate and amino acid metabolism, chlorophyll, and carotenoid biosynthesis, and interestingly many processes related to DNA replication and repair. The lack of green color in *Cnr* MG fruit could be explained by lower expression of photosynthesis and carbon fixation genes. The *nor* MG and *rin* MG fruit showed few alterations compared to WT and were mainly noted in amino acid metabolism and plant hormone signal transduction. In contrast, at the RR stage, the three ripening mutants showed significant alterations across multiple molecular pathways that range from primary and secondary metabolism to transcription, translation, and signaling processes.

We proceeded to mine the mutation-related DEGs for key genes known to affect the fruit traits evaluated in the ripening mutants: color, firmness, TSS, and acidity. We selected five carotenoid biosynthesis genes involved in fruit pigmentation, six genes encoding

cell wall degrading enzymes (CWDEs) that promote fruit softening, four genes related to sugar accumulation and transport that impact the fruit's TSS, and one gene that regulates the levels of citric acid then affecting the fruit's acidity (Table 1). At the MG stage, we observed that *Cnr* fruit showed significantly lower expression than WT for several of these key genes, consistent with our phenotypic data (Figure 2.1), including firmness related enzymes and carotenoid biosynthesis genes. MG fruit from the three ripening mutants showed significantly lower gene expression in an important invertase in fruit (SISUCR), which may contribute to the lower levels of TSS observed in all the mutants (Table 2.1; (Klann et al., 1993)). At the RR stage, most of the fruit trait-associated genes surveyed in the ripening mutants had a significantly lower expression than WT, in support of the phenotypic data and reinforced by the numerous functional enrichments among the mutation-related DEGs (Supplementary Figure 2.S3). The critical carotenoid biosynthesis gene that encodes PHYTOENE SYNTHASE 1 (PSY1) was significantly lower expressed than WT in the mutant fruit across all stages, accounting for the lack of red pigmentation at the RR stage. Also, downstream genes in the pathway encoding Lycopene  $\beta$ -cyclases (SILCY1 and SILCY2) were highly expressed in the mutants at the RR stage, suggesting that not only was less lycopene being produced but more was being metabolized. CWDEs were negatively affected across all genotypes, with *Cnr* having the most mutation-related DEGs in this category.

We were interested in examining if the *Cnr*, *nor*, and *rin* mutant fruit displayed altered ripening progression or if they were completely inhibited or delayed in ripening events. We performed another set of differential expression analyses comparing RR against MG fruit for WT and each of the mutants to reveal ripening-related DEGs. As anticipated, WT had the largest number of ripening-related DEGs (9,825), while *nor* showed almost no change between the two ripening stages with only 89 DEGs detected (Figure 2.2C). *Cnr* and *rin* had fewer ripening-related DEGs compared to WT but still exhibited significant changes during the transition between stages with 5,788 and 2,799 DEGs, respectively. Although *Cnr* showed the most differences from WT in mutation-related DEGs (Figure 2.2B), it had the largest number of ripening-related DEGs (2,454) in common with

WT fruit (Figure 2.2C). *Cnr* also displayed similar functional enrichments ( $P_{adj} \leq 0.05$ ) to WT among their respective ripening-related DEGs, including photosynthesis-related pathways, carbohydrate, and amino acid metabolism, and plant hormone signal transduction (Supplementary Figure 2.S4). Compared to *Cnr*, *rin* shared a smaller number of ripening-related DEGs (722) and functional enrichments with WT fruit (Figure 2.2C and Supplementary Figure 2.S4). The number of ripening-related DEGs shared between *nor* and WT fruit was negligible, and no functional enrichments were detected in this set of DEGs.

Similar to our previous analysis, we mined the ripening-related DEGs to determine the patterns of expression of key genes involved in fruit quality traits (Table 2.1). We observed that *Cnr* and WT showed similar gene expression of *SIPSY1*, *SILCY1*, POLYGALACTURONASE 2A (*SIPG2A*), pectate lyase (*SIPL*), PECTIN METHYLESTERASE 1 (*SIPME1*), and ACTINATE HYDRATASE (*SIACO*). Fruit from *nor* and *rin* did not have similar ripening expression patterns to WT fruit for those genes, except for the *SIPG2A* and *SIPME1* in *rin*. Altogether, these data indicate that *Cnr* fruit undergo the most similar ripening progression to WT fruit, while *nor* and *rin* fruit have moderate to minimal changes between the MG and RR stages.

#### 2.5.4 Ripening Mutants Present Alterations in Hormone Networks

Alterations in transcriptional and hormone control likely cause the extensive gene expression differences that lead to the pleiotropic ripening defects in the mutants. Our transcriptional data pointed out that both mutation-related and ripening-related DEGs were significantly enriched ( $P_{adj} \leq 0.05$ ) in functions related to hormone regulation (see Supplementary Figures 3, 4 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Thus, we decided to look closer at defects in hormone biosynthesis and signaling in the mutant fruit, with a particular focus on ethylene and ABA as they are known to promote tomato ripening (Kumar et al., 2014; Mou et al., 2016; Zhang et al., 2009).

It has been reported multiple times that ethylene production is negatively affected in

Table 2.1: Differential expression of key genes associated with tomato fruit traits in the single ripening mutants *Cnr*, *nor*, and *rin*.

| Fruit trait        | Gene accession        | Gene name                     | Mutation Comparison<br>(Log <sub>2</sub> FC mutant/WT) |               |               |               |                   |               | Ripening comparison<br>(Log <sub>2</sub> FC RR/MG) |            |            |
|--------------------|-----------------------|-------------------------------|--|---------------|---------------|---------------|-------------------|---------------|--|------------|------------|
|                    |                       |                               | <i>Cnr</i> MG  | <i>nor</i> MG | <i>rin</i> MG | <i>Cnr</i> RR | <i>nor</i> /WT RR | <i>rin</i> RR | WT   | <i>Cnr</i> | <i>nor</i> |
|                    | <i>Solyc10g080210</i> | Polygalacturonase             | -3.39  | -4.40         | -3.58         | -5.03         | -11.38            | -8.33         | 7.54   | 5.90       | 2.80       |
| Firmness (CAZymes) | <i>Solyc03g111690</i> | Pectate lyase                 |  |               |               | -2.23         | -4.14             | -4.64         | 3.05   | 1.42       |            |
|                    | <i>Solyc12g008840</i> | β-galactosidase 4             |  |               |               | -1.97         | -2.89             | -1.87         | 1.71   |            |            |
|                    | <i>Solyc07g064170</i> | Pectin methylesterase 1       | -8.27  |               |               | -8.66         |                   | -1.02         | -1.21  | -1.60      | -2.41      |
|                    | <i>Solyc07g064180</i> | Pectin methylesterase 2       | -5.53  |               |               | -7.65         |                   |               |  | -2.60      | -1.38      |
|                    | <i>Solyc01g008710</i> | Mannan endo-1,4-β-mannosidase | -9.69  | -9.11         | -4.11         | -7.63         | -5.66             |               |  |            | 3.85       |
|                    | <i>Solyc03g031860</i> | Phytoene synthase 1           | -2.41  | -1.93         | -1.81         | -4.08         | -3.97             | -5.33         | 3.32   | 1.65       | 1.29       |
| Color              | <i>Solyc03g123760</i> | Phytoene desaturase           | -0.63  |               |               | -0.70         |                   |               |  |            |            |
|                    | <i>Solyc01g097810</i> | ζ-carotene desaturase         | -0.81  |               |               | -1.45         | -1.19             | -1.02         | 0.96   |            |            |
|                    | <i>Solyc04g040190</i> | Lycopene β-cyclase            | -0.76  |               |               |               | 2.15              | 1.94          | -2.02  | -1.19      |            |
|                    | <i>Solyc10g079480</i> | Lycopene β-cyclase            |  |               |               | 4.07          | 4.15              | 4.48          | -3.73  |            |            |
|                    | <i>Solyc03g083910</i> | Sucrose accumulator           | -2.50  | -2.33         | -2.27         | -3.88         | -5.57             | -5.85         | 2.52   |            |            |
| Sugar              | <i>Solyc11g017010</i> | SISUT1                        | 1.59   |               |               | 3.67          |                   |               |  | 1.76       |            |
|                    | <i>Solyc05g007190</i> | SISUT2                        |  |               |               | 1.48          |                   |               |  | 1.08       |            |
|                    | <i>Solyc04g076960</i> | SISUT4                        | 0.93   |               |               | 2.27          | 0.86              | 2.01          | -1.29  |            | 0.62       |
| Acidity            | <i>Solyc12g005860</i> | Aconitase                     |  |               |               | -0.58         | -1.71             | -1.55         | 1.38   | 1.14       |            |

Two comparisons were performed: one to capture differences between mutant vs. wild type (WT) fruit at the mature green (MG) and red ripe (R) stages, and the other to detect differences across ripening (RR vs. MG) in the WT and mutant fruit. Only significant fold changes ( $Log_2 \leq .05$ ) are shown.

the *Cnr*, *nor*, and *rin* mutants (Giovannoni, 2007; Li et al., 2019a; Liu et al., 2015). We confirmed that the three ripening mutants do not present the ethylene burst associated with climacteric fruit ripening at any of the stages evaluated, MG, RR, and OR (Figure 2.3). However, in a one-way ANOVA and Tukey HSD test comparing all genotypes at the MG stage we noted that *Cnr* fruit produced significantly ( $P = 0.0004$ ) more ethylene at the MG stage than WT MG fruit and the other mutants at the equivalent stages. We validated these results with field-grown tomatoes across four field seasons. The results from each season can be found in Supplementary Table 2 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). To give a sense if ethylene was inhibited at an early step in biosynthesis in the mutants, we measured the accumulation of the immediate ethylene precursor ACC at the MG and RR stages. ACC accumulates typically at the RR stage in WT fruit, reflecting the increase in ethylene biosynthesis and ethylene production. Surprisingly, ACC concentrations also increased in *Cnr* and *rin* fruit during ripening, reaching values similar to WT fruit (Figure 2.3B); yet the fruit did not produce normal ethylene levels. Moreover, the ACC accumulation in *Cnr* RR fruit was the highest across all genotypes and ripening stages, significantly more than WT RR fruit. These results suggest that the low levels of ethylene in *Cnr* and *rin* RR fruit may be partially explained by inhibition of the final enzymatic step in ethylene biosynthesis.

We found ethylene biosynthesis significantly enriched ( $P_{adj} \leq 0.05$ ) among mutation-related and ripening-related DEGs in several of the mutants (Figure 3A). At the MG stage, *nor* and *rin* fruit had significantly lower expression of the primary ripening ACC synthases (*SLACS2* and *SLACS4*) and ACC oxidases (*SLACO1* and *SLACO4*). At RR, this pattern was maintained except for *SLACO4*, which was higher than WT for both mutants. *SLACS2* was significantly down-regulated across all mutants and stages compared to WT. We validated the expression patterns of *SLACS2* by RT-qPCR experiment using independent samples from WT and the mutant fruit obtained from another field season (Figure 2.3C). We included fruit at the T and OR stages in the validation experiment to capture the gene expression dynamics across fruit ripening and senescence.

In *Cnr* MG fruit, *SLACS4* was significantly lower expressed than WT, like the other

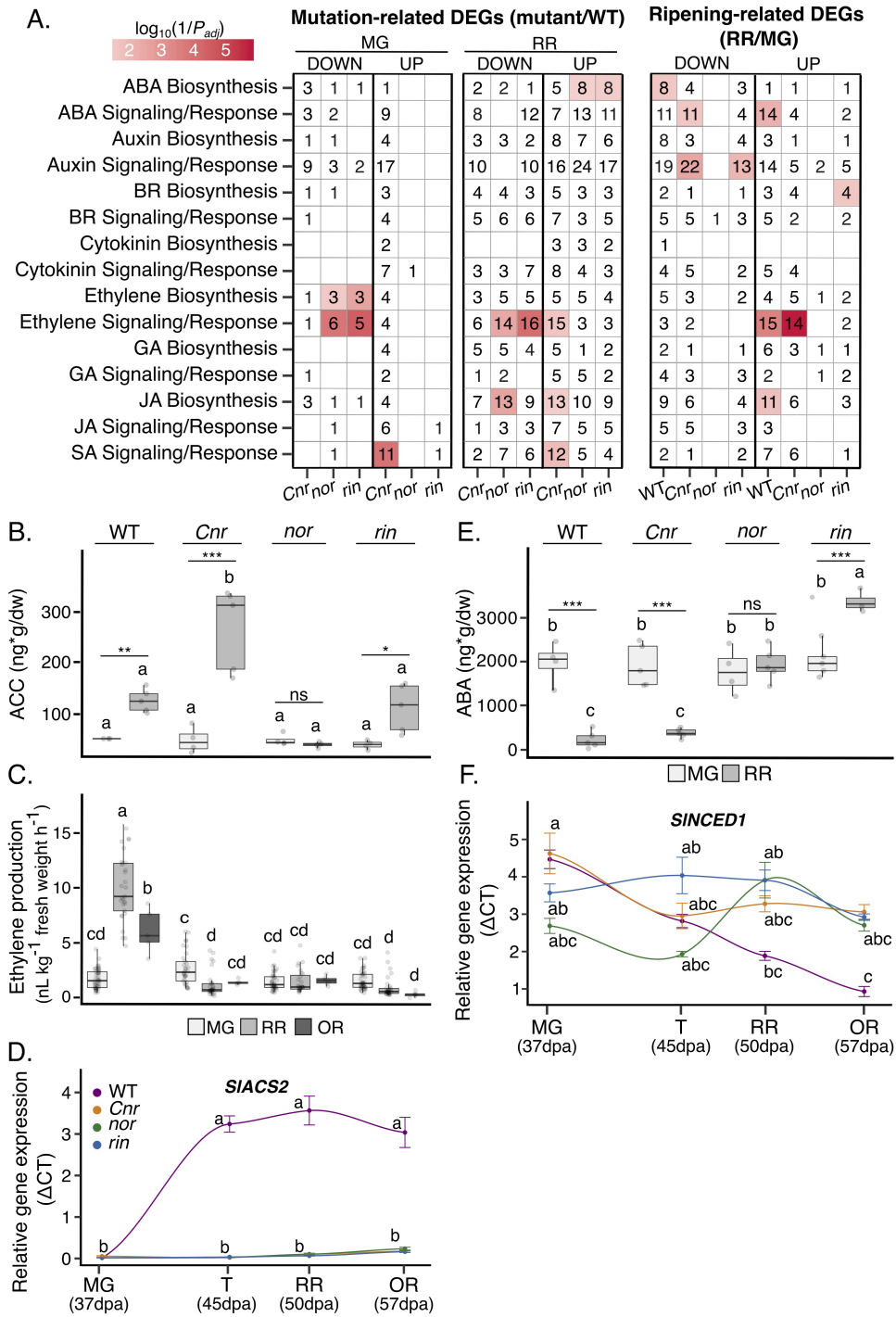


Figure 2.3: Caption presented on following page.

Figure 2.3: (A) Functional enrichments in hormone functions among differentially expressed genes (DEGs;  $P_{adj} \leq 0.05$ ) in two comparisons: mutation-related DEGs obtained when comparing each mutant to the wild type (WT) at the mature green (MG) and red ripe (RR) stages, and ripening-related DEGs when the RR stage was compared against MG for each genotype. Each comparison is separated into significant down- and up-regulated DEGs. The heat map colors indicate the significance of the functional enrichment using a  $\log_{10}(1/P_{adj})$  scale. Numbers in each tile indicate the number of DEGs within each category. Only significant ( $P_{adj} \leq 0.05$ ) functional enrichments are shown. Hormone measurements of the (B) ethylene precursor 1-aminocyclopropane-1-carboxylate (ACC) ( $n = 4-6$ ), (C) ethylene ( $n = 5-45$ ), and (E) abscisic acid (ABA) ( $n = 4-6$ ) for WT, *Cnr*, *nor*, and *rin* fruit at the MG, RR, and/or overripe (OR) stages. Relative gene expression by RT-qPCR of key hormone biosynthesis genes of (D) ethylene and (F) ABA across ( $n = 6$ ) across four ripening stages MG, turning (T), RR, and OR for each genotype. Error bars represent standard error between biological replicates of each sample. Letters in (B-F) indicate significant differences among genotypes and ripening stages calculated by ANOVA and Tukey HSD ( $P \leq 0.05$ ). Asterisks in (B,C,E) denote significant differences ( $*P \leq 0.05$ ,  $**P \leq 0.01$ ,  $***P \leq 0.001$ ) between two ripening stages within a single genotype calculated by Student's t-test.

mutants, but *SLACS2* showed no significant difference. Interestingly, *Cnr* MG fruit had higher gene expression of four ACC oxidases than WT MG fruit, including *SLACO3*, which is involved in System 1 of ethylene biosynthesis. The increased ACC oxidase expression in *Cnr* MG fruit could explain the high ethylene levels detected in these fruit (Figure 2.3B). At RR, four ACO genes had significantly higher expression than WT, except for *SLACO1* that showed no significant differences in RT-qPCR relative expression shown in Supplementary Table 6 (see <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

Ethylene signaling and response genes, including ETHYLENE INSENSITIVE 3 (EIN3) and EIN3-BINDING F-BOX (EBF) homologs, were generally higher expressed in *Cnr* than WT at MG and RR stages. *Nor* and *rin* displayed the opposite trend, with gener-

ally lower expression than WT at both stages in these genes (see Supplementary Table 4 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). These patterns were also reflected in significant enrichments ( $P_{adj} \leq 0.05$ ) of ethylene signaling and response genes at the RR stage (Figure 2.3). Interestingly, ethylene receptor encoding genes (ETRs) were lower expressed across all genotypes and stages compared to WT. In contrast, ethylene response TFs (ERFs) were generally higher expressed in all genotypes at the RR stage.

We measured ABA levels present in the WT and mutant fruit at the MG and RR stages. A decrease of ABA during ripening was found in WT, consistent with previous reports (Mou et al., 2016). This pattern was also present in *Cnr* fruit. However, in *nor* fruit, ABA remained at the same level across both stages, and *rin* showed a significant increase at the RR stage. ABA biosynthesis was significantly enriched among mutation-related DEGs in *nor* and *rin* RR fruit, consistent with the high ABA levels observed (Figure 2.3A). We looked at specific ABA biosynthesis genes enriched in *nor* and *rin* that were also down-regulated in WT at the RR stage and found *SINCED1*, encoding the 9-cis-epoxycarotenoid dioxygenase that catalyzes the rate-limiting step in ABA biosynthesis (Ji et al., 2014). We validated the expression of *SINCED1* with independent samples and additional stages (T and OR) using RT-qPCR (Figure 2.3C). We also confirmed the expression of an upstream biosynthesis gene, *SIZEP*, encoding a zeaxanthin epoxidase, which was also significantly up-regulated in *nor* at the RR stage and in *rin* at the T stage (see Supplementary Table 6 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). While *Cnr* accumulated ABA, signaling and response genes were altered in at MG and RR fruit, including higher expression in ABRE-binding protein (AREB)/ABRE binding factors (ABFs) at both stages compared to WT. *Nor* and *rin* showed alterations in signaling and response at the RR stage, such as lower expression of receptor protein (PYR/PYL) genes and higher expression of the PP2C phosphatase (see Supplementary Table 4 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

We observed changes in biosynthesis and signaling of other plant hormones impli-



cated in fruit development, such as auxins, cytokinins, jasmonic acid, and brassinosteroids (Figure 2.3A). *Cnr* MG fruit had alterations in all hormone pathways examined, further supporting the differences present in *Cnr* phenotype before ripening begins. At the RR stage, all mutants presented multiple defects in hormone metabolism compared to WT. Ripening-related DEGs with hormone functions displayed a similar expression pattern in WT and *Cnr* fruit, whereas *nor* and *rin* displayed low numbers of ripening-related DEGs from these categories.

### 2.5.5 Ripening Mutations Influence the Expression Dynamics of *CNR*, *NOR*, and *RIN* in Fruit

Another way in which the mutations in the *CNR*, *RIN*, and *NOR* may affect gene expression of ripening processes is through direct or indirect interactions with each other. We performed RT-qPCR on fruit from the MG, T, RR, and OR stages in each genotype for each of the genes encoding the ripening TFs (Figure 2.4). In WT, each TF follows a ripening pattern, peaking in expression at the T stage. Mutations in any of the three TFs led to a decrease or delay in the expression of the other TFs compared to WT. For example, *RIN* expression does not begin to show an increase until the OR stage for *nor* and *Cnr*. A similar pattern was exhibited in *CNR* expression for *nor* and *rin* and *NOR* expression in *rin* and *Cnr*. The *Cnr* fruit displayed the most dramatic decreases in expression across the TFs, while the *nor* fruit showed the most delays.

### 2.5.6 Phenotypic Differences in Double Mutants Reveal Genetic Relationships

The changes in gene expression of *CNR*, *NOR*, and *RIN* in the ripening mutants indicate that the genes are interconnected during fruit development. In addition, *Cnr* consistently showed earlier defects in fruit traits, gene expression, and hormone pathways. To characterize the combined genetic effects of the mutations on tomato fruit, we generated homozygous double mutants through reciprocal crosses of the single mutants. We then phenotyped the double mutants for fruit traits and ethylene production (Figure 5). Because the reciprocal crosses produced fruit indistinguishable from each other, we report

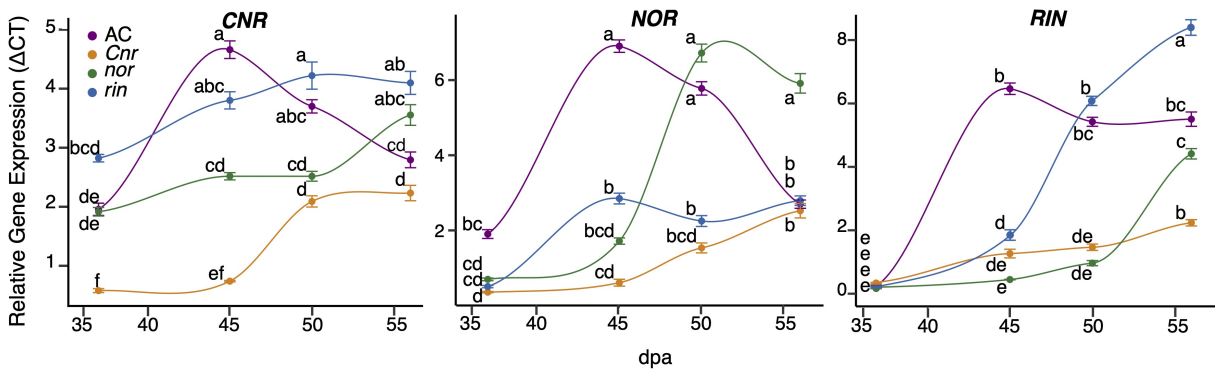


Figure 2.4: Relative gene expression of the three transcription factors across ripening in the wild type (WT), *Cnr*, *nor*, and *rin* genotypes measured by quantitative reverse transcription (qRT)-PCR. The measurements were done with fruit ( $n = 6$ ) collected at 37 days post anthesis (dpa), equivalent to the mature green stage (MG), 45 dpa, equivalent to the turning stage (T), 50 dpa, equivalent to the red ripe stage (RR), and 57 dpa, equivalent to the overripe (OR) stage. Error bars represent standard error between biological replicates of each sample. Letters indicate significant differences among genotypes and ripening stages calculated by ANOVA and Tukey HSD ( $P \leq 0.05$ ).

them as only one double mutant (see Supplementary Table 7 and Supplementary Figure 5 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Fruit of *nor/rin* double mutants were almost indistinguishable from both *nor* and *rin* fruit in appearance and external color. Fruit resulting from any cross with *Cnr* as a parent presented similar visual characteristics (Figure 5A). We also performed a PCA of the color measurements to compare the double mutants to their parental lines at the RR stage and confirmed this observation (Figure 2.5B and see Supplementary Table 7 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Based on these observations and our earlier phenotypic and transcriptional data, we confirmed that the *Cnr* mutation affects early fruit development. In contrast, the *nor* and *rin* mutations act during fruit ripening.

