



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Descifrando el comportamiento genético y
fisiológico de frutos de *Persea americana* var
americana cultivar Lorena

Camilo Andres Ruíz Avila

Universidad Nacional de Colombia
Ciencias Agrarias
Bogotá, Colombia
2018

Descifrando el comportamiento genético y fisiológico de frutos de *Persea americana* var americana cultivar Lorena

Camilo Andrés Ruíz Avila

Tesis presentada como requisito parcial para optar al título de:
Magister en Ciencias Agrarias línea de profundización en genética y fitomejoramiento

Director:

Dr. Guillermo Torres

Investigador

Institute of Clinical Molecular Biology

Codirectora:

Dra. Esperanza Torres Rojas

Profesora Asociada

Facultad de Ciencias Agrarias

Universidad Nacional de Colombia

Línea de Investigación: Productos naturales vegetales

Grupo de Investigación: Bioprocesos y bioprospección

Universidad Nacional de Colombia

Ciencias Agrarias

Bogotá, Colombia

2018

Agradecimientos

Agradezco a la Universidad Nacional de Colombia, a la facultad de Ciencias Agrarias y su cuerpo docente por la formación académica e investigativa recibida durante el desarrollo de mi Maestría en Ciencias Agrarias con énfasis en genética y fitomejoramiento. A COLCIENCIAS contrato 0459-2013, red RIFRUTBIO por la financiación para el desarrollo de mi tesis y el estímulo económico recibido para desarrollar mis actividades laborales y cursar mi posgrado.

A la profesora Coralia Osorio Roa (PhD) y al profesor Fabio Ancízar Aristizábal (PhD), por la confianza depositada al permitirme ser parte de la red RIFRUTBIO.

Al investigador y candidato a doctorado Guillermo Torres, director de tesis, y el profesor Michael Gribskov (PhD), por su asesoramiento y apoyo en el entendimiento de herramientas bioinformáticas.

A la profesora Esperanza Torres Rojas (PhD), codirectora de tesis, por su acompañamiento, disposición y entrega en este proyecto.

A Liliana Avila, Julián Ruíz, Oscar Avila, Andrea Botero y en general a mi familia y amigos por su apoyo y comprensión. A los miembros de los diferentes laboratorios de la Universidad Nacional de Colombia por su disposición y colaboración en el uso de equipos y entendimiento de técnicas; Tatiana Robayo, Wilson, Rolando, Diego y Diana. A la Universidad de Purdue por darme un trato similar a sus estudiantes y permitirme usar su servidor.

A los profesores Aníbal Herrera Arévalo, por permitirme desarrollar la medición de variables fisiológicas en su laboratorio, a la profesora Sandra Melo por su colaboración en el entendimiento de conceptos en estadística. Y a todas aquellas personas que de forma directa e indirecta colaboraron con el desarrollo de esta investigación.

Resumen

Este estudio determinó las características genéticas y fisiológicas que definen los frutos de *Persea americana* var *americana*, cultivar Lorena durante su estado poscosecha. A partir de un total de 40 frutos de aguacate de un mismo árbol ubicado en Mariquita-Tolima [5°11'13,9" Latitud Norte (N), 74°52'58,3" Longitud oeste (W)], se midieron las variables peso fresco, tasa de producción de etileno, tasa de respiración, firmeza, extensibilidad, masa seca, color, contenido y tipo de ácidos grasos. También se tomaron muestras del mesocarpo de tres frutos diferentes para extracción de ARN total y realizar la identificación de genes diferencialmente expresados durante los estados de preclimaterio, climaterio y posclimaterio mediante el análisis de su transcriptoma. Los resultados mostraron una vida media del fruto en poscosecha de 14 días, el comportamiento de las variables respiración y producción de etileno se adaptan al modelo cinético propuesto por Michaelis-Menten, su climaterio se alcanza en el día siete después de cosecha con un máximo de producción de CO₂ de $50,89 \pm 21,15$ mg kg⁻¹ h⁻¹, un día antes, el máximo de producción de etileno es de $215,38 \pm 58,83$ μL kg⁻¹ h⁻¹, las variables masa seca, peso fresco, firmeza y extensibilidad mostraron un decrecimiento lineal en el tiempo. El índice de color tuvo un incremento lineal y el contenido de ácidos grasos tuvo un ligero incremento donde los más predominantes son el ácido oleico y palmítico. Del ARN extraído se evidenció un total de 6335 scaffolds asociados a genes diferencialmente expresados en los tres estados de poscosecha y vinculados a los procesos de metabolismo de carbohidratos, metabolismo energético y metabolismo lipídico.

Palabras clave: cultivar Lorena, transcriptoma, aguacate, ARN.

Abstract

The genetics and physiological characteristics of avocado Lorena variety (*Persea americana* var *americana*) postharvest fruit were determined in this study. For that, 40 fruits were collected from a same tree that was located in Mariquita-Tolima [5°11'13,9" North latitude (N), 74°52'58,3" west longitude (W)]. Fresh weight, ethylene production rate, respiration rate, firmness, extensibility, dry mass, colour, oil content and type were measured. Total RNA extraction was obtained from mesocarp pieces of fruit in pre climacteric, climacteric and post climacteric states and differential expression genes were identified by analyzing its transcriptome. The results showed half-life of 14 days, respiration and ethylene production rate behavior is similar with Michaelis-Menten kinetic model, its climacteric state is reached in seventh day after fruit collected with a CO₂ production $50,89 \pm 21,15 \text{ mg kg}^{-1} \text{ h}^{-1}$, one day before, maximum of ethylene production rate is reached $215,38 \pm 58,83 \text{ } \mu\text{L kg}^{-1} \text{ h}^{-1}$. Variables dry mass, fresh weight, firmness and extensibility had a linear decrease. In contrast, colour index had a linear increase. In other hand, oil content had a little increase during postharvest and oleic and palmitic acid were most predominant. Genetic analyzes allowed to evidence 6335 scaffolds related with genes expressed differently during avocado postharvest and linked with carbohydrate, energetic and lipid metabolism.

Keywords: Lorena variety, transcriptome, avocado, RNA.

Contenido

	Pág.
Resumen y Abstract	VII
Lista de figuras	XI
Lista de tablas	XIII
Lista de Símbolos y abreviaturas	XV
Introducción	1
Objetivos	7
Objetivo general.....	7
Objetivos específicos	7
Marco teórico	9
Generalidades del cultivo de aguacate	9
Anatomía del fruto	9
Razas de aguacate (<i>Persea americana</i> Mill.) y distribución natural.....	10
Producción a nivel mundial y nacional	11
Frutos climatéricos y no climatéricos	13
Principales cambios durante la maduración de frutos climatéricos.....	14
Etileno.....	15
Firmeza.....	16
Ácidos grasos	16
Capítulo 1. Fisiología de la maduración de frutos de aguacate (<i>Persea americana</i> var americana) cultivar Lorena	19
1.1 Resumen.....	19
1.2 Introducción.....	20
1.3 Materiales y métodos	21
1.3.1 Material vegetal	21
1.3.2 Cuantificación de las variables fisiológicas	21
1.4 Diseño y Análisis estadísticos	23
1.5 Resultados	23
1.5.1 Tasa de producción de etileno y tasa de respiración	24
1.5.2 Peso fresco y variables fisiológicas determinadas en las etapas preclimaterio, climaterio y posclimaterio	28
1.6 Discusión	33
1.7 Conclusiones y Recomendaciones.....	37
1.8 Referencias	38

1.8.1	Suplementos de la metodología	42
1.8.2	Análisis estadístico	45
Marco teórico	55
	Genética y diversidad del aguacate	55
	Genómica y transcriptómica del aguacate	56
	Extracción de ARN	57
	Secuenciación de nueva generación (NGS, Next Generation Sequencing)	57
	Tecnología Illumina y preparación de librerías	59
	Análisis de secuencias	60
	Pre-procesamiento de datos	60
	Ensamblaje de transcriptoma	61
	Estimación de abundancia de transcritos	63
	Análisis de expresión diferencial de genes	63
	Anotación de secuencias diferencialmente expresadas	64
Capítulo 2. Expresión diferencial de genes presentes en los diferentes estadios de poscosecha de los frutos del cultivar Lorena (<i>Persea americana</i> var. <i>americana</i>)...	65	
2.1	Resumen	65
2.2	Introducción	66
2.3	Materiales y métodos	67
2.3.1	Extracción de ARN total, elaboración de librerías y secuenciación	67
2.3.2	Procesamiento de datos: control de calidad y ensamblaje	68
2.3.3	Cuantificación de transcritos y análisis estadísticos	69
2.3.4	Anotación funcional	70
2.4	Resultados	70
2.4.1	22,749 transcritos definen el transcriptoma del fruto <i>Persea americana</i> var <i>americana</i>	70
2.4.2	Los transcriptomas de climaterio y posclimaterio son al menos 90% similares	73
2.4.3	La anotación funcional identificó un programa de expresión génica específico para cada estado fisiológico	76
2.5	Discusión	79
2.6	Conclusiones	81
2.7	Bibliografía	82
2.8	Anexos	85

Lista de figuras

	Pág.
Figura 1-1: Anatomía del fruto de aguacate (Priego <i>et al.</i> , 1996).	8
Figura 1-2: Producción de etileno y CO ₂ para frutos climatéricos y no climatéricos desde antesis hasta maduración (Nath <i>et al.</i> , 2014)	11
Figura 1-3: Ruta de síntesis de etileno (Taiz & Zeiger, 2010)	13
Figura 1-4: Tasa de producción de etileno medida en 40 frutos. Barras verticales indican los valores máximos y mínimos obtenidos en cada día. Letras de color rojo simbolizan los estados que muestran diferencias significativas (p-valor < 0.05) en los valores de etileno con base en la prueba de Tukey.	23
Figura 1-5: Tasa de producción de etileno medida en 40 frutos. Línea negra continua para mediciones de la tasa de producción de etileno observada y línea negra discontinua para estimación del modelo cinético calculado. R ² de 0.88 entre los datos observados y calculados. Las barras verticales indican el error estándar.	24
Figura 1-6: Tasa de respiración mediada en 40 frutos. Barras verticales indican los valores máximos y mínimos en cada día. Letras de color rojo simbolizan los estados que mostraron diferencias significativas (p-valor < 0.05) en los valores de producción de CO ₂ con base en la prueba de Tukey.	25
Figura 1-7: Tasa de respiración tomada en 40 frutos. Línea negra continúa para mediciones de tasa de respiración observada y línea negra discontinua para estimación de respiración con base al modelo cinético propuesto. R ² de 0.92 entre los datos observados y calculados.	26
Figura 1-8: Porcentaje de pérdida de peso fresco de frutos de aguacate con respecto al día tres después de cosecha.	27
Figura 1-9: Variables fisiológicas medidas en frutos de aguacate en estado preclimatérico, climatérico y posclimatérico. Acidos grasos (a), extensibilidad (b),	

firmeza (c), índice de color (d), masa seca (e) y peso fresco (f). Barras verticales indican los valores máximos y mínimos en cada día. Diferencias significativas son mostradas con letras color rojo en la parte superior de la figura. 28

Figura 1-10: Color medido sobre 40 frutos de aguacate mediante las coordenadas del espacio de color CIELab para el día tres (pre climaterio), siete (climaterio) y doce (posclimaterio). Coordenadas a* (a), coordendas b* (b), coordendas L* (c). 30

Figura 2-1: Número de transcritos con un mínimo de 10 conteos por millón en al menos tres de las librerías. 70

Figura 2-2: Análisis de componentes principales (PCA) para los transcritos obtenidos del proceso de secuenciación de muestras de cDNA a partir de pulpa de fruta de aguacate en los estados de pre climaterio, climaterio y posclimaterio. 71

Figura 2-3 a) Porcentaje de transcritos sobre expresados (up, color rojo), sub expresados (Down, color verde) y de expresión constitutiva (NO_DE, gris). A partir de comparaciones pareadas entre estados de madurez poscosecha. b) transcritos de muestras en estado de pre climaterio, climaterio y posclimaterio. Con color rojo transcritos sobre expresados, color verde sub expresados y color negro en condiciones normales de expresión. 72

Figura 2-4: Porcentaje de transcritos asociados a procesos metabolicos durante la maduración de frutos de aguacate cultivar Lorena. 74

Lista de tablas

	Pág.
Tabla 1-1: Razas de aguacate y algunos de sus cultivares más representativos (Bernal et al., 2008).	2
Tabla 1-2: Distribución de la producción de variedades de aguacate en el territorio colombiano (Ministerio de agricultura, 2017).	10
Tabla 1-3: Ácidos grasos saturados e insaturados más comunes en plantas. El número encontrado antes de los dos puntos hace referencia a la cantidad de átomos de carbono y el número después de los dos puntos hace referencia a la cantidad de dobles enlaces (Taiz & Zeiger, 2010)	15
Tabla 1-4: Modelo cinético propuesto por Michaelis Menten (Castellanos et al., 2016) ajustado al tiempo (t) en días después de cosecha como variable independiente y las variables tasa de producción de etileno y tasa de respiración como variables dependientes.	24
Tabla 1-5: Correlación según análisis longitudinal y promedio de valores obtenidos por estado de madurez poscosecha para las variables fisiológicas evaluadas.	29
Tabla 1-6: Promedio del porcentaje del tipo de ácido graso expresado por tres frutos diferentes de aguacate cultivar Lorena para los estados de maduración pre climaterio, climaterio y postclimaterio. Relación de ácidos grasos saturados sobre insaturados (S/I)	31
Tabla 1-7: Clasificación de productos agrícolas de acuerdo a su tasa de producción de etileno modificada/tomada de (Hernandez <i>et al.</i> 2010)	33
Tabla 1-8: Curva de referencia a diferentes volúmenes de etileno comercial (99 ppm mL ⁻¹) 27 de junio de 2016	41
Tabla 2-1: Plataformas de secuenciación de nueva generación, características, ventajas y desventajas (Masoudi-Nejad <i>et al.</i> , 2013).	56

Tabla 2-2: Total de lecturas luego del proceso de secuenciación y trimming para cada una de las muestras de ARN de pulpa de fruta.	69
Tabla 2-3: Lecturas implementadas durante el proceso de ensamblaje de transcriptoma <i>De novo</i> , determinadas mediante el software Transrate.	70
Tabla 2-4: Enzimas afectadas por expresión diferencial de genes obtenidas a partir de comparaciones pareadas entre estados poscosecha.	75
Tabla 2-5: Proporción de transcritos asociados a metabolismo de almidón y sacarosa (Metabolismo de carbohidratos), fotosíntesis (metabolismo energético) y biosíntesis de ácidos grasos (metabolismo de lípidos). Usando como transcritos totales los obtenidos para Pre Vs Clim.	76

Lista de Símbolos y abreviaturas

cm	Centímetro (s)
°C	Grado (s) celsius
HR	Humedad relativa
Ha	Hectarea
Kg	Kilogramo (s)
mg	Miligramo (s)
ml	Mililitro (s)
msnm	Metros sobre el nivel del mar
min	Minuto (s)
mm	Milimetro (s)
N	Newton (s)
ppm	Partes por millón
Ton	Tonelada (s)
v/v	Volumen * volumen ⁻¹
µL	Microlitro (s)
CA	California
NJ	New jersey

Introducción

El incremento en la demanda de alimentos a nivel global debido a la sobrepoblación es una oportunidad de mercado para diversos productos agrícolas. Sin embargo, también es una problemática que puede incrementar la desnutrición en el mundo si no se tiene acceso a los alimentos con calidad nutricional. Una solución a este problema consiste en prolongar el tiempo de vida de los productos agrícolas mediante la implementación de tecnologías que permitan reducir la producción de desperdicios alimenticios debido a inapropiadas prácticas de poscosecha que comprometan la calidad de los mismos (Alamar *et al.*, 2017). De igual manera, otra salida la constituye el desarrollo de alimentos con mayores contenidos nutricionales mediante mejoramiento, edición de genomas o ingeniería genética (Saltzman *et al.*, 2017). Así, para la implementación de cualquiera de estas dos alternativas, una vez se han colectados los productos agrícolas, es importante conocer las características de su comportamiento fisiológico, su composición bioquímica y su relación con la expresión génica.

Entre los productos agrícolas, el consumo de aguacate (*Persea americana* Mill) es cada vez mayor a nivel global y para países como Colombia su importancia y proyección es tan grande que incluso se le compara con lo alcanzado por el sector floricultor. Según cifras oficiales de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), la producción mundial de aguacate creció 114% entre los años 1999 y 2015, pasando de producirse 2.491.741 Ton para el año 1999 a 5.346.968 Ton para el 2015. En Colombia, según el Ministerio de Agricultura, el aumento de producción alcanzó el 95% en ese mismo periodo (1999-2015); pasando de 158.505 Ton a 309.852 Ton de frutos (FAO, 2017). Dicha producción se enfoca principalmente en materiales de la raza Antillana como el cultivar Lorena, papelillo, criollo y de la raza Guatemalteca entre las que se destaca el cultivar Hass.

En Colombia, del cultivo de aguacate dependen y hacen parte 25.000 personas y 12.000 productores distribuidos en 18.113 unidades productivas que se encuentran ubicadas principalmente en los departamentos de Tolima, Antioquia, Caldas, Santander, Bolívar, Cesar, Valle del Cauca, Risaralda y Quindío, lugares donde se reporta la mayor área sembrada (Ministerio de agricultura, 2017). Estas cifras, han hecho que Colombia se consolide como el segundo país con mayor área cosechada de aguacate a nivel global. Así como, el tercer mayor productor en el mundo, después de México y República Dominicana y, por encima de los países de Perú, Indonesia, Kenia, Estados Unidos, Chile y Brasil, entre otros (Ministerio de agricultura de Colombia, 2017). Esta producción, es comercializada en mercados internacionales y nacionales, aumentando año tras año el total de toneladas exportadas y consiguiendo reducir la cantidad de importaciones de este producto al satisfacer la demanda interna (Ministerio de Agricultura de Colombia, 2017).

Dentro de la especie *Persea americana* Mill. existen tres razas reportadas: raza Antillana (*Persea americana* var. *americana*), raza Guatemalteca (*Persea americana* var. *guatemalensis*) y raza Mexicana (*Persea americana* var. *drymifolia*). Los principales cultivares pertenecientes a estas razas se encuentran reportados en la (Tabla 1-1).

Tabla 1-1: Razas de aguacate y algunos de sus cultivares más representativos (Bernal et al., 2008).

Raza	Principales cultivares
Antillana (<i>Persea americana</i> var. <i>americana</i>)	Butler Lorena Simmonds
Mexicana (<i>Persea americana</i> var. <i>drymifolia</i>)	Duke Puebla Susan
Guatemalteca	Hass

(<i>Persea</i>	Edranol
<i>americana</i> var.	
guatemalensis)	Pinkerton

Las razas Mexicana y Guatemalteca, han sido las más estudiadas a nivel fisiológico y genético y se han implementando herramientas de secuenciación de nueva generación para estudiar su genoma y transcriptoma (Lopez-Gomez *et al.*, 2007) (Ibarra-Laclette *et al.*, 2015) (Kilaru *et al.*, 2015). En contraste, la raza antillana cultivar Lorena, carece de estos estudios, lo que limita el conocimiento fisiológico, bioquímico y genético de esta raza.

En comparación al cultivar Lorena (raza Antillana), los cultivares Hass, Edranol (raza Guatemalteca) y el híbrido Fuerte (Guatemalteca x Mexicana) (Bernal *et al.*, 2008), son ampliamente aprovechados mediante la generación de productos con valor agregado como pulpas congeladas usadas como base de productos untables, comercialización de mitades congeladas, obtención de aceites con fines culinarios y para la industria cosmética, entre otros (Olaeta, 2003). Esto, debido a que su caracterización fisiológica y bioquímica en poscosecha es ampliamente conocida, lo cual ha permitido definir su buena tolerancia al frío con propósitos de conservación a baja temperatura. Además, ha permitido la comercialización de aceites debido a su alto contenido, logrando producir hasta 2500 L de aceite ha^{-1} , con una concentración mayor al 80% de ácidos grasos insaturados constituidos principalmente por ácido oleico (Olaeta, 2003).

Existen estudios sobre la fisiología del fruto durante los estadios de pre y poscosecha para el cultivar Hass que reportan que la pulpa, en estadios tempranos de precosecha, esta conformada principalmente por ácidos grasos saturados; luego, al finalizar esta etapa, los más abundantes son los ácidos grasos de tipo insaturado (Ozdemir & Topuz, 2004b). Estudios posteriores han concluido que el contenido de aceites y su composición, varía de acuerdo a la zona del cultivo, la variedad, el número de días que transcurren entre floración y cosecha, contenido de materia seca, la parte del fruto que es medida y a la manipulación del fruto durante su poscosecha (Ferreya *et al.*, 2016). Por otra parte, existen procesos fisiológicos particulares que sólo ocurren durante poscosecha, entre los que se encuentran la degradación de la pared celular y

consecuente ablandamiento del fruto, así como la biosíntesis de pigmentos. Estos procesos son estimulados como respuesta a la acción de la hormona etileno y su continua producción autocatalítica en frutos climatéricos como el aguacate (Bapat *et al.*, 2010).

Teniendo en cuenta estas consideraciones, es posible afirmar que los frutos del cultivar Lorena de aguacate (*Persea americana* var. *americana*), tienen un aprovechamiento limitado dada la carencia de conocimiento en su fisiología, composición bioquímica y genética, a nivel de poscosecha. Dentro de los pocos estudios reportados se encuentra el realizado por Romero (2012) sobre la caracterización del comportamiento fisiológico de la planta durante su desarrollo reproductivo. No obstante, es importante investigar el cultivar Lorena debido a la fácil adaptación del cultivo a zonas particulares del país que presentan condiciones de temperatura, altitud y humedad relativa aptas para este cultivo y no para otro, lo que impide que este pueda ser sustituido por materiales de otras razas como el cultivar Hass. A pesar de que éste es el tipo de cultivar más importante a nivel mundial (FAO, 2017) y del que se tiene mayor conocimiento, no ha sido posible producirlo eficientemente y con los mismos estándares internacionales de calidad en todas las regiones de Colombia y en especial en aquellas en las que los materiales de la raza antillana están mejor adaptados.

Por lo tanto, es necesario resolver el vacío de conocimiento que se tiene sobre el comportamiento fisiológico y genético del cultivar Lorena, raza Antillana. Este estudio busca explicar el proceso de maduración poscosecha del cultivar Lorena, identificando los cambios fisiológicos y de expresión génica relacionados con los estados de preclimaterio, climaterio y posclimaterio. Esta información es relevante para desarrollar un modelo que permita predecir el tiempo de maduración de los frutos y optimizar los tiempos de almacenamiento y mercadeo del producto, así como el aprovechamiento de las características fisiológicas del mismo, por ejemplo, conocer la composición de ácidos grasos y sus propiedades organolépticas. Adicionalmente, se brinda conocimiento básico para plantear nuevos proyectos de investigación que permitan aprovechar de forma eficiente estas características en diferentes tipos de mercados mediante el desarrollo y adopción de tecnologías que prolonguen su vida útil en poscosecha.

PREGUNTAS DE INVESTIGACION:

¿Cuál es el comportamiento de las características fisiológicas de frutos de aguacate cultivar Lorena que definen su maduración poscosecha? ¿Qué genes están asociados con la expresión de las características fisiológicas de frutos de aguacate cultivar Lorena durante su poscosecha?

Objetivos

Objetivo general

Determinar el comportamiento pos cosecha a nivel genético y fisiológico de frutos de aguacate variedad Lorena en los estados de maduración pre climaterio, climaterio y posclimaterio para entender los procesos de maduración del fruto.

Objetivos específicos

- Determinar los estados de madurez poscosecha de frutos de aguacate Lorena y describir las características fisiológicas en cada uno de ellos.
- Identificar los genes diferencialmente expresados que inciden en el proceso de maduración del fruto e inferir la función con la cual estos se encuentran relacionados.

