



Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

**ANÁLISIS DE LOS GENES LICOPENO CICLASA (LCY)
DEL GENOMA DE *Bixa orellana* L.**

Tesis que presenta

Julia Gabriela Rivero Manzanilla

En opción al título de

DOCTOR EN CIENCIAS

(Ciencias Biológicas: Opción Bioquímica y Biología Molecular)

Mérida, Yucatán, México

2024

CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.
POSGRADO EN CIENCIAS BIOLÓGICAS



RECONOCIMIENTO

Por medio de la presente, hago constar que el trabajo de tesis de Julia Gabriela Rivero Manzanilla titulado **Análisis de los genes licopeno ciclasa (LCY) del genoma de *Bixa orellana* L.**, fue realizado en la Unidad de Biología Integrativa, en la línea de investigación Metabolismo Integrativo, en el laboratorio 27 del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección de la Dra. Renata Lourdes Barbara Rivera Madrid dentro de la opción de Bioquímica y Biología Molecular, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro y la Codirección del Dr. José Alberto Narváez Zapata del Centro de Biotecnología Genómica.

Atentamente



Dra. Cecilia Hernández Zepeda
Directora de Docencia

Mérida, Yucatán, México, a 11 de abril de 2024

DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en la sección de Materiales y Métodos Experimentales, los Resultados y Discusión de este documento proviene de las actividades de experimentación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que a razón de lo anterior y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información, le pertenecen patrimonialmente al Centro de Investigación Científica, A.C., y en el mismo tenor, reconozco que si derivaran de este trabajo productos intelectuales o desarrollos tecnológicos, en lo especial, estos se registrarán en todo caso por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.



Julia Gabriela Rivero Manzanilla

Este trabajo se llevó a cabo en la Unidad de Biología Integrativa del Centro de Investigación Científica de Yucatán, A.C., y forma parte del proyecto titulado “Análisis de los genes *licopeno ciclasa* (LCY) del genoma de *Bixa orellana* L.” bajo la dirección de la Dra. Renata Rivera Madrid y el Dr. José Alberto Narváez Zapata.

AGRADECIMIENTOS

Mi sincero agradecimiento y aprecio a mi directora de tesis, la Dra. Renata Rivera Madrid por abrirme las puertas de su laboratorio, hacerme parte de su equipo de trabajo y permitirme desarrollarme como doctorante. Así como a mi coasesor de tesis, el Dr. José Alberto Narváez Zapata, por su asesoría, conocimientos y apoyo durante todo el proceso de la tesis.

A la M. C. Margarita Aguilar por ser parte esencial del desarrollo de todas y todos los estudiantes del laboratorio 27, por su apoyo, guía y enseñanzas durante esta tesis.

Al Centro de Investigación Científica de Yucatán, A. C. (CICY), por brindarme el espacio donde me desarrollé mi proceso de doctorado, al personal y especialmente a la Unidad de Biología Integrativa donde me desempeñé estos 4 años de posgrado.

Al CONAHCYT por la beca otorgada (No. CVU/Becario): 460371/760051, para la realización del doctorado en Ciencias Biológicas opción Bioquímica y Biología Molecular de Plantas en la Unidad de Biología Integrativa del CICY.

A mi comité tutorial: la Dra. Luisa Alhucema López Ochoa y Dr. Jorge Humberto Ramírez Prado por sus observaciones y consejos durante el desarrollo de esta tesis; al Dr. Felipe Vazquez Flota y Dr. Hugo Antonio Tovar Romero por sus comentarios y aportes en la versión final de la tesis, así como a la Dra. Nancy Santana Buzzy por aceptar ser parte de mi comité lector.

A mis compañeras y compañeros de laboratorio: Diana Sierra, por su apoyo, comprensión y amistad; Lupita Hernández por su compañía y guía al inicio del doctorado; Rocío Tamayo por sus consejos histológicos; Diego Torres por confiar en mí y orientarle en bioinformática; Ana Gurgua por ser compañera en bioinformática y al Dr. Rodolfo Pech por brindar carisma al laboratorio. A Gaby Vázquez por haber compartido conmigo el camino del doctorado; a Neimi Estrella por su apoyo en técnicas de laboratorio; Al Dr. Víctor Carballo por su apoyo en técnicas de RT-PCR y al Dr. Pedro Escobar por sus asesorías y apoyo con el programa R.

A Sanry Marrufo, mi más sincero agradecimiento por su invaluable guía a mi persona y mi familia durante este proceso de crecimiento y aprendizaje. Infinitas gracias.

DEDICATORIAS

A mi madre por ser mi compañera y fuerza en esta vida. Gracias por creer en mí.

A mi padre, por su apoyo, cariño y esfuerzo.

A mi hermanita Karla, por ser mi inspiración.