If defects in *Cnr* occur earlier in fruit development than those caused by *nor* or *rin*, we expected the *Cnr/rin* and *Cnr/nor* double mutants to behave similarly to *Cnr* and

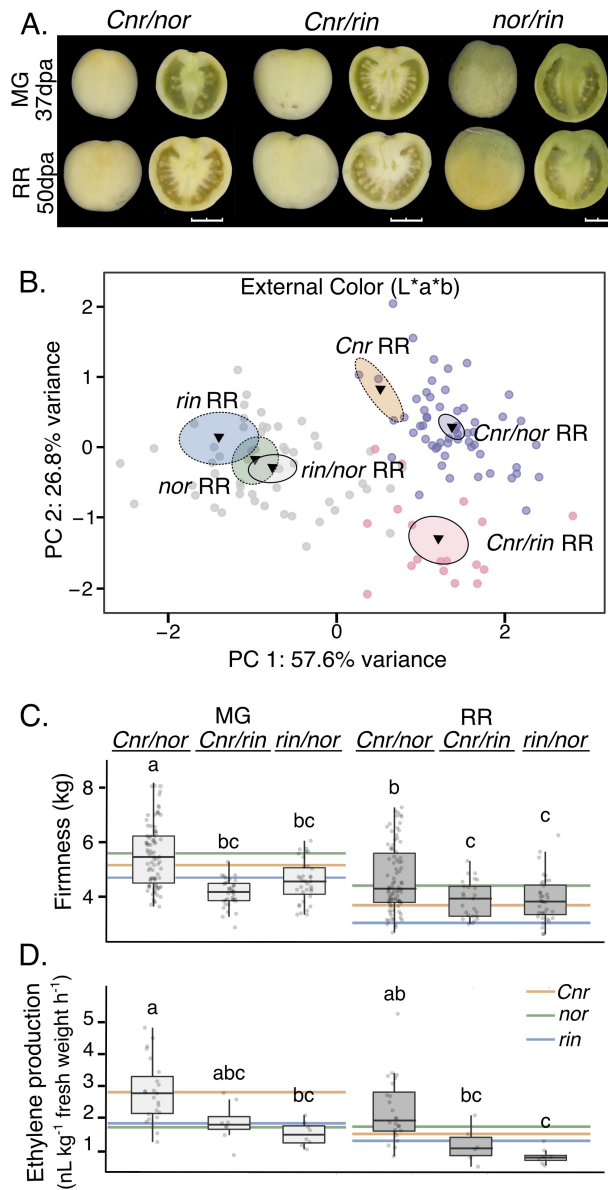


Figure 2.5: (A) Homozygous double mutants pictured at the mature green (MG) and red ripe (RR) stages. Fruit shown whole at left and in longitudinal sections at right. Images were extracted and processed with the VideometerLab instrument. Bar represents 2 cm. (B) Principal component analysis of fruit external color ( $n = 18-54$ ). The center of gravity is represented by a triangle with surrounding ellipses indicating 95% confidence interval. Dashed ellipses indicate the values of the single mutant parents. (C) Texture analysis of (caption continued on following page)

Figure 2.5: (Continued from previous page) fruit firmness at MG and RR stages ( $n = 25-86$ ). (D) Ethylene production of MG and RR fruit ( $n = 6-24$ ). Letters indicate significant differences among genotypes and stages ( $P \leq 0.05$ ). Colored lines indicate averages of the parents at each stage for comparison.

display similar phenotypes (Figures 2.5C,D). *Cnr/rin* fruit were significantly ( $P \leq 0.05$ ) less firm than either parent at the MG stage but performed most similarly to *Cnr* at the RR stage. *Cnr/nor* fruit was not distinguishable from either parent in firmness at MG but was firmer ( $P \leq 0.05$ ) than *Cnr* RR fruit. Interestingly, *Cnr/nor* fruit exhibited high ethylene production at the MG stage like the *Cnr* fruit. At the RR stage, *Cnr/nor* showed a less pronounced decrease in ethylene production, resulting in higher hormone levels than either parent. Although some phenotypic differences were detected, we verified that *Cnr/rin* and *Cnr/nor* resembled the *Cnr* parent for most of the fruit traits measured.

If *nor* and *rin* act synergistically during ripening, the *rin/nor* double mutants would have a more extreme phenotype than either on their own. At the MG stage, *rin/nor* fruit firmness was statistically similar to *rin* ( $P \leq 0.05$ ; Figure 2.5C) but became an intermediate phenotype at the RR stage. For ethylene, *rin/nor* fruit produced less than either parent at both stages, although not significant, suggesting a combined effect of both mutations.

### 2.5.7 Double Mutant *Cnr/nor* Shows Gene Expression Unique From Both Parents

The strong effect of *Cnr* in the double mutant phenotypes led us to investigate if gene expression in the fruit was altered in a similar way. We selected the *Cnr/nor* double mutant to perform an RNAseq experiment of fruit at MG and RR stages and assessed the overall transcriptional changes resulting from the two mutations combined. We conducted a PCA of total mapped reads for MG and RR fruit of *Cnr/nor* and the single mutant parents (Figure 2.6A). In this analysis, *Cnr/nor* expression appeared more similar to *Cnr* than *nor* in PC1 (66% of variance), but PC2 (20% of variance) accounted for differences between *Cnr* and *Cnr/nor*.

We analyzed mutation-related DEGs ( $P_{adj} \leq 0.05$ ) in the *Cnr/nor* fruit by comparing the gene expression patterns of the double mutant against WT at both MG and RR stages. We then determine which of these mutation-related DEGs were also differentially expressed between the single mutant parents and WT (Figure 2.6B). Similar to *Cnr*, *Cnr/nor* fruit started with a high number of mutation-related DEGs (10,643) at the MG stage, showing defects in development before the initiation of ripening. However, *Cnr/nor* MG and RR fruit showed more mutation-related DEGs than either *nor* or *Cnr* fruit, including 634 unique DEGs at the MG stage and 948 at the RR stage. These data indicate that the *Cnr/nor* fruit present additional defects than *Cnr* fruit prior to ripening.

Interestingly, *Cnr/nor* and both its parents at the RR stage shared many mutation-related DEGs (1,980) (Figure 2.6B). These shared mutation-related DEGs were significantly ( $P_{adj} \leq 0.05$ ) enriched in glycolysis, starch and sucrose metabolism, fructose and mannose metabolism, among others, suggesting that carbohydrate metabolism is altered in all three genotypes (see Supplementary Table 4 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). When we looked again at the key genes associated with fruit traits, we observed greater defects in the double mutant compared to the single mutant parents (see Supplementary Table 8 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

We only identified 272 ripening-related DEGs ( $P_{adj} \leq 0.05$ ) in *Cnr/nor* fruit, indicating that the double mutant fruit changed very little between the MG and RR stages (see Supplementary Table 4 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). This inhibition of ripening progression is similar to the ripening-related DEG patterns exhibited by *nor* (Figure 2.2), highlighting a critical difference between *Cnr/nor* and *Cnr*. Overall, the transcriptional data indicate that *Cnr/nor* have stronger alterations in fruit development and more significant inhibition of fruit ripening than the *Cnr* and *nor* fruit.

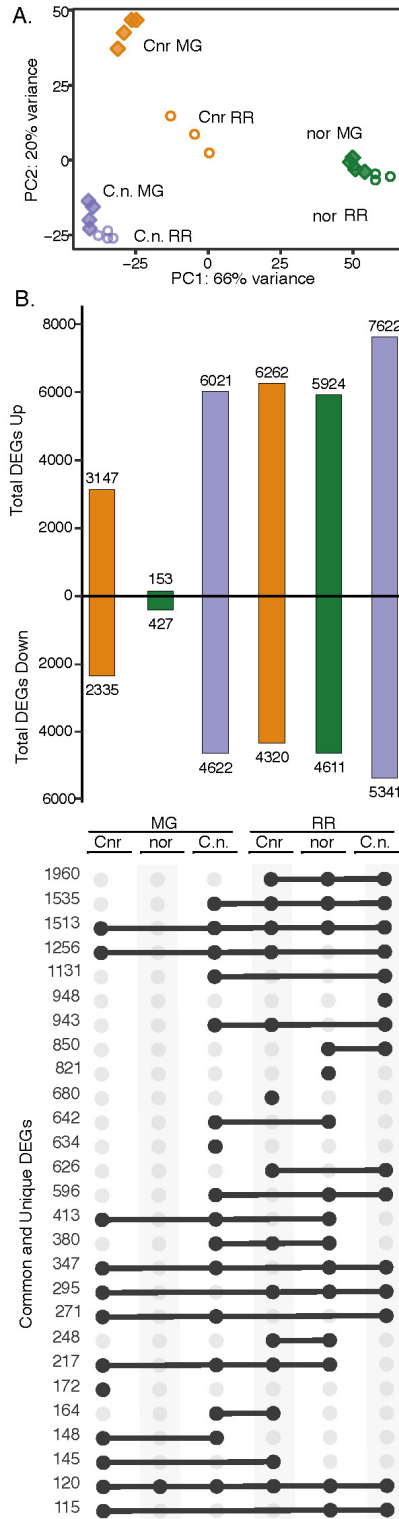


Figure 2.6: Comparison of Cnr/nor double mutant gene expression to parents.(Caption presented on following page.)

Figure 2.6: (Continued from previous page) (A) Principal component analysis of total mapped RNAseq reads for the double mutant *Cnr/nor* (C.n.), and the parents *Cnr* and *nor* at the mature green (MG) and red ripe (RR) stages. (B) Total, unique, and intersecting differentially expressed genes (DEGs) compared to wild type (WT) for each genotype (mutant/WT) visualized using UpSetR. Dots connected by lines indicate common DEGs between categories and single dots indicate unique DEGs, with the numbers of shared or unique genes at the left.

## 2.6 Discussion

The spontaneous ripening mutants, *Cnr*, *nor*, and *rin*, are essential genetic tools to untangle the complexity of climacteric fruit ripening (Chen et al., 2020) and to breed for extended shelf-life or field harvest traits in tomato (Garg et al., 2008; Kitagawa et al., 2005). However, thorough phenotyping of the fruit traits affected by these mutants using plants grown under field conditions has been neglected. Here, we produced an extensive quantitative study of fruit quality in the tomato ripening mutants and corroborated it across multiple field seasons. We were able to carefully describe physiological and molecular differences between the mutants by sampling large numbers of fruit and surveying distinct stages through ripening in ways not feasible with greenhouse experiments.

### 2.6.1 Delay or Inhibition of Ripening Events Vary in *nor* and *rin*

We determined that some ripening events in the mutants *nor* and *rin* were not completely blocked but severely delayed. By examining the OR stage, we found that the mutation in *nor* may strongly affect firmness and taste while pigment accumulation was only delayed and slightly perturbed (Figure 2.1). These phenotypes were supported by higher expression of carotenoid biosynthesis genes in *nor* RR than WT and an increase in *SLPSY1* between the MG and RR stages (Table 2.1 and Supplementary Figure 2.S3). The accumulation of pigments in *nor* fruit, particularly at late stages in development, has gone unnoticed in previous studies, but it partially resembles the CRISPR-NOR mutants (Gao et al., 2019; Wang et al., 2019). In contrast, *rin* fruit showed strong inhibition of pigment accumulation but less dramatic alterations to fruit taste-related traits, only delaying the

accumulation of sugars and decrease in acidity (Figure 2.1). The lack of upregulation of *SlPSY1* in *rin* (Table 2.1) appears to contribute to the color defects, consistent with evidence that RIN directly regulates this gene (Fujisawa et al., 2013). Both *nor* and *rin* exhibited severe delays or inhibition of ripening-related gene expression changes. While highly similar to WT at the MG stage, *nor* and *rin* fruit showed large deviations from WT at the RR stage (Figure 2.2B). In fact, the gene expression profiles of *nor* and *rin* RR fruit remained similar to those from WT MG fruit.

The physiological data generated in this study show *nor* and *rin* mutations have different impacts on fruit quality traits. Soluble solids and acid accumulation are negatively impacted in both mutants, but more dramatically in *nor* fruit. In addition, previous reports have demonstrated a similar pattern among volatile profiles of the mutants at the red ripe stage, with *rin* again showing more similarity to WT in flavor related traits (Kovács et al., 2009). This suggests *rin* fruit are less likely to hinder flavor profiles than *nor* fruit when breeding for fresh-market hybrid varieties with extended shelf-life. Although *nor* showed lower quality flavor attributes, its coloration at overripe stages was most similar to WT compared to *rin*; and thus, it can be useful in breeding hybrid varieties when coloration is a critical fruit trait, such as in the case of processing tomato varieties. Overall, this knowledge will provide valuable information on these tradeoffs of using either loci for breeding programs.

Because the *Cnr*, *nor*, and *rin* mutants never acquire equivalent colorations to WT, their ripening stages have been determined based on the fruit's age expressed as days after anthesis (dpa) or days after the breaker (BR) stage. Sometimes described as BR + 7 days, the RR stage has been the primary developmental time employed for studying the ripening mutants. As we showed here, the OR stage could provide better comparisons against WT RR fruit for mutants with delayed ripening phenotypes. We demonstrated that in the *nor* fruit, the *RIN* and *CNR* genes only begin to increase in expression in a way comparable to WT at the OR stage (Figure 2.4). This observation corresponds to over a 10-day delay for some of the ripening processes to begin. The delayed ripening events observed in the OR fruit have not been described before in the spontaneous *nor*



mutant.

### 2.6.2 *Cnr* Is More Than a Ripening Mutant

Although the *Cnr* mutant has been assumed to have normal fruit development before ripening (Lai et al., 2020), there have been indications that the *Cnr* mutant displays defects that are not ripening-specific, such as earlier chlorophyll degradation and altered expression of CWDE (Eriksson et al., 2004; Wang et al., 2020b). We showed that the *Cnr* mutation causes substantial defects in fruit prior to ripening as seen through statistically significant deviations in fruit size, color, firmness, and TA, ethylene production, and gene expression at the MG stage (see Supplementary Table 2 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full> and Figures 2.1, 2.2, 2.3). Therefore we propose *Cnr* may be more accurately described as a developmental mutant and not exclusively a ripening mutant. Further complementing these results, the *Cnr* fruit displayed large transcriptional deviations from WT that can be traced back as far as 7 dpa (Figure 2.2 and Supplementary Figure 2.S2). These early development defects are likely a result of reduced CNR expression in the mutant, which is typically expressed in locular tissue before fruit maturity (Giovannoni et al., 2017).

Our analysis of ripening-related gene expression in *Cnr* showed striking similarities to WT in the number and functions of genes changing between stages. Moreover, 69.5% of ripening-related DEGs in *Cnr* were shared with WT (Figure 2.2). These results further support the hypothesis that *Cnr* is not exclusively a ripening mutant. Instead, *Cnr* fruit undergoes gene expression changes consistent with WT “ripening.” However, the ripening-related changes in gene expression that occur in *Cnr* are not enough to compensate for the large defects accumulated in the fruit during growth and maturation. In a recent report, a knockout mutation to the gene body of CNR yielded little visible effects on fruit development and ripening (Gao et al., 2019), which suggests that the *Cnr* mutant phenotype may result from more than just a reduced expression of the CNR gene as previously reported (Manning et al., 2006). It has also been demonstrated that *Cnr* fruit have genome-wide methylation changes that inhibit ripening-related gene expression (Zhong et al., 2013). The developmental defects observed in *Cnr* are likely caused by these

methylation changes, directly or indirectly caused by the *Cnr* mutation (Chen et al., 2018). Thus, to better understand the *Cnr* mutation, more physiological data at earlier stages of development needs to be analyzed and complemented with more in-depth functional analysis of gene expression alterations at the corresponding stages. In addition, further molecular and genetic studies need to be performed and compared against complete *CNR* knockout mutants.

### 2.6.3 The *Cnr* Mutant Produces Ethylene Beyond Basal Levels

Previous reports have shown ethylene levels to be very low or even undetectable in the ripening mutants (Giovannoni et al., 2017). Our data support that the mutants never produce a burst in ethylene production, even at the OR stage where more ripening phenotypes are observed (Figure 2.3B). The orange-red pigmentation in nor OR fruit and the similarities of OR fruit in texture and taste-related attributes to WT RR fruit occur independently of an ethylene burst. These observations evidence that other regulatory mechanisms exist to initiate ripening events outside of ethylene (Li et al., 2019b).

Unlike previous reports, our data consistently showed that *Cnr* presented increased ethylene levels at the MG stage compared to WT (Wang et al., 2020b). Interestingly, *Cnr* fruit produced more of the ethylene precursor ACC than WT at the RR stage. Also, *rin* made equivalent levels to WT fruit. Ethylene biosynthesis is divided into two programs: System 1 produces basal levels of the hormone during development, and System 2 generates the climacteric rise in ethylene during ripening (MCMURCHIE et al., 1972). Each of these systems is catalyzed by a different set of ethylene biosynthetic enzymes (Liu et al., 2015). It is clear that all mutants show defects to System 2 of ethylene biosynthesis, but they also appear to have alterations specific to System 1. For example, we observed that *SLACO3*, a System 1-specific ACC oxidase, was higher expressed in *Cnr* fruit than WT (see Supplementary Table 6 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>).

## 2.6.4 ABA Biosynthesis and Accumulation Is Affected in *nor* and *rin*

The role of ABA in climacteric ripening is not as well explored but has been reported to be complementary to ethylene (Ji et al., 2014). Previous reports in WT fruit have shown that ABA increases until the breaker stage, just before the ethylene burst (Mou et al., 2016; Zhang et al., 2009). ABA has also been shown to induce ethylene production and linked to the NOR transcription factor (Mou et al., 2018). We found that *nor* and *rin* fruit did not show decreases in ABA concentration during ripening like WT did (Figure 2.3). For *nor*, the constant levels of ABA between MG and RR stages are another example of how fruit ripening events are delayed or inhibited. RIN and ABA have been demonstrated to have an inverse relationship where RIN expression is repressed with the induction of ABA (Diretto et al., 2020). The significant increase of ABA accumulation in *rin* during ripening suggests that ABA biosynthesis and metabolism are misregulated in this mutant. *rin* fruit appear to present a delayed peak in ABA levels compared to WT fruit. Our results support the indirect interaction between the TFs and ABA during ripening. More developmental stages, genetic manipulations, and exogenous hormone treatments are needed to investigate further the trends of ABA accumulation seen in the ripening mutants.

### 2.6.4.1 CNR, NOR, and RIN Act Interdependently

The interactions between the CNR, NOR, and RIN in ripening have been debated in the literature (Chen et al., 2020). The TF RIN directly interacts with NOR and CNR, binding to their respective promoters, and therefore has been proposed to be the most upstream TF among the three regulators (Fujisawa et al., 2013). Here we provided evidence that the three TFs display at least indirect effects on each other. We have argued that the *Cnr* mutant shows a wide breadth of defects across fruit development before ripening begins, and thus, we propose the *Cnr* mutation is acting before NOR or RIN. This further supports the hypothesis made in Wang et al. that *Cnr* acts epistatically to *nor* and *rin* (Wang et al., 2020b). The gene expression patterns of CNR, NOR, and RIN across ripening stages were decreased or delayed in each of the single ripening mutants.

The most substantial variation in gene expression was the downregulation of NOR and RIN expression across all stages in the *Cnr* mutant (Figure 2.4).

We present for the first time double ripening mutants, homozygous for both loci, that can be used to see the combined effects of each mutation on fruit development and quality traits. We successfully generated the double mutants by establishing reliable and high throughput genotyping protocols for each mutation and evaluating segregation of the mutant phenotypes in field trials across multiple growing seasons. We obtained double mutants from both reciprocal crosses but saw no fruit phenotypic differences between them, suggesting that the ripening mutations are not influenced by maternal or paternal effects (see Supplementary Table 7 at <https://www.frontiersin.org/articles/10.3389/fpls.2021.647035/full>). Because the *nor* and *rin* mutants look so similar, it was hard to visually determine the individual effects of each mutation on the appearance of *rin/nor* fruit. However, when specific fruit traits were measured, we could detect additive or intermediate fruit phenotypes in this double mutant, supporting the proposed relationship in Wang et al. ((Wang et al., 2020b); Figure 2.5). Thus, *nor* and *rin* appear to influence similar fruit traits and act in coordination.

The *Cnr* mutation had a significant effect on the *Cnr/nor* and *Cnr/rin* mutants resulting in fruit with similar appearance and ethylene production to the *Cnr* fruit (Figure 2.5). When analyzing the gene expression profiles of the *Cnr/nor* fruit, we also observed multiple similarities to the *Cnr* parent, but also several deviations (Figure 2.6). Surprisingly, *Cnr/nor* was also reminiscent of *nor*, as it displayed few ripening-related gene expression changes, suggesting the inhibition or delay of specific ripening events in *nor* carried over to the double mutant. Here, we proposed that the *Cnr* mutation causes defects throughout fruit development while the *nor* mutation causes defects predominantly in ripening. However, the *Cnr/nor* double mutant showed additional phenotypic and transcriptional defects before ripening than both mutant parents (Figure 2.6). These observations indicate that in combination with *Cnr*, *nor* may contribute to alterations in early fruit development and the inhibition of ripening progression.

## 2.7 Conclusion

Our study contributes new information about the spontaneous tomato ripening mutants, which have been employed to study fruit ripening for at least the past two decades. Also, given the importance of both *nor* and *rin* for tomato breeding, the fruit trait data generated in this study could be applied to improve quality in tomato hybrids or at least identify tradeoffs between fruit traits. Ultimately, our results extend knowledge of underlying genetic and molecular factors affecting fruit ripening and quality while providing insights into fruit physiological changes through ripening and senescence.

## References

- Agar, I., Abak, K., and Yarsi, G. (1994). Effect of Different Maturity Stages on the Keeping Quality of nor (non-ripening), rin (ripening-inhibitor) and Normal Type Tomatoes. *Acta Horticulturae*, (368):742–753.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.
- Blanco-Ulate, B., Vincenti, E., Powell, A. L. T., and Cantu, D. (2013). Tomato transcriptome and mutant analyses suggest a role for plant stress hormones in the interaction between fruit and *Botrytis cinerea*. *Frontiers in Plant Science*, 4:142.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120.
- Casteel, C. L., De Alwis, M., Bak, A., Dong, H., Whitham, S. A., and Jander, G. (2015). Disruption of ethylene responses by Turnip mosaic virus mediates suppression of plant defense against the green peach aphid vector. *Plant Physiology*, 169(1):209–218.
- Chen, T., Qin, G., and Tian, S. (2020). Regulatory network of fruit ripening: current understanding and future challenges. *New Phytologist*, 228(4):1219–1226.
- Chen, W., Yu, Z., Kong, J., Wang, H., Li, Y., Zhao, M., Wang, X., Zheng, Q., Shi, N., Zhang, P., Zhong, S., Hunter, P., Tör, M., and Hong, Y. (2018). Comparative WGBS identifies genes that influence non-ripe phenotype in tomato epimutant Colourless non-ripening. *Science China Life Sciences*, 61(2):244–252.
- Conway, J. R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33(18):2938–2940.
- De, F., Maintainer, M., and De Mendiburu, F. (2017). Package 'agricolae' Title Statistical Procedures for Agricultural Research. *Statistical procedures for agricultural research*.

- Diretto, G., Frusciante, S., Fabbri, C., Schauer, N., Busta, L., Wang, Z., Matas, A. J., Fiore, A., K.C. Rose, J., Fernie, A. R., Jetter, R., Mattei, B., Giovannoni, J., and Giuliano, G. (2020). Manipulation of  $\beta$ -carotene levels in tomato fruits results in increased ABA content and extended shelf life. *Plant Biotechnology Journal*, 18(5):1185–1199.
- Eriksson, E. M., Bovy, A., Manning, K., Harrison, L., Andrews, J., De Silva, J., Tucker, G. A., and Seymour, G. B. (2004). Effect of the Colorless non-ripening mutation on cell wall biochemistry and gene expression during tomato fruit development and ripening. *Plant physiology*, 136(4):4184–4197.
- Fujisawa, M., Nakano, T., Shima, Y., and Ito, Y. (2013). A large-scale identification of direct targets of the tomato MADS box transcription factor RIPENING INHIBITOR reveals the regulation of fruit ripening. *The Plant cell*, 25(2):371–86.
- Gao, Y., Wei, W., Fan, Z., Zhao, X., Zhang, Y., Jing, Y., Zhu, B., Zhu, H., Shan, W., Chen, J., Grierson, D., Luo, Y., Jemrić, T., Jiang, C. Z., and Fu, D. Q. (2020). Re-evaluation of the nor mutation and the role of the NAC-NOR transcription factor in tomato fruit ripening. *Journal of Experimental Botany*, 71(12):3560–3574.
- Gao, Y., Zhu, N., Zhu, X., Wu, M., Cai-Zhong, J., Grierson, D., Luo, Y., Shen, W., Zhong, S., Fu, D. Q., and Qu, G. (2019). Diversity and redundancy of the ripening regulatory networks revealed by the fruitENCODE and the new CRISPR/Cas9 CNR and NOR mutants. *Horticulture Research*, 6(1):39.
- Garg, N., Cheema, D. S., and Dhatt, A. S. (2008). Utilization of rin, nor, and alc alleles to extend tomato fruit availability. *International Journal of Vegetable Science*, 14(1):41–54.
- Giovannoni, J., Nguyen, C., Ampofo, B., Zhong, S., and Fei, Z. (2017). The Epigenome and Transcriptional Dynamics of Fruit Ripening. *Annual Review of Plant Biology*, 68(1):61–84.
- Giovannoni, J., Tanksley, S., Vrebalov, J., and Noensie, F. (2004). NOR gene composition and methods for use thereof.