Marco teórico

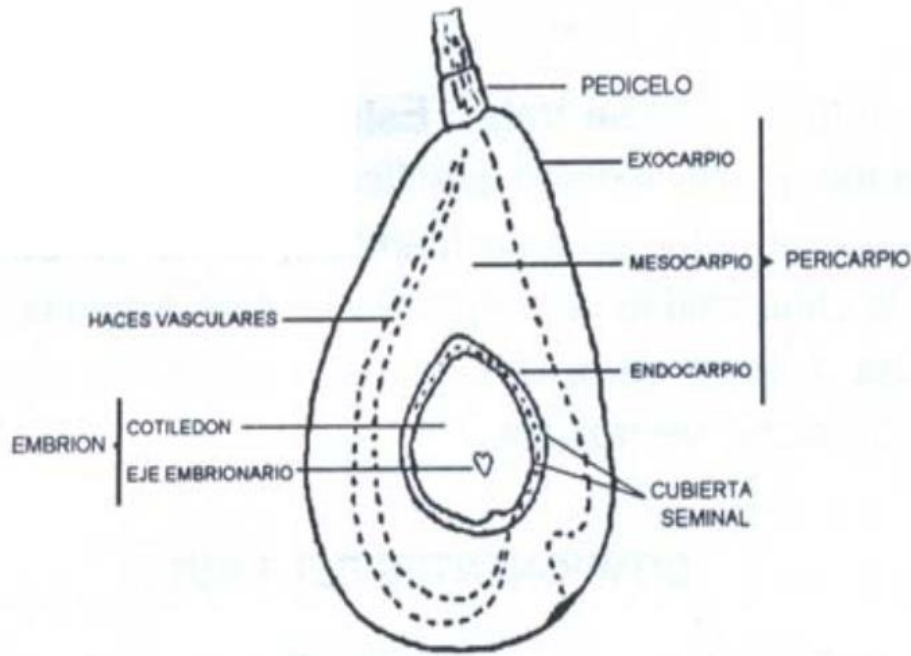
Generalidades del cultivo de aguacate

El aguacate (*Persea americana* var. *americana*), es una planta perenne que tiene un periodo vegetativo que dura entre 1,5 a 3 años. Desde el tercer año inicia la producción de frutos, alcanzando el periodo de plena producción en el quinto hasta el octavo año y se estabiliza la producción hasta los 25 años dependiendo del cultivar (Bernal *et al.*, 2008). La propagación de esta especie puede ser de tipo sexual por semilla, asexual por injertación y estacas o *in vitro*. La propagación de tipo sexual puede causar una alta variabilidad en campo dado que esta planta es alógama, por lo que comunmente este tipo de propagación es implementado para la obtención de portainjertos y desarrollo de programas de fitomejoramiento. La propagación asexual se desarrolla por injertación de tipo inglés, hendidura o según el criterio del injertador, mediante la utilización de una variedad comercial sobre un porta injerto obtenido de semilla o cultivar con condiciones aptas para el sitio de cultivo. Aunque menos popular, el otro tipo de propagación asexual de esta especie, puede darse por la siembra de estacas entre 50 a 100 cm de longitud que cuenten con la presencia de varias yemas (Bernal *et al.*, 2008).

Anatomía del fruto

El fruto de aguacate es una baya compuesta por una sola semilla cuyo tamaño, forma, color de epidermis y color de pulpa difieren según el cultivar (Bernal *et al.*, 2008). La anatomía de este fruto se compone de un pericarpio que consta de tres capas: el exocarpio, que comprende la cáscara; mesocarpio, constituido por la pulpa de la fruta y una capa interna delgada junto a la cubierta de la semilla que corresponde al endocarpio (Figura 1-1) (Priego *et al.*, 1996).

Figura 1-1: Anatomía del fruto de aguacate (Priego et al., 1996).



Razas de aguacate (*Persea americana* Mill.) y distribución natural

El aguacate es una planta de origen americano, perteneciente al orden Laurales, familia Lauraceae, con cerca de 50 géneros y entre 2500 a 3000 especies (Wolstenholme & Whiley, 2013). Su distribución natural incluye los países de México, Centroamérica, Colombia, Venezuela, Ecuador y Perú (Bernal *et al.*, 2008) (Wolstenholme & Whiley, 2013). Dentro de esta especie, es posible distinguir tres razas ecológicas (Chen *et al.*, 2009): la raza Mexicana (*Persea americana* var. *drymifolia*), la raza Guatemalteca (*Persea americana* var. *guatemalensis*) y la raza Antillana (*Persea americana* var. *americana*). Cada una de estas razas presentan características adaptativas a zonas geográficas específicas con climas diferentes (Cañas-Gutiérrez *et al.*, 2015). Los aguacates de la raza Mexicana, crecen y se desarrollan eficientemente en altitudes entre los 1.700 a 2.500 msnm y pueden cultivarse a temperaturas óptimas cercanas a 17 °C. La raza Guatemalteca está adaptada principalmente a zonas subtropicales con temperaturas entre 18 a 24 °C y altitud entre los 1.000 a 2.000 msnm. Por último, la raza Antillana se encuentra mejor adaptada a zonas tropicales por debajo de los 1000 msnm,

donde las temperaturas oscilan entre los 18 a 26 °C con alta humedad relativa (Ibarra-Laclette *et al.*, 2015; Bernal *et al.*, 2008). La raza Antillana es una especie arbórea y es distintiva de las demás razas por varios aspectos entre los que se destaca su adaptación a regímenes de precipitación de 1.150 mm y humedad relativa $82,5\% \pm 7,5\%$. Mientras que las razas Mexicana y Guatemalteca presentan unos requerimientos de precipitación que comprenden rangos más amplios que la raza antillana entre 660 a 1.500 mm y una humedad relativa menor cercana al 80%. En cuanto al peso del fruto, la raza Antillana puede alcanzar valores hasta de 1000 g, a diferencia del peso de los frutos de la raza Mexicana y Guatemalteca que pueden ser hasta 250 g. Por otro lado, el tiempo de floración a cosecha que tarda tanto la raza Antillana y Mexicana es similar (entre 5 a 8 meses), mientras que la raza Guatemalteca tarda 15 meses (Romero, 2012).

Producción a nivel mundial y nacional

Las características adaptativas del cultivo han permitido que en países como Colombia se tenga una amplia distribución de producción como lo muestra la (Tabla 1-2). La raza Antillana es la más implementada para la producción de aguacate en el país (Ministerio de agricultura, 2017) y está representada por las variedades o cultivares denominadas papelillo, que son aquellas que tienen frutos con cáscara de poco grosor y por tanto fáciles de pelar, como el cultivar Lorena. Adicionalmente, se encuentra la variedad criolla, que se caracteriza por tener frutos de cuello largo, cáscara lisa y bajo contenido de aceites (Bernal *et al.*, 2008). Las zonas de occidente y centro, también cuentan con el cultivo de materiales de la raza guatemalteca donde sobresale el cultivar Hass, que es el cultivo principal a nivel mundial, el cual se caracteriza por presentar frutos de forma ovoide, tamaño mediano, cáscara rugosa y alto contenido de aceites (cerca al 20%).

Tabla 1-2: Distribución de la producción de variedades de aguacate en el territorio colombiano (Ministerio de agricultura, 2017).

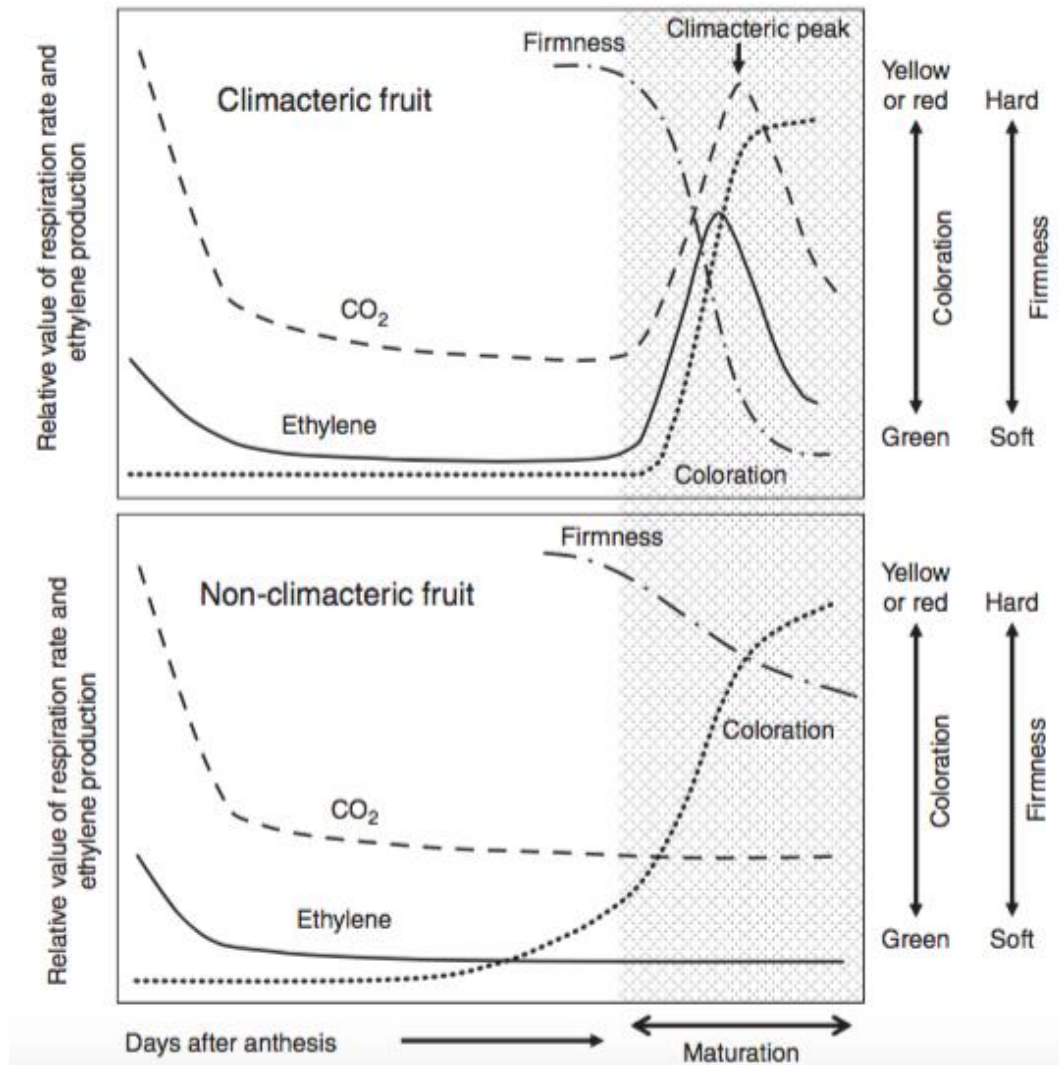
Zona	Departamentos	Hectáreas	Razas	% del área sembrada Nacional
Occidente	Antioquia, Caldas, Quindío, Risaralda, Valle del cauca.	30043	Guatemalteca (Hass) y Antillana	43,8
Centro	Tolima	14190	Guatemalteca (Hass) y Antillana	20,7
Oriente	Cundinamarca, Boyacá, Meta, Casanare y Arauca	10404	Antillana	15,1
Santanderes	Santander y norte de santander	6294	Antillana	9,1
Atlantica	Bolivar y Cesar	7729	Antillana	11,3

El cultivar Lorena, es un cultivar que proviene de un programa de fitomejoramiento realizado en 1957 en el Valle del Cauca, en la finca Lorena ubicada en un municipio de Palmira (Bernal *et al.*, 2008). Su fruto se caracteriza por tener un tamaño grande, de forma alargada y ligeramente oblicuo con contenido de ácidos grasos cercano al 9% y posee una semilla de tamaño mediano con adherencia media a la pulpa (Romero, 2012). La mayoría de los materiales de aguacate son identificados por parámetros morfológicos puesto que en la actualidad se carece de la implementación de marcadores bioquímicos o moleculares que garanticen su identidad, de igual manera, los materiales papelillo y criollo son definidos indistintamente en campo y varían según la región de la cual se colectan, aspecto que impide realizar un estudio más preciso de su comportamiento en poscosecha.

Frutos climatéricos y no climatéricos

Los frutos climatéricos, y particularmente el aguacate, estando en el árbol desarrolla procesos de división y elongación celular hasta alcanzar el punto de madurez de cosecha. Posterior a este punto y una vez los frutos son desprendidos de la planta, aumentan su tasa de respiración (Kassim *et al.*, 2013). Este efecto es promovido por una mayor síntesis de etileno y tienen como consecuencia la maduración del fruto hasta alcanzar su madurez sensorial (punto óptimo para el consumo) y posterior senescencia (Kassim *et al.*, 2013) (Figura 2-2).

Figura 1-2: Producción de etileno y CO₂ para frutos climatéricos y no climatéricos desde anthesis hasta maduración (Nath *et al.*, 2014)



Durante la maduración del fruto se presentan algunos aspectos que son perceptibles mediante el color, firmeza, peso y composición de metabolitos primarios y secundarios (Kassim *et al.*, 2013). Cuando el fruto se desprende de la planta por dehiscencia propia o por acción externa, inicia una fase definida como preclimaterio, la cual se caracteriza por expresar una baja tasa de respiración al inicio que luego aumenta; en esta etapa también ocurre un incremento en la síntesis autocatalítica de la hormona gaseosa etileno hasta alcanzar el estado climatérico en donde se destaca un máximo de producción de CO₂. Posteriormente, se presenta la fase de posclimaterio caracterizada por la declinación en la respiración del fruto y producción de etileno (Bapat *et al.*, 2010) (Kassim *et al.*, 2013). Estos procesos promueven en el fruto la reducción de su firmeza y aumento en la disponibilidad de energía libre por acción de enzimas como la β -galactosidasa, la cual, cataliza la hidrólisis de los residuos terminales galactosil de carbohidratos, glicoproteínas y galactolípidos de la pared celular (Wu & Burns, 2004). Adicionalmente, también se genera una reducción en el peso fresco por respiración, pérdida de agua, modificación de masa seca, color, contenido y tipo de aceites, entre otros, (Ozdemir & Topuz, 2004a).

Principales cambios durante la maduración de frutos climatéricos

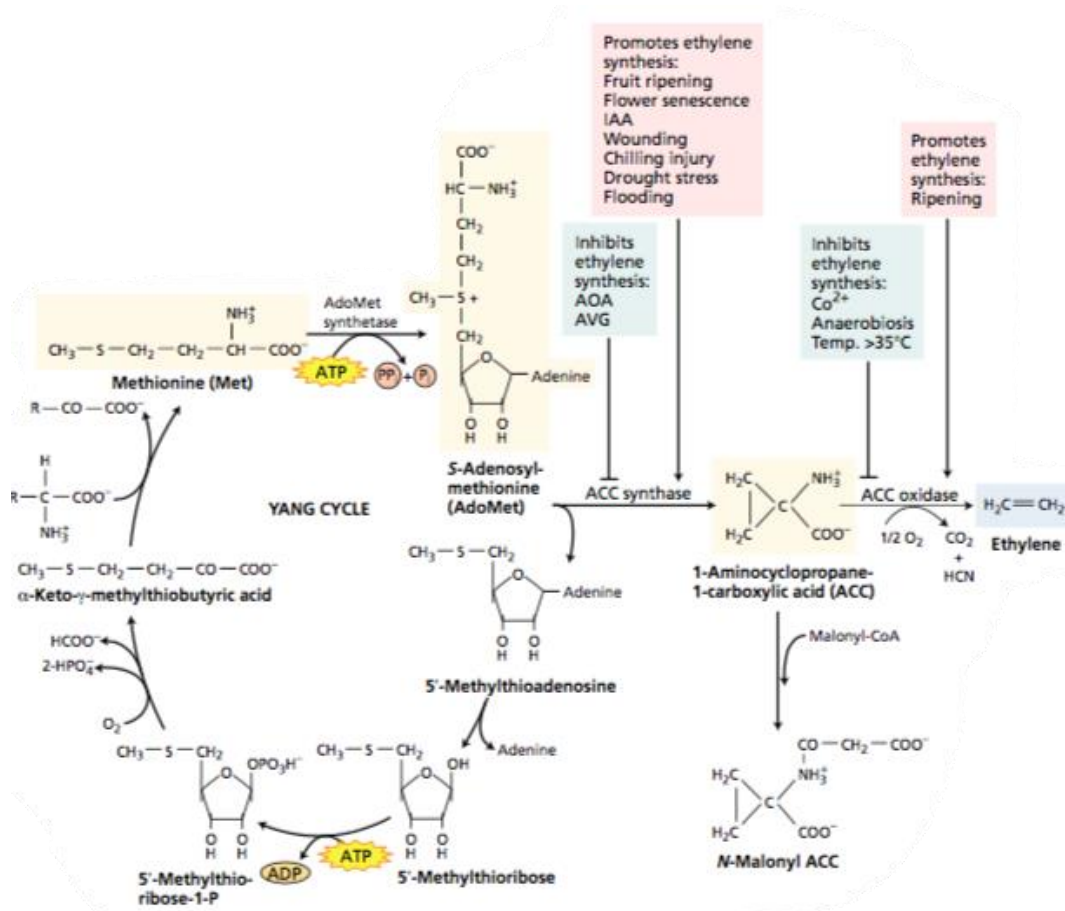
Los cambios que tienen lugar durante la maduración de los frutos climatéricos son dinámicos y se pueden determinar por medición de variables morfológicas y fisiológicas. Entre las que se encuentran textura, firmeza, color, peso fresco, masa seca, síntesis de ácidos grasos, etileno y producción de CO₂. La producción de etileno y CO₂ definen el climaterio de los frutos, así como también, las otras variables, en especial la firmeza que está asociada a procesos de ablandamiento y el contenido y tipo de síntesis de ácidos grasos saturados e insaturados. Estos aspectos son relevantes para el fruto del aguacate ya que este es uno de los pocos frutos que almacena aceites tanto en pulpa como en semilla (Kilaru *et al.*, 2015).

Etileno

El etileno es una hormona sintetizada a partir del aminoácido metionina. Este aminoácido es inicialmente convertido a S-adenosilmetionina (SAM) por adición de una adenina y el

consumo de ATP. SAM, a través de una reacción catalizada por la enzima ACC – sintasa (ACS), es convertida a 1-aminociclopropano-1-ácido carboxílico (ACC) y 5'-metiltioadenosina (MTA), el cual es reciclado a metionina. Este ACC obtenido, es convertido finalmente a etileno mediante un proceso de oxidación por la enzima ACC oxidasa (Martínez-Romero *et al.*, 2007) (Figura 1-3).

Figura 1-3: Ruta de síntesis de etileno (Taiz & Zeiger, 2010)



El etileno regula la expresión de genes por medio de factores de transcripción, así como por factores de respuesta a etileno (ERFs). Los cuales controlan la firmeza de frutos, el sabor, color y aroma (Martínez-Romero *et al.*, 2007). Por otra parte, es preciso mencionar que existen dos sistemas de biosíntesis de etileno. El sistema 1 que comparte tanto frutos climatéricos como no climatéricos, corresponde a su baja biosíntesis. El sistema 2, exclusivo de frutos climatéricos hace referencia a la síntesis autocatalítica de la hormona

etileno; razón por la cual la presencia o ausencia de este tipo de síntesis es un rasgo que permite definir a un fruto como climatérico o no (Martínez-Romero *et al.*, 2007).

Firmeza

La pared celular de la célula vegetal se encuentra formada principalmente por microfibrillas de celulosa recubiertas por hemicelulosa; así como también por pectina y lignina, esta última está presente cuando se forma la pared secundaria (Taiz & Zeiger, 2010). La celulosa está compuesta por largas cadenas de β -1,4-glucano; la hemicelulosa, conformada por xilano, glucomanano, arabinoxilano y xiloglucano, con numerosas cadenas secundarias. La pectina es un heteropolisacárido compuesto por arabinano, galactano, homogalacturona y ramnogalacturona, con numerosas cadenas secundarias (Taiz & Zeiger, 2010).

Las enzimas celulasas, pectinasas y hemicelulasas aumentan su actividad durante la maduración de frutos. Promueve una reducción en la firmeza de frutos debido a la degradación de la pared celular por acción de enzimas como: poligaracturonasa, β -galactosidasa y metilesterasa pectina (PME). Esta última, promueve la metilación del homogalacturono, afectando las uniones de calcio y consigo las propiedades físicas de la pectina (Wang *et al.*, 2017).

Ácidos grasos

Pocas especies vegetales como el aguacate además de almacenar ácidos grasos en semillas, también lo hacen en grandes proporciones en el mesocarpo de sus frutos. Entre estas, se encuentran las plantas de olivo (*Olea europaea*) y palma de aceite (*Elaeis guineensis*). En general, los ácidos grasos en tejidos vegetales son almacenados en forma de triacilglicerol (TAG), luego de ser sintetizados en el plástido y transferidos al retículo endoplasmático, seguido por una esterificación a glicerol-3-fosfato (Kilaru *et al.*, 2015).

Los triacilgliceroles (TAG) son almacenados en organelos conocidos como oleosomas que cuentan con una membrana que separa a los TAG de la fase acuosa del citoplasma. La connotación de ácido graso para este grupo de moléculas viene dada por el grupo

acil. Los grupos acil están unidos por enlaces éster a los tres grupos hidroxilo del glicerol y pueden ser moléculas de cadenas cortas conformadas por 12 carbonos y largas de hasta 20, en donde se encuentran enlaces simples y dobles para ácidos grasos insaturados y únicamente enlaces sencillos para ácidos grasos saturados (Tabla 1-3) (Taiz & Zeiger, 2010).

Tabla 1-3: Ácidos grasos saturados e insaturados más comunes en plantas. El número encontrado antes de los dos puntos hace referencia a la cantidad de átomos de carbono y el número después de los dos puntos hace referencia a la cantidad de dobles enlaces (Taiz & Zeiger, 2010).

Nombre	Estructura
ácido graso saturado	
ácido láurico (12:0)	$\text{CH}_3(\text{CH}_2)_{10}\text{CO}_2\text{H}$
ácido mirístico (14:0)	$\text{CH}_3(\text{CH}_2)_{12}\text{CO}_2\text{H}$
ácido palmítico (16:0)	$\text{CH}_3(\text{CH}_2)_{14}\text{CO}_2\text{H}$
ácido esteárico (18:0)	$\text{CH}_3(\text{CH}_2)_{16}\text{CO}_2\text{H}$
ácido graso insaturado	
ácido oleico (18:1)	$\text{CH}_3(\text{CH}_2)_7\text{CH}=\text{CH}(\text{CH}_2)_7\text{CO}_2\text{H}$
ácido linoleico (18:2)	$\text{CH}_3(\text{CH}_2)_4\text{CH}=\text{CH}-\text{CH}_2-\text{CH}=\text{CH}(\text{CH}_2)_7\text{CO}_2\text{H}$
ácido linolenico (18:3)	$\text{CH}_3\text{CH}_2\text{CH}=\text{CH}-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}_2-\text{CH}=\text{CH}-(\text{CH}_2)_7\text{CO}_2\text{H}$

La mayor proporción de ácidos grasos son sintetizados cuando el fruto se desarrolla en el árbol (Kilaru *et al.*, 2015). Debido a que el fruto alcanza la madurez de cosecha en él, pero no la madurez de consumo. No obstante, una vez los frutos son cosechados, el porcentaje de ácidos grasos saturados e insaturados puede modificarse significativamente conforme avanza su maduración y dependerá de factores climáticos y el cultivar, entre otros factores (Ozdemir & Topuz, 2004).

Capítulo 1. Fisiología de la maduración de frutos de aguacate (*Persea americana* var *americana*) cultivar Lorena

1.1 Resumen

El aguacate *Persea americana* var *americana* cultivar Lorena proviene de un programa de fitomejoramiento colombiano iniciado en 1957. Sin embargo, aún no se conoce el comportamiento fisiológico del fruto en poscosecha lo que impide optimizar su comercialización y conocer sus atributos para desarrollar productos derivados. Este estudio identificó las características fisiológicas y modeló la dinámica de la maduración del fruto de aguacate cultivar Lorena. Para ello, se usaron cuarenta frutos para cuantificar su peso fresco, firmeza, extensibilidad, masa seca, color, producción de etileno, tasa de respiración, contenido y tipo de ácidos grasos a lo largo de las tres etapas de maduración: preclimaterio, climaterio y posclimaterio. Las variables fisiológicas fueron analizadas mediante un análisis de varianza, un modelamiento de autocorrelación autorregresiva y un modelo cinético enzimático. Este estudio determinó que en el día siete después de cosecha este material alcanza el climaterio con un pico de producción de CO₂ correspondiente a $50,89 \pm 21,15 \text{ mg kg}^{-1} \text{ h}^{-1}$. Se identificó que el punto máximo de producción de etileno antecede al climaterio en el día seis después de cosecha con un valor de $215,38 \pm 58,83 \text{ } \mu\text{L kg}^{-1} \text{ h}^{-1}$. Las variables masa seca, peso fresco, firmeza y extensibilidad mostraron un decrecimiento lineal en el tiempo. El índice de color tuvo un incremento lineal y, el contenido de ácidos grasos se mantuvo constante, el ácido graso más predominante en las muestras evaluadas fue el ácido oleico seguido por el ácido palmítico. La poscosecha de frutos de aguacate cultivar Lorena tiene una vida media de 14 días y el climaterio se alcanza en el día siete. La tasa de producción de etileno y respiración son las variables que mejor explican el progreso de la maduración de estos frutos. Los frutos de este material tienen una baja proporción de ácidos grasos con

respecto a otros cultivares y tanto esta proporción como el tipo de ácidos grasos no presentaron diferencias significativas durante las etapas de poscosecha. Los resultados obtenidos proporcionan información vital para el mejor aprovechamiento de este cultivo en Colombia.