A mi hermanito Ulises, por ser un guardián silencioso.

A mi esposo Efraín, por ser mi mejor amigo, mi cómplice, mi hogar.

Los amo inmensamente.

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO I	3
ANTECEDENTES	3
1. <i>BIXA ORELLANA</i> L..	3
1.1. SISTEMÁTICA	3
1.2. DISTRIBUCIÓN Y DESCRIPCIÓN	3
1.3. USOS Y CULTIVO.	4
1.4. IMPORTANCIA CULTURAL, BIOLÓGICA Y ECONÓMICA	5
2. CAROTENOIDES	6
2.1. BIOSÍNTESIS DE CAROTENOIDES EN PLANTAS	7
2.2. RUTA DE BIOSÍNTESIS DE BIXINA	9
2.3. BIXINA	11
2.4. LICOPENO CICLASAS	12
3. GENOMA DE ACHIOTE	15
3.1. VARIACIÓN GENÉTICA	15
3.2. CITOGENÉTICA MOLECULAR	18
RECAPITULACIÓN	23
HIPÓTESIS	23
OBJETIVO GENERAL.	24
OBJETIVOS ESPECÍFICOS.	24
JUSTIFICACIÓN.	25

ESTRATEGIA EXPERIMENTAL	26
CAPÍTULO II	27
ANÁLISIS DE LOS GENES LICOPENO CICLASA (<i>Boβ-LCY1</i>, <i>Boβ-LCY2</i>, <i>Boε-LCY</i>) DEL GENOMA DE <i>Bixa orellana</i> L.	27
2.1. INTRODUCCIÓN	27
2.2. MATERIALES Y MÉTODOS	28
2.2.1. SECUENCIAS DE LOS GENES <i>BoLCY</i> Y SUS REGIONES CONSERVADAS	28
2.2.2. ANÁLISIS DE LOS ELEMENTOS REGULATORIOS <i>CIS</i> DE LOS GENES <i>BoLCY</i>	29
2.2.3. ANÁLISIS DE MIRNAS DE LOS GENES <i>BoLCY</i>	29
2.2.4. SECUENCIAS REPETITIVAS DE DNA DE LOS GENES <i>BoLCY</i>	30
2.3. RESULTADOS	30
2.3.1 SECUENCIAS DE LOS GENES <i>LCY</i> DE ACHIOTE Y SUS REGIONES CONSERVADAS	30
2.3.2. FACTORES DE TRANSCRIPCIÓN PUTATIVOS DE LOS GENES <i>BoLCY</i>	36
2.3.3. MIRNAS DE LOS GENES <i>BoLCY</i>	41
2.3.4. SECUENCIAS REPETITIVAS DE DNA DE LOS GENES <i>BoLCY</i>	41
2.4 DISCUSIÓN	44
2.4.1 CARACTERIZACIÓN DE LOS GENES <i>LCY</i> DE ACHIOTE	44
2.4.2. ELEMENTOS REGULATORIOS <i>CIS</i> DE LOS GENES <i>BoLCY</i>	45
2.4.3. MIRNAS DE LOS GENES <i>BoLCY</i>	47
2.4.4. SECUENCIAS REPETITIVAS DE DNA DE LOS GENES <i>BoLCY</i>	48
CAPÍTULO III	50
METAANÁLISIS DE LOS GENES LICOPENO CICLASA (<i>Boβ-LCY1</i>, <i>Boβ-LCY2</i>, <i>Boε-LCY</i>) DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE <i>Bixa orellana</i> L.	50

3.1. INTRODUCCIÓN	50
3.2. MATERIALES Y MÉTODOS	52
3.2.1. IDENTIFICACIÓN DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE ACHIOTE	52
3.2.2. ANÁLISIS DE MIRNAS DE LOS GENES <i>BoLCY</i>	52
3.2.3. ANÁLISIS DE EXPRESIÓN DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE <i>B. ORELLANA</i>	52
3.3. RESULTADOS	52
3.3.1. SECUENCIAS DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE ACHIOTE	52
3.3.2. MIRNAS DE LOS GENES <i>BoLCY</i>	55
3.3.3. EXPRESIÓN DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE <i>B. ORELLANA</i>	57
3.4. DISCUSIÓN	58
3.4.1. SECUENCIAS DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE ACHIOTE	58
3.4.2. MIRNAS DE LOS GENES <i>BoLCY</i>	59
3.4.3. EXPRESIÓN DE LOS GENES <i>LCY</i> DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE <i>B. ORELLANA</i>	61
3.4.3.1. β <i>LCY1</i>	61
3.4.3.2. β <i>LCY2</i>	62
3.4.3.3. ϵ <i>LCY</i>	62
CAPÍTULO IV	64
ANÁLISIS CITOGENÉTICO DE <i>Bixa orellana</i> L.	64
4.1. INTRODUCCIÓN	64
4.2. MATERIALES Y MÉTODOS	65
4.2.1. MATERIAL VEGETAL	65
4.2.2. CITOGENÉTICA DE ACHIOTE	66