- Giovannoni, J. J. (2007). Fruit ripening mutants yield insights into ripening control. *Current Opinion in Plant Biology*, 10(3):283–289.
- Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M., Shima, Y., Nakamura, N., Kotake-Nara, E., Kawasaki, S., and Toki, S. (2017). Re-evaluation of the rin mutation and the role of RIN in the induction of tomato ripening. *Nature Plants*, 3(11):866–874.
- Ji, K., Kai, W., Zhao, B., Sun, Y., Yuan, B., Dai, S., Li, Q., Chen, P., Wang, Y., Pei, Y., Wang, H., Guo, Y., and Leng, P. (2014). SINCED1 and SlCYP707A2: key genes involved in ABA metabolism during tomato fruit ripening. *Journal of experimental botany*, 65(18):5243–55.
- Karlova, R., Chapman, N., David, K., Angenent, G. C., Seymour, G. B., and De Maagd, R. A. (2014). Transcriptional control of fleshy fruit development and ripening. *Journal of Experimental Botany*, 65(16):4527–4541.
- Kassambara, A., Mundt, F., Kassambara, A., and Mundt, F. (2017). Factoextra: extract and visualize the results of multivariate data analyses. *R. Package Version*, 1(5):337–354.
- Kitagawa, M., Ito, H., Shiina, T., Nakamura, N., Inakuma, T., Kasumi, T., Ishiguro, Y., Yabe, K., and Ito, Y. (2005). Characterization of tomato fruit ripening and analysis of gene expression in F1 hybrids of the ripening inhibitor (rin) mutant. *Physiologia Plantarum*, 123(3):331–338.
- Klann, E. M., Chetelat, R. T., and Bennett, A. B. (1993). Expression of Acid Invertase Gene Controls Sugar Composition in Tomato (*Lycopersicon*) Fruit. *Plant Physiology*, 103(3):863 LP – 870.
- Kopeliovitch, E., Rabinowitch, H. D., Mizrahi, Y., and Kedar, N. (1979). The potential of ripening mutants for extending the storage life of the tomato fruit. *Euphytica*.
- Kovács, K., Fray, R. G., Tikunov, Y., Graham, N., Bradley, G., Seymour, G. B., Bovy,



- A. G., and Grierson, D. (2009). Effect of tomato pleiotropic ripening mutations on flavour volatile biosynthesis. *Phytochemistry*, 70(8):1003–1008.
- Kumar, R., Khurana, A., and Sharma, A. K. (2014). Role of plant hormones and their interplay in development and ripening of fleshy fruits.
- Lai, T., Wang, X., Ye, B., Jin, M., Chen, W., Wang, Y., Zhou, Y., Blanks, A. M., Gu, M., Zhang, P., Zhang, X., Li, C., Wang, H., Liu, Y., Gallusci, P., Tör, M., and Hong, Y. (2020). Molecular and functional characterization of the SBP-box transcription factor SPL-CNR in tomato fruit ripening and cell death. *Journal of Experimental Botany*, 71(10):2995–3011.
- Langmead, B. and Salzberg, S. (2013). Bowtie2. *Nature methods*.
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1).
- Li, S., Chen, K., and Grierson, D. (2019a). A critical evaluation of the role of ethylene and MADS transcription factors in the network controlling fleshy fruit ripening. *New Phytologist*, 221(4):1724–1741.
- Li, S., Xu, H., Ju, Z., Cao, D., Zhu, H., Fu, D., Grierson, D., Qin, G., Luo, Y., and Zhu, B. (2018). The RIN-MC fusion of MADS-box transcription factors has transcriptional activity and modulates expression of many ripening genes. *Plant Physiology*, 176(1):891–909.
- Li, S., Zhu, B., Pirrello, J., Xu, C., Zhang, B., Bouzayen, M., Chen, K., and Grierson, D. (2019b). Roles of RIN and ethylene in tomato fruit ripening and ripening-associated traits. *New Phytologist*.
- Liu, R., How-Kit, A., Stammitti, L., Teyssier, E., Rolin, D., Mortain-Bertrand, A., Halle, S., Liu, M., Kong, J., Wu, C., Degraeve-Guibault, C., Chapman, N. H., Maucourt, M., Hodgman, T. C., Tost, J., Bouzayen, M., Hong, Y., Seymour, G. B., Giovannoni, J. J., and Gallusci, P. (2015). A DEMETER-like DNA demethylase governs tomato

- fruit ripening. *Proceedings of the National Academy of Sciences of the United States of America*, 112(34):10804–10809.
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550.
- Lü, P., Yu, S., Zhu, N., Chen, Y.-R., Zhou, B., Pan, Y., Tzeng, D., Fabi, J. P., Argyris, J., Garcia-Mas, J., Ye, N., Zhang, J., Grierson, D., Xiang, J., Fei, Z., Giovannoni, J., and Zhong, S. (2018). Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening. *Nature Plants*, 4(10):784–791.
- Manning, K., Tor, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., Giovannoni, J. J., and Seymour, G. B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat Genet*, 38(8):948–952.
- MCMURCHIE, E. J., MCGLASSON, W. B., and EAKS, I. L. (1972). Treatment of Fruit with Propylene gives Information about the Biogenesis of Ethylene. *Nature*, 237(5352):235–236.
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, 35(SUPPL.2):W182–W185.
- Mou, W., Li, D., Bu, J., Jiang, Y., Khan, Z. U., Luo, Z., Mao, L., and Ying, T. (2016). Comprehensive analysis of ABA effects on ethylene biosynthesis and signaling during tomato fruit ripening. *PLoS ONE*, 11(4):e0154072.
- Mou, W., Li, D., Luo, Z., Li, L., Mao, L., and Ying, T. (2018). SIAREB1 transcriptional activation of NOR is involved in abscisic acid-modulated ethylene biosynthesis during tomato fruit ripening. *Plant Science*, 276(May):239–249.
- Osei, M. K., Danquah, A., E.T., B., Danquah, E., and Adu-Dapaah (2017). An overview

- of tomato fruit-ripening mutants and their use in increasing shelf life of tomato fruits. *African Journal of Agricultural Research*, 12(51):3520–3528.
- Osorio, S., Carneiro, R. T., Lytovchenko, A., McQuinn, R., Sørensen, I., Vallarino, J. G., Giovannoni, J. J., Fernie, A. R., and Rose, J. K. (2020). Genetic and metabolic effects of ripening mutations and vine detachment on tomato fruit quality. *Plant Biotechnology Journal*, 18(1):106–118.
- Robinson, R. and Tomes, M. (1968). Ripening inhibitor: a gene with multiple effects on ripening. *Tomato Genetics Cooperative*, 18(18):36–37.
- Shinozaki, Y., Nicolas, P., Fernandez-Pozo, N., Ma, Q., Evanich, D. J., Shi, Y., Xu, Y., Zheng, Y., Snyder, S. I., Martin, L. B., Ruiz-May, E., Thannhauser, T. W., Chen, K., Domozych, D. S., Catalá, C., Fei, Z., Mueller, L. A., Giovannoni, J. J., and Rose, J. K. (2018). High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. *Nature Communications*, 9(1):364.
- Silva, C. J., Van Den Abeele, C., Ortega-Salazar, I., Papin, V., Adaskaveg, J. A., Wang, D., Casteel, C. L., Seymour, G. B., and Blanco-Ulate, B. (2021). Host susceptibility factors render ripe tomato fruit vulnerable to fungal disease despite active immune responses. *Journal of Experimental Botany*, 72(7):2696–2709.
- Thompson, A. J., Tor, M., Barry, C. S., Vrebalov, J., Orfila, C., Jarvis, M. C., Giovannoni, J. J., Grierson, D., and Seymour, G. B. (1999). Molecular and genetic characterization of a novel pleiotropic tomato-ripening mutant. *Plant Physiology*, 120(2):383–389.
- Tieman, D., Zhu, G., Resende, M. F., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., and Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.)*, 355(6323):391–394.
- Tigchelaar, E., McGlasson, W., and Franklin, M. (1978). Natural and ethephon-stimulated ripening of F1 hybrids of the ripening inhibitor (rin) and non-ripening (nor)

- mutants of tomato (*Lycopersicon esculentum* Mill.). *Australian Journal of Plant Physiology*, 5(4):449–456.
- Tigchelaar, E., Tomes, M., Kerr, E., and Barman, R. (1973). A new fruit ripening mutant, non-ripening (nor). *Rep Tomato Genet Coop*, 23:33–34.
- Wang, R., Angenent, G. C., Seymour, G., and de Maagd, R. A. (2020a). Revisiting the Role of Master Regulators in Tomato Ripening. *Trends in Plant Science*, 25(3):291–301.
- Wang, R., Lammers, M., Tikunov, Y., Bovy, A. G., Angenent, G. C., and de Maagd, R. A. (2020b). The rin, nor and Cnr spontaneous mutations inhibit tomato fruit ripening in additive and epistatic manners. *Plant Science*, 294:110436.
- Wang, R., Tavano, E. C. d. R., Lammers, M., Martinelli, A. P., Angenent, G. C., and de Maagd, R. A. (2019). Re-evaluation of transcription factor function in tomato fruit development and ripening with CRISPR/Cas9-mutagenesis. *Scientific Reports*, 9(1):1696.
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13(1):134.
- Zhang, M., Yuan, B., and Leng, P. (2009). The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. *Journal of Experimental Botany*, 60(6):1579–1588.
- Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J. J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology*, 31(2):154–159.

## 2.8 Supplemental Material

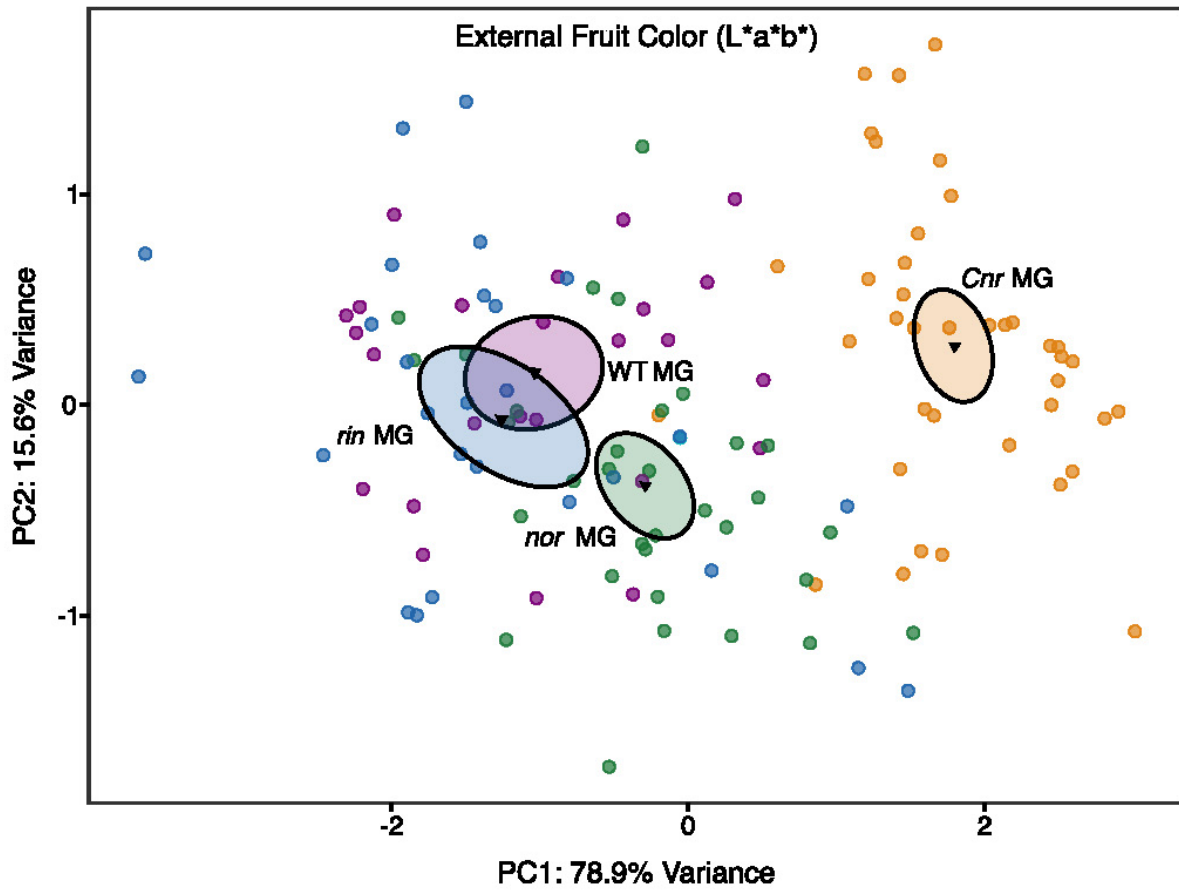


Figure 2.S1: External color of single ripening mutant fruit at the mature green (MG) stage. Principal component analysis of the external color of wild type (WT), *Cnr*, *rin*, and *nor* fruit measured on the  $L^*a^*b^*$  color scale. The center of gravity is represented by a triangle with surrounding ellipses indicating 95% confidence interval.

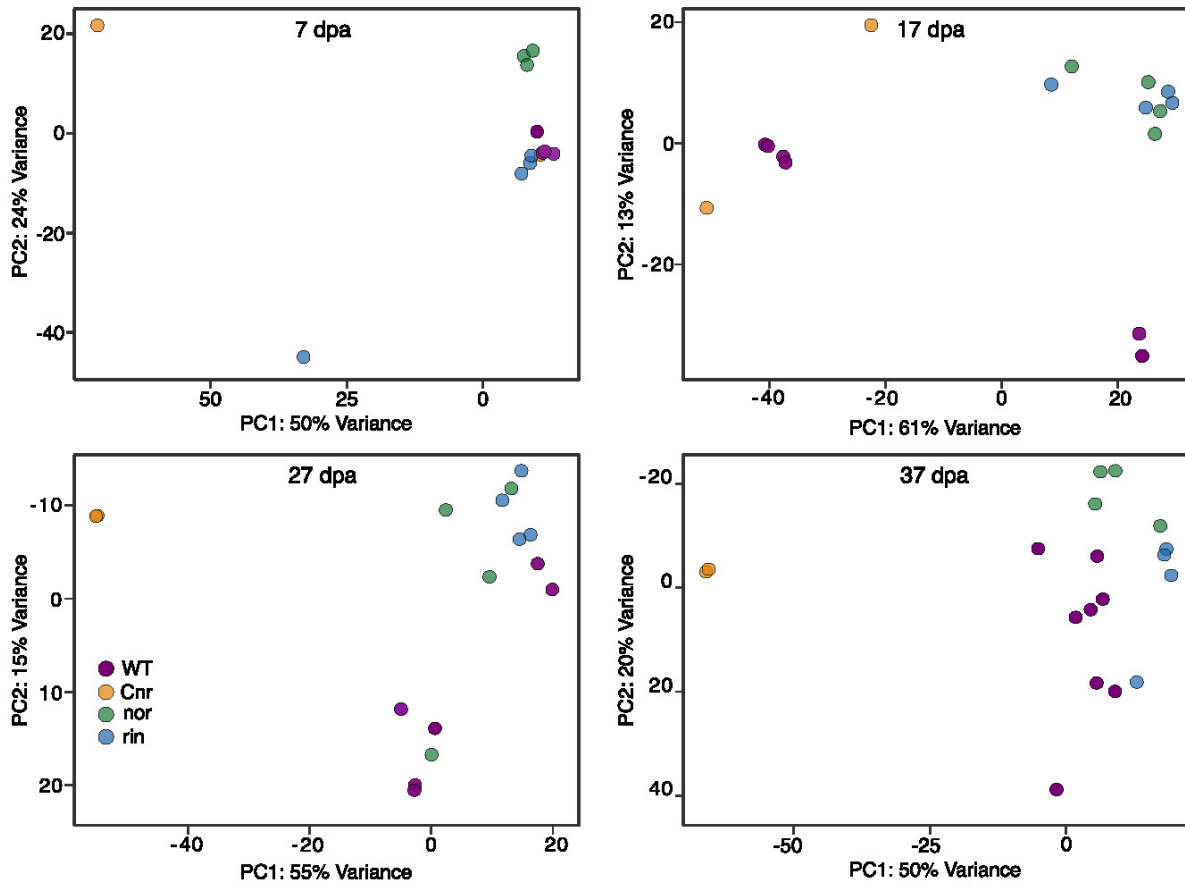


Figure 2.S2: Principal component analysis (PCA) of normalized RNAseq reads for Cnr, nor, and rin wild type fruit at immature and mature green (MG) stages. The RNAseq data of the single mutants at 7 days post anthesis (dpa), 17, 27, and 37 dpa (MG) were obtained from Lü et al. (2018) and reanalyzed using our bioinformatics pipeline (Lü et al., 2018).

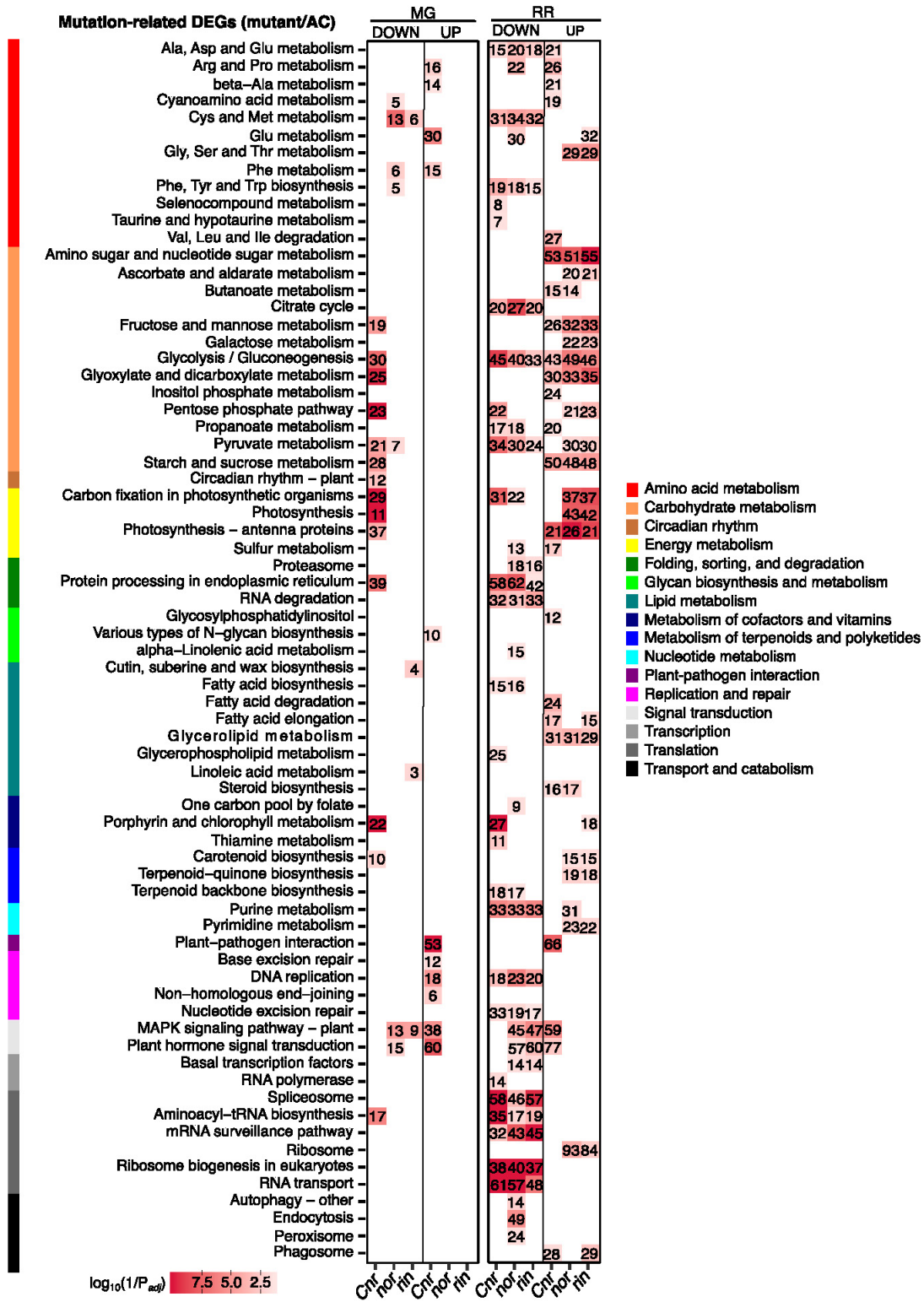


Figure 2.S3: Caption presented on following page.

Figure 2.S3: (Continued from previous page.) Functional enrichments in Kyoto Encyclopedia of Genes and Genomes (KEGG) functions among differentially expressed genes (DEGs;  $P_{adj} \leq 0.05$ ). Mutation-related DEGs were obtained by comparing each mutant to the wild type (WT) at the mature green (MG) and red ripe (RR) stages. Each comparison is separated into significant down- and up-regulated DEGs. The heat map colors indicate the significance of the functional enrichment using a  $\log_{10}(1/P_{adj})$  scale. Numbers in each tile indicate the number of DEGs within each category. Only significant ( $P_{adj} \leq 0.05$ ) functional enrichments are shown.



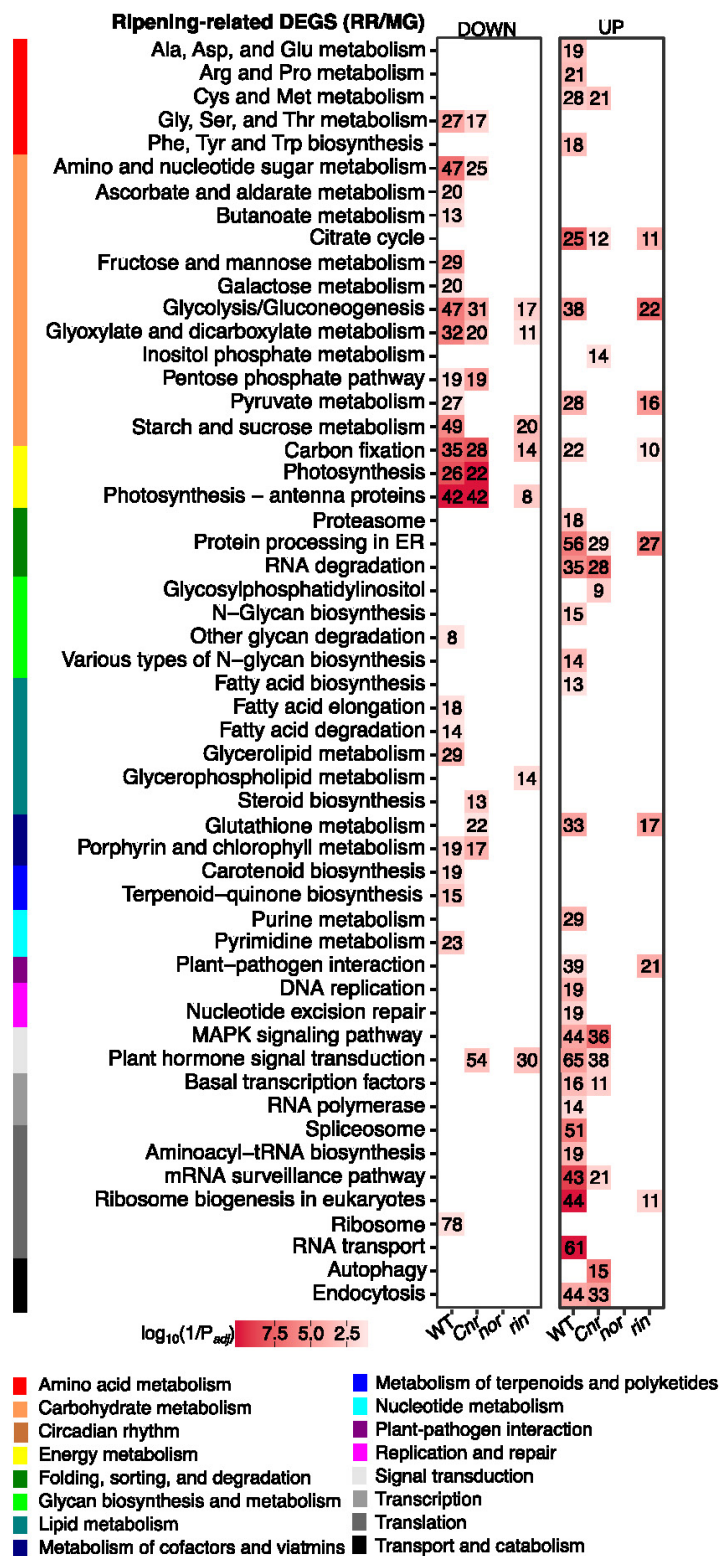


Figure 2.S4: Caption presented on following page. (Caption continued on following page.)

Figure 2.S4: (Continued from previous page.) Functional enrichments in Kyoto Encyclopedia of Genes and Genomes (KEGG) functions among differentially expressed genes (DEGs;  $P_{adj} \leq 0.05$ ). Ripening-related DEGs were obtained by comparing the RR stage against MG for each genotype. Each comparison is separated into significant down- and up-regulated DEGs. The heat map colors indicate the significance of the functional enrichment using a  $\log_{10}(1/P_{adj})$  scale. Numbers in each tile indicate the number of DEGs within each category. Only significant ( $P_{adj} \leq 0.05$ ) functional enrichments are shown.