Palabras Claves: poscosecha, frutos climatéricos, etileno, respiración.

1.2 Introducción

El cultivo de aguacate (*Persea americana* Mill) en Colombia está representado principalmente por las razas guatemalteca (*Persea americana* var. *guatemalensis*) y antillana (*Persea americana* var. *americana*) y para la raza Mexicana (*Persea americana* var. *drymifolia*) no se reportan cifras de área sembrada (Ministerio de Agricultura, 2017). La raza guatemalteca es originaria de la parte sur de Mexico (Chiapas) y Guatemala y sus principales cultivares son Hass, Lamb-Hass, Hass Carmen, Reed, Edranol, Itzama, Nabal, Linda, Pinkerton y Mayapan en donde sobresale Hass que es el principal cultivar producido a nivel mundial. Sin embargo, en Colombia la raza Antillana, la cual se cree que se originó en la parte Norte de Colombia y se distribuyó luego a las Antillas (Bernal 2008), es la más implementada (Ministerio de Agricultura, 2017) debido a su adaptación a las diferentes regiones del país que se encuentran ubicadas por debajo de los 1.000 m.s.n.m.

Dentro de las variedades o cultivares de la raza Antillana se tienen: Butler, Fuchs, fucsia, Hulumanu, Lorena, Peterson, Pinelli, Pollock, Ruehle, Russell, Simmonds, Trapp, Villacampa, Waldin, común o criollo, venezolano y Curumaní en donde la variedad común o criollo es la más consumida en Colombia (Ministerio de Agricultura, 2017). Sin embargo se desconoce sus parentales ya que no proviene de un programa de mejoramiento y se hace difícil su caracterización. A diferencia de la variedad criolla, el cultivar Lorena, proviene de un programa de fitomejoramiento iniciado en 1957 que se llevó a cabo en Palmira (Valle del Cauca - Colombia) (Bernal *et al.*, 2008).

Los frutos de aguacate son climatéricos, y a diferencia de los demás frutos climatéricos, estos no alcanzan el punto de madurez sensorial en el árbol (punto óptimo para el consumo humano). Por lo tanto, una vez los frutos son desprendidos de la planta tienen

una fisiología bastante activa. Dicha dinámica fisiológica es perceptible mediante la medición de las características físicas del fruto (color, firmeza, peso fresco, masa seca y extensibilidad), tasa de respiración, producción de etileno y composición de metabolitos primarios y secundarios hasta que el fruto alcanza la madurez sensorial (Kassim *et al.*, 2013) y posterior senescencia.

Pese a los esfuerzos realizados en el estudio del cultivar Lorena, los frutos son identificados con base en parámetros morfológicos y se carece de información de su fisiología en poscosecha lo que ha limitado su aprovechamiento con fines de exportación, elaboración de productos derivados y aumento del tiempo de vida en anaquel. Por lo tanto, la pregunta de investigación de este estudio fue: ¿cuál es el comportamiento de las características fisiológicas del fruto de aguacate cultivar Lorena que define su maduración a nivel poscosecha?

1.3 Materiales y métodos

1.3.1 Material vegetal

Cuarenta (40) frutos de aguacate de raza *Persea americana* var. *americana* cultivar Lorena, fueron colectados de un árbol de seis años en estado productivo; sin incidencia de plagas ni enfermedades e injertado sobre un material criollo (raza Antillana). El sitio de muestreo fue en la finca “La Juliana” (5°11’13,9” Latitud Norte (N), 74°52’58,3” Longitud oeste (W), temperatura promedio 27.3 °C y 443 m.s.n.m. en la vereda “La Guardia” del municipio de Mariquita (Tolima). Los frutos, posterior a su recolección, se transportaron en canastillas hasta el laboratorio de poscosecha de la Facultad de Ciencias Agrarias de la Universidad Nacional de Colombia sede Bogotá, donde se cuantificaron las variables fisiológicas.

1.3.2 Cuantificación de las variables fisiológicas

El seguimiento de la maduración de frutos se realizó mediante mediciones periódicas de las variables fisiológicas a una temperatura de 19.5 ± 0.5 °C y humedad relativa (HR) de $53 \pm 8\%$ durante 14 días después de cosechados. Las variables fisiológicas que se midieron con una periodicidad de dos días fueron: tasa de producción de etileno, tasa de respiración y pérdida de peso fresco. Las otras variables: firmeza, extensibilidad, color,

masa seca y contenido y tipo de ácidos grasos se determinaron en cada una de las etapas de maduración del fruto: preclimaterio, climaterio y posclimaterio. Estas etapas se determinaron en experimentos previos evaluando la tasa de respiración y etileno sobre frutos del cultivar Lorena. Detalle de la metodología utilizada en la cuantificación de las variables fisiológicas se presenta en el anexo 1 y se resumen a continuación.

Para determinar la tasa de producción de etileno se utilizó una muestra de 1 mL de la atmósfera contenida en cámaras de cierre hermético de 20 L que almacenaron los frutos por un período de dos horas. Dichas muestras fueron analizadas mediante cromatografía de gases en un equipo Agilent (GC/MSD 5975C,CA), a una temperatura del inyector de 70 °C, temperatura de horno de 50 °C y temperatura del detector FID de 250 °C. Usando Helio como gas de arrastre a un flujo de 7.0 mL min⁻¹ y los gases de combustión del detector FID fueron aire seco e hidrógeno con flujos de 300 y 40 mL min⁻¹, respectivamente. Para la cuantificación se empleó un patrón de etileno (AGA, Colombia) con una concentración de 99 ppm en nitrógeno. La tasa de producción de etileno de las muestras fue calculada con base en las curvas de calibración realizadas previamente (anexo 1). La tasa de respiración de cada fruto se determinó con base en la producción de CO₂ mediante el equipo Vernier (LabQuest®2, Beaverton) ajustando los valores obtenidos a un modelo lineal.

El porcentaje de pérdida de peso fresco se calculó de acuerdo a la ecuación que relaciona el peso inicial con el peso de final de la muestra en el momento de la medición. La masa seca se determinó luego de liofilizar una muestra de pulpa de peso fresco mediante el equipo LABCONCO (FreeZone 2.5, Kansas).

La firmeza y extensibilidad se calcularon mediante la implementación del texturómetro Ametek® (Lloyd LS1, Berwyn). El color se determinó mediante el colorímetro Konica Minolta (Chroma meter CR-400, NJ). Finalmente, el contenido y tipo de ácidos grasos se llevó a cabo a través de un análisis por cromatografía de gases en un equipo Agilent (GC/MSD 5975C,CA) siguiendo el protocolo AOAC 920.39 - 2003 (Serpa G *et al.*, 2014).

1.4 Diseño y Análisis estadísticos

Dado que el comportamiento de las variables fisiológicas tasa de respiración y tasa de producción de etileno es atribuido a diferentes actividades enzimáticas (Taiz & Zeiger, 2010), dicho comportamiento es modelable. El comportamiento de dichas variables se modeló mediante un análisis longitudinal balanceado con una estructura de autocorrelación de orden uno, AR (1) (Bates, 1988) y con base en el modelo cinético propuesto por Michaelis-Menten (Castellanos *et al.*, 2016).

Se realizó un análisis de varianza clásico (ANOVA) bajo un nivel de significancia del 5% para las variables fisiológicas: tasa de respiración, tasa de producción de etileno, pérdida de peso, firmeza, extensibilidad, color, masa seca y contenido de ácidos grasos. Esto con el fin de identificar si existían o no diferencias significativas entre las medias de la variable respuesta para cada uno de los tiempos evaluados (cada dos días durante 14 días), previa validación de los supuestos del modelo. Para validar normalidad se utilizó el test de Kolmogorov-Smirnov o Shapiro-Wilk y homocedasticidad el test de Bartlett. Una vez rechazada la hipótesis nula del ANOVA se realizó una prueba de comparaciones múltiples de Tukey para identificar en que días se presentaron diferencias significativas mediante el software estadístico (S.A.S, Carolina del Norte).

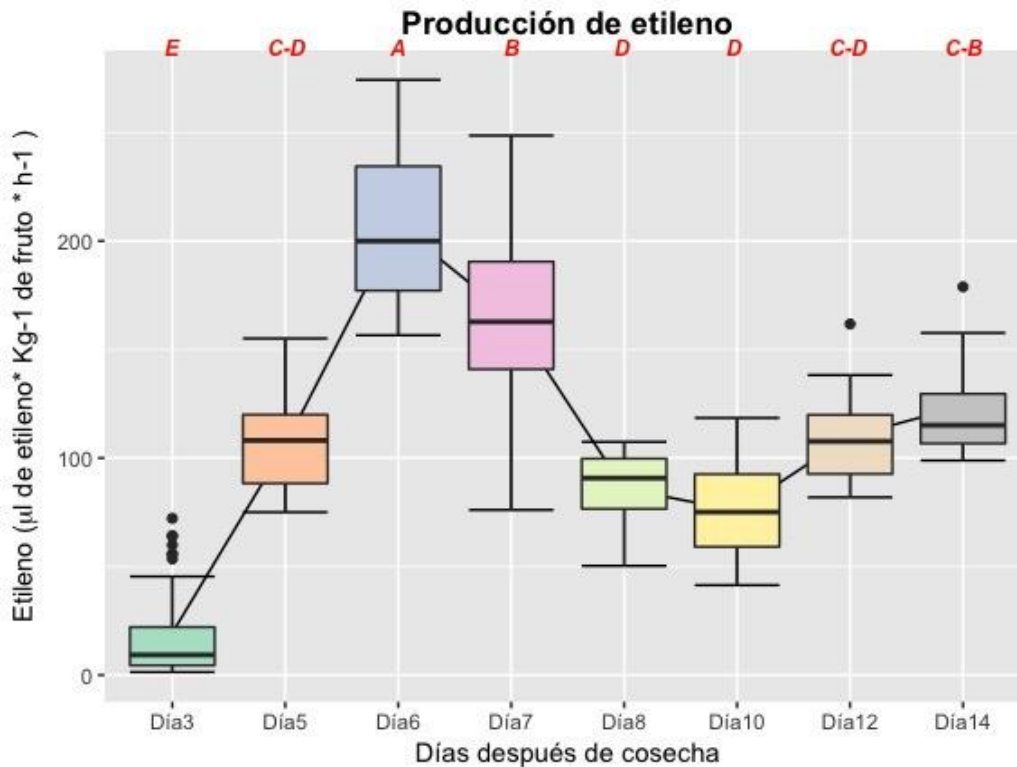
Para la determinación de textura, color, firmeza y extensibilidad se utilizaron 10 frutos para cada una de las etapas de poscosecha, para la determinación de masa seca, cantidad y tipo de ácidos grasos se utilizaron 3 frutos por etapa de poscosecha. Para peso fresco, tasa de respiración y tasa de producción de etileno se empezó con 40 frutos y a medida que se identificaron los estados de poscosecha se redujo a 30 (climaterio) y 20 (posclimaterio). No obstante, para realizar el análisis longitudinal balanceado se tomaron únicamente 10 frutos a los cuales se realizó el mismo número de mediciones a través de todo el tiempo.

1.5 Resultados

1.5.1 Tasa de producción de etileno y tasa de respiración

La producción de etileno genera una cascada de señalización en diferentes rutas metabólicas asociadas a los procesos de degradación de la pared celular, color de los frutos, hidrólisis del almidón, acumulación de azúcares y reducción de los ácidos orgánicos y compuestos fenólicos, entre otros procesos (Taiz & Zeiger, 2010) que influyen en las variables fisiológicas asociadas a la maduración del fruto. Su comportamiento presentó diferencias significativas ($p < 0.05$) durante la poscosecha de fruto de aguacate. El comportamiento de la producción de etileno muestra un aumento continuo desde la cosecha hasta el día 6 de poscosecha ($215.38 \pm 58.83 \mu\text{L kg}^{-1} \text{h}^{-1}$). Posteriormente, la síntesis de etileno disminuye de manera continua hasta el día 10 ($79.94 \pm 38.52 \mu\text{L kg}^{-1} \text{h}^{-1}$) y luego aumenta hasta el día 14 ($138.86 \pm 40 \mu\text{L kg}^{-1} \text{h}^{-1}$) (Fig. 1.1).

Figura 1-4: Tasa de producción de etileno medida en 40 frutos. Barras verticales indican los valores máximos y mínimos obtenidos en cada día. Letras de color rojo simbolizan los estados que muestran diferencias significativas (p -valor < 0.05) en los valores de etileno con base en la prueba de Tukey.



La estimación del modelo producción de etileno mediante un análisis longitudinal balanceado obtuvo un ajuste lineal con una correlación de 0.38, para un alfa de 0.05 (anexos). Con relación al modelo cinético propuesto por Michaelis-Menten los datos de producción de etileno observado en comparación con los datos calculados muestran un R^2 de 0.88 (Tabla 1-4; Figura 1-5).

Tabla 1-4: Modelo cinético propuesto por Michaelis Menten (Castellanos et al., 2016) ajustado al tiempo (t) en días después de cosecha como variable independiente y las variables tasa de producción de etileno y tasa de respiración como variables dependientes.



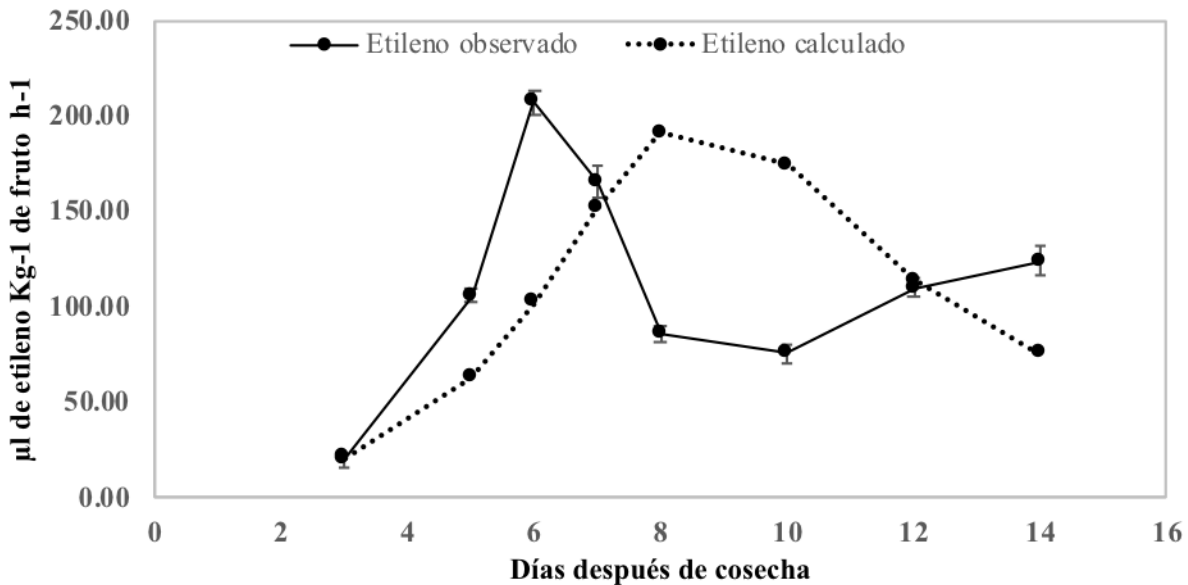
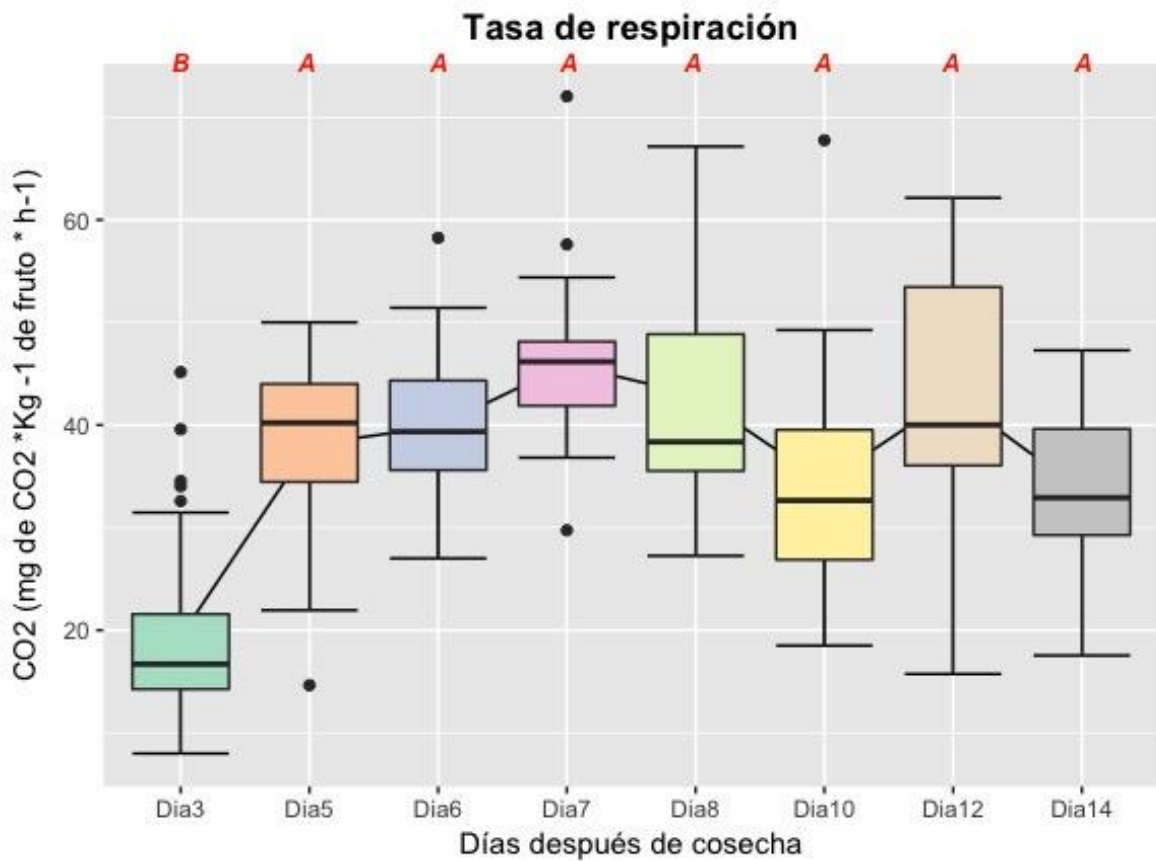
Variable dependiente, Y	Modelo obtenido	R ²
Respiración		0.92
Etileno		0.88

Figura 1-5: Tasa de producción de etileno medida en 40 frutos. R² de 0.88 entre los datos observados y calculados. Las barras verticales indican el error estándar.



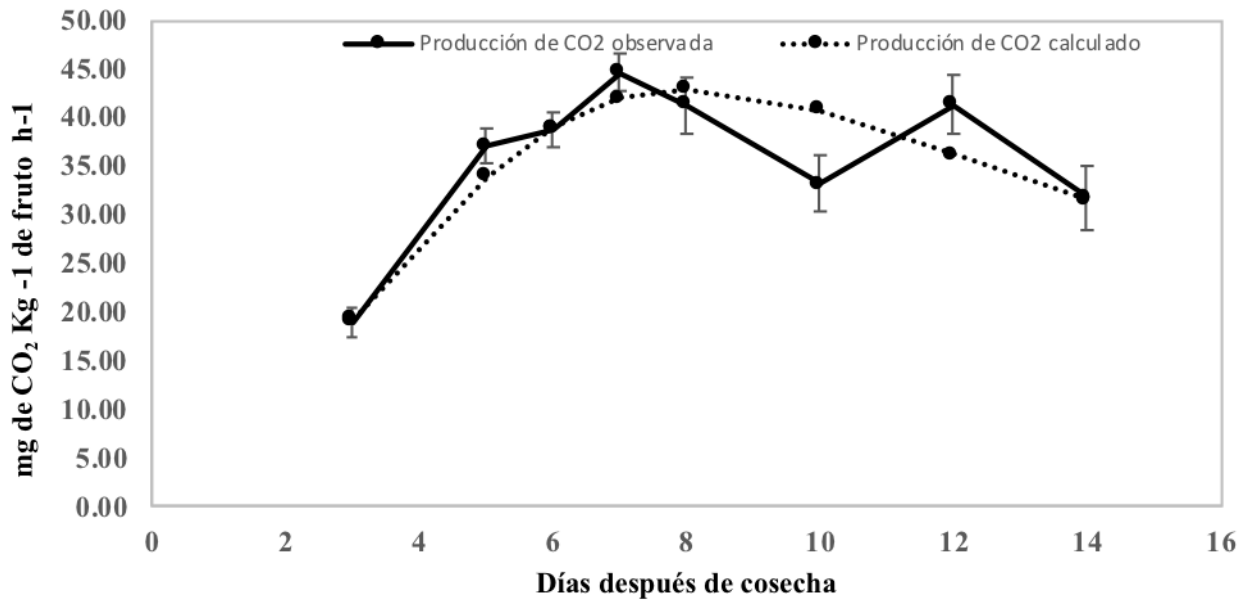
La respiración de frutos no expreso diferencias significativas para un p valor de 0.05. No obstante, se evidenció una tasa máxima de producción de CO₂ de $50.89 \pm 21.15 \text{ mg kg}^{-1} \text{ h}^{-1}$ durante el día 7 después de cosecha. Durante los días siguientes está disminuyó hasta alcanzar un valor de $32.41 \pm 14.87 \text{ mg kg}^{-1} \text{ h}^{-1}$ (Fig. 1.3) (Anexos; Figura 1-6).

Figura 1-6: Tasa de respiración mediada en 40 frutos. Barras verticales indican los valores máximos y mínimos en cada día. Letras de color rojo simbolizan los estados que mostraron diferencias significativas (p -valor < 0.05) en los valores de producción de CO_2 con base en la prueba de Tukey.



A partir de un análisis longitudinal balanceado se obtuvo un ajuste lineal con una correlación de -0.04 , par aun alfa de 0.05 (Anexos). Mediante el análisis de los datos obtenidos en relación el modelo cinético de Michaelis-Menten (Castellanos et al., 2016), la tasa de producción de CO_2 observada en relación a la tasa de producción de CO_2 calculada expresan un R^2 de 0.92 (Tabla 1-4; Figura 1-7).

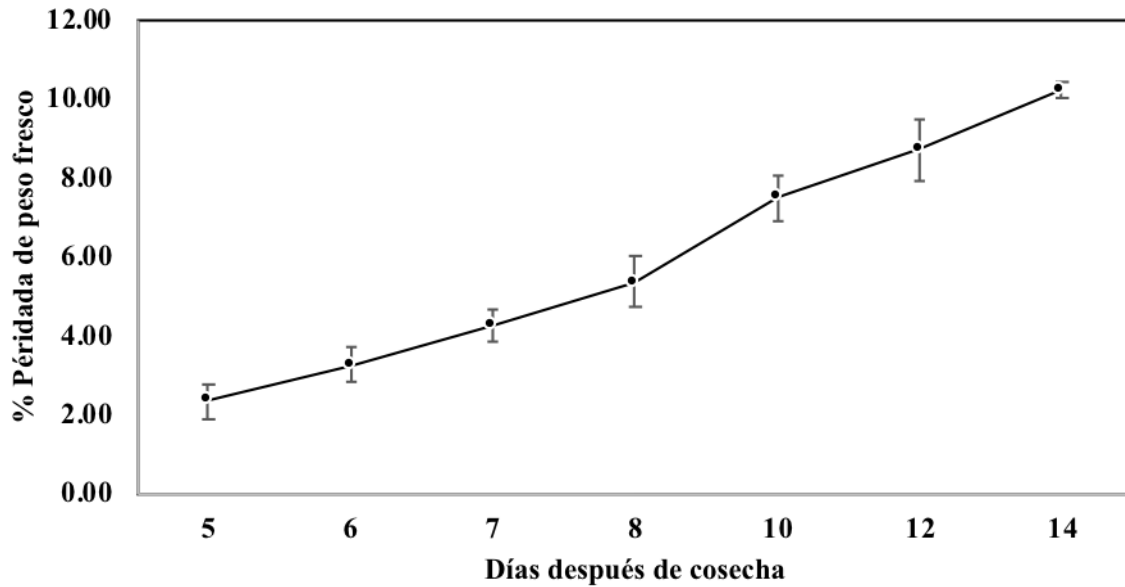
Figura 1-7: Tasa de respiración tomada en 40 frutos. R^2 de 0.92 entre los datos observados y calculados. Las barras verticales indican el error estándar.