4.2.3. CARACTERIZACIÓN DE LOS CROMOSOMAS DE <i>BIXA ORELLANA</i>	66
4.3. RESULTADOS	67
4.3.1. CITOGENÉTICA DE ACHIOTE	67
4.3.2. CARIOTIPO E IDEOGRAMA DE <i>BIXA ORELLANA</i>	67
4.4 DISCUSIÓN	68
4.4.1. CITOGENÉTICA DE ACHIOTE	68
4.4.2. CARIOTIPO E IDEOGRAMA DE <i>BIXA ORELLANA L.</i>	70
CAPÍTULO V	72
DISCUSIÓN, CONCLUSIONES Y PERSPECTIVAS	72
5.1. DISCUSIÓN	72
5.1.1. ESTRUCTURA DE LOS GENES <i>BoLCY</i>	72
5.1.2. REGULACIÓN DE LOS GENES <i>BoLCY</i>	73
5.1.3. MICRORNAS DE LOS GENES <i>BoLCY</i>	75
5.1.4. SECUENCIAS REPETITIVAS DE DNA DE LOS GENES <i>BoLCY</i>	76
5.1.5. EXPRESIÓN DE LOS GENES <i>LCY</i> DE TRES MORFOTIPOS DE <i>BIXA ORELLANA L.</i>	77
5.1.6. CITOGENÉTICA DE <i>BIXA ORELLANA L.</i>	78
5.2. CONCLUSIONES	79
5.3. PERSPECTIVAS	81
BIBLIOGRAFÍA	82
ANEXOS	103

ÍNDICE DE FIGURAS

Figura 1.1. Vía metabólica de la síntesis de carotenoides en plantas superiores	8
Figura 1.2. Modelo de regulación génica en la biosíntesis de bixina	10
Figura 1.3. Estructura de bixina (C ₂₅ H ₃₀ O ₄)	11
Figura 1.4. Ciclación del licopeno por las enzimas β-ciclasa y ε-ciclasa	13
Figura 1.5. Cariotipo e ideograma de <i>Bixa orellana</i> L.	21
Figura 1.6 Mapas genéticos parentales e integrados de <i>Bixa orellana</i>	22
Figura 1.7 Estrategia experimental para el análisis de los genes licopeno ciclasa en <i>Bixa orellana</i> L.	26
Figura 2.1. Alineamiento de secuencias de nucleótidos y aminoácidos de los genes <i>Boβ-LCY1</i> , <i>Boβ-LCY2</i> y <i>ε-LCY</i>	33
Figura 2.2. Alineamiento de regiones características de la beta licopeno ciclasa	35
Figura 2.3. Mapa de calor de categorías de los elementos <i>cis</i> regulatorios de los genes <i>LCY</i> de <i>Bixa orellana</i>	36
Figura 2.4. Mapa de calor de las categorías principales de elementos <i>cis</i> y sus factores de transcripción putativos asociados	37
Figura 2.5. Estructura de los genes <i>BoLCY</i> en Ugene	38
Figura 2.6. Sitios de unión de los factores de transcripción putativos para los genes <i>BoLCY</i> , 1,500 pb río arriba	40
Figura 2.7. Secuencias miRNAs identificadas para los genes <i>BoLCY</i> , 1,000 pb río arriba	40
Figura 2.8. Motivos LTR y dominios a proteína de los genes <i>BoLCY</i>	42
Figura 2.9. Estructura de los genes <i>LCY</i> de achiote	43
Figura 3.1. Alineamiento de secuencias de nucleótidos obtenidas de las bases de datos del	

genoma y transcriptoma de los genes *Boβ-LCY1*, *Boβ-LCY2* y *ε-LCY* 53

Figura 3.2. Expresión relativa de los genes *BoLCY* del transcriptoma de tres morfotipos de *Bixa orellana* 59

Figura 3.3. Modelo de expresión de las *BoLCY* de *B. orellana* en tejidos fotosintéticos (hojas) y no fotosintéticos (estadios de desarrollo de la semilla) 63