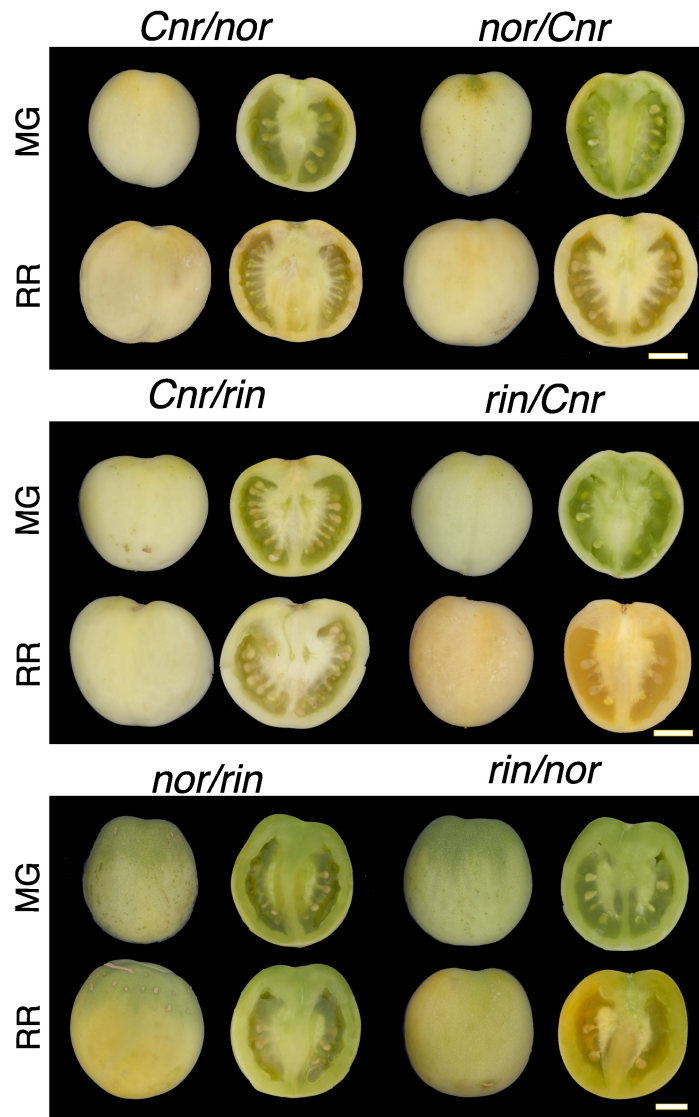


Figure 2.S5: Caption presented on following page.

Figure 2.S5: (Cont. from previous page) Representative fruit from the reciprocal crosses of the double mutants. The maternal genotype is listed first for each double mutant. Fruit are pictured at the mature green (MG) and red ripe (RR) stages. Fruit shown whole at left and in longitudinal sections at right. Images were extracted and processed with the VideometerLab instrument. Bar represents 1.5 cm.

## Chapter 3

# Defining pistachio fruit ripening and its impacts on quality

### 3.1 Abstract

Pistachio (*Pistacia vera*) is an important global nut crop that continues to increase in popularity due to its nutritional benefits to consumers and tolerance to environmental stress. Pistachios are among a few crop species for which the fruit and seed develop asynchronously, providing a unique opportunity to investigate molecular processes behind fruit ripening and seed maturation. We hypothesized that significant transcriptional reprogramming occurs during fruit ripening and seed maturation, resulting in the gain or establishment of fruit quality traits. Here we produced an extensive quantitative study of pistachio development combining multiple approaches: physiology, biochemistry, and genomics. To facilitate our study, we first generated the most complete genome assembly of pistachio to date, corresponding to a 561.79 Mb genome of *Pistacia vera* cv. Kerman with 38,395 genes. Using Omni-C technology combined with manual curation, 96.89% of the primary contig assembly was assigned to 15 chromosomes together with an 98.6% BUSCO assessment score suggests the completeness of Kerman genome. We then measured physiological changes in the hull, shell, and kernel tissue throughout multiple growing seasons and locations to redefine previous developmental stages into four distinct periods. We found the seed (kernel) growth was initiated in Stage III only after fruit (shell and hull) tissues have reached their peak of growth in Stages I-II and created a new stage (Stage IV) representing fruit ripening. To further characterize the stages, we utilized the new

genomic resources and integrated substantial transcriptomic data extending across 15 weeks of development and three fruit tissues (hull, shell, kernel) to identify networks of coexpressed genes associated with each stage. Later, we focused on defining the molecular changes occurring at the onset and during Stage IV in each tissue that explain fruit traits, such as hull softening and kernel fat accumulation. We discovered that hull softening is likely driven by pectin-degrading enzymes, specifically one pectate lyase (PL) gene highly expressed in the hull. High gene expression of a fatty acid desaturase (FAD2) enzyme that synthesizes linoleic acid explains the increased accumulation of this monounsaturated fatty acid in the kernel. Overall, our study provides an unparalleled analysis of pistachio fruit development and genetics and establishes a basis for further developmental and genomic research in pistachio and other fruit crops.

## 3.2 Introduction

Pistachio (*Pistacia vera*,  $2n=30$ ) world production has more than doubled over the past two decades, with over 1 million metric tons of pistachio fruits produced in 2021 (FAO; <https://www.fao.org/faostat/en/#search/pistachio>). The United States produced 47% of the world's pistachios in 2021, with California comprising 99% of the national total (CDFA 2022; <https://www.cdfa.ca.gov/Statistics/> ; Figure 3.1A-C). The nutritional and pharmacological attributes of pistachio contribute to its growing popularity. Pistachios provide numerous health benefits to humans including decreased risk of coronary heart disease and can aid in weight loss (Mandalari et al., 2022). In addition to its economic and nutritional importance, pistachio is highly adaptive to stress and grows well in arid climates, making it a sustainable option for fruit production with continued climate changes (Moazzam Jazi et al., 2017). Pistachio belongs to the Anacardiaceae family, which includes mango, cashew, and sumac from other genera. Among the 11 species in the *Pistacia* genus, *P. vera* is the only edible and commercially used species. The Kerman cultivar has been considered the industry standard in the United States with the highest proportion of acreage in California (Kallsen et al., 2020).

Pistachio fruits are dehiscent drupes composed of three main tissues: a leathery exo-

mesocarp (hull), a stony endocarp (shell), which encloses a seed (kernel) (Figure 1A-B). Three stages of fruit growth have been previously proposed (Goldhamer and Beede, 2004). Cell division and expansion (Stage: Fruit growth) of the hull and shell was reported to occur rapidly from April to late-May. As soon as the fruits reach the final size, fruit growth appeared to be reduced from late-May to early-July and lignin accumulates in the shell (Stage: Shell hardening). The remaining part of the season (from July forward) was believed to be dominated by the growth of the kernel (Stage: Embryo growth). While this understanding of pistachio growth has guided research and production in the past, it was defined using a few parameters, like fruit and embryo size. Thus, to date, a comprehensive analysis of the dynamics of the pistachio fruit and embryo development is lacking.

The coordinated development of pistachio fruit tissues (shell, hull, and kernel) is required for high-quality, marketable fruits. Quality attributes include a soft detachable hull without signs of deterioration, hard and split shells, and a flavorful kernel free from damage. Pistachio fruit physiological and compositional studies have described aspects of pistachio quality with particular emphasis on mature fruits at or near the time of harvest (Mandalari et al., 2022; Zarei et al., 2014). Like many fruit, color changes in the hull can serve as an index of harvest, changing from green-yellow to shades of pink at the end of the growing season (Ferguson et al., 2005). Pistachio shell quality relies upon the endocarp hardening and splitting before harvest, as well as being free from stains caused by the deteriorating hull. Pistachio kernels are naturally rich in antioxidants, vitamins, and unsaturated fatty acids essential for the human diet (Mandalari et al., 2022). Both mono- and poly-unsaturated fatty acids accumulate in the kernel (Polari et al., 2019) and are important for the human diet, as polyunsaturated fats cannot be synthesized by the body. To date, there is no knowledge of the genetic basis of these quality traits in the fruit and how these compounds are accumulated in the kernel.

While the pistachio industry has grown exponentially over the past 50 years in the US, there is a limited body of research in fruit development and the mechanisms behind key quality traits to inform breeding, production and management. Here, we present a study integrating physiological, genomic, transcriptomic, and biochemical data to define pista-

chio fruit development and identify biological processes that contribute to fruit quality, such as hull softening, shell hardening, and kernel fatty acid composition. A high-quality genome and annotation were fundamental for analyses in this study; however, while improvements have been made over the past three years the existing genome is incomplete spanning . We, therefore, assembled a reference-quality chromosome-scale genome for *P. vera* cv. Kerman using PacBio HiFi long reads and chromatin interaction information with Dovetail Genomics Omni-C data. We integrated Isoform sequencing data from multiple tissue types (leaf, fruit, bud, flower) and RNA sequencing data from two fruit tissues (hull and shell) across multiple time points to obtain the most complete genome annotation for pistachio to date. We assessed physiological and metabolic changes in the hull, shell, and kernel tissues with a comprehensive 24-week study encompassing the entire growing season and generated gene expression data for hull, shell, and kernel tissues for 15 of those weeks. The datasets produced allowed us to construct networks of coexpressed genes for each developmental stage and connect genes and molecular pathways to specific fruit traits. Overall, our study provides the first understanding of genetic networks governing pistachio fruit development and defines the processes occurring during fruit ripening that determine fruit quality traits.

### **3.3 Materials and Methods**

#### **3.3.1 Sample collection and preparation for genome sequencing**

High molecular weight genomic DNA was extracted from young leaves of ‘Kerman’ tree using Circulomics Nanobind Plant nuclei kit (<http://circulomics.com>). The library construction and PacBio Hifi sequencing were completed on a Sequel II system at Gentyane facilities in the French National Institute of Agronomy (INRAE). For Omni-C data, young leaf tissue samples were collected, directly snap-frozen in liquid nitrogen, stored at -80°C, and sent to Dovetail genomics for the Omni-C library construction and sequencing.

For Iso-seq sequencing, samples of fruits, leaves, inflorescences, and dormant and developing buds were collected and immediately stored in liquid nitrogen. 100 mg of material for each sample was pulverized with a pestle and mortar in liquid nitrogen. Total RNA

was extracted by using the Spectrum Plant Total RNA (Sigma) Kit, according to the protocol recommended for "difficult" species. The quantity and purity of the total RNA were checked with the Nanodrop (ND1000 Thermo Fisher Scientific). Iso-seq sequencing was performed using RNA samples with 400 ng/ul at Gentyane facilities in INRAE.

### **3.3.2 Genome size estimation**

The genome size, heterozygosity, and repeat content were estimated based on k-mer frequency analysis with PacBio HiFi reads. We used Jellyfish v2.2.10 (Marçais and Kingsford, 2011) to count 21-mers with the maximum k-mer depth of 1e6, which takes repetitive regions into account. The resulting histogram from Jellyfish was subjected to Genomescope v1 web to estimate genome size, levels of heterozygosity, and repeat content (Vurture et al., 2017) .

### **3.3.3 Genome assembly and chromosome construction**

A *de novo* assembly of PacBio HiFi reads into contigs was performed using Hifiasm v0.16.0 with default parameters (Cheng et al., 2021). We then manually filtered out organelle (plastid and mitochondrial) origin contigs from the primary contig assembly using Kerman plastid and mitochondrial sequences in Geneious Prime (<https://www.geneious.com>). The high-coverage Omni-C data was first quality-checked with FastQC toolkit and aligned to filtered primary contig assembly using Juicer v1.6 (Andrews and Others, 2010; Durand et al., 2016). These aligned read pairs were utilized to scaffold the assembly into 15 chromosomes based on the Omni-C chromatin interaction data with 3D-DNA and manual correction on Juicebox (Dudchenko et al., 2017; Robinson et al., 2018). We scaffolded filtered primary contig assembly again with 3D-DNA scaffolds as a reference using RagTag v2.1.0 followed by manual correction by comparing both RagTag and 3D-DNA scaffolds (Alonge et al., 2022). The completeness of assemblies, scaffolds, and 15 chromosomes was assessed using BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.4.4 with the embryophyta-db10 database (Simão et al., 2015).



### 3.3.4 Genome annotation

Annotation of transposable elements (TE) was accomplished with primary contig assembly using Extensive de-novo TE Annotator (EDTA) v2.1.0, which generates a non-redundant TE library and TEs classified into the superfamily level (Ou et al., 2019). We then filtered out any sequences from the non-redundant TE library overlapping with the filtered plant protein database using protExcluder v1.2 (<https://www.canr.msu.edu/hrt/uploads/535/78637/ProtExcluder1.2.tar.gz>) to hinder exclusion of genes in the gene prediction analysis. Output from protExcluder was employed to re-annotate TEs in the primary contig assembly using EDTA. The assembly was softmasked for further gene prediction analysis.

To annotate genes, softmasked ‘Kerman’ assembly, RNA sequencing (RNA-seq) and Isoform sequencing (Iso-seq) data were used. We first assembled PacBio Iso-seq data from five different tissues using the IsoSeq3 pipeline (Pacific Biosciences), which involved the following steps: i) demultiplexing and primer removal using lima v2.0.0, ii) removing of poly(A) tails and concatemers, iii) clustering isoforms, and iv) mapping to the assembly using pbmm2 v1.9.0 and collapsing isoforms. The quality of RNA-seq data from pistachio fruit (hull and shell) was assessed using FastQC v0.11.9 (Andrews and Others, 2010). Raw RNA-seq reads were filtered using Trimmomatic v0.39 and clean reads were mapped to the assembly using Hisat2 v2.2.1 (Bolger et al., 2014; Kim et al., 2019). We then assembled both aligned RNA- and Iso-seq reads into transcripts using StringTie v2.2.1 (Pertea et al., 2015). Before ab initio gene prediction, the Iso-seq data and softmasked assembly served as a training set for Braker2 v2.1.2 (Br ° una et al., 2021). The ab initio prediction was carried out using Augustus v3.1.0 with training files, softmasked assembly, and exon hints from Iso-seq data (Keller et al., 2011). All data including assembled RNA- and Iso-seq data and Augustus output gene models were combined to find gene models using EvidenceModeler (EVM) v2.0.0 with different weights for each input data (7, 4, and 1 for Iso-, RNA-seq transcripts, and Augustus gene models, respectively) (Haas et al., 2008). Finally, 5’ and 3’ UTR regions and different isoforms were updated from EVM gene models using PASApipeline v2.5.2 (Haas et al., 2008). The final gene annotation on

primary contig assembly was lifted over to 15 chromosomes using liftoff v1.6.3 (Shumate and Salzberg, 2020). We assess the quality of the final gene models using BUSCO (embryophyta\_db10 set) with protein sequences extracted by using GffRead v0.12.7 in GFF Utilities (Pertea and Pertea, 2020; Simão et al., 2015). We used the RIdeogram R package to plot gene and TE density on the 15 chromosomes in the ‘Kerman’ genome (Hao et al., 2020).

### **3.3.5 Genome synteny analysis**

To conduct a comparative analysis, we first downloaded the complete genome and annotation of *Mangifera indica* (mango) in the same family, Anacardiaceae, and *Citrus sinensis* (sweet orange) in the same order, Sapindales (Wang et al., 2020; Wu et al., 2014). Synteny of the chromosome-level genome of ‘Kerman’ was compared with that of Mango and Sweet Orange using GENESPACE v0.9.4, which takes orthology information into account using OrthoFinder v2.5.4 (Emms and Kelly, 2019; Lovell et al., 2022). The manual reformation of gene annotation and protein files was needed for GENESPACE input.

### **3.3.6 Sample collection for physiological data and RNAseq experiments**

Pistachio physiological data collected across three field seasons (2019, 2020, 2021) were collected for evaluation from an experimental pistachio orchard (cv. Kerman grafted onto UCB1 rootstock) with 30 year old trees at the Kearney Agricultural Research and Extension Center in 2019. The results were validated in a commercial orchard with 10 years old trees of the same cultivar in Woodland, CA (Yolo County) in 2020 and again in 2021 in a separate commercial orchard with trees 10 years old located in Three Rocks, CA (Fresno County) (Supplementary Data–Models). In each orchard the Kerman female was cross pollinated by the Peters variety male trees. In each study, trees were randomly selected across the orchard and were continuously sampled throughout the season. Four whole rachis clusters yielding about 50 fruits of uniform maturity were collected per tree at each sampling. Environmental temperature was recorded with the use of the HOBOware instrument placed in the orchards at three separate locations. Accumulated heat units are

expressed as growing degree days (GDD) and were calculated with the following equations to normalize development based on accumulation of temperature to the system (Zhang et al., 2021):

$$T_{avg} = ((T_{max} + T_{min})/2)$$

$$GDD = \sum (T_{avg} - T_b) - 7$$

In 2019, samplings were conducted weekly after fruit set on April 25th (120 GDD) continuing through October 14th (2760 GDD), corresponding to two weeks after harvest maturity (70% shell split). In 2020 samplings occurred weekly from April 19th (24 GDD) through September 10th (2452 GDD, commercial harvest), and in 2021 samplings occurred weekly from July 15 (1531 GDD) to September 22nd (2828 GDD, commercial harvest).

### 3.3.7 Physiological measurements

To assess fruit area and color, fruits were imaged longitudinally using a VideometerLab 3 (Videometer, Denmark) facilitated by Aginnovation LLC. VideometerLab 3 software was utilized for image analysis in both 2019 and 2020. Color measurements were taken on the L\*a\*b\* color scale as an average across the entire fruit area. For both color and area, 10-30 fruits per tree (n=12) were sampled weekly for 25 weeks. Growth was determined through fresh weight and dry weight measurements of the whole fruit and the kernels. Fresh weight was taken the day of harvest of 10 fruits per tree (n=12) and the average per fruit weight was calculated. Fruits were cut open and separated into kernels and the remaining tissues (hull and shell). Fruits were put in a drying oven at 80°C for 2 days until all moisture was evaporated and measured. The average per whole fruit and per kernel weights per cluster (n= 7-12 per sampling) were calculated from the total weight and total number of fruits. Shell split was measured as the incidence of fruits with any degree of separation between sides and taken as a proportion of the total fruits. Destructive texture measurements were obtained with the use of a TA.XT2i Texture Analyzer (Texture Technologies, United States) using a TA52 2mm probe with a trigger force of 5 g and test speed of 2.00 mm/sec with Exponent software (Texture Technologies Corporation, United States). The probe punctured through the hull, shell and kernel tissues and peaks of each

tissue were distinguished and recorded by the software. Measurements were reported as kilograms (kg) of force. 20 to 60 fruits were assessed for each sampling for 25 weeks. Fat content was obtained from oven-dried kernels, as described in (Polari et al., 2020). Briefly, dried pistachio kernels ( $5.0 \pm 0.1$  g) were ground in a coffee grinder (JavaPresse Coffee Company, WY), weighed into a cellulose extraction thimble, placed in the Soxhlet extractor, and extracted using n-hexane for 6 h. The solvent was distilled and residual solvent eliminated in an oven at 105 °C for 3 h. Fat content was expressed as grams of fat per 100 grams.

### 3.3.8 Statistical analysis of physiological parameters

Physiological parameters including kernel and fruit dry weights (g), kernel and fruit areas ( $\text{mm}^2$ ), fruit colors ( $L^*$ ,  $a^*$ ,  $b^*$ ), and kernel, shell, and fruit textures were modeled against heat accumulation (measured in growing degree days, °C). Various Box-Cox transformations (i.e. square, square root, and logarithm transformation, etc.) on traits data using MASS package (Ripley, B., Venables, B., Bates, D. M., ca 1998, K. H., 2021) in R, were made before modeling to ensure an approximate normal distribution of traits and roughly equal variance of the error terms. Outliers were removed prior to model fitting for trait nut area. Initial model fitting started with a linear model of polynomial 3 of heat accumulation, and stepAIC function (from MASS package) was used in R with criterion BIC (Claeskens and Hjort, 2008). The upper and lower bounds of the model were polynomial 7 and 1, respectively. We chose the model with the lowest possible BIC value. The final model selection results were capped at polynomial 3 to avoid model complexity, overfitting, and difficulty for interpretation. As a result, linear models and linear mixed models were fitted for each trait with a polynomial of 2 or 3 as a function of heat accumulation, using lme4 and lmerTest packages in R (Bates et al., 2015; Kuznetsova et al., 2017). Random intercepts were added in linear mixed models. Random effects included cluster, tree, and year depending on the models (Supplemental Table: Models).

### **3.3.9 RNA extraction of fruit tissues**

Fruit tissues from 15 of the 24 weeks assessed for physiological data were separated into the hull, shell and kernel and were each flash frozen on the day of sampling. Four replicates composed of 12 fruits from three separate trees, which were from the same fruits used for physiological measurements, were frozen. Frozen tissues were ground into a fine powder with the Retsch Mixer Mill MM 400 (Verder Scientific, Netherlands). RNA extractions were performed on samplings from 865-2564 GDD for hull tissues, 865-2139 GDD for shell tissues, and 1106-2564 GDD for kernel tissues. One gram of ground tissue was used for RNA extraction as described in (Blanco-Ulate et al., 2013). RNA concentrations were quantified with Nanodrop One Spectrophotometer (Thermo Scientific, United States) and Qubit 3 (Invitrogen, United States). RNA integrity was then assessed on an agarose gel.

### **3.3.10 cDNA library preparation, RNA sequencing, and sequencing data processing**

cDNA libraries were prepared with Illumina TruSeq RNA Sample Preparation Kit v.2 (Illumina, United States) from the extracted RNA. The quality of the barcoded cDNA libraries was assessed with the High Sensitivity DNA Analysis Kit in the Agilent 2100 Bioanalyzer (Agilent Technologies, United States) and then sequenced (50 bp single-end reads) on the Illumina HiSeq 4000 platform by the DNA Technologies Core at UC Davis Genome Center.

Raw reads were trimmed for quality and adapter sequences using Trimmomatic v0.39 (Bolger et al., 2014) with the following parameters: maximum seed mismatches = 2, palindrome clip threshold = 30, simple clip threshold = 10, minimum leading quality = 3, minimum trailing quality = 3, window size = 4, required quality = 15, and minimum length = 36. Trimmed reads were then mapped using Bowtie2 (Langmead and Salzberg, 2012) to the pistachio transcriptome assembly. Count matrices were made from the Bowtie2 results using sam2counts.py v0.913. A summary of all read mapping results can be found in Supplementray Table 1.

### 3.3.11 Functional annotation and enrichments

To annotate transcripts with GO terms and Pfam, the best blast hit from the NR database was imported to OmicsBox software (BioBam, v 2.2.4). Likewise, the predicted protein sequences were searched against Pfam, ProDom, ProSiteProfiles by InterProScan in OmicsBox to retrieve conserved domains/motifs and corresponding GO terms. The Kyoto Encyclopedia of Genes and Genome (KEGG) orthology (KO) and pathways were identified using the Automatic Annotation Server (KAAS; <http://www.genome.jp/kegg/kaas/>) with bi-directional best hit (BBH) and BLAST as search program against closely related organisms, mallow, rose, mustard, and papaya. The carbohydrate active enzyme (CAZy) was annotated using trinotate software (v 3.2.2) hmmscan function by searching the proteome against the Carbohydrate Active Enzymes database (<http://www.cazy.org/>) with default settings (Drula et al., 2022). The plant Transcription factor Protein Kinase Identifier and Classifier (iTAK) online tool ([http://itak.feilab.net/cgi-bin/itak/online\\_itak.cgi](http://itak.feilab.net/cgi-bin/itak/online_itak.cgi)) was used to identify transcription factor families with default settings (website, ref). Enrichment analysis for all functional annotations was performed using a Fisher test. The P-values obtained from the Fisher test were adjusted with the Benjamini and Hochberg method (Benjamini and Hochberg, 1995).

### 3.3.12 Coexpression Analysis

Count matrices were normalized with the variance stabilizing transformation function from the DESeq2 package in R. The normalized reads were then used for weighted gene co-expression networking analysis by WGCNA package in R (Langfelder and Horvath, 2008). The WGCNA was conducted for each tissue separately. Transcripts were checked by the goodSamplesGenes function, and transcripts that had no expression across all time points were removed. The dendrogram plots were made to identify and remove sample outliers. The network construction was done step by step. The soft thresholding power was selected to obtain the approximate scale-free topology, and adjacency was calculated with power equals selected soft power and type is signed type. TOM was calculated with signed type. The initial module eigengenes were detected by default hierarchical clustering function with default parameters other than the appropriate minModuleSize selected for

each tissue, then final module eigengenes were obtained by merging close modules with  $\text{cutHeight} = 0.25$ . The module eigengenes values were then used to identify modules that are significantly associated with measured physiological traits by calculating correlations using Pearson's product-moment correlation coefficient. The intramodular connectivity was assessed using the edge file output based on TOM. Each transcript connectivity was calculated by how many toNode each fromNode connects, and transcripts that had the top 5% highest number of connections in each module were identified as having high connectivity.

### 3.3.13 Volatile measurements

Volatile profile determination methodology as described in (Polari et al., 2020). Briefly, fresh pistachio kernels were ground to a paste in a mortar, and a sample ( $1.5 \pm 0.1$  g) was spiked with dodecane (2.5 mg/kg fruitmeat). Deionized water (3 mL) was added, and the vial was sealed with a PTFE/silicon septum (Supelco, United States). After 15 min at 50 °C, a solid-phase microextraction (SPME) fiber (DVB/CAR/PDMS, Sigma-Aldrich, United States) was exposed to the sample headspace for 45 min for volatile extraction. The volatile compounds analysis was performed with a Varian 450-GC equipped with a Varian 220-MS ion trap (Agilent Technologies, United States). A DB-5MS capillary column (30 m · 0.25 mm · 0.25 mm, Agilent Technologies, United States) was used for compound separations. After sampling, the fiber was thermally desorbed in the GC injector for 5 min at 260 °C. Helium was the carrier gas at a flow rate of 1 mL/min. Oven temperature started at 40 °C during the 5-min desorption process and was ramped at 3.5 °C/min to 120 °C, then ramped at 10 °C/min to 220 °C and held for 10 min. Ionization energy of 70 eV was adopted, and the ions were analyzed in the m/z range from 40 to 400. The data were recorded and analyzed using MS Workstation v. 6.9.3 (Agilent Technologies, United States) software. Volatile compounds were identified by their mass spectra and Kovats retention index (KI). Results were expressed as nanograms of dodecane per kilogram of sample.