1.5.2 Peso fresco y variables fisiológicas determinadas en las etapas preclimaterio, climaterio y posclimaterio

Los resultados muestran que los frutos pierden peso de manera continua a través del tiempo, logrando una pérdida del 10.22% para el día 14 después de cosecha, con respecto al día 3 (Figura 1-8). Mediante un análisis longitudinal balanceado se encontró que los datos de peso fresco medidos a través del tiempo tienen una correlación de 0.99 para un alfa de 0.05 con una estructura de autocorrelación autoregresiva de orden I.

Figura 1-8: Porcentaje de pérdida de peso fresco de frutos de aguacate con respecto al día tres después de cosecha. Las barras verticales indican el error estándar.



Con relación a las variables peso fresco, masa seca y contenidos de ácidos grasos no se evidencian diferencias estadísticamente significativas para los valores obtenidos en cada una de las etapas poscosecha. En contraste, las variables índice de color, firmeza y extensibilidad sí mostraron diferencias significativas para cada una de las etapas (Figura 1-9). Por otro lado, se encontró que hay una mayor dispersión de los datos para la variable peso fresco con respecto a las variables textura y extensibilidad, las cuales presentaron la menor dispersión de todas las variables estudiadas (Figura 1-9). Los promedios de estos valores obtenidos para cada estado se presentan en la (Tabla 1-5). junto con los resultados obtenidos de correlación mediante un análisis longitudinal balanceado con un nivel de significancia de 0.05 (Tabla 1-5).

Figura 1-9: Variables fisiológicas medidas en frutos de aguacate en estado preclimático, climático y posclimático. Ácidos grasos (a), extensibilidad (b), firmeza (c), índice de color (d), masa seca (e) y peso fresco (f). Barras verticales indican los valores máximos y mínimos en cada día. Diferencias significativas son mostradas con letras color rojo en la parte superior de la figura.

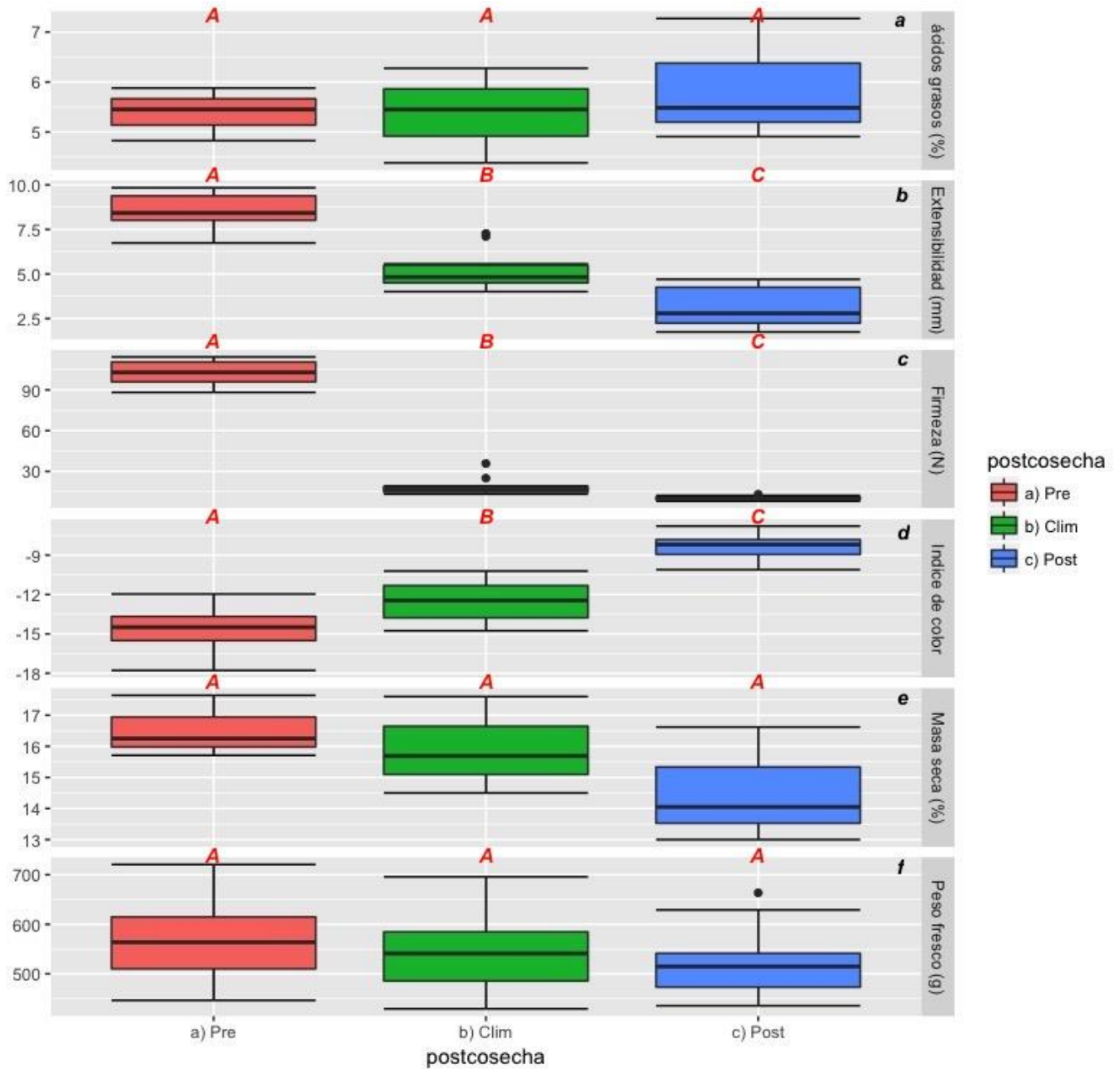
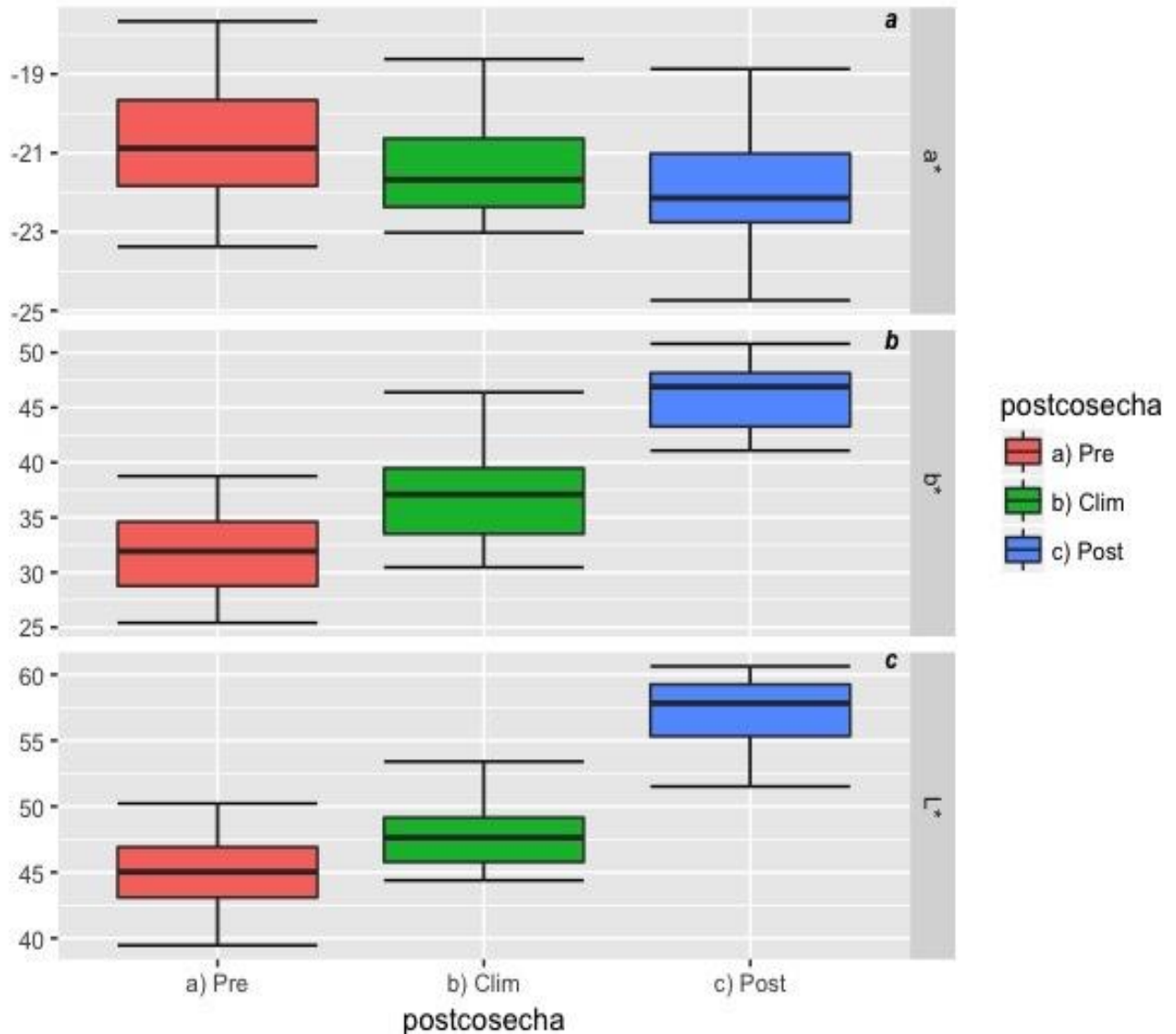


Tabla 1-5: Correlación según análisis longitudinal y promedio de valores obtenidos por estado de madurez poscosecha para las variables fisiológicas evaluadas.

Variable	Poscosecha			Correlación según análisis longitudinal
	Preclimaterio	Climaterio	Posclimaterio	
ácidos grasos (%)	5.36 ± 0.52 ^a	5.34 ± 0.95 ^a	6.09 ± 1.18 ^a	0.75
extensibilidad (mm)	8.29 ± 1.55 ^a	5.65 ± 1.64 ^b	3.23 ± 1.48 ^c	0.07
fimeza (N)	101.24 ± 13.04 ^a	24.53 ± 11.26 ^b	10.61 ± 2.56 ^c	0.4
Índice de color (IC)	-14.87 ± 2.91 ^a	-12.49 ± 2.28 ^b	-8.45 ± 1.66 ^c	0.57
masa seca (%)	16.68 ± 0.97 ^a	16.05 ± 1.55 ^a	14.81 ± 1.81 ^a	0.66
peso fresco (g)	572.6 ± 126.12 ^a	562.63 ± 133.17 ^a	549.6 ± 113.95 ^a	0.99

Con respecto a la variable color, según los datos obtenidos por mediciones realizadas dentro del espacio de color CIELab, se observa que el fruto de aguacate adquiere mayor luminosidad con valores promedio de 45 en estado preclimaterio a 57 en posclimaterio (Figura 1-10c). En cuanto a la variación del color entre el espectro verde y rojo, el fruto adquiere un color verde conforme avanza su climaterio y un tono amarillo para el espectro entre el azul y amarillo (Figura 1-10a-b).

Figura 1-10: Color medido sobre 40 frutos de aguacate mediante las coordenadas del espacio de color CIELab para el día tres (pre climaterio), siete (climaterio) y doce (posclimaterio). Coordenadas a^* (a), coordenadas b^* (b), coordenadas L^* (c).



Por otro lado, el perfil de ácidos grasos cuantificados en las muestras de aguacate es presentado en la (Tabla 1-6). El comportamiento expresado por los ácidos grasos encontrados en las muestras de pulpa para las diferentes etapas poscosecha evidencia que el ácido oleico C18:1 es alrededor de un 3% más alto en preclimaterio que en los otros estados y el contenido de Ac. linoleico C18:2 incrementa 2.75% en promedio a lo largo de la evolución fisiológica del fruto. De igual manera, es posible afirmar que los

ácidos grasos oleico C18:1 y palmítico C16:0 son los más abundantes en todas las etapas poscosecha y la relación entre ácidos grasos saturados (S) e insaturados (I), se mantiene constante para los tres estados poscosecha.

Tabla 1-6: Promedio del porcentaje del tipo de ácido graso expresado por tres frutos diferentes de aguacate cultivar Lorena para los estados de maduración pre climaterio, climaterio y postclimaterio. Relación de ácidos grasos saturados sobre insaturados (S/I)

	Ac. palmítico C16:0	Ac. palmitoléico C16:1	Ac. esteárico C18:0	Ac. oleico C18:1	Ac. linoleico C18:2	Ac. linoléico C18:3	S/I
Preclimaterio	33.62 ± 0.16	7.64 ± 0.28	0.99 ± 0.02	38.11 ± 0.41	14.33 ± 0.20	1.21 ± 0.08	0,56
Climaterio	33.31 ± 0.55	7.5 ± 0.20	1.01 ± 0.01	35.4 ± 0.30	16.98 ± 0.50	1.7 ± 0.12	0,56
Postclimaterio	33.08 ± 0.45	7.27 ± 0.19	0.97 ± 0.02	35.61 ± 1.16	17.08 ± 0.89	1.77 ± 0.22	0,55

1.6 Discusión

El estudio fisiológico sobre los frutos de aguacate cultivar Lorena realizado en este trabajo identificó que la tasa de producción de etileno alcanza su máximo el día 6 después de cosecha con un valor de $215.38 \pm 58.83 \mu\text{L kg}^{-1} \text{ h}^{-1}$, el cual permite identificarlo dentro de la lista de frutos con mayor tasa de producción de esta hormona junto la chirimoya, maracuyá, zapote, gulupa y cholupa, que presentan tasa mayores a $100 \mu\text{L kg}^{-1} \text{ h}^{-1}$, ubicándose en la clase “muy alto” (Tabla 1-6) (Hernández *et al.* 2010).

La tasa máxima de respiración (la cual define el punto de climaterio), se alcanza en el día 7 para el cultivar Lorena, un día después de alcanzar el pico de producción de etileno. Estos resultados contrastan con lo reportado en otros cultivares de aguacate como Simmonds (raza Anillana) y el híbrido Booth 7 (raza Antillana x raza Guatemalteca) en donde este estado se alcanza en el día cinco y nueve respectivamente (Maftoonazad & Ramaswamy, 2005). No obstante, estos materiales tienen en común que el máximo de

producción de etileno antecede por un día al pico de producción de CO₂ (Pereira *et al.*, 2013).

Estas mediciones permitieron definir el preclimaterio de este cultivar para el tiempo que transcurre desde su cosecha hasta el día seis, el climaterio para el día siete y el posclimaterio desde el día ocho en adelante para condiciones de temperatura 19.5 ± 0.5 °C y humedad relativa (HR) de $53 \pm 8\%$. Para ambas variables se observó que en la etapa de posclimaterio se presentó una disminución desde el día siete hasta el día 10, luego un incremento para el día 12 en ambas variables y posterior disminución de las mismas. Este efecto es debido a la variabilidad biológica de las muestras, las cuales a pesar de haber sido cosechadas de un mismo árbol manifiestan diferente edad de fruto desde antesis.

La dinámica del etileno y respiración evaluada mediante un análisis longitudinal balanceado mostró una estructura de autocorrelación de orden uno, AR (1). Esta estructura significa que los valores obtenidos a través del tiempo presentan una correlación, es decir que los datos registrados a través del tiempo no son independientes por lo que no se recomienda realizar un análisis de varianza clásico (Anova). Esto se confirmó cuando la variable respiración se analizó mediante un Anova y se presentó una suma de cuadrados mayor para el error que para la suma de cuadrados del modelo (anexos) impidiendo probar si hay diferencias significativas para esta variable.

El modelo cinético enzimático propuesto por Michaelis-Menten (Castellanos *et al.*, 2016) tiene un ajuste (R^2) de 0.88 para tasa de producción de etileno y 0.92 para tasa de respiración, lo que permite explicar la maduración de los frutos de aguacate a través de estas dos variables y la ecuación desarrollada a partir de este análisis y reportada por primera vez para este cultivar (Tabla 1-4). En procesos de poscosecha es importante identificar el comportamiento de estas variables para predecir su tiempo de vida en anaquel y así garantizar un mejor seguimiento para su comercialización y elaboración de productos con valor agregado.

Tabla 1-7: Clasificación de productos agrícolas de acuerdo a su tasa de producción de etileno modificada/tomada de (Hernandez *et al.* 2010)

Clase	Rango a 20 °C ($\mu\text{L de C}_2\text{H}_4 \text{ kg}^{-1} \text{ h}^{-1}$)	Especie
Muy bajo	Menor a 0.1	Alcachofa, uva, espárrago, coliflor, cereza, fresa, granada, cítricos, raíces, papa, la mayoría de flores de corte, ají, cocona.
Bajo	0.1 - 1.0	Mora, arándano, melón, cocombro, oliva, pimienta, piña, tamarindo, papa criolla.
Moderado	1.0 - 10.0	Banano, guayaba, higo, melón honeydew, mango, lychee, plátano, tomate, copoazu.
Alto	10.0 - 100.0	Manzana, albaricoque, melón cantaloupe, feijoa, Kiwi, nectarín, papaya, durazno, pera, ciruela, araza, uchuva.
Muy Alto	Mayor de 100.0	Chirimoya, maracuyá, zapote, gulupa, cholupa.

Las variables color, firmeza, extensibilidad, peso fresco, masa seca y contenido de ácidos grasos también mostraron autocorrelación de orden uno, AR (1) al ser analizadas por un modelo longitudinal. Estas variables también presentan cambios a través de la maduración del fruto debido a su interacción, es así como la reducción en el peso fresco es favorecida por el proceso de respiración que describe la oxidación de 12 moléculas de carbono y la reducción de 12 moléculas de O_2 , para una reacción neta que conlleva a la producción de 12 moléculas de CO_2 y 11 moléculas de H_2O mediante los procesos: de glicólisis, ciclo del ácido cítrico, ruta de las pentosas fosfato y fosforilación oxidativa (Taiz & Zeiger, 2010) en donde ocurre una pérdida de agua, consumo de moléculas de reserva para la producción de energía, conversión de metabolitos y mantenimiento de estructuras (Balaguera, 2011).

La firmeza, asociada a la resistencia del fruto a la penetración y la extensibilidad como la capacidad del fruto para soportar la perturbación generada por un factor externo sin

romperse, son variables que disminuyen conforme avanza la poscosecha de frutos debido al consumo de moléculas de reserva que incluyen el desensamble de redes de polisacáridos, degradación de pectinas, hidrólisis de almidón, celulosa y hemicelulosa contenidas en la pared celular, así como, pérdida de agua por efecto de la transpiración que conlleva a un cambio en la presión de turgencia (Castellanos *et al.*, 2016). Por otra parte, los frutos de aguacate presentan sensibilidad a concentraciones superiores a $0.4 \mu\text{L L}^{-1}$, de etileno, lo que desencadena cambios en su color, firmeza, sabor y aroma (Martínez-Romero *et al.*, 2007); características que fueron diferencialmente significativas para este fruto durante sus tres etapas climatéricas. El color tuvo una progresión hacia el incremento en intensidad al amarillo debido posiblemente a una mayor síntesis de carotenoides, como xantófilas y la reducción de clorofila durante la maduración del fruto (Kang *et al.* 2010).

En este estudio se encontró que los frutos de la raza Antillana, cultivar Lorena, el ácido oleico (36.37%) y palmítico (33.3%) tienen proporciones muy similares lo que correspondería a los frutos pertenecientes a esta raza que están adaptados principalmente a zonas cálidas. Según Ferreyra *et al.* (2016) una menor temperatura, entre $16.5 \pm 4.5 \text{ }^\circ\text{C}$ en el sitio de cultivo para el cultivar Hass, permite una mayor síntesis de ácidos grasos insaturados y menores niveles de ácidos grasos saturados. Woolf *et al.*, (2000) también soporta la idea de que materiales adaptados a zonas más frías como el material Hass de la raza Guatemalteca poseen diferencias entre el nivel de ácidos grasos insaturados y saturados. Los frutos de Lorena usados en este estudio provienen de una región con una temperatura mayor a lo reportado por Ferreyra y colaboradores ($27.3 \text{ }^\circ\text{C}$ de temperatura promedio) y por eso posiblemente se encontraron valores similares en el contenido de éstos ácidos grasos.

En cuanto al contenido de ácidos grasos, se obtuvo un ligero incremento para el día 12 (posclimaterio) con respecto a los días tres (preclimaterio) y siete (climaterio) Tabla 1-5. Sin embargo, este efecto puede atribuirse a la deshidratación del fruto y a la degradación de la pared celular por enzimas celulasas y poligaracturonasas que facilitan la extracción de aceites (Hernández *et al.*, 2016). Aunque autores como Ferreyra *et al.* (2016) conciben que durante la poscosecha de frutos de aguacate se sintetizan ácidos grasos en menor proporción con respecto al peso total del fruto. Durante la maduración del fruto

en la planta, los niveles de ácidos grasos saturados e insaturados cambian, presentándose mayores concentraciones de ácidos grasos saturados en estados iniciales que decrecen hasta el punto de cosecha y luego se mantienen en estado de poscosecha; los ácidos grasos insaturados presentan un comportamiento inverso (Ferreyra *et al.* 2016).

Con los resultados obtenidos en este trabajo se aporta conocimiento básico sobre la composición de ácidos grasos, caracterización de parámetros fisiológicos en poscosecha de variables fisiológicas, modelación del comportamiento de la maduración de uno de los cultivares pertenecientes a una de las razas más implementadas para la producción de aguacate en zonas tropicales por debajo de los 1000 msnm. Aspectos que permitirán proponer un mejor aprovechamiento de las características de frutos de esta raza, así como también, plantear metodologías en poscosecha que permiten su comercialización por periodos de tiempo más prolongados. De igual manera, existen otros compuestos que deben ser evaluados para el material Lorena dado que aún es muy poco el conocimiento que se tiene sobre ellos. Este es el caso de los esteroides, moléculas muy apetecidas por la industria alimenticia y farmacéutica debido a que son estructuras análogas al colesterol que carecen de sus efectos negativos, adicionalmente son implementadas para el tratamiento de enfermedades como la osteo-artritis y para la regeneración de tejidos cutáneos al favorecer la síntesis de colágeno (Robayo, 2017).

1.7 Conclusiones y Recomendaciones

En el fruto de aguacate cultivar Lorena la tasa de producción de etileno y tasa de respiración son las variables que mejor explican el proceso de maduración del fruto en etapa poscosecha. Las variables firmeza, extensibilidad y color presentan diferencias significativas en cada una de las etapas de poscosecha lo que permite identificar los estadios de preclimaterio, climaterio y posclimaterio.

Este estudio generó por primera vez un modelo que explica el proceso de maduración del fruto de la raza Antillana utilizando las variables tasa de producción de etileno y respiración mediante un modelo cinético enzimático cuya dinámica se explica mediante un polinomio de grado dos. Con base en este modelo se estableció que el estado climaterio se alcanza para el día 7 de poscosecha.

El contenido de ácidos grasos se incrementó a lo largo de la maduración del fruto, esta dinámica no muestra cambios estadísticamente significativos; lo cual sugiere que una extracción de aceites puede llevarse a cabo independiente del estado poscosecha.

En futuros estudios de poscosecha sobre frutos del cultivar Lorena, se recomienda evaluar diferentes condiciones de temperatura y humedad relativa, para así determinar su efecto en la longevidad de este material. Así como también, evaluar compuestos con implicaciones positivas a la salud humana y de los cuales aún hoy en día se tienen poco conocimiento. Este es el caso de los esteroides, moléculas muy apetecidas por la industria alimenticia y farmacéutica debido a que son estructuras análogas al colesterol que carecen de sus efectos negativos, adicionalmente son implementadas para el tratamiento de enfermedades como la osteo-artritis y para la regeneración de tejidos cutáneos al favorecer la síntesis de colágeno.