Figura 4.1. Fases mitóticas de achioté, morfotipo B (flor rosa) 67

Figura 4.2. Ideograma somático de *Bixa orellana*, morfotipo B (flor rosa) 69

ÍNDICE DE CUADROS

Cuadro 1.1. Tres morfotipos caracterizados de la especie <i>Bixa orellana</i> L. en Yucatán	15
Cuadro 1.2. Registros de números cromosómicos para <i>Bixa orellana</i> L.	20
Cuadro 2.1. Secuencias de los genes <i>BoLCY</i> obtenidas del genoma de <i>Bixa orellana</i> , morfotipo B	30
Cuadro 2.2. Estructura de los genes <i>LCY</i> de <i>B. orellana</i>	32
Cuadro 3.1. microRNAs transcritos de los genes <i>BoLCY</i> de achiote	56
Cuadro 3.2. miRNAs conservados de tomate en transcriptomas de Achiote	57
Cuadro 4.1. Longitud de los pares de cromosomas de <i>B. orellana</i>	68

ABREVIATURAS

μm	Micrómetro
2n	Número diploide de los cromosomas
A.N.	Apertura numérica
ABA	Ácido abscísico
Ag-NOR	Tinción de plata para las regiones organizadores nucleolares
APE	APETALA, factor de transcripción
AR	Relación del brazo
BoALDH	Aldehído deshidrogenasa de <i>Bixa orellana</i>
BoALDH	Aldehído deshidrogenasas de <i>Bixa orellana</i>
BoBADH	Deshidrogenasa de <i>Bixa orellana</i>
BoBMT	Norbixina metiltransferasa de <i>Bixa orellana</i>
BoCCD	Carotenoides dioxigenasas de <i>Bixa orellana</i>
BoCCD	Dioxigenasa de escisión de caroteno de <i>Bixa orellana</i>
BoLCD	Licopeno dioxigenasa de <i>Bixa orellana</i>
BoLCY	Licopeno ciclasa de <i>Bixa orellana</i>
BoMT	Metiltransferasas de <i>Bixa orellana</i>
BoSABATH	Metiltransferasas de <i>Bixa orellana</i>
Boβ-LCY1	Licopeno beta ciclasa 1 de <i>Bixa orellana</i>
Boβ-LCY2	Licopeno beta ciclasa 2 de <i>Bixa orellana</i>
Boε-LCY	Licopeno épsilon ciclasa de <i>Bixa orellana</i>
C	Tinción de la heterocromatina constitutiva
cDNA	ADN complementario
CL	Longitud del cromosoma
CRTISO	Caroteno <i>cis-trans</i> isomerasa

DMAPP	Dimetilalil difosfato
DNA	Ácido desoxirribonucleico
DXS2	1-Desoxi-D-xilulosa-5-fosfato sintasa 2
FAO	Food and Agriculture Organization
GADP	Gliceraldehido 3 fosfato
GGDS	Geranilgeranil difosfato sintasa
GGPP	Geranilgeranil difosfato sintasa
GL	Grupo de enlace
HCl 1N	Ácido clorhídrico
IPP	Isopentinil pirofosfato
L	Longitud de brazo largo
LCY	Licopeno ciclasa
LTR	Retrotransposones de repetición de terminal larga
m	Metacéntrico
MEP	Vía del metileritritol fosfato (2C-metil-D-eritrol 4-fosfato)
miRNA	Micro ácido ribonucleico
n	Número haploide de los cromosomas
NCBI	National Center for Biotechnology Information
NGS	Secuenciación de nueva generación
PacBio	Pacific Bioscience
PDS	Fitoeno desaturasa
Pg	Picogramos
PYS	Fitoeno sintasa
S	Longitud de brazo corto
SINAFERI	Sistema Nacional de Recursos Fitogenéticos para la Alimentación y la Agricultura
SL	Estrigolactonas

sm	Submetacéntrico
SNP	Polimorfismo de un solo nucleótido
SRAP	Polimorfismo amplificado de secuencias relacionadas
SSCP	Polimorfismo de conformación de cadena simple
TF	Factor de transcripción
UBI	Unidad de Biología Integrativa
URT	Región no traducida
Z-ISO	15-cis- ζ -caroteno isomerasa
<i>β-LCY1</i>	Licopeno beta ciclasa 1
<i>β-LCY2</i>	Licopeno beta ciclasa 2
<i>ϵ-LCY</i>	Licopeno épsilon ciclasa