### 3.3.14 Fatty acid profiles

Fatty acid profile was determined according to (Polari et al., 2020), as described. Briefly, the sample ( $0.010 \pm 0.001$  g) was weighed and dissolved in toluene (0.4 mL). Methanol (3 mL) and methanol/HCl (0.6 mL, 80:20, v/v) were added. The vial was heated at 80 °C for 1 h. Hexane (1.5 mL) and deionized water (1 mL) were added. Anhydrous sodium sulfate (0.5 g) was added to the decanted upper hexane layer containing the methyl esters. GC analysis was conducted on a Varian 450-GC (Agilent Technologies, Santa Clara, CA) equipped with a flame ionization detector (FID). Helium was the carrier gas at a flow rate of 1.5 mL/min. Fatty acid methyl esters (1 ml) were injected onto a DB-23 capillary column (60 m · 0.25 mm · 0.25 mm (Agilent Technologies, United States). A mix of FAME standard was used as references for peak identification by retention times; relative fatty acid proportions were determined by peak area normalization. Results were expressed as normalized area percentage.

## 3.4 Results

### 3.4.1 A complete pistachio genome assembly as a valuable resource for fruit development research

Due to limitations in the previous assembly of the *P.vera*, we assembled the first reference genome for *P. vera* c.v Kerman to facilitate analyses of fruit development and identify biochemical pathways associated with fruit quality. The *Pistacia vera* cv. Kerman genome sequencing yielded a total of 25.18 Gb PacBio HiFi reads, which is roughly 42X coverage of the previously estimated genome size of 600 Mb ((Ziya Motalebipour et al., 2016), (Table3.1). The mean length of the PacBio HiFi reads was 16,305.5 bp. The initial genome survey using the k-mer analysis ( $k = 21$ ) with PacBio HiFi reads estimated the Kerman genome size of 521 Mb with moderate heterozygosity (0.755%) and repetitiveness (54.1 %) (Figure 3.1D). After filtration of organelle-origin contigs from the primary contig assembly, the number of contigs was reduced to 102 with a size of 579.8 Mb with high contiguity (Table 3.1). The scaffold assembly was obtained and 15 pseudochromosomes were constructed using Juicer, 3D-DNA, and JuiceBox pipeline with high-coverage Omni-



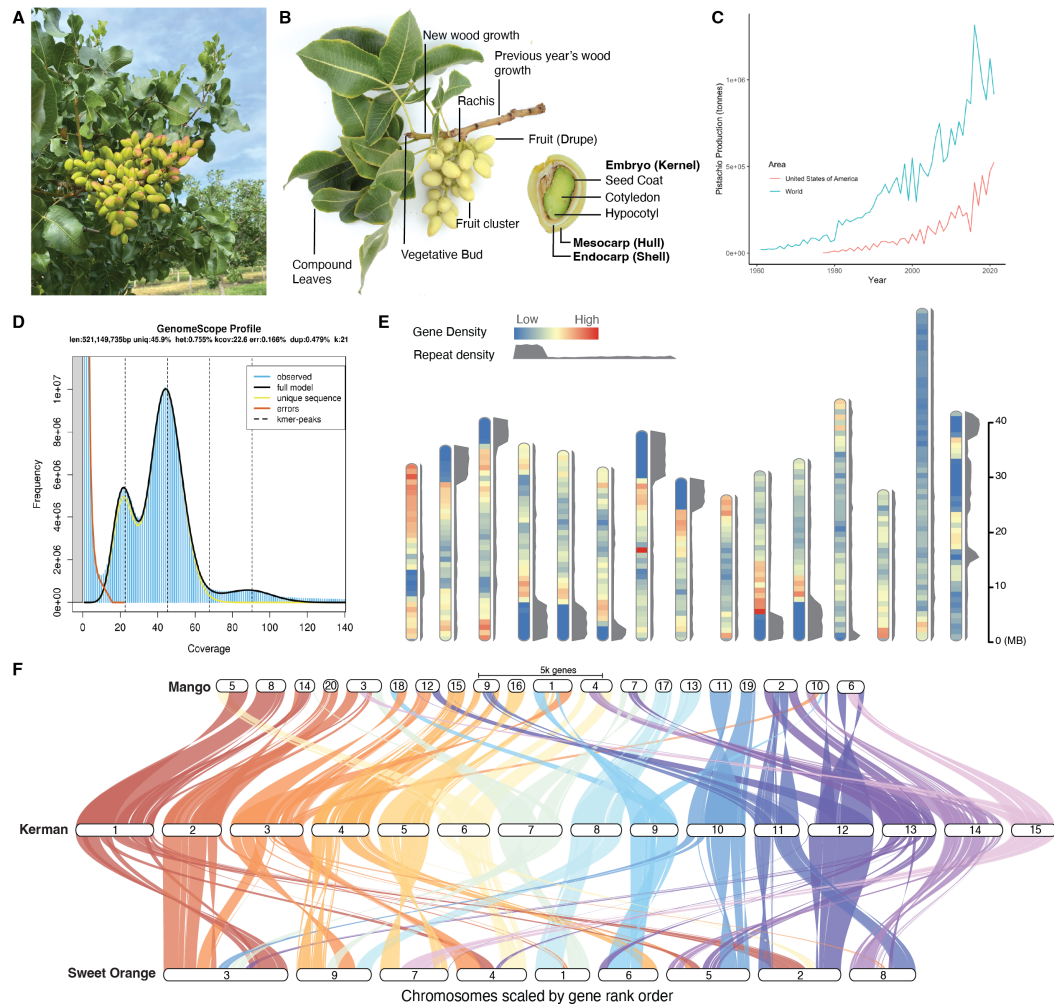


Figure 3.1: Chromosome-scale genome assembly of *Pistacia vera* cv. Kerman. Pistachio growth pattern on a tree (A) and on a branch (B) with anatomical features and fruit tissues labeled. C) Pistachio production increases from 1960 to 2022 in the United States compared to the world. D) Estimation of genome size, heterozygosity, and repetitiveness using GenomeScope. E) Ideogram with gene and repeat density in 1 Mb window size on 15 chromosomes of ‘Kerman’ genome. Gene density is represented in each chromosome in heatmap style and repeat density is plotted to the side of each chromosome in orange. The scale bar for chromosome size is indicated on the right in light gray. The highest gene density in chromosome 7 in red color was manually changed from 369 to 200 for better visualization of all other regions. F) Syntenic analysis of ‘Kerman’ chromosomes compared to Mango and Sweet orange genomes.

Table 3.1: Statistics of PacBio HiFi sequencing and assembly of *Pistacia vera* cv. Kerman.

| <i>P. vera</i> cv. Kerman    |                                |                      |                             |                      |
|------------------------------|--------------------------------|----------------------|-----------------------------|----------------------|
|                              | <b>PacBio HiFi Reads</b>       |                      |                             |                      |
| Number of reads              | 1,544,028                      |                      |                             |                      |
| Mean read length (bp)        | 16,305.50                      |                      |                             |                      |
| Total base (Gbp)             | 25.18                          |                      |                             |                      |
| <b>Assembly</b>              | <i>Primary Contig Assembly</i> |                      | <i>15 Pseudochromosomes</i> |                      |
| Number                       | 102                            |                      | 15                          |                      |
| Size (bp)                    | 579,837,297                    |                      | 561,794,941                 |                      |
| Max length (bp)              | 49,785,438                     |                      | 62,545,632                  |                      |
| N50 (bp)                     | 25,977,840                     |                      | 36,813,375                  |                      |
| L50 (bp)                     | 9                              |                      | 7                           |                      |
| L90 (bp)                     | 21                             |                      | 13                          |                      |
| Number of Gaps               | -                              |                      | 14                          |                      |
| %GC                          | -                              |                      | 36.67                       |                      |
| <b>Annotation</b>            | <i>Primary Contig Assembly</i> |                      | <i>15 Chromosomes</i>       |                      |
| Number of genes              | 38,315                         |                      | 38,315                      |                      |
| Total length of genes (bp)   | 109,586,159                    |                      | 108,462,206                 |                      |
| % of genome covered by genes | 19                             |                      | 19.3                        |                      |
| Mean length of genes (bp)    | 2,860                          |                      | 2,890                       |                      |
|                              | <i>Without isoforms</i>        | <i>With isoforms</i> | <i>Without isoforms</i>     | <i>With isoforms</i> |
| Number of exons              | 174,192                        | 322,815              | 172,048                     | 320,956              |
| Total length of exons (bp)   | 49,738,000                     | 91,915,250           | 49,184,597                  | 90,975,896           |
| Mean length of exons (bp)    | 285                            | 284                  | 285                         | 283                  |
| Number of introns            | 135,877                        | 268,010              | 134,527                     | 266,990              |
| Total length of introns (bp) | 58,573,770                     | 114,572,561          | 58,688,059                  | 114,188,459          |
| Mean length of introns (bp)  | 431                            | 427                  | 436                         | 427                  |

C data. After RagTag scaffolding against 3D-DNA scaffolds, the manually corrected 15 chromosomes were 561.79 Mb in size with only 14 gaps (Figure 3.1E and Table 3.1). The primary contig assembly and 15 chromosomes were evaluated with BUSCO (Benchmarking Universal Single-Copy Orthologs), revealing 98.76% and 98.6% of complete conserved land plant genes (Table 3.2).

We detected over 0.83 million repetitive elements constituting 376.56 Mb of sequence

Table 3.2: The summary of BUSCO (Benchmarking Universal Single-Copy Orthologs) assessment of Kerman primary contig assembly and 15 chromosomes.

|                       | Genome                  |         |                |         | Gene Annotation         |         |                |         |
|-----------------------|-------------------------|---------|----------------|---------|-------------------------|---------|----------------|---------|
|                       | Primary contig assembly |         | 15 Chromosomes |         | Primary contig assembly |         | 15 Chromosomes |         |
|                       | #BUSCOS                 | %BUSCOs | #BUSCOS        | %BUSCOs | #BUSCOS                 | %BUSCOs | #BUSCOS        | %BUSCOs |
| Complete              | 1594                    | 98.76   | 1591           | 98.6    | 1597                    | 98.95   | 1596           | 98.88   |
| Complete, single-copy | 1540                    | 95.42   | 1543           | 95.6    | 1557                    | 96.47   | 1558           | 96.53   |
| Complete, duplicated  | 54                      | 3.35    | 48             | 3       | 40                      | 2.48    | 38             | 2.35    |
| Fragmented            | 11                      | 0.68    | 11             | 0.7     | 8                       | 0.50    | 9              | 0.56    |
| Missing               | 9                       | 0.56    | 12             | 0.7     | 9                       | 0.56    | 9              | 0.56    |
| Total                 | 1614                    | 100     | 1614           | 100     | 1614                    | 100     | 1614           | 100     |

space, which represents approximately 65% of the primary contig assembly length (Table 3.3). The most abundant transposable element (TE) group belonged to long terminal repeat retrotransposons (LTR-RTs), taking up 283.82 Mb (48.95%) of the assembly, with two major types, Ty1/copia and Ty3/Gypsy, accounting for 14.39% and 29.1% of the assembly, respectively (Table 3.3). Given that about 75.37% of repetitive sequences in the genome consisted of Ty1/copia and Ty3/Gypsy, the majority of these TE types were densely distributed on one arm region of 11 chromosomes and the submetacentric region of chromosome 1 (Figure 3.1E). The Class II DNA transposons contributed only about 10% with the highest proportion of Mutator (4.09%) and hAT (3.36%), followed by unclassified repeat regions (6.35%) (Table 3.3).

The annotation of protein-coding regions was performed with primary contig assembly using transcript evidence from RNA- and Iso-seq data plus *ab initio* predictions trained on RNA- and Iso-seq data. In the transcriptome assembled with Iso-seq data from all five tissue types (leaves, developing and dormant buds, fruits, and inflorescence), a total of 16,366 transcripts were identified compared to the final gene models and the number of transcripts with tissue-specific expression was variable between different tissues, with the highest number in developing buds (450), followed by leaves (398), dormant buds (376), fruits (326), and flowers (265) (Table 3.4). The initial *ab initio* gene prediction using Augustus with the training set evidenced by Iso-seq data was followed by EVM with weighted evidence. After the final updates with the addition of different isoforms and UTR regions using PASA pipeline, in total, 38,315 protein-coding genes were annotated,

Table 3.3: Summary statistics of repeat content in 'Kerman' primary contig assembly.

| Order  | Kerman        |             |             |                |
|--|---------------|-------------|-------------|----------------|
|  | Subfamilies   | Number      | Length (bp) | Proportion (%) |
| Class I (retrotransposons)                       | Total         | 619,084     | 283,826,501 | 48.95          |
|  | Copia         | 84,483      | 83,466,391  | 14.39          |
| LTR  | Gypsy         | 478,927     | 168,707,676 | 29.1           |
|  | Unknown       | 54,675      | 31,099,485  | 5.36           |
| DIRS   | DIRS          | N/A         | N/A         | N/A            |
| PLE  | Penelope      | N/A         | N/A         | N/A            |
| LINE   | LINE-element  | 762 499,192 | 0.09        |                |
|  | Unknown       | 237         | 53,757      | 0.01           |
| Class II (DNA transposons) - Subclass 1<br>(TIR) | Total         | 115,731     | 50,439,981  | 8.7            |
|  | CACTA         | 15,588      | 5,329,171   | 0.92           |
|  | Mutator       | 54,603      | 23,717,139  | 4.09           |
|  | PIF-Harbinger | 4,016       | 1,118,701   | 0.19           |
|  | Tc1-Mariner   | 2,902       | 821,046     | 0.14           |
|  | hAT           | 38,622      | 19,453,924  | 3.36           |
|  | Polinton      | N/A         | N/A         | N/A            |
| Class II (DNA transposons) - Subclass 2          | Total         | 14,445      | 5,449,359   | 0.94           |
|  | Helitron      | Helitron    | 14,445      | 5,449,359      |
| Other  | Repeat region | 89,355      | 36,846,549  | 6.35           |
| Total  |               | 838,615     | 376,562,390 | 64.94          |

of which 37,521 (97.9%) were lifted over onto the 15 chromosomes (Table 3.1). Genic regions constituted 19.3% of the genome with, on average, 65.94 genes per Mb, gene size of 2,890 bp, exon size of 285 bp, and intron size of 436 bp (Table 3.1). The final gene models on the 15 chromosomes contained 1,596 (98.88%) of BUSCOs out of 1,614 genes in the Embryophyta-db10 database (Table 3.2).

We performed a comparative analysis of orthology-constrained synteny between the pistachio, mango (*Mangifera indica*), and sweet orange (*Citrus sinensis*) genomes, all belonging to the Spinadales order. Overall, extensive structural variation among the three genomes was detected (Figure 3.1F). More collinearity from macrosynteny was found between pistachio and mango chromosomes although one region of pistachio chromosomes

Table 3.4: Statistics of the number of transcripts in different tissue types.

| Tissue type                      | Fruits        | Leaves | Inflorescences | Dormant bud | Developing bud |
|----------------------------------|---------------|--------|----------------|-------------|----------------|
| # of transcripts                 | 12,758        | 12,844 | 13,186         | 13,103      | 13,704         |
| # of missing transcripts         | 3,608         | 3,522  | 3,180          | 3,263       | 2,662          |
| # of tissue-specific transcripts | 326           | 398    | 265            | 376         | 450            |
| Total # of transcripts           | <b>16,366</b> |        |                |             |                |

was aligned to two different regions of mango chromosomes, which supports the lineage-specific whole genome duplication event in mango. The new genome annotations and the insights to the relationship between pistachio and its close relatives further motivated analysis of the pistachio fruit developmental program.

### 3.4.2 Pistachio fruits develop in four distinct stages

Pistachio fruit tissues (hull and shell) and the seed (kernel) develop asynchronously. To redefine the developmental stages occurring in pistachio, we sampled fruit clusters (c.v. Kerman) weekly for 24 weeks from one week after fruit set to two weeks after commercial harvest would have occurred in the 2019 season (Figure 3.2C). Because heat accumulation regulates enzymatic activities that influence many areas of metabolism including fruit growth (Corelli-Grappadelli and Lakso, 2004; Marino et al., 2018), we use accumulated heat expressed as growing degree days (GDD) and calculated as described in the Section 3.3.6, in addition to calendar time in our analysis to allow for more transferable results. Approximately 48 clusters were collected per time point from 12 trees using a randomized block design. Across sampling, we focus on three main tissues—the mesocarp (hull), endocarp (shell), and embryo (kernel). To capture environmental variability and provide a comprehensive timeline of fruit development we assessed a large sample size of individual fruit data ( $n = 663\text{--}6,805$ ) and validated our results across two additional independent field seasons (2020 and 2021) in distinct orchards across multiple geographical locations in the California Central Valley.

By assessing physiological characteristics of fruit and kernel growth we redefined the stages of fruit development in relation to the previous proposed stages of (1) growth and division of hull/shell, (2) shell hardening, and (3) kernel growth. From late April through

May, the shell and hull tissues grew in a logarithmic growth pattern plateauing at 500 GDD (Figure 3.2B). During this time (Stage I), the hull and shell tissues were fused together and became increasingly green, losing the initial red pigmentation. Stage II was redefined as a period before embryo development, where the fruit does not experience any expansion but continues to accumulate dry weight. The initiation of embryo growth began after 900 GDD (late June), marking the beginning of Stage III. Kernels accumulated fats during Stage III while the shell simultaneously became increasingly woody and firm (3.2B). While the kernel had reached its maximum size at the end of Stage III, it continued to accumulate dry weight up until the end of the season. During Stage IV, the hull underwent final changes consistent with fruit ripening. Hull softening occurred concurrently with changes in red and yellow coloration. (Figure 3.2C, D). Shells began to split during Stage IV with most fruits splitting by 2600 GDD.

We fitted linear and linear-mixed polynomial models for each physiological trait as a function of accumulated heat (GDD) to describe their behavior across the season using tree as a random variable. We decided to test if the behavior of these traits was consistent across multiple seasons and could be anticipated across different environmental conditions. To do this, we integrated multiple years datasets and fitted models and found that fruit size and weight could be accurately modeled across years and separate locations (Supplemental Table: Models). Although color and texture changes followed similar patterns across locations/years, they were more subject to environmental variation.

### **3.4.3 Pistachio developmental stages have distinct patterns of gene expression**

To investigate the molecular events occurring during pistachio ripening, we performed transcriptomic analysis of the hull, shell, and kernel tissues across up to 15 weeks of development using the same samples used for the physiology experiment. Samplings correspond to the end of Stage II, 886 GDD, through Stage IV, 2564 GDD. RNA was extracted and sequenced from at least three replicates per time point and tissue. The resulting reads were mapped to the predicted transcriptome (38,315 annotated genes, 3.1) producing an average of 76.41% mapping percentage across all samples. On average, 25,941 genes total

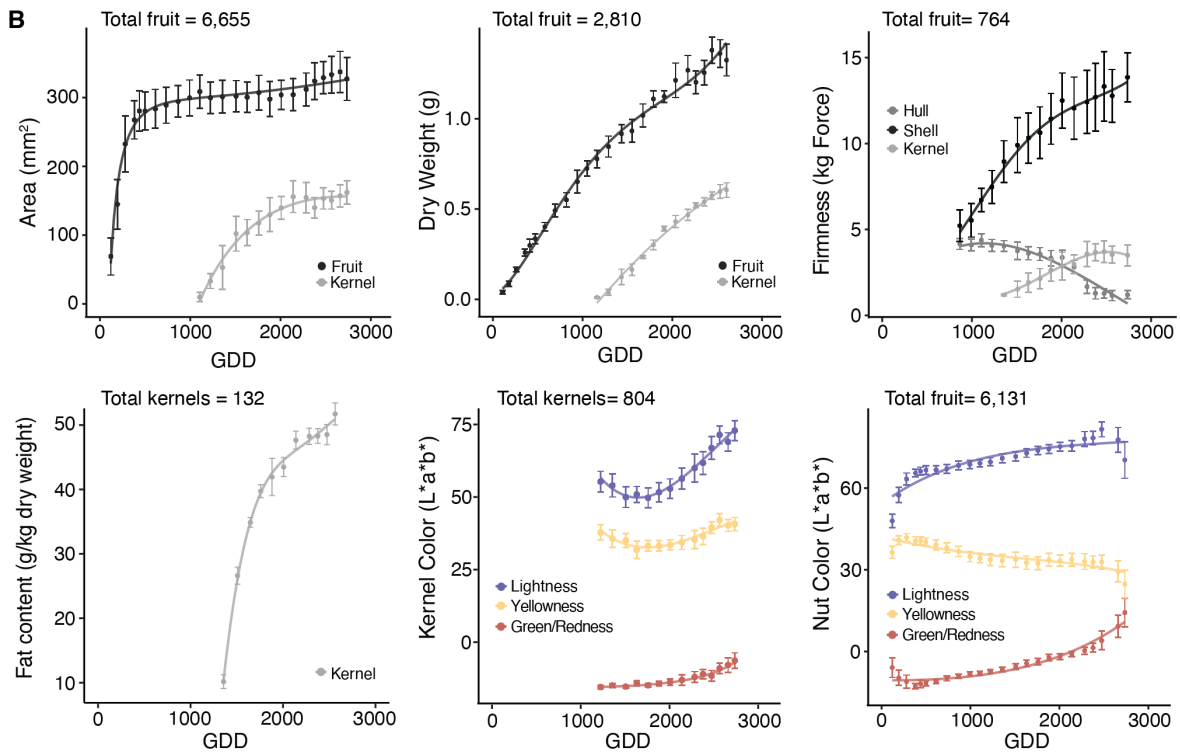
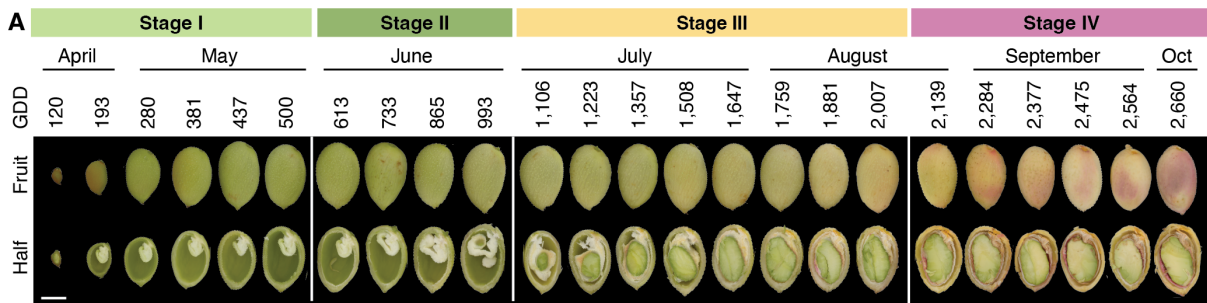


Figure 3.2: Defining pistachio development stages with physiological measurements. C) The fruits assessed from April through September and the corresponding growing degree days (GDD) for the whole fruit, halved fruit, and kernel. D) Graphs of physiological data for fruit and kernel area (mm<sup>2</sup>), fruit and kernel dry weight (g), hull, shell, and kernel texture (kg of Force), kernel fat content (g/kg dry weight), hull color (L\*, or lightness, a\* or redness, b\* or yellowness), and kernel color. Lines show linear models of growth fit as a function of heat accumulation (GDD).

were expressed across each tissue. We performed a principal component analysis (PCA) with the normalized reads of all samples (Figure 3.3A). The PCA revealed that the main driver of separation was tissue type with the kernel samples being very distinct from the hull and shell (maternal) samples (PC1, 76% variance). The developmental stage contributed to separating the samples across PC2 (8% of variance). An additional PCA was performed for maternal tissues only, which showed that the shell and hull tissues were separated by tissue type (PC1, 52% variance), increasing in similarity with age (Figure 3.3B).

We performed a weighted gene co-expression analysis (WGCNA) for each tissue type to identify gene modules associated with specific pistachio developmental stages. The WGCNA yielded 19 modules for hull samples, 17 modules for shell samples, and 22 modules for kernel samples (Supplemental Figure 3.S1). In each set of modules, we focused on patterns of expression corresponding to the stages of development identified in the assessment of the fruit physiology for each tissue (Figure 3.3B). We then performed trait-module correlations to identify module trends significantly correlated to the patterns observed in hull color and texture, shell texture, and kernel firmness, weight and fat content. Highly correlated modules were further investigated to identify regulatory factors or molecular pathways related to each trait.