1.8 Referencias

- Alamar, M. C., Collings, E., Cools, K., & Terry, L. A. (2017). Impact of controlled atmosphere scheduling on strawberry and imported avocado fruit. *Postharvest Biology and Technology*, 134(September), 76–86. <https://doi.org/10.1016/j.postharvbio.2017.08.003>
- Balaguera López, H. E. (2011). Estudio del crecimiento y desarrollo del fruto de champa (*Campomanesia lineatifolia* R & P) y determinación del punto óptimo de cosecha, 152.
- Bapat, V. A., Trivedi, P. K., Ghosh, A., Sane, V. A., Ganapathi, T. R., & Nath, P. (2010). Ripening of fleshy fruit : Molecular insight and the role of ethylene. *Biotechnology Advances*, 28(1), 94–107. <https://doi.org/10.1016/j.biotechadv.2009.10.002>
- Bernal, J., Díaz, C., Tamayo, A., Córdoba, O., Londoño, M., Tamayo, P., & Londoño, M. (2008). *Tecnología para el cultivo del aguacate*. Corpoica.
- Cañas-Gutiérrez, G. P., Galindo-López, L. F., Arango-Isaza, R., & Saldamando Benjumea, C. I. (2015). Diversidad genética de cultivares de aguacate (*Persea americana* Mill) en Antioquia, Colombia. *Agronomía Mesoamericana*, 26(1), 129–143. <https://doi.org/10.15517/am.v26i1.16936>

- Castellanos, D. A., Polanía, W., & Herrera, A. O. (2016). Development of an equilibrium modified atmosphere packaging (EMAP) for feijoa fruits and modeling firmness and color evolution. *Postharvest Biology and Technology*, *120*, 193–203. <https://doi.org/10.1016/j.postharvbio.2016.06.012>
- Chen, H., Morrell, P. L., Ashworth, V. E. T. M., De La Cruz, M., & Clegg, M. T. (2009). Tracing the geographic origins of major avocado cultivars. *Journal of Heredity*, *100*(1), 56–65. <https://doi.org/10.1093/jhered/esn068>
- Ferreira, R., Sellés, G., Saavedra, J., Ortiz, J., Zúñiga, C., Troncoso, C., ... De, B. G. (2016). Identification of pre-harvest factors that affect fatty acid profiles of avocado fruit (*Persea americana* Mill.) cv. 'Hass' at harvest. *104*, 15–20. <https://doi.org/10.1016/j.sajb.2015.10.006>
- Hernández, I., Fuentealba, C., Olaeta, J. A., Lurie, S., Defilippi, B. G., Campos-Vargas, R., & Pedreschi, R. (2016). Factors associated with postharvest ripening heterogeneity of 'Hass' avocados (*Persea americana* Mill.). *Fruits*, *71*(5), 259–268. <https://doi.org/10.1051/fruits/2016016>
- Ibarra-Laclette, E., Méndez-Bravo, A., Pérez-Torres, C. A., Albert, V. A., Mockaitis, K., Kilaru, A., ... Herrera-Estrella, L. (2015). Deep sequencing of the Mexican avocado transcriptome, an ancient angiosperm with a high content of fatty acids. *BMC Genomics*, *16*(1), 599. <https://doi.org/10.1186/s12864-015-1775-y>
- Kassim, A., Workneh, T. S., & Bezuidenhout, C. N. (2013). A review on postharvest handling of avocado fruit. *African Journal of Agricultural Research*, *8*(21), 2385–2402. <https://doi.org/10.5897/AJAR12.1248>
- Kilaru, A., Cao, X., Dabbs, P. B., Sung, H.-J., Rahman, M. M., Thrower, N., ... Ohlrogge, J. B. (2015). Oil biosynthesis in a basal angiosperm: transcriptome analysis of *Persea Americana* mesocarp. *BMC Plant Biology*, *15*, 203. <https://doi.org/10.1186/s12870-015-0586-2>
- Lopez-Gomez, R., Torres-Cardenas, Y., Chavez-Mochezuma, M., Salgado-Garciglia, R., Jimenez-Moraila, B., Corona-Armenta, G., & Herrera-Estrella, L. (2007). Genomics of Avocado Criollo Fruit (*Persea americana* Mill. VAR. *Drymifolia*). *Proceedings VI World Avocado Congress*, (978).

- Maftoonazad, N., & Ramaswamy, H. S. (2005). Postharvest shelf-life extension of avocados using methyl cellulose-based coating. *LWT - Food Science and Technology*, 38(6), 617–624. <https://doi.org/10.1016/j.lwt.2004.08.007>
- Manera, F. J., Legua, P., Melgarejo, P., Brotons, J. M., Hernández, F. C. A., & Martínez, J. J. (2013). Determination of a colour index for fruit of pomegranate varietal group “Mollar de Elche.” *Scientia Horticulturae*, 150, 360–364. <https://doi.org/10.1016/j.scienta.2012.11.036>
- Martínez-Romero, D., Bailén, G., Serrano, M., Guillén, F., Valverde, J. M., Zapata, P., ... Valero, D. (2007). Tools to maintain postharvest fruit and vegetable quality through the inhibition of ethylene action: A review. *Critical Reviews in Food Science and Nutrition*, 47(6), 543–560. <https://doi.org/10.1080/10408390600846390>
- Olaeta, J. a. (2003). Industrialización Del Aguacate: Estado Actual Y Perspectivas Futuras. *Actas V Congreso Mundial Del Aguacate*, 749–754.
- Ozdemir, F., & Topuz, A. (2004a). Changes in dry matter, oil content and fatty acids composition of avocado during harvesting time and post-harvesting ripening period. *Food Chemistry*, 86(1), 79–83. <https://doi.org/10.1016/j.foodchem.2003.08.012>
- Ozdemir, F., & Topuz, A. (2004b). Changes in dry matter , oil content and fatty acids composition of avocado during harvesting time and post-harvesting ripening period, 86, 79–83. <https://doi.org/10.1016/j.foodchem.2003.08.012>
- Pereira, M. E. C., Tieman, D. M., Sargent, S. A., Klee, H. J., & Huber, D. J. (2013). Volatile profiles of ripening West Indian and Guatemalan-West Indian avocado cultivars as affected by aqueous 1-methylcyclopropene. *Postharvest Biology and Technology*, 80, 37–46. <https://doi.org/10.1016/j.postharvbio.2013.01.011>
- Pravendra Nath, Mondher Bouzayen, Autar K. Mattoo, J. Claude P. (2014). *Fruit Ripening: Physiology, Signalling and Genomics*. Retrieved from <https://books.google.com/books?id=hceWBAAAQBAJ&pgis=1>
- Priego, B., Villanueva, E. G., & García, E. A. (1996). Anatomía Del Fruto De Aguacate, ¿Drupa o Baya?, 2(1942), 189–198.
- Robayo, M. A. T. (2017). Caracterización fisicoquímica de diferentes variedades de

- aguacate, *Persea americana* Mill. (Lauraceae) e implementación de un método de extracción del aceite de aguacate como alternativa de industrialización, 111.
Retrieved from <http://www.bdigital.unal.edu.co/56953/>
- Rodríguez-Carpena, J. G., Morcuende, D., & Estévez, M. (2012). Avocado, sunflower and olive oils as replacers of pork back-fat in burger patties: Effect on lipid composition, oxidative stability and quality traits. *Meat Science*, *90*(1), 106–115.
<https://doi.org/10.1016/j.meatsci.2011.06.007>
- Saltzman, A., Birol, E., Oparinde, A., Andersson, M. S., Asare-Marfo, D., Diressie, M. T., ... Zeller, M. (2017). Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. *Annals of the New York Academy of Sciences*, *1390*(1), 104–114. <https://doi.org/10.1111/nyas.13314>
- Serpa G, A. M., Echeverri L, A., Lezcano C, M. P., Velez A, L. M., Rios, A. F., & Adolfo H, G. (2014). Extracción de Aceite de Aguacate Variedad “Hass ” (*Persea americana* Mill) Liofilizado por Prensado en Frío. *Revista Investigaciones Aplicadas*, *8*(2), 113–123.
- Taiz, L., & Zeiger, E. (2010). Plant Physiology, Fifth Edition. *Cell*, *1*, 782. Retrieved from <http://www.amazon.com/dp/0878938567>
- Wang, M. M., Zhu, Q. G., Deng, C. L., Luo, Z. R., Sun, N. J., Grierson, D., ... Chen, K. S. (2017). Hypoxia-responsive ERFs involved in postdestringency softening of persimmon fruit. *Plant Biotechnology Journal*, *15*(11), 1409–1419.
<https://doi.org/10.1111/pbi.12725>
- Wolstenholme, N., & Whiley, A. (2013). *The avocado botany, production and uses*. (B. Schaffer, Ed.) (2da ed.). Croydon. Retrieved from www.cabi.org
- Woolf, A. B., Roughan, G., Schroeder, R., & Young, H. (2000). Content and Fatty Acid Composition of ‘ Hass ’ Avocados Report to the NZ Avocado Industry Council, 1–27.
- Wu, Z., & Burns, J. K. (2004). A b -galactosidase gene is expressed during mature fruit abscission of ‘ Valencia ’ orange (*Citrus sinensis*), *55*(402), 1483–1490.
<https://doi.org/10.1093/jxb/erh163>

Anexos

1.8.1 Suplementos de la metodología

Tasa de respiración: La tasa de respiración de frutos se estimó cada dos días mediante un equipo Vernier (LabQuest®2, Beaverton) el cual registra la producción de CO₂ mediante un sensor. Para este proceso se ubicaron los frutos de forma individual dentro de cámaras herméticas y oscurecidas mediante la implementación de bolsas plásticas de color negro durante un periodo de diez minutos. Los valores de CO₂ resultantes fueron ajustados a un modelo lineal a partir del cual se calculó la pendiente (m) de la recta estimada y con ella los valores de tasa de respiración (mg de CO₂ kg⁻¹ hora⁻¹) a partir del peso del fruto y de la implementación de la fórmula:

$$\text{mg de CO}_2 \text{ kg}^{-1} \text{ hora}^{-1} = [((3.6 \times (2 - \text{Kg de fruto})) \times (m \times \text{Kg de fruto}))].$$

Tasa de producción de etileno: La tasa de producción de etileno de frutos de aguacate se midió cada dos días, luego de almacenarlos por un período de dos horas en cámaras con cierre hermético de 20 L. Transcurrido este tiempo, se extrajo una muestra de 1 mL del gas contenido en la cámara mediante una jeringa de 1 mL. El gas extraído se inyectó en un equipo de cromatografía de gases Agilent (GC/MSD 5975C,CA) el cual reporta el área de la muestra de etileno por fruto y esta área sirve para determinar la tasa de producción de etileno $\mu\text{L Kg}^{-1} \text{ h}^{-1}$ con base en la curva de calibración previamente establecida usando un mL de etileno comercial (99 ppm). Las curvas de calibración relacionan el volumen de etileno en mL, las partes por millón (ppm) de etileno y el área bajo la curva obtenida en el equipo. La (Tabla 1-7) y (Tabla 1-8) muestra la curva de referencia obtenida a diferentes volúmenes de etileno comercial en dos tiempos diferentes.

Tabla 1-7: Curva de referencia a diferentes volúmenes de etileno comercial (99 ppm mL⁻¹) 27 de junio de 2016.

ml de etileno	ppm de etileno	Área
0.05	4.95	9.21
0.10	9.90	15.29
0.20	19.80	24.06
0.30	29.70	32.49
0.40	39.60	38.20
0.50	49.50	52.41
0.70	69.30	76.60
1.00	99.00	134.88

Tabla 1-8: Curva de referencia a diferentes volúmenes de etileno comercial (99 ppm mL⁻¹) 30 de junio de 2016.

ml de etileno	ppm de etileno	Área
0.05	4.95	8.30
0.10	9.90	15.45
0.20	19.80	31.85
0.40	39.60	39.09
0.50	49.50	47.01
0.60	59.40	49.07
0.70	69.30	60.99
1.00	99.00	85.80

Pérdida de peso fresco: El peso de la fruta fue tomado cada dos días. A partir de los valores obtenidos se calculó el porcentaje de peso perdido durante la poscosecha mediante la ecuación:

$$\% \text{ pérdida de peso fresco} = 100 * [(\text{peso inicial} - \text{peso final})/\text{peso inicial}].$$

Masa seca: La masa seca se determinó luego de pesar una porción de pulpa de fruta liofilizada y almacenada en bolsas de papel cuyo peso fresco fue previamente estimado. Este proceso se realizó mediante la implementación del liofilizador LABCONCO

(FreeZone 2.5, Kansas), sobre tres muestras para cada una de las etapas de maduración poscosecha. El porcentaje de masa seca fue estimado mediante la fórmula:

% de masa seca de aguacate = $100 * (\text{gramos de masa seca} / \text{gramos de masa fresca})$

Color: El color se determinó mediante un colorímetro Konica Minolta (Chroma meter CR-400, NJ) para cada etapa de maduración. Mediante un análisis cuantitativo que ubica los valores obtenidos en el espacio de color CIELab compuesto por tres variables L^* , a^* y b^* . La variable L^* define la luminosidad del fruto, siendo los valores más cercanos a cero para negro y 100 para blanco. La variable a^* , mide el color del fruto con valores negativos cercanos a -60 para el color verde y valores positivos cercanos a +60 para el color rojo. La variable b^* , con valores positivos cercanos a +60 para el color amarillo y negativos cercanos a -60 para el color azul (Manera *et al.*, 2013) (Rodríguez-Carpena, Morcuende, & Estévez, 2012).

Firmeza y extensibilidad: Las variables firmeza y extensibilidad fueron calculadas mediante la implementación del texturometro Ametek® (Lloyd LS1, Berwyn) que cuenta con un dispositivo que permite introducir un penetrómetro a una presión de 56000 Newtons y a una tasa de 10 cm min^{-1} en una profundidad de 10 mm. Se realizaron tres mediciones sobre el plano ecuatorial en tres frutos con cáscara, para cada estado poscosecha. Los resultados arrojados de firmeza se expresaron en Newtons (N) y la extensibilidad obtenida en (mm).

Contenido y tipo de ácidos grasos: El contenido de aceites se determinó para cada uno de los tres estados de maduración poscosecha mediante el protocolo AOAC 920.39 - 2003 (Serpa G *et al.*, 2014). Se realizó una extracción continua durante 12 horas empleando un extractor Soxhlet, bencina de petróleo 60 – 80 °C para arrastrar los aceites contenidos en muestras de 2 g de mesocarpo de aguacate, previamente liofilizadas y contenidas en bolsas de papel. La estimación del porcentaje del contenido de aceites se realizó con base al peso fresco y el peso de la muestra luego de retirar los aceites con bencina de petróleo.

Las muestras liofilizadas se maceraron para poder llevar a cabo la determinación del perfil de ácidos grasos en los tres estados de maduración poscosecha del aguacate. A cada uno de los productos macerados, se le adicionó 12 mL de solución cloroformo-

hexano relación v/v 4:1, luego de almacenarlos en frascos de vidrio de manera independiente. Dicha mezcla se llevó a baño de maría y sonicación durante un período de 5 min para reducir el tamaño de partícula del producto liofilizado y liberar los ácidos grasos. Los aceites liberados fueron recolectados por filtración a través del uso de una membrana. La solución obtenida de este proceso se dejó secar durante un período de 24 h a una temperatura promedio de 20 °C y HR del 60%. 500 mg del extracto seco se depositaron en tubos de ensayo y se agregó trifluoruro de boro en metanol al 14%. Dicha solución fue calentada a 30 °C durante una hora. Luego se adicionó un mL de n-heptano. El sobrenadante orgánico se transfirió a tubos de ensayo nuevos donde se le adicionó sulfato de sodio. Sobre el nuevo sobrenadante se realizó la identificación de aceites para cada una de las muestras mediante cromatografía de gases usando un espectrómetro de masas Agilent (GC/MSD 5975C,CA).

1.8.2 Análisis estadístico

Tasa de producción de etileno

ANOVA ETILENO MUESTRA DE 80 DATOS

Sistema SAS

Procedimiento GLM

Variable dependiente: ETILENO

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	7	205384.7906	29340.6844	31.19	<.0001
Error	72	67736.6689	940.7871		
Total corregido	79	273121.4595			

Prueba de Tukey

Sistema SAS

Procedimiento GLM

Prueba del rango estudentizado de Tukey (HSD) para ETILENO

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	72
Error de cuadrado medio	940.7871
Valor crítico del rango estudentizado	4.41490
Diferencia significativa mínima	42.822

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	203.09	10	6
B	154.72	10	7
B			
C	123.90	10	14
C			
C	105.20	10	12
C			
C	101.41	10	5
D			
D	80.13	10	8
D			
D	73.57	10	10
E	24.05	10	3

Tasa de respiración

ANOVA PARA RESPIRACIÓN CON LOS 80 DATOS

Sistema SAS

Procedimiento GLM

Variable dependiente: RESPIRACION

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	7	4293.827339	613.403906	7.95	<.0001
Error	72	5552.379710	77.116385		
Total corregido	79	9846.207049			

Prueba de Tukey

Prueba del rango estudentizado de Tukey (HSD) para RESPIRACION

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	72
Error de cuadrado medio	77.11638
Valor crítico del rango estudentizado	4.41490
Diferencia significativa mínima	12.26

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	44.269	10	12
A			
A	44.095	10	7
A			
A	41.360	10	8
A			
A	38.714	10	6
A			

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	37.371	10	5
A			
A	34.896	10	10
A			
A	33.564	10	14
B	20.022	10	3

Peso fresco y variables fisiológicas determinadas en las etapas preclimaterio, climaterio y posclimaterio

Peso fresco

ANOVA CON LOS 80 DATOS PARA PESO

Procedimiento GLM

Variable dependiente: PESO

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	7	23292.3403	3327.4772	0.54	0.8030
Error	72	445454.8436	6186.8728		
Total corregido	79	468747.1840			

Prueba de Tukey para PESO

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	72
Error de cuadrado medio	6186.873
Valor crítico del rango estudentizado	4.41490
Diferencia significativa mínima	109.81

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	543.93	10	3
A			
A	533.62	10	5
A			
A	528.77	10	6
A			
A	523.49	10	7
A			
A	518.43	10	8
A			
A	506.91	10	10
A			
A	501.89	10	12
A			
A	488.27	10	14

Masa seca

ANOVA masa seca

Variable dependiente: MASASECA

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	2	6.15726667	3.07863333	1.34	0.3309
Error	6	13.81153333	2.30192222		
Total corregido	8	19.96880000			

Prueba de Tukey para Masa seca

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	6

Error de cuadrado medio	2.301922
Valor crítico del rango estudentizado	4.33917
Diferencia significativa mínima	3.8009

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	16.533	3	3
A			
A	15.930	3	7
A			
A	14.557	3	12

Color

ANOVA indice de color

Sistema SAS

Procedimiento GLM

Variable dependiente: IC

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	3	331.0170152	110.3390051	92.05	<.0001
Error	36	43.1543205	1.1987311		
Total corregido	39	374.1713357			

Prueba de Tukey para indice de color

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	36
Error de cuadrado medio	1.198731
Valor crítico del rango estudentizado	3.80880
Diferencia significativa mínima	1.3187

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	-7.5146	10	14
A			
A	-8.5932	10	12
B	-11.9523	10	7
C	-14.8170	10	3

Textura
ANOVA TEXTURA

Sistema SAS

Procedimiento GLM

Variable dependiente: TEXTURA

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	2	51958.68342	25979.34171	550.73	<.0001
Error	27	1273.65005	47.17222		
Total corregido	29	53232.33347			

Prueba Tukey

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	27
Error de cuadrado medio	47.17222
Valor crítico del rango estudentizado	3.50633
Diferencia significativa mínima	7.6154

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	102.446	10	3
B	18.662	10	7
C	10.265	10	12

Extensibilidad

ANOVA EXTENSIBILIDAD

Variable dependiente: EXTENSIBILIDAD

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	2	149.8870867	74.9435433	63.33	<.0001
Error	27	31.9532100	1.1834522		
Total corregido	29	181.8402967			

Prueba de Tukey para extensibilidad

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	27
Error de cuadrado medio	1.183452
Valor crítico del rango estudentizado	3.50633
Diferencia significativa mínima	1.2062

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	8.5540	10	3
B	5.2070	10	7
C	3.1280	10	12

Contenido de ácidos grasos
ANOVA porcentaje de aceites

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	2	0.52046667	0.26023333	0.29	0.7579
Error	6	5.37693333	0.89615556		
Total corregido	8	5.89740000			

Prueba de Tukey para porcentaje de aceites

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	6
Error de cuadrado medio	0.896156
Valor crítico del rango estudentizado	4.33917
Diferencia significativa mínima	2.3716

**Medias con la misma letra
no son significativamente
diferentes.**

Tukey Agrupamiento	Media	N	DIA
A	5.8900	3	12
A			
A	5.3867	3	3
A			
A	5.3733	3	7

Marco teórico

Genética y diversidad del aguacate

El aguacate (*Persea americana* Mill) es una planta alógama diploide de 24 cromosomas ($2n = 24$) altamente heterocigota (Wolstenholme & Whiley, 2013). Dentro de esta especie es posible encontrar una alta diversidad favorecida por la diferencia en el tiempo de maduración de sus órganos reproductivos (dicogamia) ubicados en sus flores hermafroditas. Los órganos femeninos maduran primero que los masculinos (protógina) lo que favorece una polinización cruzada y fecundación entre individuos genéticamente diferentes. Adicionalmente, existe un mayor cuajado de frutos ante condiciones de polinización abierta .

La planta de aguacate pertenece a la división Magnoliophyta, clase Magnoliopsida, orden Laurales, familia Lauraceae, género *Persea*. Este género reúne cerca de 50 especies entre las que sobresale la especie americana por su amplio uso en la producción de este fruto (Wolstenholme & Whiley, 2013) y para *Persea americana* Mill, se han identificado tres razas ecológicas: Antillana (*Persea americana* var. *americana*), Guatemalteca (*Persea americana* var. *guatemalensis*) y Mexicana (*Persea americana* var. *drymifolia*) (Ibarra-Laclette *et al.*, 2015). La relación genética entre los cultivares y razas se ha explorado a través de diferentes marcadores moleculares tipo RFLPs (Polimorfismos de Fragmentos de Longitud de Restricción), VNTRs (Número de Variables de Repeticiones en Tandem); AFLPs (Polimorfismos en la Longitud de Fragmentos Amplificados), SSRs (Secuencias Simples Repetidas) y SNPs (Polimorfismos de un Solo Nucleótido). Marcadores que también han permitido el desarrollo de mapas físicos y de ligamiento implementados en el estudio de su diversidad y de la segregación de loci de rasgos cuantitativos (QTLs) que pueden ser implementados en programas de mejoramiento (Chanderbali *et al.*, 2008).

Se ha comprobado que los cultivares y variedades pertenecientes a una raza pueden

contener en su genoma una proporción determinada del genoma de una, dos e incluso las tres razas ecológicas, esto debido principalmente a la polinización abierta de esta especie (Chen *et al.*, 2009) y de los parentales utilizados en el programa de fitomejoramiento. Este es el caso del cultivar Hass, donde a través de un estudio de diversidad usando marcadores SNP's, se encontró que su genoma contiene en mayor medida secuencias provenientes de la raza Guatemalteca (58%) y en menor proporción de la raza Mexicana (42%) (Chen *et al.*, 2009), por eso, este cultivar es considerado de la raza Guatemalteca. En contraste, un estudio de diversidad realizado a partir de la implementación de marcadores moleculares AFLP's determinó que el cultivar Lorena posee un genoma enteramente de la raza Antillana y que este material tiene una mayor relación con los cultivares y variedades Trapica, Trapp y papelillo (Ocampo *et al.*, 2006).

Genómica y transcriptómica del aguacate

El genoma de un organismo permite conocer la organización de los diferentes elementos génicos que se encuentran presentes en él; por ejemplo, facilita la identificación de secuencias polimórficas y la localización de genes específicos asociados a un rasgo de interés agronómico. En el NCBI, se reporta el primer borrador del genoma para el cultivar Hass pero solo el 50% está ensamblado (GCA_002908915.1). Por otro lado, Arumuganathan & Earle (1991), con base en un estudio realizado por citometría de flujo, reportan que el tamaño del genoma de *Persea americana* es pequeño de aproximadamente 883 Mbp.