RESUMEN

Bixa orellana L. es una especie nativa del neotrópico, domesticada y ampliamente distribuida, conocida como achiote. Se caracteriza por presentar una tonalidad naranja-rojiza en sus semillas debido al compuesto bixina; un apocarotenoide que representa más del 80% de los pigmentos presentes en la semilla, con el cual se elaboran colorantes para la industria de alimentos, farmacéutica y cosmética. Actualmente, su producción se ha impulsado debido a la sustitución de los colorantes sintéticos por los naturales. Sin embargo, la especie presenta características que afectan su producción a mayor escala. Al ser de polinización cruzada, la especie es altamente heterocigota, generando fenotipos contrastantes en flores, frutos y producción de bixina. Se ha documentado que los genes licopeno ciclasa son importantes marcadores moleculares, al reportar una asociación fenotipo-genotipo entre contenido de bixina y SNPs funcionales del gen *Boβ-LCY1*. En este contexto, el análisis de los genes *licopeno ciclasas* (*Boβ-LYC1*, *Boβ-LYC2*, *Boε-LYC*) del proyecto de secuenciación del genoma de achiote (aún no publicado) así como del proyecto del transcriptoma de tres morfotipos con diferente producción de bixina; apoyaría a la elucidación de la variación intraespecífica de estos genes, mediante el estudio de su localización, estructura y posibles secuencias regulatorias que podrían aportar información sobre su relación con la producción de bixina en la especie para futuros trabajos de mejoramiento genético.

ABSTRACT

Bixa orellana L. is a native, domesticated and widely distributed species in the Neotropics known as achiote. It is characterized by an orange-reddish hue of its seeds due to the compound bixin; an apocarotenoid that represents more than 80% of the pigments in the seed, which are used to prepare colorants for the food, pharmaceutical and cosmetic industries. Currently, its production has been boosted due to the substitution of synthetic dyes for natural ones. However, the species presents characteristics that affect its production on a larger scale; being cross-pollinated, it is highly heterozygous, generating contrasting phenotypes in flowers, fruits and bixin production. Lycopene cyclase genes have been documented to be important molecular markers, reporting a phenotype-genotype association between bixin content and functional SNPs of the *Boβ-LCY1* gene. In this context, the analysis of the lycopene cyclase genes (*Boβ-LYC1*, *Boβ-LYC2*, *Boε-LYC*) from the annatto genome sequencing project (not yet published), as well as from the transcriptome project of three morphotypes with different production of bixin; would support the elucidation of the intraspecific variation of these genes, by studying their location, structure and possible regulatory sequences that could provide information on their relationship with bixin production in the species for future genetic improvement work.

INTRODUCCIÓN

Bixa orellana L. es una especie nativa del neotrópico, domesticada y ampliamente distribuida, popularmente conocida como achiote (Avedaño-Arrazate *et al.*, 2018; Pinzón López *et al.*, 2018; Cárdenas Conejo *et al.*, 2015; Rivera Madrid *et al.*, 2006; Arce, 1999). La especie se caracteriza por la tonalidad naranja-rojiza de sus semillas debido al compuesto bixina; un apocarotenoide que representa más del 80% de los pigmentos presentes en la semilla (Avedaño-Arrazate *et al.*, 2018; Pérez *et al.*, 2003; Arce, 1999), con el cual se elaboran colorantes para la industria de alimentos, farmacéutica y cosmética. Actualmente, se ha impulsado su producción a mayor escala debido a la sustitución de los colorantes sintéticos por los naturales (Carballo-Uicab *et al.*, 2019; Pinzón López *et al.*, 2018; Rivera Madrid *et al.*, 2016; Cárdenas Conejo *et al.*, 2015; Leal y Michelangeli, 2010; Valdez Ojeda *et al.*, 2008; Rivera Madrid *et al.*, 2006; Pérez *et al.*, 2003), por lo que la especie cuenta con un gran potencial para su explotación industrial.

En México, se encuentra distribuida principalmente en el sureste del país (Avedaño-Arrazate *et al.*, 2018), principalmente el estado de Yucatán que cuenta con las condiciones agroecológicas, culturales y de demanda apropiadas para el cultivo de achiote (Pérez *et al.*, 2003). Sin embargo, la especie presenta algunas características que afectan su producción a mayor escala; al ser el achiote de polinización cruzada, lo hace altamente heterocigoto, generando fenotipos contrastantes en el color de sus flores, características en frutos y, por consiguiente, en la calidad y cantidad de producción del pigmento entre sus poblaciones (Romero *et al.*, 2019; da Silva *et al.*, 2018; Pech-Hoil *et al.*, 2017; Trujillo-Hdz., *et al.*, 2016; Rivera-Madrid *et al.*, 2006).

Al ser *Bixa orellana* un cultivo con gran potencial para su explotación agroindustrial; el equipo de trabajo de la Dra. Rivera Madrid se ha dedicado a implementar estudios bioquímicos, genético-moleculares y actualmente, técnicas de secuenciación de última generación, para identificar genotipos con caracteres deseables, con el fin del desarrollo y diseño de programas de mejoramiento genético para la especie. Sus recientes estudios en genes implicados en la ruta de los carotenoides, han revelado a los genes licopeno ciclasa como importantes marcadores moleculares para direccionar la búsqueda de estos genotipos, al reportar una asociación fenotipo-genotipo entre el carácter deseable, contenido de bixina y SNPs (polimorfismo de un solo nucleótido) funcionales del gen licopeno beta ciclasa (*Boβ-LCY1*). En este contexto, el presente trabajo se centrará en el análisis de los genes *licopeno ciclasas* (*Boβ-LYC1*, *Boβ-LYC2*, *Boε-LYC*) del proyecto de secuenciación del genoma de achiote (aún no publicado), así como el

transcriptoma de tres morfotipos con diferente producción de bixina, nos ayudaría a dilucidar la variación intraespecífica; mediante el estudio de la estructura y sus elementos regulatorios en el genoma de la especie, con el fin de generar información sobre su relación con la producción de bixina para futuros trabajos de mejoramiento genético.