#### **3.4.4 Hormone-related gene expression changes in a tissue- and temporal-specific manner**

We hypothesized that plant hormones may regulate the initiation and progression of developmental changes in pistachio fruits. We analyzed the expression of various hormone biosynthesis and signaling genes (Figure 3.4). We focused on hormones known to be related to fruit ripening because of the relevance of Stage IV for the gain of fruit quality attributes. ABA biosynthesis was highest during the beginning of Stage IV in the hull and shell as ripening began, and was also significantly enriched in the H-IV-2 module composed of genes with elevated expression during Stage IV. Among biosynthesis genes, the rate-limiting step NCED had the highest expression and drives the pattern observed in Figure 3.4A. Similar to ABA, auxin and cytokinin related genes increased in expression



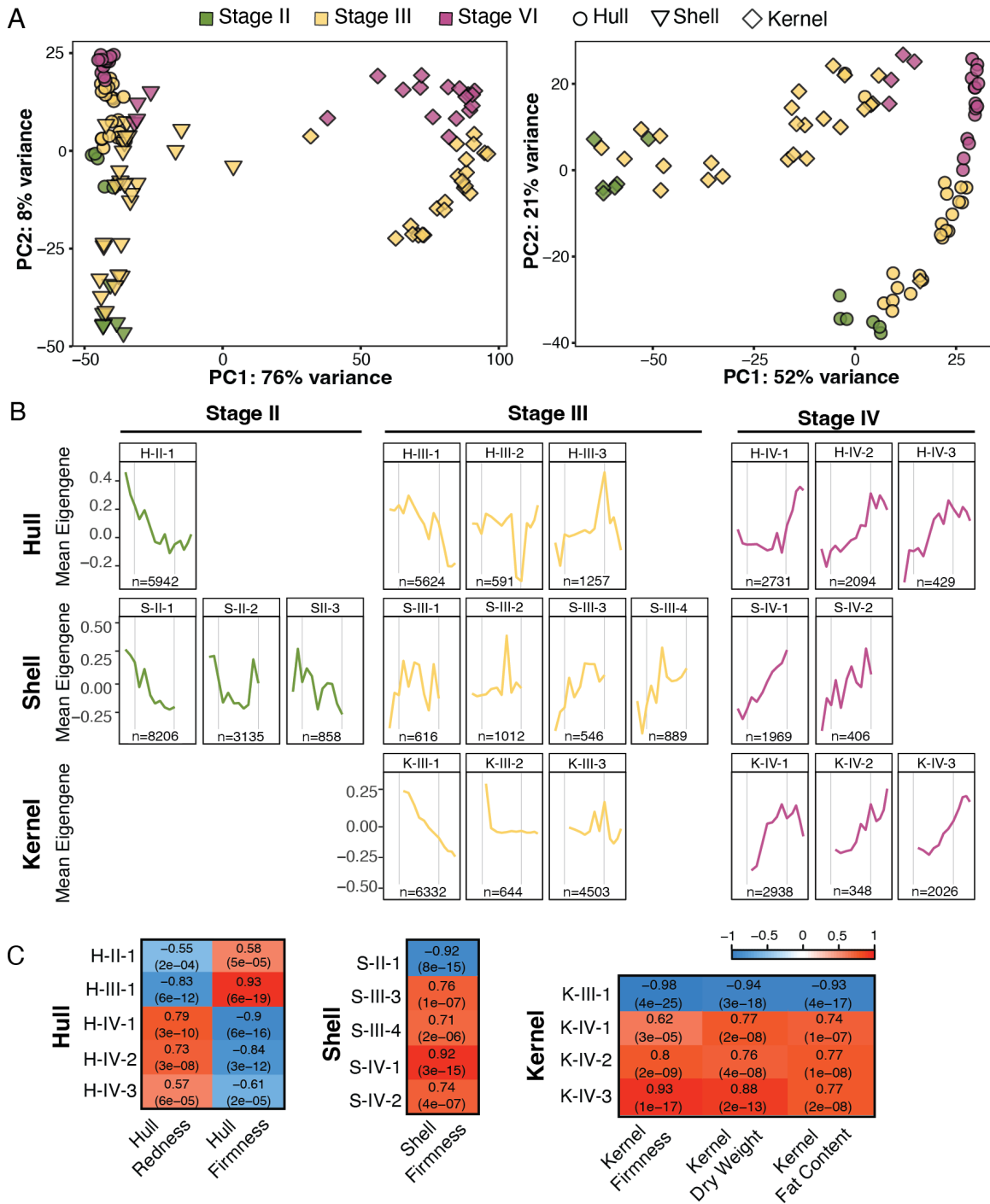


Figure 3.3: Coexpressed gene expression module patterns and their correlations to traits. A) Principal component analysis (PCA) of total gene expression (normalized reads) for all samples, (left) marked by stage (color) and tissue (shape). Right, depicts the distribution of just the shell and hull tissues through development. Weighted gene co-expression (Continued on following page)

Figure 3.3: (Continued from previous page) network analysis conducted for each tissue produced modules of genes with similar expression patterns. B) Selected modules with elevated expression (mean eigengene value) at each developmental stage. Gray lines indicate the transitions between stages. C) Modules were correlated against physiological trait data. Correlations  $R^2 \geq 0.6$  with significance  $P \leq 0.01$  are shown for each tissue.

at the beginning of ripening. The kernel showed the highest levels of GA compared to the hull and shell, peaking at the beginning of ripening. ABA signaling was also higher in the kernel during Stage IV and was also significantly enriched ( $P_{adj} \leq 0.05$ ) in the kernel modules associated with Stage IV (K-IV-1 and K-IV-3) (Figure 3.4B).

We found a significant ( $P_{adj} \leq 0.05$ ) enrichment of ethylene biosynthesis genes in the H-III-3 module (Figure 3.4B). While many of these genes were expressed in this module, the total normalized reads of biosynthesis genes remained constant during Stage III and Stage IV, meaning that ethylene was not likely involved in fruit ripening. Instead, the highest total normalized reads for ethylene biosynthesis in the hull and shell occurred at the beginning of Stage II. JA biosynthesis and signaling genes were also very highly expressed (total normalized reads) at the same time as ethylene in the hull, corresponding to a significant ( $P_{adj} \leq 0.05$ ) enrichment of JA signaling genes in the H-II-1 (Figure 3.4B). The signal transduction network for JA biosynthesis including homologs for JAZ TFs (7 identified), MYC2 TF (4 identified), and COI-1(1 identified) were also coexpressed in the H-II-1 module along with many of the JA biosynthesis genes.

### 3.4.5 Peak in volatiles production preceding ripening

Because hull integrity is important for maintaining the quality of the shell and kernel, we were interested in changes in the chemical composition of the hull during Stage III and IV that may influence ripening and hull breakdown. We measured volatile organic compounds (VOC) weekly from 1106 GDD to 2564 GDD from the same samples collected and analyzed for physiological data and gene expression analysis. At the transition between Stage III and IV, we observed a rise in total VOC production (Figure 6A). Terpenes, the main volatile fraction components in pistachio hulls, were quantified across

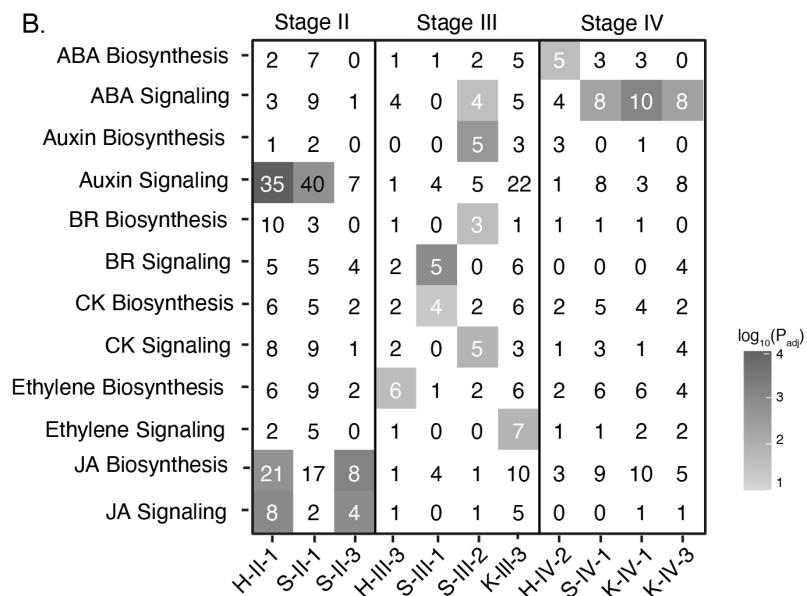
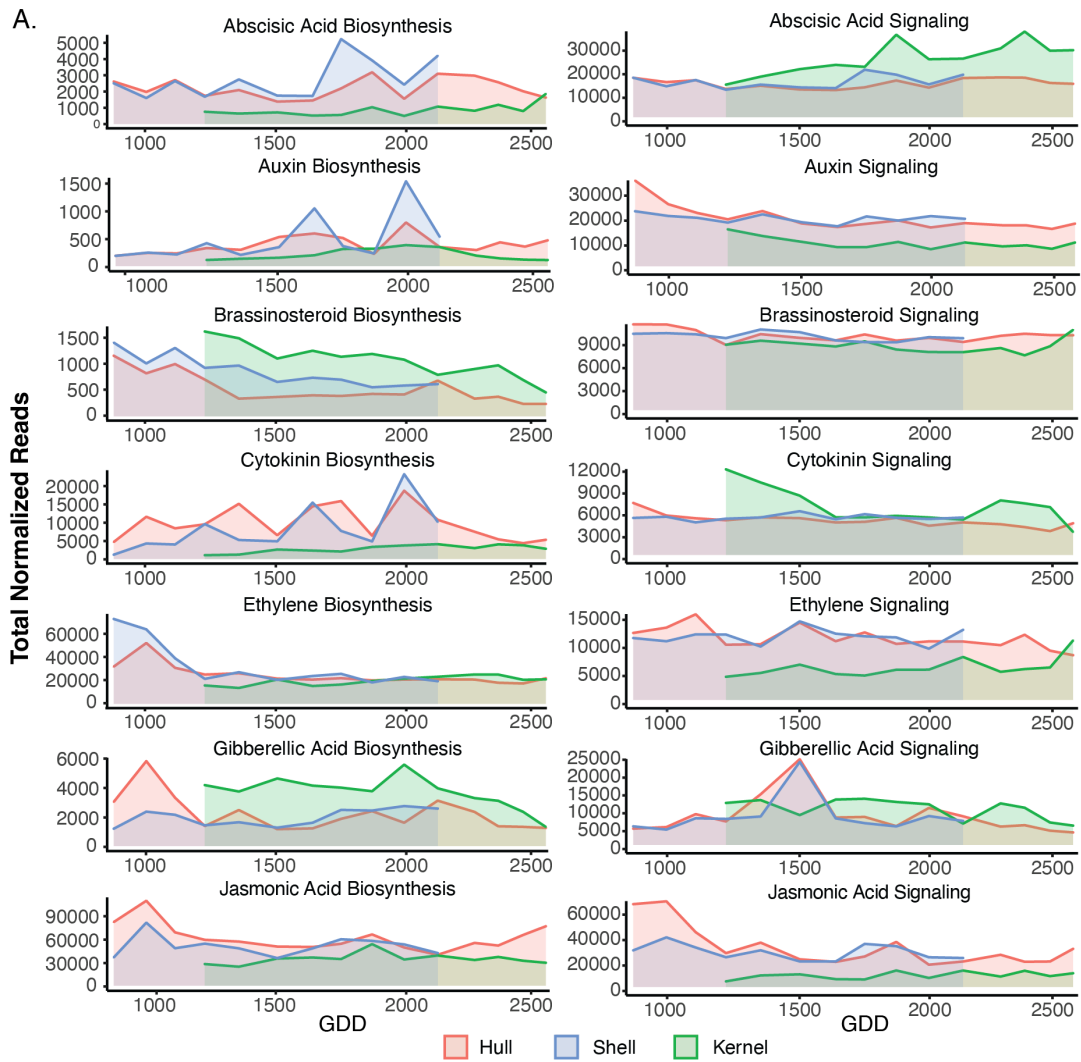


Figure 3.4: Caption presented on following page.

Figure 3.4: (Continued from previous page.) Transcriptional induction of hormone pathways. A) Hormone gene expression through development for abscisic acid (ABA), auxin, brassinosteroid (BR), cytokinin (CK), ethylene, and jasmonic acid (JA). Normalized reads of all annotated biosynthesis genes for each hormone were added together and graphed across sampling time points, represented as the growing degree days (GDD) at sampling. The same was done for annotated hormone signaling genes. C) Functional enrichments of hormone annotated coexpressed genes modules. The heat map colors indicate the significance of the functional enrichment using a  $\log_{10}(1/P_{adj})$  scale. Numbers in each tile indicate the number of genes within each category for each module. Only significant ( $P_{adj} \leq 0.05$ ) functional enrichments are shown.

Stage III and IV and all metabolites exhibited the same pattern peaking at the transition between stages (Figure 3.5B). Limonene and  $\alpha$ -terpinolene were the highest produced monoterpenes, which exhibited the strongest patterns consistently (Supplemental Table: Volatiles). To investigate the events leading to this accumulation of monoterpenes, we performed a Fisher's exact test to identify enriched KEGG pathways in the modules. Terpenoid backbone biosynthesis was significantly enriched in the H-II-1 module (Supplemental Table: KEGG enrichments). Figure 3.4 depicts the MEP terpene backbone biosynthesis producing GPP that leads to monoterpene biosynthesis. A flux of terpene biosynthesis occurred at the end of Stage II and the beginning of Stage III, which indicates precursors for monoterpene metabolism were being synthesized (Figure 3.5C). From the WGCNA, we also identified the top 5% highest connected genes within the module network. In the H-II-1 module, the top highly connected genes included the gene encoding HDR (*ptg00000.1l268*) and a limonene synthase (*ptg000016l.829*) in the terpene biosynthesis pathway.

### 3.4.6 Fruit color changes are associated with carotenoid biosynthesis gene expression

Color changes are a characteristic of fruit ripening. To further define ripening in the pistachio hull, we investigated the underlying biological cause of the change in fruit coloration

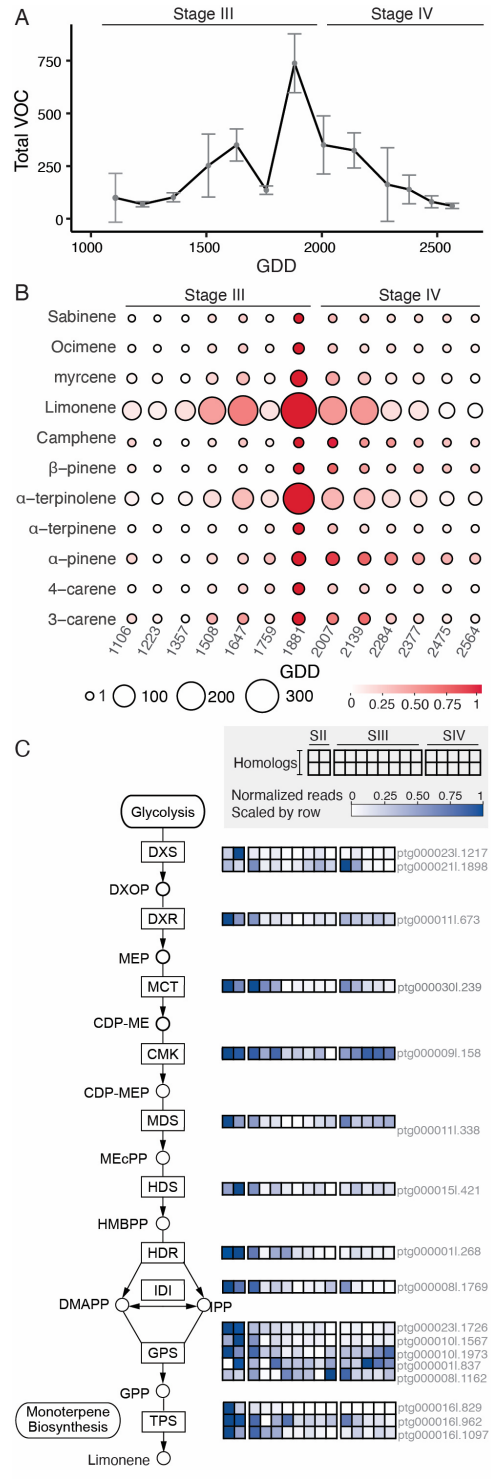


Figure 3.5: Caption presented on following page.

Figure 3.5: (Continued from previous page.) Hull volatile biosynthesis during fruit development. A) Total volatile organic compounds (VOCs) in pistachio hull tissues through time, measured in growing degree days (GDD) through Stage III and IV of fruit development. B) Monoterpene volatile metabolite accumulation through time (GDD). Color of the circles indicates relative changes through time, while size indicates amount of the compound in mg/kg dry weight. C) Pathway of backbone terpene biosynthesis along the MEP pathway leading to the synthesis of monoterpenes. The changes in gene expression are represented in the boxes, with each box representing a sampling date from Stage II through Stage IV and each row representing a gene. Samplings for gene expression include dates from 865 GDD to 2564 GDD. Gene expression was measured in normalized read counts. DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS, (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase; IDI, isopentenyl-diphosphate Delta-isomerase; GPS, geranyl diphosphate synthase; TPS, terpene synthase.

from green-yellow to hues of red-pink observed in the hull during Stage IV (Figure 3.2B). We found a significant correlation between red coloration increase ( $a^*$ ) in the hull with the H-IV-1 ( $R^2=0.79$ ) and H-IV-2 ( $R^2=0.79$ )= $0.73$ ) modules during ripening. This was further supported by a Fisher's exact test for enrichments of KEGG pathways in each module. The H-IV-2 module was significantly enriched ( $P_{adj} \leq 0.05$ ) for the carotenoid biosynthesis pathway. The B-carotene hydrolase (K15746) was the highest expressed carotenoid gene in this module and is annotated to be involved in the production of lutein and zeaxanthin. We examined the highest connectivity genes in this module and among them was a phytoene synthase gene with 887 connections, the rate-limiting step in the carotenoid pathway. Because pink coloration often comes from anthocyanins we also looked at anthocyanin biosynthesis in the hull. While expression was present in the phenylpropanoid and flavonoid pathways, expression was low in the steps exclusive to

anthocyanin biosynthesis. We also found a high negative correlation between the hull redness and the H-III-1 module corresponding to a loss of green coloration. A significant enrichment ( $P_{adj} \leq 0.05$ ) of photosynthesis genes in the same H-III-1 module (Supplemental Table: KEGG enrichments), meaning gene expression of photosynthesis genes decreased after Stage III when fruit became less green.

### 3.4.7 Expression of pectin degrading enzymes increase during hull softening

Consistent with fruit ripening, hull softening occurs during Stage IV coinciding with color changes (Figure 3.2B). To investigate the genetic factors contributing to softening we correlated the hull softening trait with hull module peaking during Stage IV. We found a significant and high negative correlation between hull firmness and the H-IV-1 ( $R^2 = -0.9$ ) and H-IV-2 ( $R^2 = -0.84$ ) modules, respectively (Figure 3.3C). H-IV-1 contains 73 and 42 potential cell wall degrading enzymes (CWDE) in the H-IV-1 and H-IV-2 modules, respectively, which can be associated with fruit softening. We identified the top 5% highest expressed CWDE genes in each module Figure 3.4A, including a pectate lyase and a B-galactosidase, reaching a maximum of 12,585 and 18,209 normalized reads, respectively, during Stage IV. We also identified the top 5% highest connected CWDE within the modules. This produced a total of four genes across both modules that included the degradation of hemicellulose, cellobiose, and pectin (Figure 3.6A). In H-IV-1, the gene *ptg000004l.1160* encoding an  $\alpha$ -L-arabinofuranosidase was among both the most highly connected and the most highly expressed. Selected genes with high correlation between expression pattern and fruit softening ( $1/\text{Firmness}$ ) included a pectin lyase (*ptg000022l.160*), a B-galactosidase (*ptg000002l.1298*),  $\alpha$ -galactosidase (*ptg000002l.930*), and  $\alpha$ -L-arabinofuranosidase (*ptg000004l.1160*) acting synergistically for the hydrolysis of pectins (Figure 3.6B). Together, this indicates that pectin-degrading enzymes are prominent during hull maturation and may play a role in hull softening .

| A             |                 | Module          | Gene                 | CAZY                            | Function | Connectivity | Max Expression |
|---------------|-----------------|-----------------|----------------------|---------------------------------|----------|--------------|----------------|
| Top Expressed | H-IV-1          | ptg000002l.930  | GH36                 | $\alpha$ -galactosidase         | 718      | 16,631.83    |                |
|               |                 | ptg000005l.173  | GH36                 | $\alpha$ -galactosidase         | 608      | 10,973.79    |                |
|               |                 | ptg000004l.1160 | GH51                 | $\alpha$ -L-arabinofuranosidase | 1163     | 10,594.93    |                |
|               | H-IV-2          | ptg000002l.1298 | GH35                 | $\beta$ -galactosidases         | 463      | 18,209.68    |                |
| Top Connected | H-IV-1          | ptg000004l.1160 | GH51                 | Alpha-L-arabinofuranosidase     | 1163     | 10,594.93    |                |
|               |                 | ptg000010l.1331 | GH1                  | $\beta$ -glucosidase            | 915      | 1,415.47     |                |
|               | ptg000010l.1172 | GH1             | $\beta$ -glucosidase | 860                             | 315.78   |              |                |
|               | H-IV-2          | ptg000005l.99   | GH28                 | Polygalacturonase               | 639      | 293.96       |                |

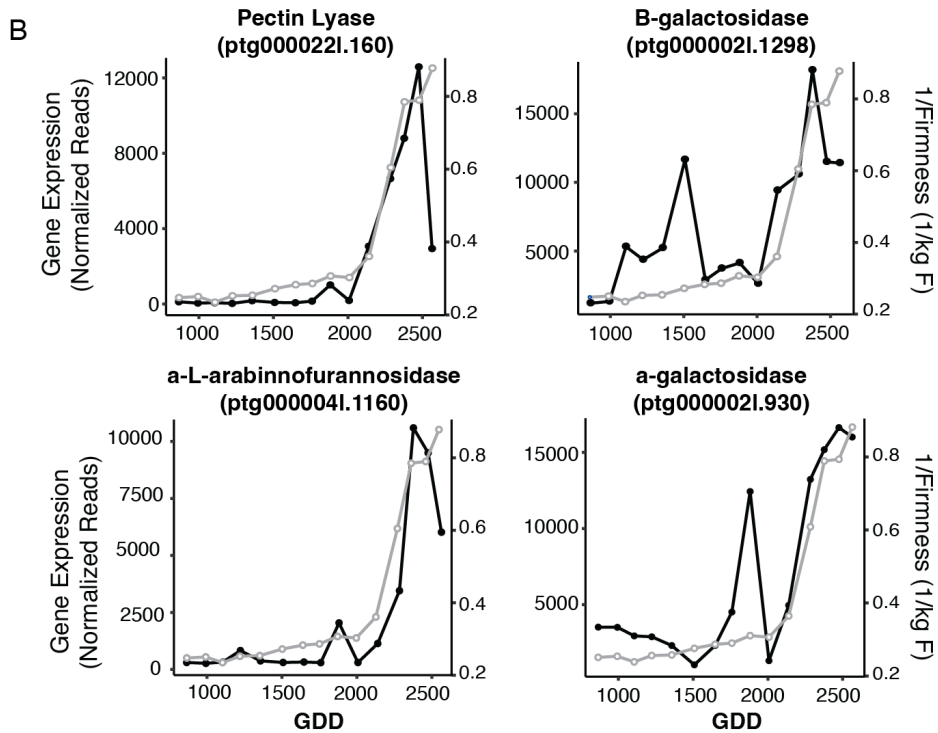


Figure 3.6: Expression of pectin degrading enzymes is correlated with hull softening. A) Table of highest expressed genes and highly connected genes encoding cell wall degrading enzymes. The maximum expression across all developmental time points (865 growing degree days (GDD) to 2564 GDD) is noted as the normalized read count. The top connectivity represents genes with highly connected genes among the H-IV-1 and H-IV-2 coexpression modules (i.e. within the top 5% of each module). Connectivity represents the number of nodes connected to the gene as determined in coexpression network analysis.



Figure 3.6: (Continued from previous page). Four genes from the table with the highest correlation to the firmness trait were graphed. Expression in normalized read counts (left axis, black line) and the inverse of firmness in kg Force (right axis, open circle grey line) was plotted against the GDD at each sampling on the were plotted against.

### 3.4.8 Shell hardening corresponds to an increase in expression of the lignin biosynthesis pathway

We hypothesized shell hardening across Stage III and IV occurs in part due to the lignification of the endocarp (Figure 3.2B). In order to determine the genes contributing to the shell hardening we analyzed module-trait relationships. The increase in firmness in the shell was negatively correlated with the Stage II module S-II-1 ( $R^2=-0.92$ ) and positively correlated with Stage IV modules S-IV-1 ( $R^2=0.92$ ). The woody texture of the shell suggests lignin drives shell hardening. Two distinct patterns of expression represented by the modules emerge along the pathway—genes decreasing after Stage II, represented predominantly by S-II-1, and genes increasing during Stage III and IV, including S-IV-1, as shown by the color key in Figure 3.7. S-II-1 was significantly ( $P_{adj} \leq 0.05$ ) enriched with phenylpropanoid biosynthesis leading to monolignol biosynthesis (Supplemental Table: KEGG Enrichments), depicted in Figure 3.7. Four lignin biosynthesis genes were among the top 5% of gene connectivity in this module (Figure 3.7). Among them, the gene encoding the CAD enzyme (*ptg000001l.1274*) had the most connectivity and was among the top 1% of connections in the module with 6,216 connections. Overall, this suggests that initial lignin deposition occurs early on while other copies of the genes may function later on when shell hardening plateaus. TFs connected to the lignin biosynthesis genes in the S-II-1 module were also of interest as they provide some indication of regulation occurring in the pathway. Among the top 5% we found nine TFs that had connections to genes in the pathway, including MYB (*ptg000001l.187*, *ptg000004l.1378*, *ptg000011l.746*), MADS (*ptg000003l.874*), and C3H (*ptg000004l.863*).