Con el fin de conocer la relación entre la fisiología de frutos de aguacate y sus secuencias expresadas, existen estudios enfocados en estudiar la ruta de la síntesis de ácidos grasos a partir de la elaboración de un transcriptoma *De novo* en la raza Guatemalteca (cultivar Hass) usando ARN extraído de diferentes partes del fruto en estado poscosecha (epicarpo, mesocarpo y semilla) (Kilaru *et al.*, 2015). Adicionalmente, se ha estudiado la expresión diferencial de genes a partir del transcriptoma de la raza Mexicana obtenido a partir del ARN extraído de raíz, tallo, hojas, flores y frutos en los diferentes estados de poscosecha (Ibarra-Laclette *et al.*, 2015). Estos estudios han permitido conocer los genes implicados en la ruta de síntesis de ácidos grasos y en el proceso de maduración del fruto. Sin embargo, hasta la fecha no existen reportes de este tipo de estudios para la raza Antillana.

Extracción de ARN

Para analizar el transcriptoma de un organismo se requiere extraer y purificar ARN de buena calidad, libre de contaminantes y con buena integridad. Para ello, se recomienda partir de tejido fresco previamente conservado en un reactivo como RNAlater® de SIGMA-ALDRICH o congelado, a una temperatura de -80 °C, inmediatamente después de coleccionar. La extracción puede obtenerse a través de kits o protocolos de laboratorio que garanticen la inactivación de RNAsas debido a que el ARN se degrada más fácil que el ADN. Sin embargo, al igual que el proceso de extracción de ADN, este comprende las etapas de lisis celular, separación del ARN de agentes contaminantes, precipitación, limpieza de sales y solubilización del ARN en agua estéril libre de RNAsas. Para verificar que el ARN total no esté contaminado con ADN es necesario tratar la muestra con DNAsas antes de utilizarlo.

La concentración y pureza del ARN total extraído se ha cuantificado tradicionalmente mediante espectrofotometría a través de equipos como Nanodrop (Massachusetts, USA) determinando las relaciones 260/280 y 260/230 y visualizando el ARN mediante geles de agarosa. También existen otras metodologías para determinar la concentración basadas en la utilización de agentes fluorescentes que se unen al ARN y son más precisos en la cuantificación como el Qubit y más recientemente Agilent Technologies (Santa clara, USA), ha integrado la evaluación de integridad, calidad y concentración mediante el desarrollo de equipos como Bioanalyzer y TapeStation. Estos equipos combinan diferentes técnicas (microfluidos, electroforesis capilar y fluorescencia) con un software que permiten cuantificar la integridad del ARN extraído a través del RIN (Número de Integridad del ARN). El RIN se clasifica en 10 categorías en donde la 1 corresponde a una mala extracción y 10 describe un ARN de buena calidad (Schroeder et al., 2006). Para proyectos de RNAseq, usualmente se usan valores superiores a 7.

Secuenciación de nueva generación (NGS, Next Generation Sequencing)

La nueva generación de secuenciadores permiten obtener la secuenciación de genomas o transcriptomas a partir de su fragmentación, ligación de adaptadores e índices, amplificación por PCR y posterior secuenciación, incrementando masivamente el

volumen de datos obtenidos a un menor costo y tiempo (Hurtado, 2014). Anteriormente, el valor por cada nucleótido secuenciado estaba cercano a los 10 dólares y en la actualidad está a menos de 0.000001 dólares (Feng *et al.*, 2018).

Entre las plataformas de secuenciación vigentes se encuentran: 454 Life sciences (Basel, Suiza) (2005), cuya tecnología se fundamenta en la pirosecuenciación; Illumina (CA, USA) (2007) basada en la secuenciación por síntesis química en donde se emplea didioxinucleótidos marcados con fluorocromos (ddNTPs); Life technologies (CA, USA) SOLiD™ basado en secuenciación por ligación; y la secuenciación en tiempo real de molécula sencilla de ADN por Pacific Bioscience (CA, USA) (2004) (Masoudi-Nejad *et al.*, 2013). En la (Tabla 2-1) se presenta las principales ventajas y desventajas de estas tecnologías.

Tabla 2-1: Plataformas de secuenciación de nueva generación, características, ventajas y desventajas (Masoudi-Nejad *et al.*, 2013).

Plataforma	Longitud de lecturas (nucleótidos)	Número de lecturas	Salida (Gb)	Número de muestras	Tiempo de corrida	Ventajas	Desventajas
Roche 454 GS FLX+	700	1 X 10 ^{^6}	0.7	192	23 h	lecturas largas, corto tiempo de corrida	Costoso, error de homopolímeros
Illumina Hiseq 2000	100	3 X 10 ^{^9}	600	384	11 días	alto rendimiento	Limitado número de index
Life technologies SOLiD 5500xl	75	15 X 10 ^{^9}	180	1152	14 días	Corrección de error intrínseco	lecturas cortas
Roche 454 GS junior	400	1 X 10 ^{^5}	0.035	132	9 h	lecturas largas	Costoso, error de homopolímeros
Illumina miSeq	150	5 X 10 ^{^6}	15	96	27 h	corto tiempo de corrida y uso fácil	Costoso por base
Ion torrent PGM ion 316 chip	> 100	1 X 10 ^{^6}	0.1	16	2 h	corto tiempo de corrida y bajo costo de reactivos	No está bien evaluado
Helicos biosciences HeliScope	35	1 X 10 ^{^9}	35	4800	8 días	Secuenciación de molécula simple, ARN	lecturas cortas y alta tasa de error
Pacific bioscience PacBio RS	> 1000	1 X 10 ^{^5}	0.1	1	90 min	Secuenciación de molécula simple, lecturas largas, corto tiempo de corrida	alta tasa de error y bajo rendimiento

Con el avance de estas tecnologías, ha sido posible estudiar el transcriptoma de un organismo mediante el análisis de ARNseq. Este análisis permite obtener una “fotografía” de todos los transcritos que se están expresando en un organismo en un espacio y tiempo determinado. Se puede cuantificar el número de transcritos tanto codificantes o no codificantes, detectar transcritos raros e isoformas génicas así como realizar un análisis de expresión diferencial de genes (SAGE: Serial Analysis of Gene Expression) (Xu, Wang, & Qiao, 2009).

Tecnología Illumina y preparación de librerías

La tecnología de secuenciación desarrollada por Illumina requiere de tres grandes etapas que corresponden a la preparación de librerías, generación de clusters y secuenciamiento. La preparación de librerías consiste en extraer el ARN mensajero (ARNm) del ARN total extraído por captura de su extremo poly A mediante la implementación de perlas oligo dT. Posteriormente, el ARNm es fragmentado por métodos físicos (nebulización, ultrasonicación, fuerzas acústicas) o químicos (enzimas de restricción), seguido de la síntesis de la primera y segunda hebra de cDNA, ligación de adaptadores y amplificación por PCR punto final para adicionar los índices y primers de secuenciación.

La generación de clusters implica la amplificación de fragmentos de hasta 150 pb sobre un soporte sólido o celda de flujo. Para esta amplificación las hebras de cADN que contienen los adaptadores, índices y primers de secuenciación se deben fijar a la celda de flujo por complementariedad de los primers de secuenciación con los oligos contenidos en ella. La amplificación se realiza de manera vertical formando puentes desde los cuales se sintetiza la nueva hebra de cADN, proceso que se repite por varios ciclos hasta saturar la celda de flujo (aproximadamente 1000 copias de fragmentos de ADN de cadena sencilla); teniendo en cuenta que en cada paso de amplificación, se remueve la hebra con orientación 3´ a 5´. Finalmente, el secuenciamiento se realiza por síntesis en donde se adicionan ddNTPs (dideoxinucleótidos) marcados con fluorocromos que al incorporarse el equipo son detectados por medio de una cámara CCD y posteriormente se remueven y el ciclo de síntesis se repite (Masoudi-Nejad *et al.*, 2013).

Análisis de secuencias

Para realizar los análisis de ARNseq es necesario procesar los datos crudos obtenidos en el proceso de secuenciación considerando las siguientes etapas: (i) pre-procesamiento de datos; (ii) ensamblaje de transcriptoma; (iii) estimación de la abundancia de los transcriptos; (iii) análisis de expresión diferencial de genes; y (iv) anotación estructural y funcional.

Pre-procesamiento de datos

El pre-procesamiento de datos consiste en realizar una descripción de la calidad de los datos obtenidos por secuenciación mediante programas como FastQC (Babraham institute, Cambridge) y una eliminación de los datos con baja calidad y lecturas de ribosomales mediante programas bioinformáticos. El programa FastQC genera un reporte en formato .html que contiene las estadísticas básicas del secuenciamiento (nombre y tipo del archivo, plataforma de secuenciación, número de secuencias totales, secuencias con baja calidad, longitud de secuencias y porcentaje de guaninas y citosinas (%GC). Adicionalmente, brinda información sobre la calidad obtenida por bases de secuencia; es decir, se define la exactitud con la cual un nucleótido es identificado en una secuencia mediante la determinación del nivel de calidad Phred (Quality Score Phred, Q). Por ejemplo, un nivel de Q20 significa que la probabilidad de que una base sea incorrecta es de 1 entre 100 (99%), valores superiores a Q20 indican una mayor precisión de que la base este correctamente identificada.

El programa FastQC también arroja el puntaje obtenido por calidad de secuencia, lo que permite determinar el número de lecturas o secuencias que poseen una calidad determinada (QS); el contenido de bases por secuencia describe el lugar que ocupa cada uno de los nucleótidos en las secuencias (en porcentaje) con el fin de determinar si existe una posible contaminación o secuenciación parcial de la librería; el contenido de GC tiene en cuenta el conjunto de datos y la relación a la distribución de estos dos nucleótidos en una secuencia y debe presentar una distribución normal si no existe contaminación de la elaboración de las librerías; el contenido de bases no identificadas por secuencia que se designa con una letra N cuyo valor ideal es aquel que sea cercano a cero; la distribución de la longitud de las secuencias, indica el tamaño de los

fragmentos de secuencias con los que se construyeron las librerías, lo más conveniente es que estas tengan un tamaño homogéneo; los niveles de duplicación de las secuencias, indican el nivel de enriquecimiento de las librerías por procesos de amplificación; las secuencias sobrerrepresentadas, permiten evidenciar si hay contaminación de la muestra o si corresponde a una secuencia con importancia biológica; y por último, el contenido de adaptadores.

El proceso de limpieza de datos elimina lecturas de baja longitud, remueve secuencias de adaptadores, primers, secuencias ribosomales y secuencias o parte de secuencias de baja calidad. Para ello, se usan programas como Trimmomatic (http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic/TrimmomaticManual_V0.32.pdf) y ribopicker (<http://ribopicker.sourceforge.net/manual.html>).

Ensamblaje de transcriptoma

El ensamblaje de secuencias de transcritos de cADN provenientes del ARN extraído de un organismo permite elucidar las secuencias expresadas de su genoma (Tjaden, 2015), por lo cual se omite las secuencias intrónicas. Este ensamblaje se realiza de manera organizada agrupando las secuencias que se sobrelapan en contigs y luego en scaffolds mediante el método de extremos apareados (mate pairs) (Schliesky *et al.*, 2012). El transcriptoma se puede ensamblar por tres vías: por referencia, *De novo* y combinado.

El ensamblaje por referencia consiste en realizar un alineamiento de las lecturas de los transcritos obtenidos por secuenciación a un genoma de referencia (mapping), obteniéndose un transcriptoma más preciso ya que se contrasta contra una referencia que ya ha sido evaluada (Martin & Wang, 2011); para ello, se emplean algoritmos como los métodos de semilla y de transformación “Burrows-Wheeler” donde el alineamiento de lecturas a un genoma de referencia no permite grandes intervalos (gaps). Programas como Bowtie permiten este tipo de alineamiento (Martin & Wang, 2011).

El ensamblaje *De novo* se logra a través del uso de grafos de Bruijn que permite el sobrelape de las lecturas obtenidas por secuenciación y ensamble de estas en transcritos (Martin & Wang, 2011). En esta estrategia, se debe determinar primero la

longitud de la secuencia que se repite más veces en el proceso de secuenciación (K-mer) a través de programas como KmerGenie (Chikhi & Medvedev, 2014). El tamaño de K-mer seleccionado se usa para encontrar solapamientos entre las lecturas y obtener secuencias consenso llamadas “contigs”. La unión de estos contigs permite obtener el transcriptoma. Entre los programas más utilizados para ensamblar un transcriptoma *De novo* se encuentran: Trinity, Oases, SOAP-denovo-Trans, TransABySS and BinPacker (Wuang & Gribkov 2016).

Finalmente, la obtención de un transcriptoma puede darse por combinación de estas dos metodologías (ensamblaje combinado), en donde se parte de un transcriptoma *De novo* y luego éste es alineado a un genoma de referencia; de esta manera, se reduce el número de gaps o intervalos encontrados en los transcritos obtenidos por secuenciación (Martin & Wang, 2011).

Debido a que la calidad del ensamblaje afecta el análisis de expresión diferencial de genes, es necesario evaluar su calidad a través de diferentes métricas como la exactitud (% de bases y genes que están ensambladas correctamente con relación a un genoma de referencia), la cobertura de los transcritos ensamblados y el número de quimeras, entre otros (Wuang & Gribkov 2016). Esta evaluación se puede determinar mediante el uso de programas bioinformáticos como DETONATE (*De novo* Transcriptome rNa-seq Assembly with or without the Truth Evaluation) que presenta dos componentes: RSEM-EVAL, herramienta que no se guía por el N50 que corresponde a longitud de la lectura más larga del grupo de lecturas con que se ensamblaron al menos el 50% de las secuencias, sino que asigna un puntaje al ensamblaje total y a cada contig usado para producirlo mediante la implementación de un modelo probabilístico que depende sólo del ensamblaje y las lecturas obtenidas por RNA-seq (Li *et al.*, 2014). El segundo componente corresponde a REF-EVAL, herramienta que provee un puntaje para evaluar el transcriptoma obtenido con base a una referencia. No obstante, al comparar estas dos herramientas es posible observar que el puntaje asignado al ensamblaje obtenido correlaciona muy bien (Li *et al.*, 2014). Se han reportado otros programas como TransRate (Smith-Unna, 2016) y rnaQUAST que utiliza el genoma de referencia y bases de datos de genes (Bushmanova *et al.*, 2015).

Estimación de abundancia de transcritos

Para analizar la expresión diferencial de genes de un organismo es necesario estimar la abundancia de transcritos y predecir el número de isoformas de un gen producto de un empalme alternativo (splicing alternative). Esta predicción es difícil de calcular con exactitud debido principalmente a que las isoformas provenientes de un mismo gen se superponen y una lectura corta puede asignarse a más de una isoforma de transcripción (Huang et al 2013). Sin embargo, existen propuestas como RSEM (Li & Dewey, 2011) y Cufflinks (Trapnell 2010), que son programas informáticos que pueden usarse sobre datos producidos por secuenciación por solo un extremo (single-end) o extremos emparejados (paired-end) y calculan la abundancia relativa de los transcritos mediante el mapeo de los mismos a un genoma de referencia mediante el uso del programa Bowtie, para después estimar la abundancia de genes e isoformas basándose en las lecturas mapeadas y en una normalización previa de los conteos de las lecturas, la longitud de los transcritos y el número total de lecturas mapeadas en la muestra (RPKM: reads per kilobase of transcript per million mapped reads, o FPKM: fragments per kilobase of transcript per million mapped reads cuando los datos se secuenciaron por extremos emparejados).

Análisis de expresión diferencial de genes

Sobre la tabla de conteos generada por el paquete RSEM, es posible realizar una estimación bayesiana empírica, normalización y el uso de un modelo de distribución binomial negativa que tiene en cuenta la variabilidad biológica y el efecto lote por diferencias en la profundidad de secuenciación entre las muestras a través de paquetes como DESeq2 y edgeR del software R, para así determinar las lecturas asociadas a genes que cuentan con una expresión diferencial (Wang & Gribkov, 2017 ; Varet *et al.*, 2016).

Aunque el análisis de expresión diferencial de genes a partir de datos de ARNseq también puede realizarse a través de una distribución de Poisson, la cual tiene la desventaja de que es aplicada sobre réplicas técnicas y trabaja sobre un solo parámetro determinado por su media; mientras que la implementación de una distribución binomial negativa (BN) es más precisa para analizar las réplicas biológicas ya que tiene en cuenta la variabilidad biológica de las muestras (Soneson & Delorenzi, 2013). Por esta razón, los

paquetes edgeR y DEseq2 contemplan un análisis de distribución binomial y normalización usando valores como Trimmed Mean Values (TMM, en EdgeR) o sizeFactor (DEseq2) para determinar la expresión diferencial de genes para un bajo número de réplicas biológicas (Robison *et al.*, 2007) sobre tablas de conteos generadas por RSEM.

Anotación de secuencias diferencialmente expresadas

El proceso de anotación consiste en asociar metadatos a una colección de datos que han sido previamente identificados y propiamente definidos (Danchin *et al.*, 2018). Existen dos tipos de anotaciones definidas como estructural y funcional. La anotación funcional busca determinar la ubicación de genes y sus componentes, así como, sitios de splicing, regiones promotoras y reguladoras, regiones no traducidas (UTR), entre otros. Mientras que la anotación funcional busca asignar una función putativa de los genes con base a su homología con otras secuencias de genes y proteínas, cuya acción está vinculada a un proceso biológico previamente descrito (Hurtado, 2014).

A través del software R y mediante la elaboración de una base de datos obtenida desde NCBI o Uniprot; para genes o proteínas, a partir de la misma especie ó especie cercana, organismo modelo ó nivel taxonómico superior y posterior alineamiento local por homología de secuencias a través del algoritmo BLASTN (para nucleótidos) ó BLASTX (para proteínas), entre la base de datos generada y los genes representados por scaffold como diferencialmente expresados. Se genera como salida un archivo en formato XML que puede ser analizado por Blast2GO para identificar, por gene ontology (GO), el tipo de proceso biológico, función molecular y tipo de componente celular en el que actúan los genes diferencialmente expresados (Páez *et al.*, 2015).⁹

Capítulo 2. Expresión diferencial de genes presentes en los diferentes estadios de poscosecha de los frutos del cultivar Lorena (*Persea americana* var. *americana*)

2.1 Resumen

El aguacate (*Persea americana* var. *americana*) es una planta alógama diploide de 24 cromosomas ($2n = 24$), con un tamaño de genoma de aproximadamente 883 Mpb. Es altamente heterocigota, aspecto que podría estar relacionado con el hecho de poseer flores hermafroditas que exhiben protoginia, lo que favorece una polinización cruzada y una fecundación entre individuos genéticamente diferentes. Con el propósito de identificar aquellos genes relacionados a la maduración del aguacate Lorena durante los diferentes estadios de poscosecha: preclimaterio, climaterio y posclimaterio. Se realizó un análisis de expresión diferencial de genes mediante el desarrollo de un transcriptoma *De novo*. El transcriptoma de este cultivar fue secuenciado mediante el equipo NovaSeq 6000 de la plataforma illumina (Illumina®, CA), luego de producir nueve librerías de cADN con tamaño de fragmento promedio de 150 pb a partir de distintas réplicas biológicas, destinando tres muestras por estado poscosecha (pre climaterio, climaterio y posclimaterio). De este proceso se obtuvo un total de 574'375.160 lecturas que permitieron obtener un promedio de 50 ± 31 millones de lecturas por muestra; para una cobertura aproximada de 89,22 X. Sobre los datos crudos obtenidos se realizó una limpieza, elaboración de transcriptoma *De novo* con 22,749 transcritos, alineamiento de transcritos, determinación de transcritos diferencialmente expresados y posterior anotación, que llevo a identificar 6335 scaffolds asociados a genes diferencialmente expresados en los tres estados de poscosecha. Evidenciándose que el metabolismo de carbohidratos, metabolismo energético y metabolismo lipídico son los procesos más afectados por la expresión diferencial de genes, en donde hay una disminución

progresiva en el número de transcritos y enzimas involucrados al inicio de la poscosecha de frutos (preclimaterio) con respecto al final (posclimaterio).

Palabras Claves: transcriptoma *De novo*, RNAseq, Illumina, aguacate, secuenciación, anotación.

2.2 Introducción

La raza antillana de aguacate (*Persea americana* var. *americana*) es una de las variedades con mayor implementación en zonas con altitud inferior a 1.000 m.s.n.m y es popularmente conocido por productores y consumidores debido a su buena aceptación en Colombia (Ministerio de agricultura, 2017). No obstante, existe un escaso conocimiento en poscosecha de sus atributos fisiológicos y en mayor medida de su relación con la expresión de genes asociados a los estados preclimaterio, climaterio y posclimaterio del fruto, aspectos que pueden aportar en el entendimiento de la dinámica de estos materiales durante su vida en anaquel, para desarrollar marcadores funcionales y proponer nuevas formas de aprovechamiento mediante la elaboración de productos con valor agregado; como lo demuestran investigaciones realizadas en cultivares de otras razas tales como Guatemalteca (Kilaru *et al.*, 2015) y Mexicana (Ibarra-Laclette *et al.*, 2015).

La disponibilidad de implementar herramientas de secuenciación de nueva generación para caracterizar el transcriptoma de un organismo permitió por primera vez relacionar las secuencias expresadas con los diferentes estados fisiológicos de maduración del fruto de uno de los materiales más representativos de la raza antillana. En este capítulo se presenta un estudio de RNAseq a partir de muestras de pulpa de fruta extraídas en pre climaterio, climaterio y posclimaterio de aguacate cultivar Lorena perteneciente a la raza antillana, con el fin de determinar globalmente el grupo de genes diferencialmente expresados para cada uno de estos estadios poscosecha y su relación con rutas metabólicas.

Pese a que la disponibilidad de información y elementos puntuales como un genoma o transcriptoma de referencia es limitado para el desarrollo de esta metodología, el análisis de los transcritos se puede realizar a partir del ensamblaje *De novo* del transcriptoma, lo

que permite un mayor entendimiento de los procesos fisiológicos que ocurren durante el estado poscosecha y se puede responder al interrogante: ¿cuántos y cuáles son los genes expresados diferencialmente que están relacionados en el proceso de maduración de fruto de aguacate y a qué principales rutas metabólicas están relacionados?.

2.3 Materiales y métodos

2.3.1 Extracción de ARN total, elaboración de librerías y secuenciación

Las extracciones de ARN total se realizaron de acuerdo al protocolo reportado por Chang *et al.* (1993). Para cada uno de los estados preclimatérico, climatérico y posclimatérico se trabajaron tres réplicas biológicas diferentes y de cada una de ellas se extrajo ARN total a partir de porciones del mesocarpo del fruto sin cáscara y semilla, de aproximadamente 700 mg de peso. De esta manera, para preclimaterio se obtuvo las muestras pre1, pre2 y pre3; climaterio las muestras cli1, cli2 y cli3; y posclimaterio las muestras pos1, pos2 y pos3. Posterior a su obtención, el ARN extraído se mantuvo a -80°C . La concentración y pureza del ARN extraído se evaluó mediante espectrofotómetro NanoDrop (2000, Thermo Fisher Scientific™, Massachusetts). Su integridad se determinó mediante un gel de agarosa al 1% con corrida de 100 voltios durante 20 minutos. El RNA se visualizó mediante el documentador de geles Gel Doc™ XR+ (Bio-Rad laboratories Inc, CA).

Posteriormente, las muestras fueron tratadas con ADNasa I, amplification grade (Thermo Fisher Scientific™, Massachusetts) con base en los valores de concentración obtenidos por nanodrop (2000, Thermo Fisher Scientific™, Massachusetts) y enviadas al centro de Genómica y Bioinformática de la Universidad de Purdue en Lafayette - EE. UU asegurando cadena de frío mediante el uso de hielo seco. Una vez allí, se evaluaron nuevamente las nueve muestras mediante el uso de Bioanalyzer (Agilent® Technologies, CA), se prepararon las nueve librerías (3 estadios de poscosecha X 3 réplicas biológicas) de cADN de las muestras de aguacate mediante el uso del kit TruSeq RNA (Agilent® Technologies, CA). La secuenciación del cADN se realizó bajo una aproximación de extremos apareados (paired end 2X150 pb) sobre un canal multiplexado del equipo NovaSeq 6000 de la plataforma Illumina (Illumina®, CA). Se determinó que la cantidad

de lecturas secuenciadas tuviera al menos una cobertura de 30 veces (30X) el transcriptoma de referencia más cercano (mínimo sugerido para RNAseq). La cobertura fue estimada con base a la (fórmula 1).