CAPÍTULO I

ANTECEDENTES

1. *Bixa orellana* L.

1.1. Sistemática

Bixa orellana es una especie nativa y domesticada en el neotrópico, posiblemente del suroeste de la Amazonia, popularmente conocida como achiote (Avedaño-Arrazate *et al.*, 2018; Pinzón López *et al.*, 2018; Cárdenas Conejo *et al.*, 2015; Rivera Madrid *et al.*, 2006; Arce, 1999). Esta planta corresponde a la división de Angiospermas, Clase Magnoliopsida, Subclase Ranunculidae, Orden Malvales, pertenece a la familia Bixaceae, la cual se caracteriza por contener tanto en hojas como en el tallo una savia rojiza, siendo *Bixa* el único miembro de este género y *B. orellana* la única especie cultivada (Avedaño-Arrazate *et al.*, 2018; Pinzón López *et al.*, 2018; Moreira *et al.*, 2015; Akshatha *et al.*, 2011; Leal y Michelangeli, 2010). A pesar de ser domesticada en el neotrópico, aún se desconoce la población silvestre de la cual proviene, sin embargo, se tiene evidencia que podría tratarse de *Bixa orellana* var. *urucurana* Willd, (Dequigiovanni *et al.*, 2018; Rivera Madrid *et al.*, 2016; Moreira *et al.*, 2015); por esto, diversos autores mencionan como centro de origen la Amazonía brasileña, en donde se encuentran poblaciones contienen la mayor diversidad de la especie (Gomes-Pacheco *et al.*, 2019; Moreira *et al.*, 2015; Clement *et al.*, 2010).

1.2. Distribución y descripción

Bixa orellana está ampliamente distribuida en el neotrópico, desde México hasta Ecuador, Brasil y Bolivia (Avedaño-Arrazate *et al.*, 2018). México, es considerado un centro de domesticación de la especie (Acevedo *et al.*, 2009) y se encuentra en el sureste del país, en el trópico húmedo y subhúmedo de Oaxaca, Tabasco, Chiapas, Veracruz y la península de Yucatán. Se distribuye entre los 100 y 1000 msnm, a veces hasta los 1500 de altitud formando parte del bosque tropical caducifolio, subcaducifolio y perennifolio (SIRE, 2003). El achiote es un arbusto o arbolillo perenne de entre 3 a 8 m de altura con un tronco oscuro que se ramifica a distancia corta del suelo formando una copa amplia de hojas simples y alternas, de haz membranosa color verde oscuro de unos 10-25cm de largo y 10-20cm de ancho, pecíolos de unos 6-12cm de largo. Tiene de entre 20-30 flores, hermafroditas, regulares y actinomorfas, en panículas de 5-10cm de largo y 4-8 cm de ancho; posee 5 sépalos color rojizos o blanquecinos de unos 8-10mm de largo y 5 pétalos con colores rosados a púrpura, blanco o amarillentos de unos 25mm de largo y 12mm de

ancho, anteras blancas con granos de polen pulverulento de 1mm de largo y 0.7mm de ancho. El fruto con forma tipo cápsula loculicisa, polispérmica, ovoidea a esferoidal de unos 30-50mm de largo y 25-40mm de ancho con (o sin espinas) de 5-15mm de largo de color verde, amarillo, rojizo o púrpura; cuando estos maduran adquieren un color marrón con o sin dehiscencia apical; con 30-60 semillas turbinadas de unos 5mm de largo y 4mm de ancho, la testa es pigmentada color naranja, roja o marrón (Leal y Michelangeli, 2010; Aliphath, 2009).

Las mejores condiciones para su crecimiento son regiones bajas (de 100 a 500 msnm) con temperatura de entre 25-35°C con una marcada temporada seca y una precipitación anual de 1000-1200 mm. Se puede dar en suelos pobres, pero bien drenados (Rivera-Madrid *et al.* 2006). Su propagación puede ser por semilla o por estacas; las plántulas deben de crecer a la sombra y ser trasplantadas a campo abierto para evitar plagas y enfermedades. La especie se cosecha a los 3 o 4 años de manera anual y su producción puede extenderse hasta 15 años. En condiciones normales, el rendimiento de la planta es de 500 a 600kg por hectárea que genera entre 50 a 60 kg del extracto de achiote (Aliphath, 2009).