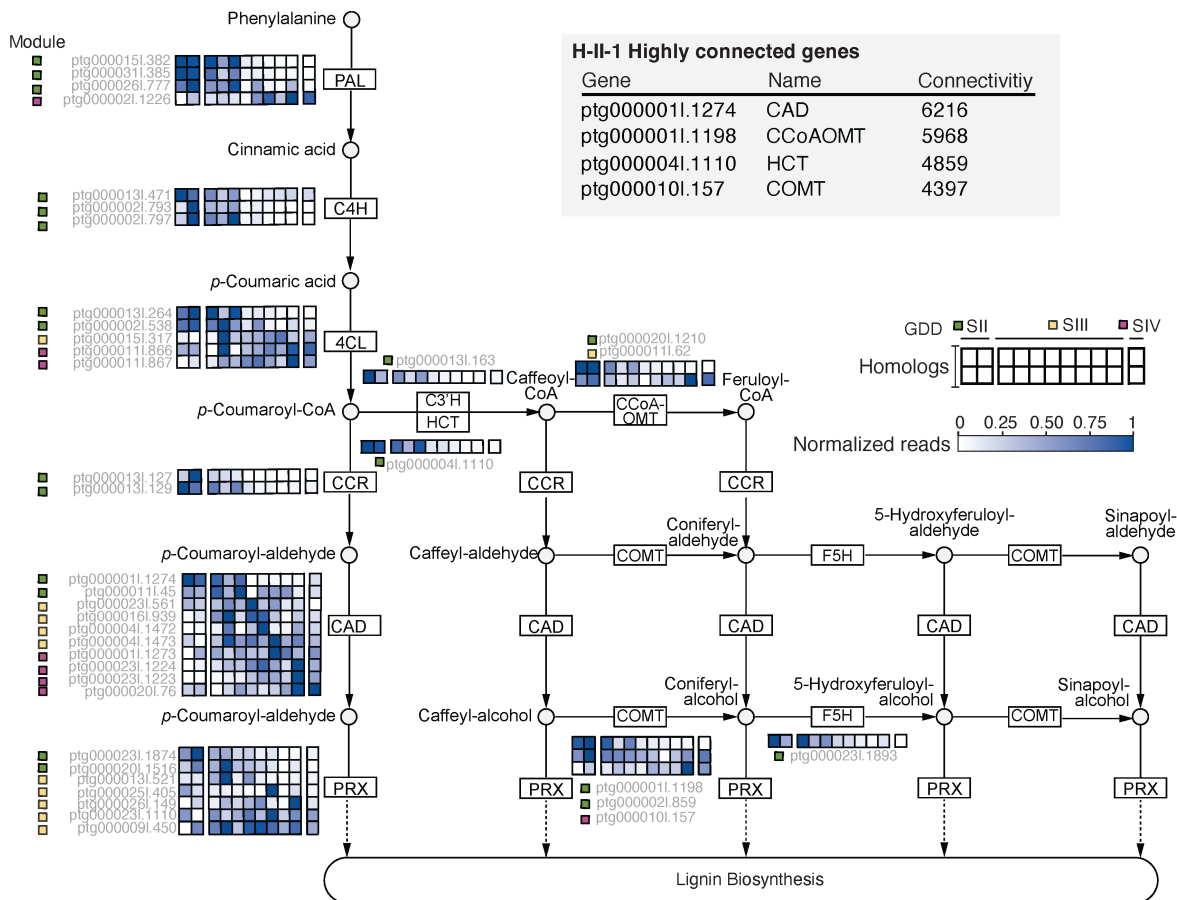


Figure 3.7: Lignin biosynthesis pathway. The phenylpropanoid pathway leading to the synthesis of mono-lignols. The changes in gene expression are represented in the boxes, with each box representing a sampling date from Stage II through Stage IV and each row representing a gene. Samplings for gene expression include dates from 865 growing degree days (GDD) to 2139 GDD. Gene expression was measured in normalized read counts. The table inset represents the most highly connected genes among the S-II-1 coexpression module (i.e. within the top 5% of the module) that were also in the pathway, connectivity represents the number of nodes connected to the gene as determined in coexpression network analysis. PAL, phenylalanine ammonia-lyase; C4H, trans-cinnamate 4-monooxygenase; 4CL, 4-coumarate-CoA ligase; HCT, 5-O-(4-coumaroyl)-D-quinic acid 3'-monooxygenase; C3'H, shikimate O-hydroxycinnamoyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl-alcohol dehydrogenase; COMT, caffeic acid 3-O-methyltransferase; F5H, ferulate-5-hydroxylase; PRX, peroxidase

### 3.4.9 Kernel fatty acids accumulate at a maximum during ripening

Pistachio kernels contain a high proportion of fatty acids and reach their maximum fat content as the kernel matures during ripening (Figure 3.2B). To further understand the composition of the fat content, we measured unsaturated and saturated fatty acids across six timepoints during Stage III and IV of kernel development (Figure 3.8A). Unsaturated fatty acids made up 87% of the total fatty acids present when the fruits were ready to be consumed (i.e., when the commercial harvest would occur). We confirmed that the unsaturated fatty acids were composed of a higher ratio of mono-unsaturated (MUFA) to poly-unsaturated (PUFA) (Figure 3.8A). This ratio changed through time, such that by ripening MUFA were the predominant class of fatty acids present in the fruit.

We determined alterations of metabolites within each class of fatty acid contributed to the changes in MUFA and PUFA ratios during maturation (Figure 3.8B). To further understand what causes these alterations, we examined gene expression of kernel in gene modules associated with the increase in fat content. The module-trait relationships indicated that the increase in fat content was highly and significantly correlated with the K-III-1 ( $R^2 = -0.94$ ) module, along with K-IV-1, K-IV-2, and K-IV-3 (Figure 3.3C). This same relationship was also evident for these same modules and the proportion of unsaturated fatty acids through time (Supplemental Figure 3.S2). We performed an enrichment of KEGG pathway annotations in kernel modules and found that fatty acid biosynthesis was significantly ( $P_{adj} \leq 0.05$ ) enriched in the K-III-1 module (Supplemental Table: KEGG enrichments). The high expression of biosynthesis genes during Stage III indicates that fatty acids are produced early on at the start of kernel development, and taper off at the beginning of Stage IV (Figure 3.8C). Within this module, 19 genes encoding fatty acid biosynthesis were found including key genes FAB2 and FAD2 which desaturates steric acid into oleic acid and oleic acid into linoleic acid, respectively (He et al., 2020). The FAB2 and FAD2 genes were the highest expressed genes in the pathway, and were among the top 5% of genes in the module. FAB2 peaked in expression with 12,500 normalized reads at 1508 GDD while FAD2 peaked with 14,700 normalized reads at 1749 GDD. Con-

sistent with the expression data, the metabolite data also showed that oleic and linoleic acid were the top two produced fatty acids, throughout development. Interestingly, the concentration of linoleic acid decreased over time while oleic acid increased, which was not evident in the expression data.

## 3.5 Discussion

### 3.5.1 The most complete pistachio genome to date

Defining the biological events occurring during pistachio fruit development that lead to traits of interest can allow for breeding and management strategies to improve fruit quality. Further, a high-quality reference genome has been lacking, as previous genomes are incomplete and fragmented. Therefore, in order to facilitate molecular breeding and broaden the understanding of nut tree crop fruit developmental processes, we present for the first time an assembled 561 Mb reference-quality chromosome-scale genome of *P. vera* cv. Kerman. Based on k-mer ( $k = 21$ ) distribution analysis with PacBio HiFi reads, the Kerman genome showed a moderate heterozygosity estimate (0.755%) in comparison with other outcrossing highly heterozygous crops, such as pear 1.6% and grape 1.6-1.7% (Patel et al., 2018; Vurture et al., 2017). This is unexpected because the previously-reported heterozygosity levels of pistachio genome were higher, 1% (Ziya Motalebipour et al., 2016) and 1.72% , which was attributed to the nature of outcrossing by wind pollination and dioecy of pistachio trees (Zeng et al., 2019). In addition, the genome size estimate of 521 Mb in our study was smaller than the first attempt for pistachio genome size estimate (596 Mb) with 26.77 Gb whole genome sequencing (WGS) data using 17-mers (Ziya Motalebipour et al., 2016). However, later, the genome size estimated with the larger amount of data (58.80 Gb) using 21-mers was rather similar in size ( 519.17 Mb) to our assessment (Zeng et al., 2019).

In the final genome assembly, the Kerman genome size (561 Mb) was larger than the estimated size but smaller than previously published genome assemblies of different pistachio cultivars, Batoury (671 Mb) and Siirt (596 Mb) and Bagyolu (623.4 Mb) (Kafkas et al., 2022; Zeng et al., 2019). Although the size variation of the estimated and assem-

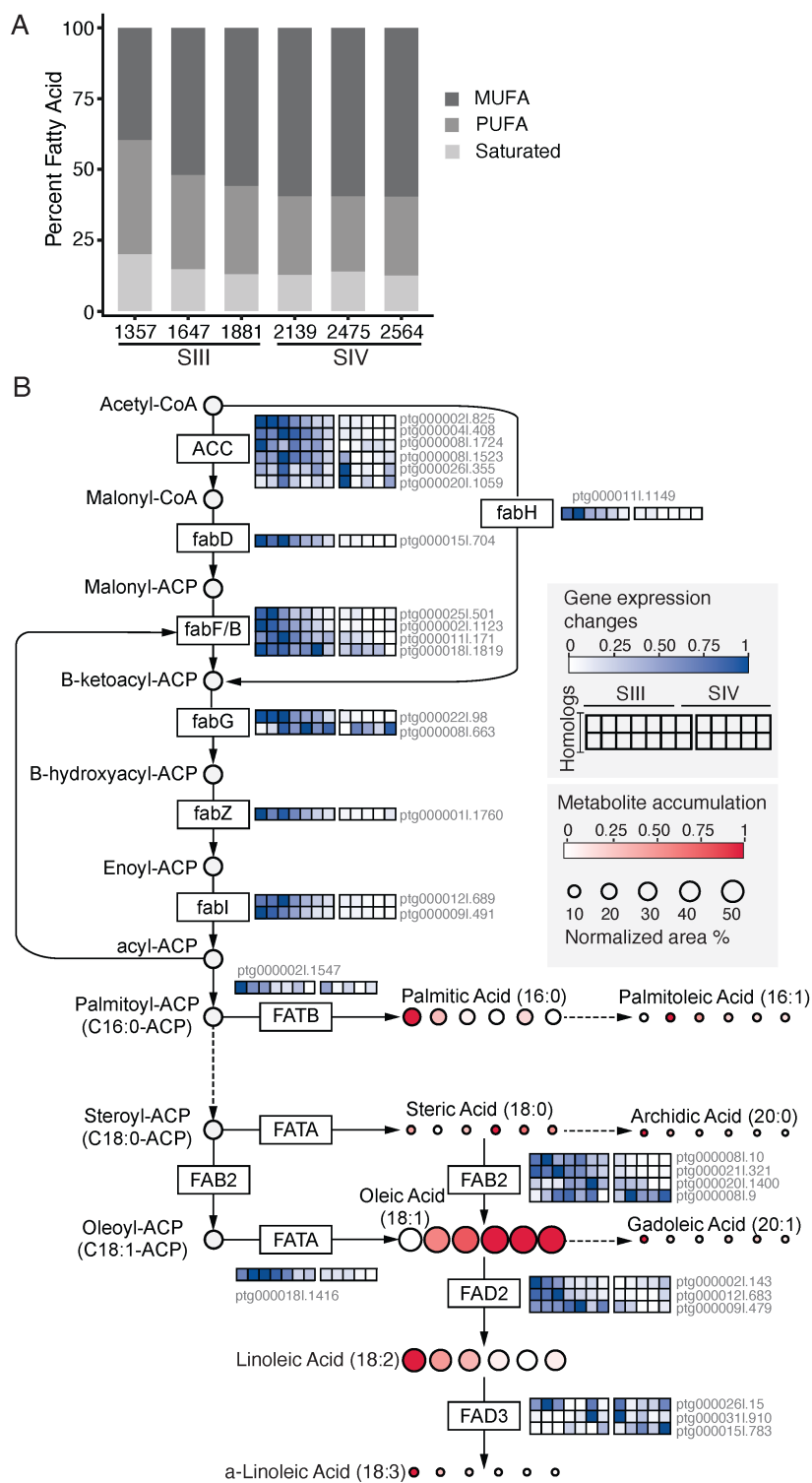


Figure 3.8: Caption presented on following page.

Figure 3.8: (Continued from previous page.) Kernel fatty acid accumulation. A) Proportions of saturated fatty acids, mono-unsaturated fatty acids (MUFA) and poly-unsaturated fatty acids (PUFA) across kernel development, measured in growing degree days (GDD). B) Fatty acid biosynthesis pathway leading to the accumulation of different fatty acid metabolites. The changes in gene expression are represented in the boxes, with each box representing a sampling date from Stage III and Stage IV and each row representing a gene. Samplings for gene expression include dates from 1223 GDD to 2564 GDD. Gene expression was measured in normalized read counts. Fatty acids are represented by colored circles. Colors indicate the changes in metabolites through time and size of each circle represents the amount of the metabolite, measured in normalized area percentages (%). Three timepoints from Stage III (1357, 1647, 1881) and Stage IV (2139, 2475, 2564) were used for all fatty acid data shown. ACA, acetyl-CoA carboxylase; fabD, S-malonyltransferase; fabF, 3-oxoacyl-[acyl-carrier-protein] synthase II; fabH, 3-oxoacyl-[acyl-carrier-protein] synthase III; fabG, 3-oxoacyl-[acyl-carrier protein] reductase; fabZ, 3-hydroxyacyl-[acyl-carrier-protein] dehydratase; fabI, enoyl-[acyl-carrier protein] reductase I; FATB, fatty acyl-ACP thioesterase B; FATA, fatty acyl-ACP thioesterase A; FAB2, acyl-[acyl-carrier-protein] desaturase; FAD2, omega-6 fatty acid desaturase; FAD3, acyl-lipid omega-3 desaturase.

bled pistachio genome assemblies could be explained by possible genome size variation across different cultivars as documented in other plant genomes, it is likely that pseudoduplication in the assemblies, especially from the highly repetitive regions in chromosome arms in the case of pistachio, is the primary cause of assembly size variation (Lovell et al., 2021; Ortega-Ortega et al., 2019; Yu et al., 2019). In 2015, Sola-Campoy and colleagues characterized massive enrichment of 180 bp repeat (PIVE-180) on one arm of 11 chromosomes in pistachio, which was also observed in the Kerman genome (Sola-Campoy et al., 2015). The largest region with dense distribution of 180 bp repeats reached about 9 Mb in chromosome 7, where no protein-coding gene was annotated (Figure 3.1B). These extremely repetitive regions could have been the major issues of accurate pistachio genome assembly and chromosome construction. Although Omni-C reads are known to offer uni-

form coverage across the genome without RE sites over represented, it was observed that the overall coverage of Omni-C reads was significantly lower in those regions in Omni-C analysis, likely due to the limitation of mapping capability (Figure 3.1C). Therefore, more careful validation on those regions is needed to improve the pistachio genome.

The annotation of intact and fragmented transposable elements in the Kerman assembly resulted in about 11% higher genome coverage (65%) than the estimated repetitiveness (54%) (Figure 3.1A and Table 3.3). As discussed in genome size and estimate, this is likely caused by the pseudo-duplication in the assembly or misestimation due to exceptionally repeat-dense regions in chromosome arms. Among 65% of repetitive regions, nearly 49% of the genome was composed of LTRs, which have been widely known as the dominant TE groups in plants (Hawkins et al., 2009) and can play a major role in adaptation and evolution by introducing novel genetic material. The protein-coding gene annotation shows high completeness based on BUSCO assessment with almost 99% (Table 2.6). However, minor improvements can still be made by filtering out false-positive gene models and recovering missing BUSCO genes.

Macrosynteny patterns between *Pistacia vera* cv. Kerman, *Mangifera indica* (mango), and *Citrus sinensis* (sweet orange) provided evidence that *P. vera* has not experienced a lineage-specific whole genome duplication event (WGD), unlike the recent WGD which occurred in the mango genome as described in (Wang et al., 2020) (Figure 3.1F). The synteny between mango and pistachio genomes and the similarity between their fruit morphology (both drupes) and growth patterns provides an interesting evolutionary comparison within the Anacardiaceae family.

### **3.5.2 Pistachios exhibit a unique asynchronous development of fruit and seed**

During the growing season, pistachios undergo a unique asynchronous development of the kernel (embryo) and maternal tissues. The hull and shell develop together in the first months, marking Stages I and II (Figure 3.2), while embryo development takes place during Stages III and IV. In contrast to previous reports, we found that shell hardening (formerly Stage II) continues to take place with kernel growth (formerly Stage III)

starting in late June at approximately 1000 GDD through late August at approximately 2000 GDD (Goldhamer and Beede, 2004). The asynchronous developmental pattern between the fruit and embryo has not been well described in the literature for other tree crops. While peaches appear to exhibit a similar pattern in seed development, this trait does not seem to have been studied in a crop whose seed is consumed (Bonghi et al., 2011). Carbohydrate dynamics in the tree may offer some explanation of the asynchrony. Carbohydrates reserved from the prior year are utilized by the tree to produce buds and develop fruit in early spring, through Stage I (Tixier et al., 2020). The lull in fruit growth identified as Stage II may serve as a transition between a net carbon loss and a net carbon gain in photosynthesis (Marino et al., 2022) leading to the growth of the kernel.

The RNAseq experiment assessed genetic changes through time and tissue type during fruit development. The shell and hull have the most similar gene expression patterns (Figure 3.3A). This was obvious in the expression of hormone-related gene expression. The shell and hull tissues exhibited very similar expression patterns for each hormone biosynthesis pathway, while the kernel expression patterns were distinct (Figure 3.4A). Interestingly, the hull and shell total gene expression became more similar over time (Figure 3.3A). This contrasts with the morphology of the tissues, which early on in development are physically fused together and appear to become increasingly different through time as shells become woody and split, and the hull and shell tissues separate during ripening (Polito and Pinney, 1999). The similarity in gene expression may be due to both tissues undergoing terminal developmental programs. This occurs earlier in the shell when the tissue reaches its peak firmness at the beginning of Stage IV, while in the hull this occurs at the end stages of Stage IV as ripening finishes.

Shell lignification was previously reported to start as early as May-June, falling in Stage I-II (Polito and Pinney, 1999). While the secondary cell walls become lignified, the shells are green and flexible at this point. However, as described above, the texture of the shell continues to change through Stage III leading to a woody tissue that then splits (Figure 3.2B). The shell tissues appear to senesce and be fully lignified at around 2100 GDD, as RNA content became very low in shell tissues after this point. Our gene expres-



sion analysis found a proportion of the genes involved in the phenylpropanoid pathway leading to monolignols to be expressed highest at Stage II (S-II-1) followed by a sharp decline, marking the initial lignification (Figure 3.7). The genes exhibiting this pattern were among the highest expressed homologs; however, other copies of the genes displayed patterns with peak expression later on during Stage III or IV indicating lignin was still being produced, contributing to the increased firmness of the shell. This suggests that the lignification process does not complete until the shell reaches peak firmness, as has been described in walnuts (Li et al., 2021). While continued lignification may be a factor leading to shell firmness changes, other factors such as cell wall modifications likely also contribute, but require further investigation. Overall, understanding the composition and alterations in the shell tissue will be important to ascertaining the underlying mechanisms leading to shell split for a higher quality nut.

### **3.5.3 Pistachio fruit exhibit a non-climacteric ripening pattern**

Although ripening has not previously been well explored in fruit tree crops, early reports suggest that pistachios are non-climacteric fruit (Labavitch et al., 1982.). We confirmed ethylene is not produced in a climacteric pattern during ripening and remains at constant low levels, as shown through biosynthesis gene expression (Figure 3.4A). In conjunction with this we found evidence that abscisic acid (ABA) may be involved in regulating ripening in pistachio. ABA has been found to be the primary hormone involved in non-climacteric (ethylene independent) ripening (Forlani et al., 2019). NCED is the rate limiting enzyme in ABA biosynthesis (Forlani et al., 2019). We found that a primary copy was expressed in the shell and hull tissues right before ripening changes began to occur, i.e., the transition between Stage III and Stage IV. This corresponded to an increase in ABA signaling genes such as, PYLs, PP2C, SNRK2, and ABFs, suggesting ABA is active at the onset of ripening (Supplemental Table: All Expressed Genes). In addition to ABA, other hormone biosynthesis genes exhibited a similar expression such as auxin and cytokinin, both of which have been implicated in non-climacteric fruit maturation and ripening (McAtee et al., 2013). Thus, pistachio should be considered a non-climacteric fruit.

While ethylene was not important for ripening in pistachio, ethylene may be more critical prior to kernel development at the end of Stage II where a rise in biosynthesis occurs (Figure 3.4A). The exact function of the hormone at this stage of development is unknown but has previously been suggested to be involved in bud abscission in alternate bearing years and needs further investigation (Vemmos et al., 1994). JA-related genes were also elevated during Stage II and were among the highest expressed hormone-related genes. JA is best known as a stress hormone involved in many responses to abiotic and biotic stress but can also function in fruit development (Gapper et al., 2013). JA biosynthesis genes were previously shown to be differentially expressed in pistachio vegetative tissues that underwent a salinity treatment, compared to the control (Zeng et al., 2019). However, there were no known stresses occurring during our samplings to explain elevated expression. This suggests JA is a critical hormone in pistachio fruit Stage II development prior to kernel initiation, or other events occurring concurrently. As pistachio is known to be tolerant to environmental stress, JA and ethylene levels may play a role evolutionarily to adapt to these extreme environments.

### **3.5.4 Volatiles produced at the onset of ripening offer insights to fruit-disperser relationships**

The hull functions as a protective tissue encapsulating the shell and kernel. The breakdown of the hull caused by senescence can lead to a lower quality commodity; for example, the shell becomes more vulnerable to staining from the hull and the kernel becomes more accessible to pests. The hull is rich in volatile compounds mainly composed of terpenes (Arjeh et al., 2020; Chahed et al., 2008). We examined the dynamic changes of volatile compounds in the hull during Stage III and IV to gain insight into the events leading to hull ripening and senescence. We saw a rise of volatile compounds at the onset of ripening that may offer a signal of developmental changes occurring (Figure 3.5.) We observed limonene and alpha-terpinolene to be the highest produced monoterpenes consistently across years. Alpha-terpinolene had also been found in high proportions of Tunisian pistachio variety but did not have as high of concentrations of limonene, which could be due to varietal differences (Chahed et al., 2008). Limonene has been shown to accumulate

in orange peel with ripening in order to attract insects and pathogens (Rodríguez et al., 2011). This relationship between limonene and other organisms was proposed to have evolved to facilitate seed dispersal, opening up the fruit to expose the seeds. This could also be the case for pistachio, in which the volatile production at the start of ripening signals that the kernel is mature and attracts seed dispersing organisms. Volatile signals further define the events leading to ripening in pistachio and have additional implications for management practices to time treatments against insects.

### **3.5.5 Hulls undergo physiological and gene expression changes consistent with fruit ripening**

Like canonical fleshy fruit, pistachio fruit quality is determined during fruit ripening. Ripening changes in the hull coincide with important quality traits (i.e. shell split, kernel maturation) and can be used to anticipate the best harvest time. Therefore, understanding the timing and relationship between the hull and kernel during Stage IV allows for increased quality. Our study integrates multiple approaches, including physiology, biochemistry, and genomics, to provide the most thorough understanding of pistachio fruit development to date. Fruit ripening in pistachio to our knowledge has not been previously explored. The hull undergoes changes in composition preceding harvest consistent with fruit ripening including, softening and color change which we define as an additional stage, Stage IV (Figure 3.2B). These changes are important attributes that help determine harvest time and maximize fruit quality. For example, hull softening allows fruits to be detached from the tree, however, if overripe the hull senesces and the degradation can cause shell stain and make kernels vulnerable to pests and disease decreasing its nutritional and market value.

We integrated gene expression data with the observed physical changes to inform the events occurring leading to fruit ripening. Color change is a characteristic of ripening and provides a visual indication of when fruit are ready for harvest. Previous studies have identified anthocyanin, carotenoids, and chlorophyll compounds in pistachio hulls (Arjeh et al., 2020; Grace et al., 2016; Mandalari et al., 2022). However, the proportion of these compounds present depended on the stage sampled and variety, with measure-

ments limited to pistachio green hulls prior to ripening. Thus, it is unclear which specific compounds lead to the pink colorations in the hull during ripening. Our gene expression analysis indicated that the flavonoid pathways were active in the hull, however there was not a strong expression of anthocyanin reductase genes, the critical final step for anthocyanin production, giving the compounds their pigmentation (Supplemental Table: All Gene Expression; (Gu et al., 2019)). It is clear that the fruit produce some anthocyanins because they have been identified in the purple colored seed coat surrounding the kernel (Mandalari et al., 2022). We found stronger gene expression of carotenoid biosynthesis. Among the pigments identified in Grace et al., lutelin was the highest measured in pistachio hulls (Grace et al., 2016). We observed high expression of *ctrZ* which is annotated to act in the step leading to lutein biosynthesis. The change in green coloration to yellow (Figure 3.2) at the end of Stage III also indicates chlorophyll degradation may be occurring. We identified several chlorophyll degradation genes expressed in modules with ripening patterns, such as STAY GREEN (*ptg000011l.1194*), a chlorophyllide reductase that regulates chlorophyll protein degradation (Hörtensteiner, 2009). Thus, from our analysis we hypothesize that hull color becomes yellow from chlorophyll degradation and shades of pink from carotenoids.