Fórmula 1:

$$\text{Cobertura} = (\text{Total de lecturas producidas por muestra} \times \text{tamaño promedio de fragmentos}) / \text{tamaño de transcriptoma}$$

2.3.2 Procesamiento de datos: control de calidad y ensamblaje

Los datos crudos fueron procesados de manera que se redujeron los errores de secuenciamiento. Se empleó el programa Trimmomatic (Usadel Lab, Aachen) para eliminar las secuencias adaptadoras, nucleótidos de baja calidad en los extremos 5' y 3' (Phred score < 10) y los primeros 10 nucleótidos del extremo 5', los cuales presentan una variabilidad atípica debido a la preparación de la librería. Las lecturas de menos de 30 pb junto las huérfanas que representan solo uno de los lados del inserto secuenciado también se eliminaron.

La calidad de las lecturas crudas posterior al proceso de secuenciación y al proceso de limpieza fueron analizadas mediante el programa FastQC (Babraham Institute, Cambridge).

El transcriptoma del fruto de aguacate para la raza Antillana se generó usando una aproximación *De novo*. Esta aproximación se seleccionó debido a i) La cantidad de lecturas generadas satisfacían los requerimientos para generar un transcriptoma *de novo* de buena calidad, ii) El transcriptoma de referencia más cercano pertenecía al fruto en poscosecha de la raza Guatemalteca (Kilaru *et al.*, 2015) y iii) la referencia no pertenece a la raza de estudio, lo que puede reducir el número de lecturas mapeadas; esto limita el potencial informativo de los datos secuenciados. El ensamblaje se realizó por medio del programa SOAPdenovo (Beijing genomics institute, Beijing) utilizando todas las librerías secuenciadas (posterior al control de calidad). El parámetro de longitud de k-mers como entrada para SOAPdenovo se determinó usando KmerGenie (Chikhi & Medvedev, 2014). La calidad del transcriptoma generado se evaluó mediante dos aproximaciones. i)

Empleando el programa TransRate (Smith-Unna *et al.*, 2016) se analizaron las estadísticas de mapeo de las lecturas sobre los transcritos generados y ii) con BUSCO v3 (Simão *et al.*, 2015) usando dos bases de datos de ortólogos (Embriofitos y Eucotiledoneos) se evaluó la completitud del transcriptoma desde un punto de vista biológico.

2.3.3 Cuantificación de transcritos y análisis estadísticos

El programa RSEM (RNA-seq by Expectation-Maximization) se usó con cada una de las librerías (posterior al control de calidad) para cuantificar la expresión de los transcritos. De esta manera, se generó la tabla de conteos que contenía la cuantificación de los transcritos para cada una de las muestras y por ende para cada uno de los estados fisiológicos del fruto. La tabla de conteos fue procesada para eliminar transcritos falsos y/o de baja calidad. Por lo tanto, para los siguientes análisis solo se consideraron transcritos que tuvieran como mínimo diez conteos por millón en al menos tres de las librerías (Kilaru *et al.*, 2015), ya que los genes con bajo número de conteos proveen poca evidencia para detectar genes diferencialmente expresados. Además, desde el punto de vista estadístico, la reducción en el número de transcritos atenúa el efecto ocasionado por los métodos post-hoc de ajuste de pruebas múltiples, y desde un punto de vista biológico aquí sólo se consideran transcritos que han sobrepasado un mínimo de expresión antes de que este pueda ser traducido a proteína o sea biológicamente importante.

La tabla de conteos filtrada se utilizó para identificar los transcritos con expresión diferencial entre los estados fisiológicos. Para esto se empleó el paquete estadístico DESeq2 (Love, Huber, & Anders, 2014) en ambiente R (R Studio, Boston). Aquí, la expresión de cada transcrito se ajustó a un modelo de regresión lineal generalizado cuya función de ligamiento fue una binomial negativa usada para modelar datos de conteos (Sánchez Santana, 2015) Adicionalmente, en el modelo se incluyó el efecto lote (batch effect) como variable 'confounder' (variables que influyen la asociación entre las variables predictor y respuesta). El efecto lote es el error introducido por diseño experimental y el diseño del secuenciamiento. Así, la aproximación de subrogadas integrada en el paquete de SVA (Leek *et al.*, 2011) se usó para calcular el efecto lote.

Los genes diferencialmente expresados entre estados fisiológicos se determinaron con base en un test Wald. El p-valor fue corregido por pruebas múltiples usando “false discovery rate” (FDR) (Benjamini & Hochberg, 1995).

Para visualizar los resultados se realizó un análisis de componentes principales (PCA) y un agrupamiento jerárquico usando el algoritmo K-means entre los perfiles de expresión de los transcritos de los diferentes estados de maduración del fruto.

2.3.4 Anotación funcional

Finalmente, se generó una base de datos de proteínas descargada desde Uniprot pertenecientes a Magnoliophyta en formato fasta a través de Blast. Esta, fue comparada con la lista de genes diferencialmente expresados en aguacate para los tres estados de madurez. Obteniendo por homología, los genes ortólogos y las rutas metabólicas de los genes expresados diferencialmente mediante la herramienta BlastX. Estos resultados se visualizaron mediante el programa MEGAN 5, el cual compara las lecturas obtenidas con la base de datos de acceso libre de 2011 de Kyoto encyclopedia of genes and genomes (KEGG).

2.4 Resultados

2.4.1 22,749 transcritos definen el transcriptoma del fruto *Persea americana* var *americana*.

Para generar el primer transcriptoma del fruto de aguacate del cultivar Lorena se utilizaron nueve librerías de secuenciación paired end que al secuenciarse produjeron un total de 50 ± 31 millones de lecturas por muestra, para un total de 574'375.160 lecturas (Tabla 2-2). Luego de verificar la calidad de las secuencias obtenidas (Anexos; Tabla 2-3, Figura 2-1, Figura 2-2). El transcriptoma *De novo* fue generado con 573'514.384 secuencias de buena calidad (Tabla 2-2). De este volumen de datos, el transcriptoma del fruto de aguacate se definió con 22,749 transcritos para un tamaño de transcriptoma de 84'061.299, lo que permitió una profundidad de secuenciación aproximada de 89,22X por librerías (Formula 1).

Tabla 2-2: Total de lecturas luego del proceso de secuenciación y trimming para cada una de las muestras de ARN de pulpa de fruta.

Muestra	Lecturas crudas después de secuenciación	Lecturas después de Trimming
Pre1	82,024,456	81,905,810
Pre2	18,954,572	18,872,478
Pre3	54,776,900	54,658,400
Cli1	71,098,868	70,992,280
Cli2	60,759,246	60,671,892
Cli3	78,949,272	78,827,584
Pos1	67,597,226	67,530,350
Pos2	76,113,214	76,034,526
Pos3	64,101,406	64,021,064

Fórmula 1. Formula implementada para la determinación de la cobertura (coverage) del proceso de secuenciación.

Cobertura = (Total de lecturas producidas por muestra X tamaño promedio de fragmentos) / tamaño de transcriptoma

$$\text{Cobertura} = (50'000,000 \times 150) / (84'061.299)$$

$$\text{Cobertura} = 89,22$$

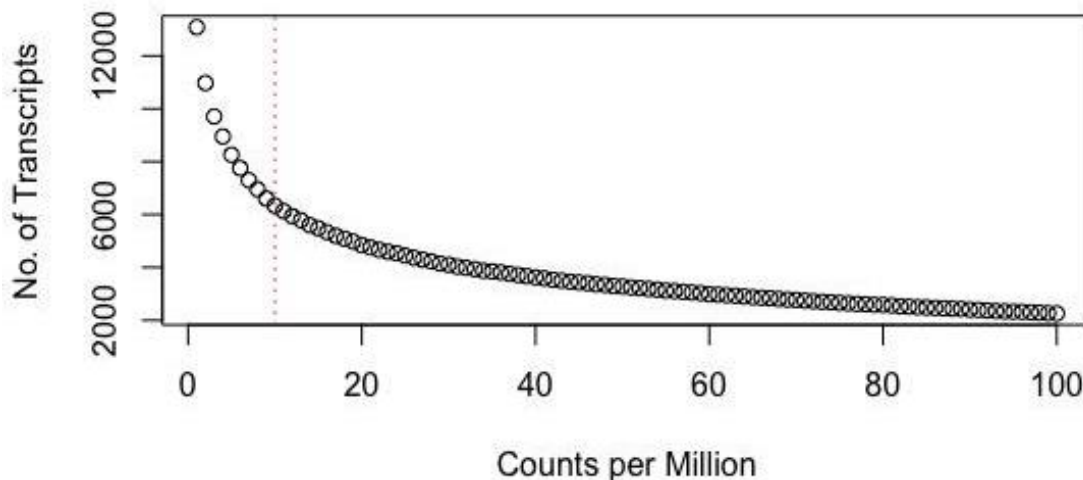
El 71.67% de las lecturas fueron mapeadas correctamente sobre el transcriptoma generado. El transcriptoma se caracterizó por un %GC de 42.06, una longitud media de transcritos de 463.40 y Tabla 2-3.

Tabla 2-3: Lecturas implementadas durante el proceso de ensamblaje de transcriptoma *De novo*, determinadas mediante el software TransRate.

Ensamblaje	Total
Transcritos ensamblados	22,749
Número de bases	84'061,299
% GC	42.06
N50	1,474
Longitud máxima	23,290
Longitud mínima	100
Longitud media (transcritos)	463.4
Lecturas mapeadas (correctamente emparejadas) usando Bowtie	201,003,370
Lecturas correctamente emparejadas (%)	71.67

Del transcriptoma se seleccionaron los transcritos que pudieran ser relevantes desde el punto de vista estadístico y biológico (ver materiales y métodos). Mediante esta selección los transcritos obtenidos fueron 6335 (Figura 2-1), los cuales definieron el potencial genético del fruto de aguacate.

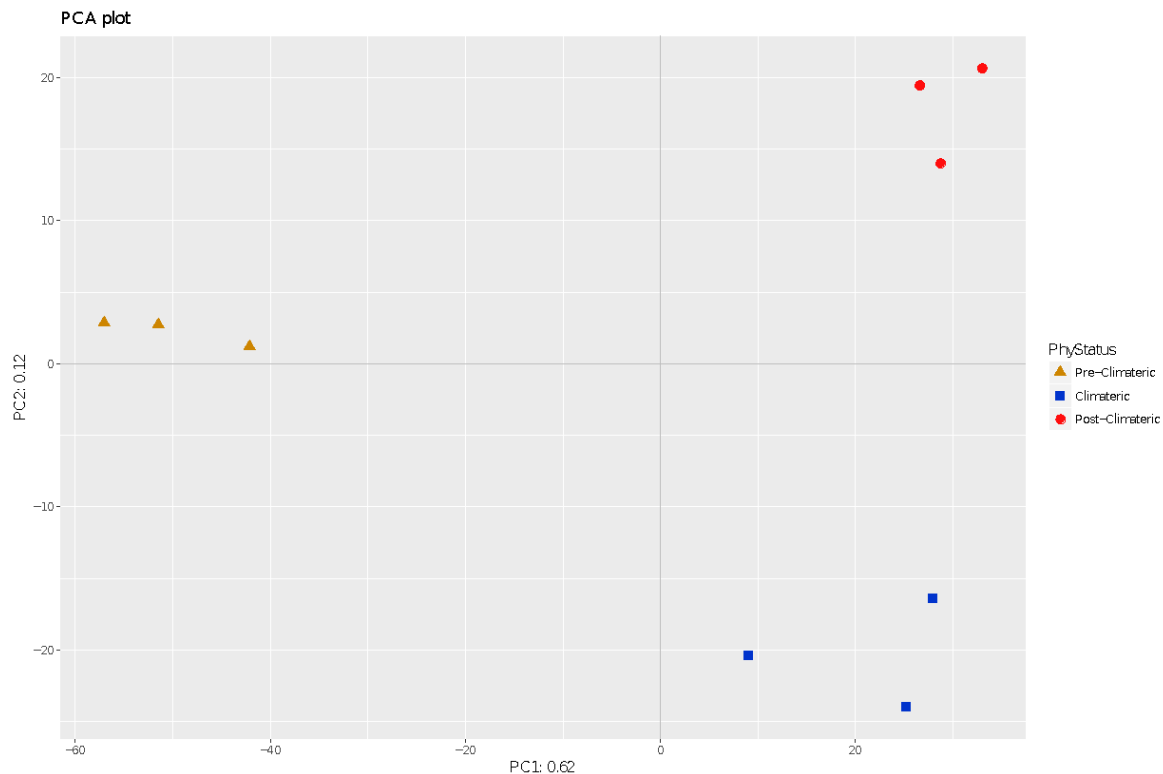
Figura 2-1: Número de transcritos con un mínimo de 10 conteos por millón en al menos tres de las librerías.



2.4.2 Los transcriptomas de climaterio y posclimaterio son al menos 90% similares

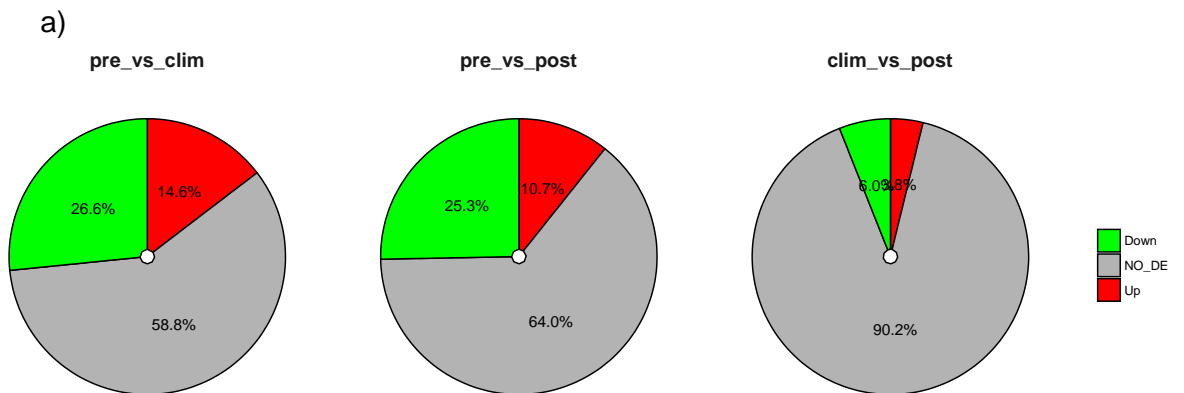
Para determinar si los transcriptomas de cada estado de maduración del fruto son diferentes, se usó un análisis de componentes principales (PCA). El PCA mostró que el 74% de la variabilidad en expresión génica está explicada por solamente dos componentes (Figura 2-2). El componente 1 (PC1) explica el 62% de la variabilidad de los datos y separa claramente al grupo de preclimaterio de los otros dos estados, mientras que el PC2 explica el 12% de la variabilidad y separa los grupos climaterio de posclimaterio. Por lo tanto, el PCA sugiere que todos los estados poscosecha son diferentes entre sí. Sin embargo, existe una mayor diferencia entre el transcriptoma en preclimaterio con respecto a los transcriptomas de climaterio y posclimaterio.

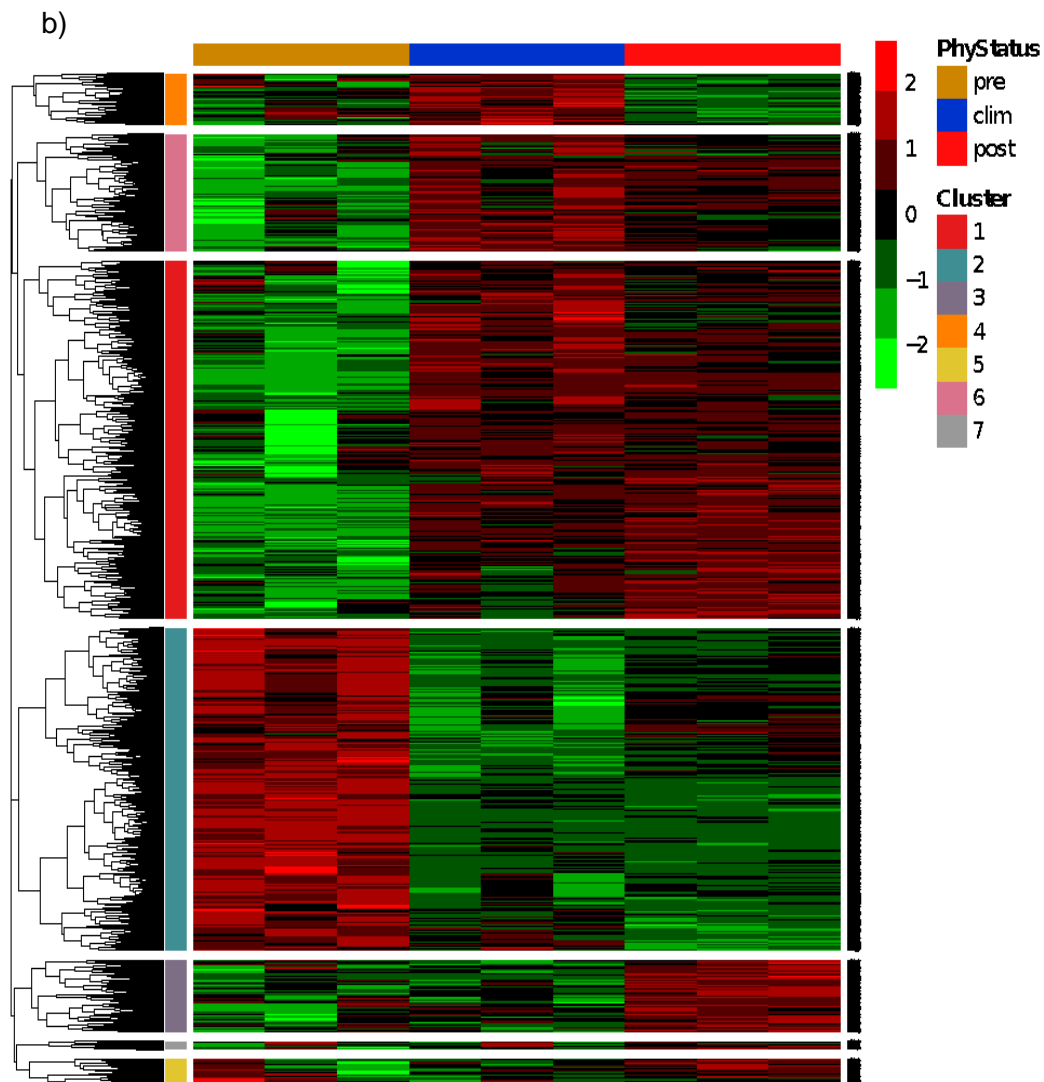
Figura 2-2: Análisis de componentes principales (PCA) para los transcritos obtenidos del proceso de secuenciación de muestras de cDNA a partir de pulpa de fruta de aguacate en los estados de pre climaterio, climaterio y posclimaterio.



Para determinar la proporción de transcritos que contribuyeron a explicar la diferencia entre los estados de maduración se realizaron comparaciones pareadas entre las expresiones relativas de los transcritos en cada uno de los estados poscosecha. El porcentaje de transcritos que contribuyen significativamente ($P < 0.01$) a la diferenciación entre el estado preclimaterio y climaterio fue de un 40,2% equivalente a 2612 transcritos; entre preclimaterio y posclimaterio fue de 36% (2279 transcritos) y entre climaterio y posclimaterio fue de apenas el 14.09% (619 transcritos) Figura 2-3a. Notablemente, los transcritos diferencialmente expresados entre los estados de maduración mostraron ciertos patrones de expresión los cuales fueron descritos en 7 grupos Figura 2-3b. En donde los resultados más contrastantes se observan en el cluster uno, dos y seis. En el cluster dos se agrupan un total de 1082 transcritos altamente expresados en pre climaterio y sub expresados en climaterio y pos climaterio. En contraste, para el cluster uno y seis, tenemos un total de 1200 y 395 transcritos sub expresados para el estado de preclimaterio con respecto a los estados de climaterio y posclimaterio, respectivamente.

Figura 2-3 a) Porcentaje de transcritos sobre expresados (up, color rojo), sub expresados (Down, color verde) y de expresión constitutiva (NO_DE, gris). A partir de comparaciones pareadas entre estados de madurez poscosecha. b) transcritos de muestras en estado de pre climaterio, climaterio y posclimaterio. Con color rojo transcritos sobre expresados, color verde sub expresados y color negro en condiciones normales de expresión.

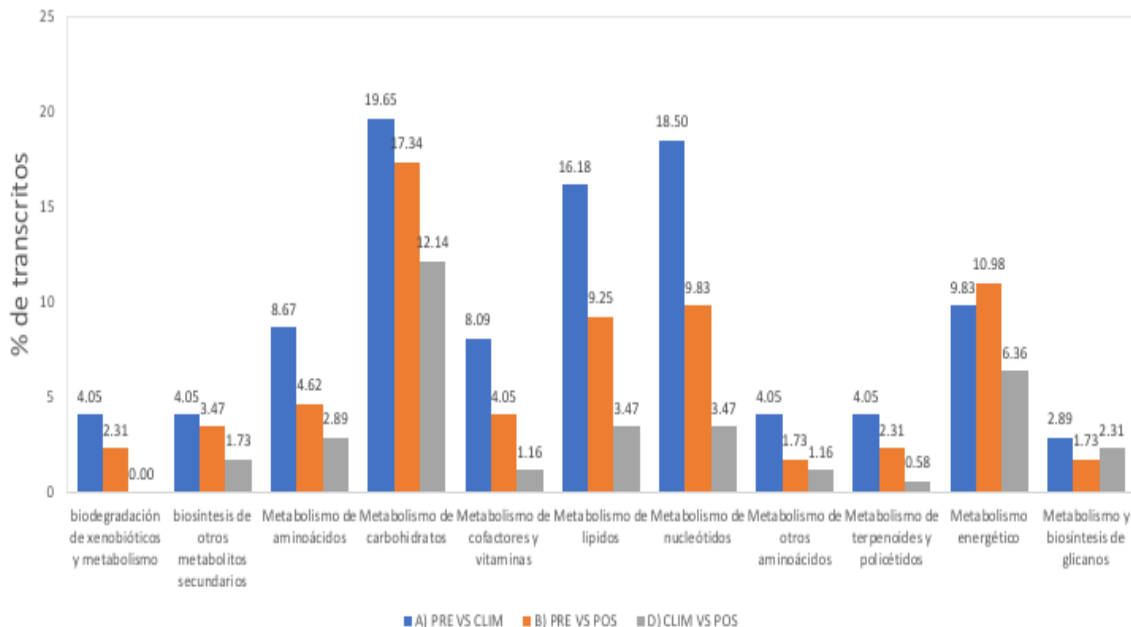




2.4.3 La anotación funcional identificó un programa de expresión génica específico para cada estado fisiológico

Para identificar los procesos metabólicos asociados a cada estado de maduración, se anotó los transcritos diferencialmente expresados utilizando la base de datos KEGG. Al realizar comparaciones pareadas entre el número de transcritos con expresión diferencial para los diferentes estados de maduración poscosecha, se encontró que los procesos metabólicos son los más afectados (Figura 2-4). Dentro de los tipos de procesos metabólicos que se ven influenciados por cambios en el número de transcritos asociados a ellos según el estado de maduración se encuentran: metabolismo energético, de nucleótidos, de carbohidratos y lipídico (Figura 2-4).

Figura 2-4: Porcentaje de transcritos asociados a procesos metabólicos durante la maduración de frutos de aguacate cultivar Lorena.



En el metabolismo de carbohidratos, los procesos de glicólisis y gluconeogénesis están afectados por transcritos diferencialmente expresados que codifican para las enzimas glucosa 6 fosfato isomerasa, fructosa bisfosfato hidrolasa, piruvato quinasa y aldehído deshidrogenasa al inicio de poscosecha (preclimaterio) (Tabla 2-4). Al final de la poscosecha (posclimaterio) los transcritos encontrados como diferencialmente expresados solamente se relacionan con la enzima 6 fosfofructoquinasa 1. Para el proceso de fotosíntesis vinculado al metabolismo energético, existe actividad por la subunidad delta y b de las enzimas ATPasa transportadora de H⁺ tipo F e hidrolasas durante preclimaterio y nula actividad enzimática durante posclimaterio. En el metabolismo lipídico existe el mismo comportamiento al reportado por los otros dos procesos metabólicos mencionados anteriormente, en donde las enzimas FatB Acil-ACP Tioesterasa B, acil-ACP Tioesterasa B y delta 7 esterol 5 desaturasa actúan durante preclimaterio y en posclimaterio sólo hay presencia de la actividad de acetil CoA carboxilasa.