1.3. Usos y cultivo

B. orellana posee una diversidad de usos tradicionales desde tiempos prehispánicos como tinturas decorativas, repelente de insectos, cicatrizante, así como colorante en sus comidas (Leal y Michelangeli, 2010). En la actualidad, sus diversos tejidos presentan una serie de propiedades documentadas como antiinflamatorio, antioxidantes, antibiótica, anticancerígenas, hipoglucémicas, gastrointestinal y antifúngico en la industria farmacológica (Zarza García *et al.*, 2017; Rivera Madrid *et al.*, 2016; Lourido y Martín, 2010).

Sin embargo, son las semillas las más apreciadas, ya que es cultivada principalmente para la extracción del pigmento colorante natural annatto (Giuliano *et al.*, 2003), del cual pueden extraerse pigmentos tanto solubles en agua (norbixina) como solubles en aceite (bixina) producido en la cubierta de sus semillas con solo variar el disolvente de extracción (da Silva *et al.*, 2018; Raddatz-Mota *et al.*, 2017; Lourido y Martín, 2010; Giuliano *et al.*, 2003). La bixina es el compuesto más cotizado, un apocarotenoide que representa más del 80% de los pigmentos presentes en la semilla (da Silva *et al.*, 2018; Avedaño-Arrazate *et al.*, 2018; Zarza García *et al.*, 2017; Chisté *et al.*, 2011; Akshatha *et al.*, 2011; Gómez-Ortiz *et al.*, 2010; Lourido y Martín, 2010; Pérez *et al.*, 2003; Pineda y Calderón, 2002; Arce, 1999).

En México, la Red de Achiote del programa SINAREFI (2018) lo reporta como un cultivo reducido, distribuido en 14 estados donde sobresalen la Península de Yucatán, Oaxaca, Veracruz, Guerrero y Chiapas. Yucatán es el estado donde se encuentra la mayor proporción de poblaciones de la especie (18.13%) (Pinzón López *et al.*, 2018), ya que cuenta con las condiciones agroecológicas, culturales y de demanda apropiadas para el cultivo de achiote (Pérez *et al.*, 2003). Se usa ampliamente en la cocina regional y, en consecuencia, se cultiva de manera extensiva en plantaciones comerciales, pero también se encuentra en milpas y huertos familiares (Pech-Hoil *et al.*, 2017; Valdez Ojeda *et al.*, 2008); presentando cultivos con marcadas diferencias morfológicas en el color de flores y características del fruto (Avedaño-Arrazate *et al.*, 2018).

1.4. Importancia cultural, biológica y económica

En México, especialmente en Yucatán existe un conocimiento tradicional de gran valor sobre la planta; forma parte de rituales religiosos, de la medicina tradicional, así como la coloración de productos y alimentos desde tiempos prehispánicos (Leal y Michelangeli, 2010). En la actualidad, es parte primordial de una gran variedad de platillos distintivos de la gastronomía yucateca (Pérez *et al.*, 2003); también sigue vigente el valor de sus propiedades medicinales por saberes mayas y resulta ser un cultivo valioso al controlar la erosión de los suelos y puede ser usado como cerco vivo en sistemas como los huertos familiares, donde puede significar un ingreso económico para las familias al manejarlo en menor escala entre sus demás cultivos frutales (Pinzón López *et al.*, 2018; Pérez *et al.*, 2003; Pech-Hoil, *et al.*, 2017).

No obstante, su mayor importancia económica recae en la producción del colorante *bixina*, un apocotenoide que se caracteriza por tener una tonalidad atractiva (naranja-rojiza) para la elaboración de colorantes en la industria de alimentos, farmacéutica y cosmética. Actualmente, su producción se ha impulsado a mayor escala debido a la sustitución de colorantes sintéticos por naturales, por lo cual, las industrias han optado por este pigmento con ausencia tóxica según la secretaría de salud en México (NOM-119-SSA1-1994) y organismos mundiales como la FAO (Food and Agriculture Organization) (Carballo-Uicab, 2019; Pinzón López *et al.*, 2018; Avedaño-Arrazate *et al.*, 2018; Rivera-Madrid *et al.*, 2016; Cárdenas-Conejo *et al.*, 2015; Raddatz-Mota *et al.*, 2017; Rodríguez *et al.*, 2014; Leal y Michelangeli, 2010; Valdez-Ojeda *et al.*, 2008; Rivera Madrid *et al.*, 2006; Agner *et al.*, 2004; Pérez *et al.*, 2003), por lo que la especie cuenta con un gran potencial para su explotación agroindustrial. Sin embargo, la especie presenta algunas características que afectan su producción a mayor escala. Al ser una planta de polinización

cruzada, muestra una alta heterocigocidad, generando una gran variación en la calidad y producción del pigmento bixina (Romero *et al.*, 2019; da Silva *et al.*, 2018; Pech-Hoil *et al.*, 2017; Akshatha *et al.*, 2011; Rivera-Madrid *et al.*, 2006); por lo que no se han logrado plantaciones con caracteres deseables y homogéneas para la especie.