It is well known that fruit softening is mediated by cell wall degrading enzymes in both climacteric and non-climacteric fruits, such as tomato and strawberry (Adaskaveg and Blanco-Ulate, 2023). Cell wall degrading enzymes acting on the backbone of pectin molecules, such as polygalacturonase and pectin lyase, are highly expressed in other fruit and exhibit a ripening-specific gene expression pattern (Wang et al., 2018). Further,  $\alpha$ -L-AFase is highly expressed and acts as a catalyst with other coexpressed cell wall degrading enzymes (Shi et al., 2023). Thus, we were interested if these enzymes were expressed in a ripening-specific pattern. Pectate lyase was among the highest expressed CWDEs annotated from the CAZy database in pistachio and began to rise in expression at Stage IV (Figure 3.5A). Consistent with this, pectins have been measured in pistachio hulls and were proposed as a potential source of commercial pectins (Arjeh et al., 2020; Kazemi et al., 2019). The presence of a large proportion of pectins in green (unripe) hulls and

the high expression of pectin-degrading enzymes suggest these enzymes promote pectin degradation and softening of the tissue. The mechanisms involved in this still need to be explored. Overall, knowing how and when hull softening occurs during the growing season can help advise the optimal time for harvest.

### **3.5.6 Seed maturity coincides with fruit ripening**

Kernel growth during Stage III leads to the maturity of the seed and ripening of the hull. Understanding when the kernel is most desirable for consumption and when the fruit is ready for harvest can improve management practices and fruit quality. Maturity can be observed when kernels reach their maximum size and fat content at the start of Stage IV, as ripening progresses (Figure 3.2B). From our gene expression and metabolite data we see that fatty acid biosynthesis occurs early on in kernel development and is primarily composed of unsaturated fatty acids, with much of the production reaching its maximum during ripening (Figure 3.8). Further, the kernel shows hormonal indications of seed maturity with an increase of GA at the start of Stage IV and ABA increasing throughout Stage IV.

Pistachio kernels are consumed for their unique flavor and nutritional benefits. Kernels are made up of primarily unsaturated fats, including both poly- and mono- unsaturated fatty acids (Figure 3.8A) . Although PUFAs provide essential nutrients to the human diet, they make kernels more vulnerable to rancidity, reducing their shelf life (Dar et al., 2017). Thus, the ratios of PUFA and MUFA are important for considering nutritional benefits and shelf life. We identified important enzymes in our gene expression data explaining the accumulation of specific unsaturated fats (Figure 3.8B). Interestingly, our study showed fluctuations in the composition of unsaturated fatty acids through kernel development similar to a previous study (Polari et al., 2019). The mono-unsaturated fatty acid (MUFA) oleic acid increases through time while the poly-unsaturated (PUFA) linoleic acid decreases. These fluctuations were not completely explainable with our expression data, but are likely caused by other fluxes in the fatty acid metabolism downstream of these compounds.

## 3.6 Conclusions

In this study, we provided a complete timeline of fruit development in pistachio across the growing season and validated it across three independent locations and years. Our resulting models of fruit growth provided new insights into pistachio fruit growth. We divided its development into four distinct stages which can be used to help advise growers to have a better estimate of when critical events occur, based on accumulated heat. The high quality reference genome and annotations created for this study is the first of its kind for *P. vera* and will serve as a reference for all future genomics work in pistachio as well as, gene functional analysis and evolution studies. We confirmed that physical changes in the fruit over the growing season can be explained by changes in gene expression and identified specific pathways associated with relevant quality traits. Our transcriptomics work validates previous studies and expands gene annotations and functional annotations for fruit specific gene expression (Moazzam Jazi et al., 2017; Zeng et al., 2019). This study not only provides further avenues to study specific processes or events critical to understanding tree fruit development, but will aid in informing management practices that will increase fruit quality, such as determining the best harvest times.

## References

- Adaskaveg, J. A. and Blanco-Ulate, B. (2023). Targeting ripening regulators to develop fruit with high quality and extended shelf life. *Current Opinion in Biotechnology*, 79:102872.
- Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., Wang, X., Lippman, Z. B., Schatz, M. C., and Soyk, S. (2022). Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.*, 23(1):258.
- Andrews, S. and Others (2010). FastQC: a quality control tool for high throughput sequence data.

- Arjeh, E., Akhavan, H. R., Barzegar, M., and Carbonell-Barrachina, Á. A. (2020). Bioactive compounds and functional properties of pistachio hull: A review.
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.
- Blanco-Ulate, B., Vincenti, E., Powell, A. L. T., and Cantu, D. (2013). Tomato transcriptome and mutant analyses suggest a role for plant stress hormones in the interaction between fruit and *Botrytis cinerea*. *Frontiers in Plant Science*, 4:142.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120.
- Bonghi, C., Trainotti, L., Botton, A., Tadiello, A., Rasori, A., Ziliotto, F., Zaffalon, V., Casadoro, G., and Ramina, A. (2011). A microarray approach to identify genes involved in seed-pericarp cross-talk and development in peach. *BMC Plant Biology*, 11(1):1–14.
- Br ° una, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom Bioinform*, 3(1):lqaa108.
- Chahed, T., Dhifi, W., Hosni, K., Msaada, K., Kchouk, M. E., and Marzouk, B. (2008). Composition of tunisian pistachio hull essential oil during fruit formation and ripening. *Journal of Essential Oil Research*, 20(2):122–125.
- Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods*, 18(2):170–175.
- Claeskens, G. and Hjort, N. L. (2008). *Model selection and model averaging*. Cambridge University Press.

- Corelli-Grappadelli, L. and Lakso, A. N. (2004). Fruit development in deciduous tree crops as affected by physiological factors and environmental Conditions. In *Acta Horticulturae*, volume 636, pages 425–441. International Society for Horticultural Science.
- Dar, A. A., Choudhury, A. R., Kancharla, P. K., and Arumugam, N. (2017). The FAD2 gene in plants: Occurrence, regulation, and role. *Frontiers in Plant Science*, 8:1789.
- Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Research*, 50(D1):D571–D577.
- Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., and Aiden, E. L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*, 356(6333):92–95.
- Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S. P., Huntley, M. H., Lander, E. S., and Aiden, E. L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*, 3(1):95–98.
- Emms, D. M. and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.*, 20(1):238.
- Ferguson, L., Polito, V., and Kallsen, C. (2005). The pistachio tree; botany and physiology and factors that affect yield. *Pistachio production manual, 4th ed. Davis, CA, USA, University of California Fruit Nut Research Information Center*, pages 31–39.
- Forlani, S., Masiero, S., and Mizzotti, C. (2019). Fruit ripening: the role of hormones, cell wall modifications, and their relationship with pathogens. *Journal of Experimental Botany*, 70(11):2993–3006.
- Gapper, N. E., McQuinn, R. P., and Giovannoni, J. J. (2013). Molecular and genetic regulation of fruit ripening. *Plant Molecular Biology*, 82(6):575–591.



- Goldhamer, D. A. and Beede, R. H. (2004). Regulated deficit irrigation effects on yield, nut quality and water-use efficiency of mature pistachio trees. *Journal of Horticultural Science and Biotechnology*, 79(4):538–545.
- Grace, M. H., Esposito, D., Timmers, M. A., Xiong, J., Yousef, G., Komarnytsky, S., and Lila, M. A. (2016). Chemical composition, antioxidant and anti-inflammatory properties of pistachio hull extracts. *Food Chemistry*, 210:85–95.
- Gu, K. D., Wang, C. K., Hu, D. G., and Hao, Y. J. (2019). How do anthocyanins paint our horticultural products? *Scientia Horticulturae*, 249:257–262.
- Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., and Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol.*, 9(1):R7.
- Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., and Chen, J. (2020). RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the ideograms. *PeerJ Comput Sci*, 6:e251.
- Hawkins, J. S., Proulx, S. R., Rapp, R. A., and Wendel, J. F. (2009). Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. *Proc. Natl. Acad. Sci. U. S. A.*, 106(42):17811–17816.
- He, M., Qin, C. X., Wang, X., and Ding, N. Z. (2020). Plant Unsaturated Fatty Acids: Biosynthesis and Regulation.
- Hörtensteiner, S. (2009). Stay-green regulates chlorophyll and chlorophyll-binding protein degradation during senescence.
- Kafkas, S., Ma, X., Zhang, X., Topçu, H., Navajas-Pérez, R., Wai, C. M., Tang, H., Xu, X., Khodaeiaminjan, M., Güney, M., Paizila, A., Karıcı, H., Zhang, X., Lin, J., Lin, H., de la Herrán, R., Rejón, C. R., Garc´-Zea, J. A., Robles, F., Muñoz, C. D. V., Hotz-Wagenblatt, A., Min, X. J., Özkan, H., Motalebipour, E. Z., Gozel, H., Çoban,

- N., Kafkas, N. E., Kilian, A., Huang, H., Lv, X., Liu, K., Hu, Q., Jacygrad, E., Palmer, W., Michelmore, R., and Ming, R. (2022). Pistachio genomes provide insights into nut tree domestication and ZW sex chromosome evolution. *Plant Commun*, page 100497.
- Kallsen, C. E., Parfitt, D. E., and Maranto, J. (2020). UC pistachio cultivars show improved nut quality and are ready for harvest earlier than 'Kerman'. *California Agriculture*, 74(2):86–93.
- Kazemi, M., Khodaiyan, F., Labbafi, M., Saeid Hosseini, S., and Hojjati, M. (2019). Pistachio green hull pectin: Optimization of microwave-assisted extraction and evaluation of its physicochemical, structural and functional properties. *Food Chemistry*, 271:663–672.
- Keller, O., Kollmar, M., Stanke, M., and Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics*, 27(6):757–763.
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.*, 37(8):907–915.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13):1–26.
- Langfelder, P. and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1):559.
- Li, P., Wang, H., Liu, P., Li, Y., Liu, K., An, X., Zhang, Z., and Zhao, S. (2021). The role of JrLACs in the lignification of walnut endocarp. *BMC Plant Biology*, 21(1):1–17.
- Lovell, J. T., Bentley, N. B., Bhattarai, G., Jenkins, J. W., Sreedasyam, A., Alarcon, Y., Bock, C., Boston, L. B., Carlson, J., Cervantes, K., Clermont, K., Duke, S., Krom, N., Kubenka, K., Mamidi, S., Mattison, C. P., Monteros, M. J., Pisani, C., Plott, C., Rajasekar, S., Rhein, H. S., Rohla, C., Song, M., Hilaire, R. S., Shu, S., Wells, L.,

- Webber, J., Heerema, R. J., Klein, P. E., Conner, P., Wang, X., Grauke, L. J., Grimwood, J., Schmutz, J., and Randall, J. J. (2021). Four chromosome scale genomes and a pan-genome annotation to accelerate pecan tree breeding. *Nat. Commun.*, 12(1):4125.
- Lovell, J. T., Sreedasyam, A., Schranz, M. E., Wilson, M., Carlson, J. W., Harkess, A., Emms, D., Goodstein, D. M., and Schmutz, J. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *Elife*, 11.
- Mandalari, G., Barreca, D., Gervasi, T., Roussel, M. A., Klein, B., Feeney, M. J., and Carughi, A. (2022). Pistachio nuts (*Pistacia vera* l.): Production, nutrients, bioactives and novel health effects.
- Marçais, G. and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27(6):764–770.
- Marino, G., Guzmán-Delgado, P., Caruso, T., and Marra, F. P. (2022). Modeling seasonal branch carbon dynamics in pistachio as a function of crop load. *Scientia Horticulturae*, 296(November 2021).
- Marino, G., La Mantia, M., Caruso, T., and Marra, F. P. (2018). Seasonal dynamics of photosynthesis and total carbon gain in bearing and nonbearing pistachio (*Pistacia vera* L.) shoots. *Photosynthetica*, 56(3):932–941.
- McAtee, P., Karim, S., Schaffer, R., and David, K. (2013). A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening.
- Moazzam Jazi, M., Seyedi, S. M., Ebrahimie, E., Ebrahimi, M., De Moro, G., and Botanga, C. (2017). A genome-wide transcriptome map of pistachio (*Pistacia vera* L.) provides novel insights into salinity-related genes and marker discovery. *BMC Genomics*, 18(1):1–21.
- Ortega-Ortega, J., Ram´-Ortega, F. A., Ruiz-Medrano, R., and Xoconostle-Cázares, B. (2019). Analysis of Genome Size of Sixteen *Coffea arabica* Cultivars Using Flow Cytometry. *HortScience*, 54(6):998–1004.

- Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J. R. A., Hellinga, A. J., Lugo, C. S. B., Elliott, T. A., Ware, D., Peterson, T., Jiang, N., Hirsch, C. N., and Hufford, M. B. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.*, 20(1):275.
- Patel, S., Lu, Z., Jin, X., Swaminathan, P., Zeng, E., and Fennell, A. Y. (2018). Comparison of three assembly strategies for a heterozygous seedless grapevine genome assembly. *BMC Genomics*, 19(1):57.
- Pertea, G. and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Res.*, 9(304):304.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.*, 33(3):290–295.
- Polari, J. J., Ferguson, L., and Wang, S. C. (2020). Pistachio kernel composition of ‘Kalehghouchi’, ‘Pete 1’, and ‘Lost Hills’ in California. *HortScience*, 55(5):666–669.
- Polari, J. J., Zhang, L., Ferguson, L., Maness, N. O., and Wang, S. C. (2019). Impact of Microclimate on Fatty Acids and Volatile Terpenes in “Kerman” and “Golden Hills” Pistachio (*Pistacia vera*) Kernels. *Journal of Food Science*, 84(7):1937–1942.
- Polito, V. S. and Pinney, K. (1999). Endocarp dehiscence in pistachio (*Pistacia vera* L.). *International Journal of Plant Sciences*, 160(5):827–835.
- Ripley, B., Venables, B., Bates, D. M., ca 1998, K. H., A. G. (2021). Support Functions and Datasets for Venables and Ripley’s MASS [Computer software].
- Robinson, J. T., Turner, D., Durand, N. C., Thorvaldsdóttir, H., Mesirov, J. P., and Aiden, E. L. (2018). Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. *Cell Syst*, 6(2):256—258.e1.
- Rodríguez, A., Andrés, V. S., Cervera, M., Redondo, A., Alquézar, B., Shimada, T., Gadea, J., Rodrigo, M., Zacarías, L., Palou, L., López, M. M., Castañera, P., and

- Peña, L. (2011). The monoterpene limonene in orange peels attracts pests and microorganisms. *Plant Signaling and Behavior*, 6(11):1820–1823.
- Shi, Y., Li, B.-J., Grierson, D., and Chen, K.-S. (2023). Insights into cell wall changes during fruit softening from transgenic and naturally occurring mutants. *Plant Physiology*.
- Shumate, A. and Salzberg, S. L. (2020). Liftoff: accurate mapping of gene annotations. *Bioinformatics*.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19):3210–3212.
- Sola-Campoy, P. J., Robles, F., Schwarzacher, T., Ruiz Rejón, C., de la Herrán, R., and Navajas-Pérez, R. (2015). The Molecular Cytogenetic Characterization of Pistachio (*Pistacia vera* L.) Suggests the Arrest of Recombination in the Largest Heteropycnotic Pair HC1. *PLoS One*, 10(12):e0143861.
- Tixier, A., Guzmán-Delgado, P., Sperling, O., Amico Roxas, A., Laca, E., and Zwierniecki, M. A. (2020). Comparison of phenological traits, growth patterns, and seasonal dynamics of non-structural carbohydrate in Mediterranean tree crop species. *Scientific Reports*, 10(1).
- Vemmos, S. N., Pontikis, C. A., and Tolza-Marioli, A. P. (1994). Respiration rate and ethylene production in inflorescence buds of pistachio in relation to alternate bearing. *Scientia Horticulturae*, 57(1-2):165–172.
- Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., and Schatz, M. C. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, 33(14):2202–2204.
- Wang, D., Yeats, T. H., Uluisik, S., Rose, J. K., and Seymour, G. B. (2018). Fruit Softening: Revisiting the Role of Pectin. *Trends in Plant Science*, 23(4):302–310.

- Wang, P., Luo, Y., Huang, J., Gao, S., Zhu, G., Dang, Z., Gai, J., Yang, M., Zhu, M., Zhang, H., Ye, X., Gao, A., Tan, X., Wang, S., Wu, S., Cahoon, E. B., Bai, B., Zhao, Z., Li, Q., Wei, J., Chen, H., Luo, R., Gong, D., Tang, K., Zhang, B., Ni, Z., Huang, G., Hu, S., and Chen, Y. (2020). The genome evolution and domestication of tropical fruit mango. *Genome Biol.*, 21(1):1–17.
- Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Takita, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., Chapman, J., Grimwood, J., Tadeo, F. R., Estornell, L. H., Muñoz-Sanz, J. V., Ibanez, V., Herrero-Ortega, A., Aleza, P., Pérez-Pérez, J., Ramón, D., Brunel, D., Luro, F., Chen, C., Farmerie, W. G., Desany, B., Kodira, C., Mohiuddin, M., Harkins, T., Fredrikson, K., Burns, P., Lomsadze, A., Borodovsky, M., Reforgiato, G., Freitas-Astúa, J., Quetier, F., Navarro, L., Roose, M., Wincker, P., Schmutz, J., Morgante, M., MacHado, M. A., Talon, M., Jaillon, O., Ollitrault, P., Gmitter, F., and Rokhsar, D. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology*, 32(7):656–662.
- Yu, J., Golicz, A. A., Lu, K., Dossa, K., Zhang, Y., Chen, J., Wang, L., You, J., Fan, D., Edwards, D., and Zhang, X. (2019). Insight into the evolution and functional characteristics of the pan-genome assembly from sesame landraces and modern cultivars. *Plant Biotechnol. J.*, 17(5):881–892.
- Zarei, M., Davarynejad, G., Abedi, B., Kafi, M., and Biabani, A. (2014). Changes in physical properties, chemical composition and antioxidant activity of four pistachio cultivars at ten maturity stages. *Advances in Environmental Biology*, 8(10):106–115.
- Zeng, L., Tu, X.-L., Dai, H., Han, F.-M., Lu, B.-S., Wang, M.-S., Nanaei, H. A., Tajabadipour, A., Mansouri, M., Li, X.-L., Ji, L.-L., Irwin, D. M., Zhou, H., Liu, M., Zheng, H.-K., Esmailizadeh, A., and Wu, D.-D. (2019). Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biol.*, 20(1):79.

Zhang, L., Laca, E., Allan, C. J., Mahvelati, N. M., and Ferguson, L. (2021). Nonlinear model selection for fruit and kernel development as a function of heat in pistachio. *HortScience*, 56(7):769–779.

Ziya Motalebipour, E., Kafkas, S., Khodaeiaminjan, M., Çoban, N., and Gözel, H. (2016). Genome survey of pistachio (*Pistacia vera* L.) by next generation sequencing: Development of novel SSR markers and genetic diversity in *Pistacia* species. *BMC Genomics*, 17(1):998.

## 3.7 Supplemental Material



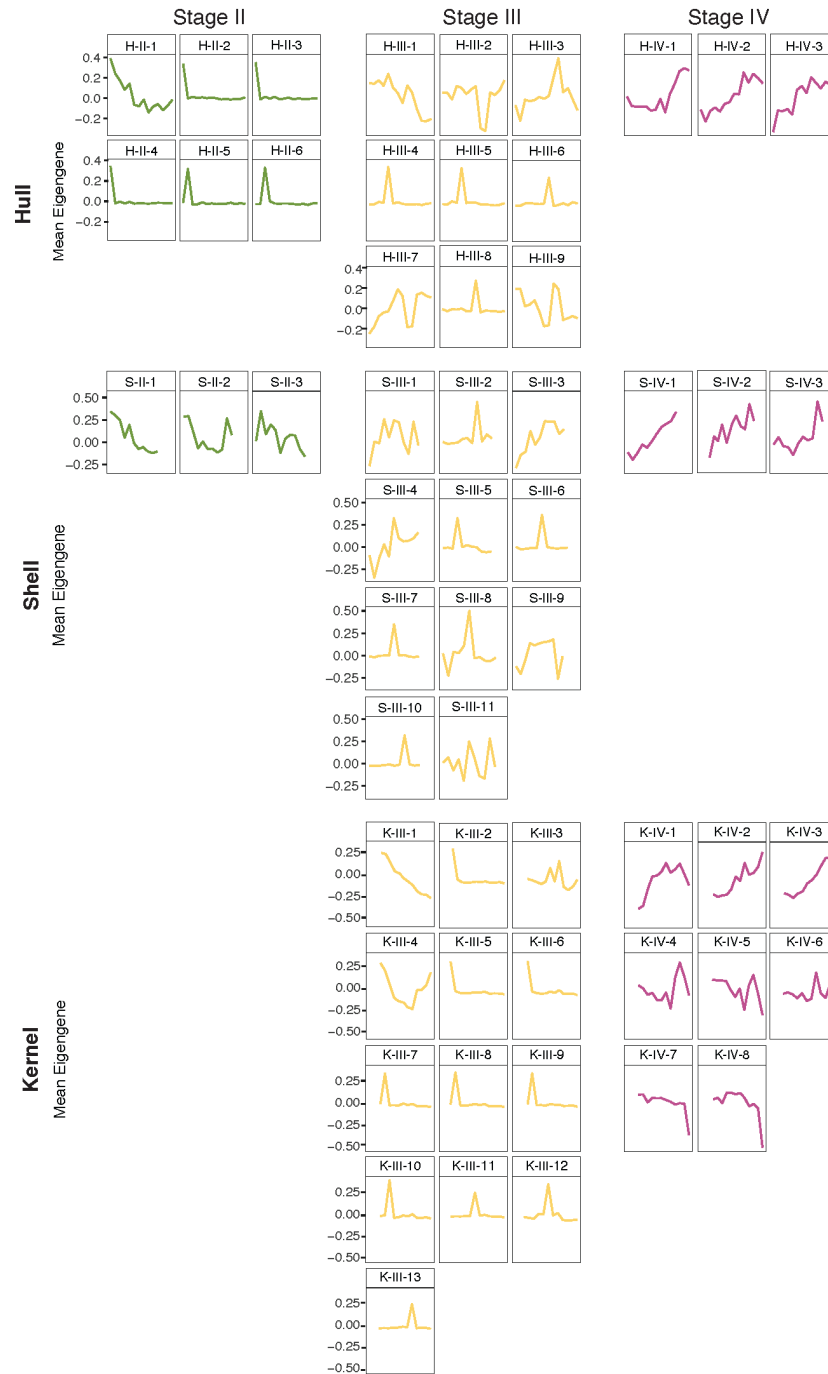


Figure 3.S1: Weighted gene correlation network analysis (WGCNA) modules. (Caption continued on next page)

Figure 3.S1: (Caption continued from previous page.) WGCNA conducted for each tissue produced modules of genes with similar expression patterns. All modules produced by the analysis are shown here organized by the developmental stage in which elevated expression (mean eigengene value) occurs. Modules are named by tissue (letter), stage (roman numeral), and number in the combination of tissue and stage (Arabic number).

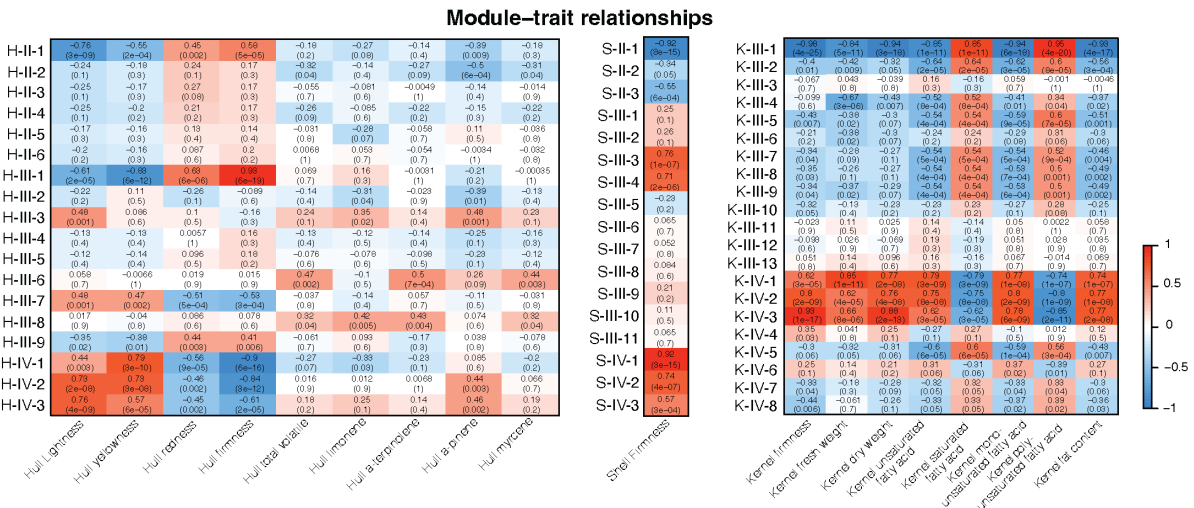


Figure 3.S2: Weighted gene correlation network analysis (WGCNA) modules-trait relationships. Modules produced by the WGCNA analysis were correlated against physiological trait data. All correlations against all traits measured are shown for each tissue.