Tabla 2-4: Enzimas afectadas por expresión diferencial de genes obtenidas a partir de comparaciones pareadas entre estados poscosecha.

Metabolismo	Enzimas diferencialmente expresadas por estado de maduración por código KEGG			
	Pre vs Clim	Pre Vs Pos	Clim Vs Pos	
Carbohidratos	Glicolisis y gluconeogénesis	5.3.1.9 glucosa 6 fosfato isomerasa	2.7.1.11 6 fosfofructoquinasa 1	2.7.1.11 6 fosfofructoquinasa 1
		4.1.2.13 fructosa bisfosfato hidrolasa	4.1.2.13 fructosa bisfosfato hidrolasa	
		2.7.1.40 Piruvato quinasa	2.7.1.40 Piruvato quinasa	2.7.1.40 Piruvato quinasa
	Ciclo de citrato	1.2.1.3 Aldehido deshidrogenasa		
		1.1.1.37 Malato deshidrogenasa	1.1.1.37 Malato deshidrogenasa	2.3.3.8 ATP citrato
		6.2.1.4 Succinil CoA sintetasa sub unidad alfa		
	Pentosas fosfato	6.2.1.5 Succinil CoA sintetasa sub unidad alfa		
		5.3.1.9 glucosa 6 fosfato isomerasa	2.7.1.11 6 fosfofructoquinasa 1	2.7.1.11 6 fosfofructoquinasa 1
		4.1.2.13 fructosa bisfosfato hidrolasa	4.1.2.13 fructosa bisfosfato hidrolasa	3.1.1.31 6 fosfogluconolactonasa
	interconversión de pentosa y glucuronato		2.2.1.1 Transquetolasa	
				4.2.2.2 Pectato liasa
	Metabolismo de fructosa y manosa	4.1.2.13 fructosa bisfosfato hidrolasa	2.7.1.11 6 fosfofructoquinasa 1	2.7.1.11 6 fosfofructoquinasa 1
			4.1.2.13 fructosa bisfosfato hidrolasa	
	Metabolismo de galactosa	3.2.1.20 glucoamilasa maltasa	2.7.1.11 6 fosfofructoquinasa 1	2.7.1.11 6 fosfofructoquinasa 1
	Ascorbato y aldarato	1.2.1.3 Aldehido deshidrogenasa		
	Almidón y sacarosa	2.4.1.43 poligaracturonato 4-alfa-galacturonosiltransferasa	3.2.1.21 beta glucosidasa	5.1.3.6 glucuronato 4 epimerasa
		3.2.1.20 glucoamilasa maltasa	2.4.1.43 alfa 1,4-galacturonosiltransferasa	3.2.1.28 alfa alfa trehalasa
		3.2.1.21 beta glucosidasa		
		5.3.1.9 glucosa 6 fosfato isomerasa		
	metabolismo de aminoácidos y nucleótidos	2.4.1.21 almidón sintasa		
		5.3.1.9 glucosa 6 fosfato isomerasa	2.4.1.43 alfa 1,4-galacturonosiltransferasa	5.1.3.6 glucuronato 4 epimerasa
		2.4.1.43 poligaracturonato 4-alfa-galacturonosiltransferasa	1.6.2.2 citocromo b5 reductasa	
	Metabolismo de piruvato	1.6.2.2 Citocromo b5 reductasa		
		2.7.1.40 Piruvato quinasa	2.7.1.40 Piruvato quinasa	2.7.1.40 Piruvato quinasa
		1.1.1.37 Malato deshidrogenasa	1.1.1.37 Malato deshidrogenasa	6.4.1.2 acetil CoA carboxilasa
	Glicoxilato y dicarboxilato	1.2.1.3 Aldehido deshidrogenasa	6.4.1.2 acetil CoA carboxilasa	6.4.1.2 acetil CoA carboxilasa
		1.1.1.37 Malato deshidrogenasa	1.1.1.37 Malato deshidrogenasa	
Propanoato	6.2.1.4 Succinil CoA sintetasa sub unidad alfa			
	6.2.1.5 Succinil CoA sintetasa sub unidad alfa			
Fosfato inositol		2.7.1.68 1-fosfatidilinositol-4-fosfato quinasa	2.7.1.68 1-fosfatidilinositol-4-fosfato quinasa	
		2.7.8.11 CDP diacilglicerol-inositol 3-fosfatidiltransferasa		

Energético	fosforilación oxidativa	COX 17 subunidad 17 de la proteína de union al citocromo	1.6.5.3 deshidrogenasa	3.6.3.14 hidrolasa	
		g subunidad g de la ATPasa transportadora de H+ tipo F	ND2 NADH ubiquinona oxidoreductasa	3.6.3.6 ATPasa transportador de hidrogeniones	
		3.6.1.1 pirofosfatasa inorgánica	QCR7 Ubiquinol citocromo c reductasa subunidad 7	I ATPasa transportador subunidad I	
		3.6.3.14 hidrolasa	1.10.2.2 Ubiquinol citocromo c reductasa subunidad hierro sulfuro		
			3.6.3.14 hidrolasa		
	Fotosíntesis	delta subunidad delta de la ATPasa transportadora de H+ tipo F		3.6.1.1 pirofosfatasa inorgánica	
		delta subunidad b de la ATPasa transportadora de H+ tipo F		3.6.3.14 hidrolasa	
		3.6.3.14 hidrolasa	delta subunidad b de la ATPasa transportadora de H+ tipo F		
	proteínas de antena de la fotosíntesis		Lhca1 proteína 1 de unión entre clorofila a/b		
	Fijación de carbono	1.1.1.37 Malato deshidrogenasa	1.1.1.37 Malato deshidrogenasa	2.7.1.40 Piruvato quinasa	
		2.7.1.40 Piruvato quinasa	2.7.1.40 Piruvato quinasa		
		2.7.1.19 fosforibuloquinasa	2.7.1.19 fosforibuloquinasa		
		1.1.1.37 Malato deshidrogenasa	2.2.1.1 Transquetolasa		
		4.1.2.13 fructosa bisfosfato hidrolasa	4.1.2.13 fructosa bisfosfato hidrolasa		
	Metano	1.11.1.7 oxidoreductasa	1.11.1.7 oxidoreductasa	2.1.2.1 hidroximetiltransferasa glicina	
			2.7.1.11 6 fosfofructoquinasa 1	2.7.1.11 6 fosfofructoquinasa 1	
	Nitrógeno	1.7.1.3 nitrato reductasa	6.3.5.4 Asparagina sintasa	1.7.1.3 nitrato reductasa	
		6.3.5.4 Asparagina sintasa		1.4.1.13 sintasa glutamato	
			1.4.1.14 sintasa glutamato NADH		

Lipídico	Biosíntesis de ácidos grasos	FatB Acil-ACP Tioesterasa B	6.4.1.2 acetil CoA carboxilasa	6.4.1.2 acetil CoA carboxilasa
		3.1.2.14 Acil-ACP Tioesterasa B	1.14.19.2 Delta 7 esterol 5 desaturasa	
		1.14.19.2 Delta 7 esterol 5 desaturasa		
	Biosíntesis de esteroides	1.3.1.72 Delta 24 esterol reductasa	1.3.1.72 Delta 24 esterol reductasa	
		1.14.13.70 14 alfa esterol demetilasa	1.14.13.70 14 alfa esterol demetilasa	
		DFW1 delta 24 esterol reductasa	1.3.1.72 Delta 24 esterol reductasa	
		CYP51G1 alfa esterol demetilasa	CYP51G1 alfa esterol demetilasa	
			DWF1 Delta 24 esterol reductasa	
	Biosíntesis de la hormona esteroide	1.14.14.1 citocromo p 450 familia 1 subfamilia A	1.14.14.1 citocromo p 450 familia 1 subfamilia A	
	Glicerolipidos			2.3.1.51 1 acil-sn-glicerol-3-fosfato aciltransferasa
	Glicerofosfolipidos	1.2.1.3 Aldehido deshidrogenasa	2.3.1.15 glicerol 3 fosfato o aciltransferasa	2.3.1.51 1 acil-sn-glicerol-3-fosfato aciltransferasa
		2.3.1.15 glicerol 3 fosfato o aciltransferasa		
	lipidos eter	LPCAT1 lisofosfatidilcolina aciltransferasa		
		LPCAT2 lisofosfatidilcolina aciltransferasa		
		3.1.4.4 fosfolipasa		
ácido linoléico	1.14.14.1 citocromo p 450 familia 1 subfamilia A			
ácido linoléico alfa	5.3.99.6 oxido ciclasa aleno	5.3.99.6 oxido ciclasa aleno	4.2.1.92 hidroperoxido dehidratasa	
	OPCL1 CoA ligasa 1	OPCL1 CoA ligasa 1		
	MFP2 enoil CoA hidratasa			
	2.1.1.141 jasmonato o metil transferasa			

2.5 Discusión

El estudio transcriptómico de la poscosecha del fruto de aguacate cultivar Lorena permitió generar un transcriptoma *De novo* del fruto el cual se caracterizó por tener 22,749 transcritos. Las etapas poscosecha de fruto (preclimaterio, climaterio y posclimaterio) fueron caracterizadas por un total de 6335 transcritos diferencialmente expresados. Interesantemente, el estado climaterio y posclimaterio comparten un 90% del potencial de expresión génica, mientras que dichos estados eran alrededor de un 40% a 35% diferentes con relación al estado preclimaterio, respectivamente (Figura 2-3a). Desde el se debieron principalmente a procesos del metabolismo energético, de nucleótidos, de carbohidratos y lipídico (Figura 2-4).

Se ha reportado que la variabilidad en la actividad de enzimas como glucosa 6 fosfato isomerasa, fructosa bisfosfato hidrolasa, piruvato quinasa y aldehído deshidrogenasa (Taiz & Zeiger, 2010) pueden regular la conversión de almidón a acetil CoA, y por consiguiente los procesos de glicólisis y gluconeogénesis (Taiz & Zeiger, 2010). En los resultados obtenidos en el presente trabajo se observa que las enzimas glucosa 6 fosfato isomerasa, fructosa bisfosfato hidrolasa, piruvato quinasa y aldehído deshidrogenasa disminuyen su expresión a lo largo de la poscosecha del fruto. Esto sugiere entonces, que existe una mayor conversión de almidón a acetil CoA durante el estado preclimaterio con respecto al posclimaterio, disminuyendo los procesos de glicólisis y/o gluconeogénesis. Adicionalmente, el acetil CoA es promotor de la síntesis de ácidos grasos (Taiz & Zeiger, 2010). Por lo tanto, su aumento estimula la producción de ácidos grasos en el fruto, conllevando a un incremento de su concentración. Este aumento en la concentración de ácidos grasos a lo largo de la poscosecha es similar al reportado para frutos de aguacate Hass (Ferreira *et al.*, 2016).

Moléculas como la ATPasa transportadora de hidrogeniones tipo f, relacionadas con la síntesis de ATP en la membrana tilacoidal de cloroplastos (Kochegarov, 2001), muestran una actividad cada vez menor en el desarrollo del proceso de fotosíntesis por disminución del número de cloroplastos totales, lo que es perceptible por un aumento en el color amarillo de frutos durante su poscosecha. De igual manera, la fijación de CO₂,

evidencia una menor actividad conforme la maduración de frutos que se desarrolla en poscosecha, lo que es evidente por la sobre expresión de genes asociados con la síntesis de enzimas malato deshidrogenasa, piruvato quinasa, fosforibuloquinasa y fructosa bisfosfato hidrolasa; únicamente durante preclimaterio y climaterio (Tabla 2.4).

En concordancia con lo reportada por Ferreyra *et al.* (2016) para frutos de aguacate Hass, quienes manifiestan que los ácidos grasos son en su mayoría sintetizados cuando el fruto se encuentra en el árbol y una menor cantidad es producida en pocosecha. La sobre expresión de transcritos asociados a la síntesis de enzimas como Acil-ACP tioesterasa B que permiten el enlace de los ácidos grasos de la proteína transportadora acil (ACP) durante la síntesis de lípidos como el ácido palmítico (Pérez *et al.*, 2012) y la sobre expresión de transcritos que codifican para la enzima acetil CoA carboxilasa regulador de la síntesis de ácidos grasos y su oxidación (Knowles, 1989). Apoyan la afirmación de Ferreyra desde el punto de vista genético. Lo que indica que existe un remanente de ácidos grasos que son sintetizados en el fruto una vez este se encuentra fuera del árbol, siendo mayor su síntesis durante preclimaterio y menor para posclimaterio (Tabla 2-4) cómo se evidenció en la evaluación fisiológica de frutos del primer capítulo en donde el contenido de ácidos grasos, obtuvo un ligero incremento para el día 12 (posclimaterio) con respecto a los días tres (preclimaterio) y siete (climaterio) Tabla 1-5. Cómo también lo reporta Ferreyra *et al.* (2016) para el cultivar Hass de aguacate quién concibe que durante la poscosecha de estos materiales se sintetizan ácidos grasos en menor proporción con respecto al peso total del fruto.

2.6 Conclusiones

- De los 22,749 transcritos que definen el transcriptoma de frutos de aguacate Lorena, solamente 6,335 fueron biológica y estadísticamente significantes en su expresión diferencial asociada a los procesos de metabolismo de carbohidratos, metabolismo energético y metabolismo lipídico principalmente.
- Existe una diferenciación entre estados de madurez poscosecha a partir de los transcritos diferencialmente expresados. Siendo mayor la proporción de transcritos en preclimaterio con respecto a posclimaterio, ya que por medio de

comparaciones pareadas se encontró que existe 14,6% de genes sobre expresados en preclimaterio con respecto a climaterio, 10,7% en preclimaterio con respecto a posclimaterio y 3,8% entre climaterio versus posclimaterio.

- Los frutos de aguacate cultivar Lorena expresan genes asociados a síntesis de ácidos grasos durante toda su poscosecha lo que permite un aumento en el contenido de ácidos grasos no solamente mientras los frutos se desarrollan en el árbol, sino también una vez son colectados.
- Este trabajo reporta por primera vez el ensamblaje de un transcrito *De novo* sobre la raza antillana de aguacate. Así como también, una descripción de los rasgos fisiológicos que son afectados durante su madurez poscosecha y los genes diferencialmente expresados que contribuyen con algunos de estos procesos.

2.7 Bibliografía

- Chanderbali, A. S., Albert, V. A., Ashworth, V. E. T. M., Clegg, M. T., Litz, R. E., Soltis, D. E., & Soltis, P. S. (2008). *Persea americana* (avocado): Bringing ancient flowers to fruit in the genomics era. *BioEssays*, 30(4), 386–396. <https://doi.org/10.1002/bies.20721>
- Chen, H., Morrell, P. L., Ashworth, V. E. T. M., De La Cruz, M., & Clegg, M. T. (2009). Tracing the geographic origins of major avocado cultivars. *Journal of Heredity*, 100(1), 56–65. <https://doi.org/10.1093/jhered/esn068>
- Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, 30(1), 31–37. <https://doi.org/10.1093/bioinformatics/btt310>
- Danchin, A., Ouzounis, C., Tokuyasu, T., & Zucker, J.-D. (2018). No wisdom in the crowd: genome annotation in the era of big data - current status and future prospects. *Microbial Biotechnology*, 11(4), 588–605. <https://doi.org/10.1111/1751-7915.13284>
- Feng, K., Costa, J., & Edwards, J. S. (2018). Next-generation sequencing library construction on a surface. *BMC Genomics*, 19(1), 416. <https://doi.org/10.1186/s12864-018-4797-4>
- Ferreira, R., Sellés, G., Saavedra, J., Ortiz, J., Zúñiga, C., Troncoso, C., ... De, B. G. (2016). Identification of pre-harvest factors that affect fatty acid profiles of avocado fruit (*Persea americana* Mill) cv. 'Hass' at harvest, 104, 15–20.

- <https://doi.org/10.1016/j.sajb.2015.10.006>
- Hurtado, U. A. (2014). Análisis transcriptómico de la respuesta de defensa de *Microcyclus Hevea brasiliensis* frente a *Microcyclus ulei* bajo condiciones controladas, 121.
- Ibarra-Laclette, E., Méndez-Bravo, A., Pérez-Torres, C. A., Albert, V. A., Mockaitis, K., Kilaru, A., ... Herrera-Estrella, L. (2015). Deep sequencing of the Mexican avocado transcriptome, an ancient angiosperm with a high content of fatty acids. *BMC Genomics*, 16(1), 599. <https://doi.org/10.1186/s12864-015-1775-y>
- Kilaru, A., Cao, X., Dabbs, P. B., Sung, H.-J., Rahman, M. M., Thrower, N., ... Ohlrogge, J. B. (2015). Oil biosynthesis in a basal angiosperm: transcriptome analysis of *Persea Americana* mesocarp. *BMC Plant Biology*, 15, 203.
- Knowles, J. R. (1989). Biotin-Dependent Enzymes, 195–221. <https://doi.org/10.1186/s12870-015-0586-2>
- Kochegarov, A. A. (2001). Modulators of ion-transporting ATPases. *Expert Opinion on Therapeutic Patents*, 11(5), 825–859. <https://doi.org/10.1517/13543776.11.5.825>
- Leek, J., Johnson, W. E., Jaffe, A., Parker, H., & Storey, J. (2011). Supplementary Analysis For : The sva package for removing batch effects and other unwanted variation in high-throughput experiments Setting up the data from an ExpressionSet, 1–11.
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA - Seq data with or without a reference genome. *BMC Bioinformatics*, 12(323), 1–16. <https://doi.org/10.1186/1471-2105-12-323>
- Li, B., Fillmore, N., Bai, Y., Collins, M., Thomson, J. A., Stewart, R., & Dewey, C. N. (2014). Evaluation of de novo transcriptome assemblies from RNA-Seq data. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0553-5>
- Lopez-Gomez, R., Torres-Cardenas, Y., Chavez-Mochezuma, M., Salgado-Garciglia, R., Jimenez-Moraila, B., Corona-Armenta, G., & Herrea-Estrella, L. (2007). Genomics of Avocado Criollo Fruit (*Persea americana* Mill. VAR. *Drymifolia*). *Proceedings VI World Avocado Congress*, (978).
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nature Reviews Genetics*, 12(10), 671–682. <https://doi.org/10.1038/nrg3068>
- Masoudi-Nejad, a, Narimani, Z., & Hosseinkhan, N. (2013). *Next generation sequencing and*

- sequence assembly: methodologies and algorithms*. <https://doi.org/10.1007/978-1-4614-7726-6>
- Páez, U. A. H., Romero, I. A. G., Restrepo, S. R., Gutiérrez, F. A. A., & Castaño, D. M. (2015). Assembly and analysis of differential transcriptome responses of *Hevea brasiliensis* on interaction with *Microcyclus ulei*. *PLoS ONE*, *10*(8), 1–22. <https://doi.org/10.1371/journal.pone.0134837>
- Pérez, E., Otero, M., & Alvarez, A. (2012). producción de biodiesel a partir de microorganismos oleaginosos. Una fuente de energía renovable. *Redalyc*, *46*(3), 26–35. <https://doi.org/10.1002/wea.23>
- Sánchez Santana, S. del C. (2015). Análisis de datos de RNA-Seq comparación de métodos para el estudio de expresión génica diferencial. Retrieved from <https://idus.us.es/xmlui/handle/11441/40809>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Smith-Unna, R., Bournsnel, C., Patro, R., Hibberd, J. M., & Kelly, S. (2016). TransRate: Reference-free quality assessment of de novo transcriptome assemblies. *Genome Research*, *26*(8), 1134–1144. <https://doi.org/10.1101/gr.196469.115>
- Taiz, L., & Zeiger, E. (2010). Plant Physiology, Fifth Edition. *Cell*, *1*, 782. Retrieved from <http://www.amazon.com/dp/0878938567>
- Tjaden, B. (2015). De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biology*, *16*(1), 1–10. <https://doi.org/10.1186/s13059-014-0572-2>
- Varet, H., Brillet-Guéguen, L., Coppée, J. Y., & Dillies, M. A. (2016). SARTools: A DESeq2- and edgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS ONE*, *11*(6), 1–9. <https://doi.org/10.1371/journal.pone.0157022>
- Wang, S., & Gribskov, M. (2017). Comprehensive evaluation of de novo transcriptome assembly programs and their effects on differential gene expression analysis. *Bioinformatics (Oxford, England)*, *33*(3), 327–333. <https://doi.org/10.1093/bioinformatics/btw625>
- Wolstenholme, N., & Whiley, A. (2013). *The avocado botany, production and uses*. (B. Schaffer, Ed.) (2da ed.). Croydon. Retrieved from www.cabi.org
- Xu, W. J., Wang, Z. X., & Qiao, Z. D. (2009). Modified PCR methods for 3' end amplification from serial analysis of gene expression (SAGE) tags. *FEBS Journal*, *276*(10), 2657–

2668. <https://doi.org/10.1111/j.1742-4658.2009.06981.x>

2.8 Anexos

El ARN total extraído, mediante el protocolo reportado por Chang *et al.* (1993), fue suficiente en concentración, calidad e integridad (Tabla 2-7 ; Figura 2-5), para la generación de librerías con calidad para producir un alto volumen de secuencias a través del equipo NovaSeq 6000 Tabla 2-3. (anexos)

Tabla 2-6: Concentración y relación 260/280 y 260/230 para ARN total extraído por muestra con duplicado.

Sample ID	Nucleic Acid Conc.	Unit	260/280	260/230	Sample Type
Pre1	734.6	ng/ μl	2.08	2.05	RNA
Pre2	1038.8	ng/ μl	2.09	2.19	RNA
Pre3	1555	ng/ μl	2.11	2.18	RNA
Cli1	2908.1	ng/ μl	2.11	2.23	RNA
Cli2	765.1	ng/ μl	2.11	2.23	RNA
Cli3	1206.6	ng/ μl	2.1	2.3	RNA
Pos1	1522.2	ng/ μl	2.1	2.18	RNA
Pos2	993.7	ng/ μl	2.11	2.21	RNA
Pos3	2112.4	ng/ μl	2.07	2.15	RNA
Pre1a	1731.9	ng/ μl	2.1	2.16	RNA
Pre2a	1008.1	ng/ μl	2.16	2.21	RNA
Pre3a	1925.5	ng/ μl	2.1	2.14	RNA
Cli1a	1907.5	ng/ μl	2.09	2.23	RNA
Cli2a	1040.6	ng/ μl	2.11	2.24	RNA
Cli3a	1038.1	ng/ μl	2.13	2.25	RNA
Pos1a	1244.3	ng/ μl	2.12	2.21	RNA
Pos2a	1799.3	ng/ μl	2.11	2.21	RNA
Pos3a	2375.4	ng/ μl	2.02	2.03	RNA

Figura 2-5: Gel de agarosa al 1% con muestras de ARN total sin tratamiento con DNAsa proveniente de pulpa de fruta de aguacate con un tiempo de corrida de 20 minutos y 100 voltios.

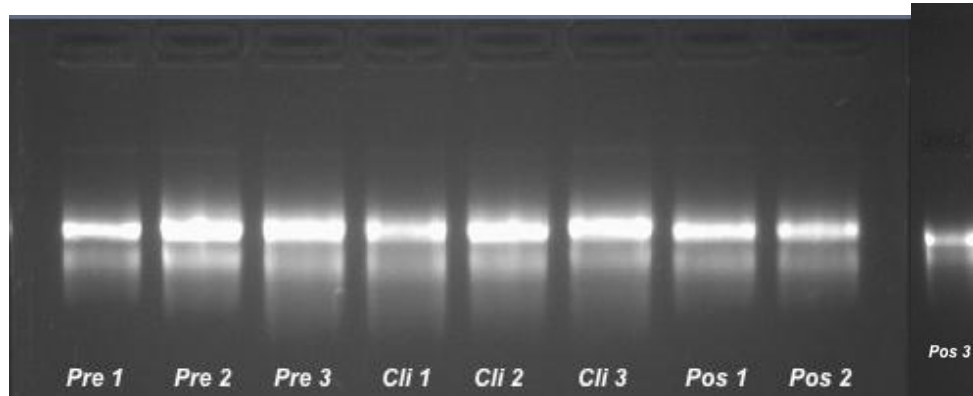


Figura 2-6: Reporte generado por FastQC. Figura de la izquierda muestra los datos crudos obtenidos del proceso de secuenciación, Figura de la derecha los datos procesados luego del uso de trimmomatic.