2. Carotenoides

Los carotenoides son pigmentos de tonalidades amarillo, naranja y rojo presentes en todos los organismos fotosintéticos y en algunas bacterias y hongos, por lo que, están ampliamente distribuidos en la naturaleza (Aguilar-Espinosa *et al.*, 2020; Maoka, 2020; Escobar-Turriza, 2017; Cunningham y Gantt, 1998). Estos pigmentos son metabolitos secundarios que consisten de una cadena principal de hidrocarburos de 40 carbonos que consta de 8 unidades de isopreno (C_5H_8) (Alcaíno *et al.*, 2016; Rosas-Saavedra y Stange, 2016; Lu y Li, 2008), alternando cuerpos dobles y simples, lo que le confiere propiedades fotoquímicas (Maoka, 2020; Nisar *et al.*, 2016).

En las plantas, los carotenoides se sintetizan y acumulan en los plástidos fotosintéticos y no fotosintéticos, principalmente cloroplastos y cromoplastos (Maoka, 2020; Nisar *et al.*, 2006; Cunningham y Gantt, 1998). En los tejidos fotosintéticos, los pigmentos carotenoides son componentes esenciales de las membranas, recolectan energía lumínica y la transfieren a las clorofilas empleada durante la fotosíntesis (Maoka, 2020; Cunningham y Gantt, 1998). De igual manera, absorben el exceso de energía de las clorofilas y la cual es liberada por la vibración del cromóforo de polieno, parte de la molécula del carotenoide que determina las propiedades de absorción y captación de luz; siendo el responsable del color y la acción fotoprotectora a longitudes de onda entre los 400 y 600nm (Maoka, 2020). En los tejidos no fotosintéticos se encuentran en frutos, semillas, raíces y flores; donde funcionan como fotoprotectores, antioxidantes, proporcionando olores atrayentes y al ser precursores de hormonas vegetales (Maoka, 2020; Cunningham y Gantt, 1998). Por lo tanto, en las plantas son primordiales para procesos como la fotosíntesis en la absorción de luz (380 a 500nm), fotoprotección, antioxidantes, componentes estructurales de la membrana, y precursores de fitohormonas como ácido abscísico y estrigolactonas, involucradas en procesos de desarrollo y defensa a diversos factores de estrés abióticos y bióticos (Kössler *et al.*, 2021; Zhao *et al.*, 2020; Alcaíno *et al.*, 2016; Fraser y Bramley, 2004).

La diversidad de estos compuestos, con más de 750 estructuras químicas diferentes, se debe a las modificaciones posteriores como escisión oxidativa, ciclaciones, oxidación y desaturaciones

de los diferentes organismos a la cadena principal $C_{40}H_{56}$ (Alcaíno *et al.*, 2016; Lu y Li, 2008). En la escisión oxidativa, sea enzimática o no enzimática, se acorta la cadena principal C_{40} por la eliminación de fragmentos de uno o ambos extremos de la molécula carotenoide, generando productos llamados apocarotenoides. Estas moléculas tienen un valor comercial al aportar pigmentos, sabores y aromas en la industria de alimentos y cosmética (Nisar *et al.*, 2016). En las plantas, apocarotenoides como el ácido abscísico (ABA) y las estrigolactonas (SL), son importantes en la regulación del crecimiento y desarrollo, como señales de comunicación entre los organismos y respuesta al estrés biótico y abiótico (Dzib-Cauich *et al.*, 2020; Alcaíno *et al.*, 2016; Rosas-Saavedra y Stange, 2016). No obstante, los apocarotenoides también desempeñan un rol ecológico importante; como efecto indirecto, contribuyen al aroma y sabor de flores y frutos, así como una diversidad de colores atractivos para los insectos y animales para su polinización y dispersión de semillas (Alcaíno *et al.*, 2016; Rosas-Saavedra y Stange, 2016; Lu y Li, 2008; Cunningham y Gantt, 1998); por lo cual, son importantes en las plantas asegurando su reproducción (Aguilar-Espinosa *et al.*, 2020; Dzib-Cauich *et al.*, 2020). En los animales y seres humanos, son compuestos benéficos para la salud por sus propiedades antioxidantes. Por esta razón, los carotenoides tienen aplicación en la industria farmacológica, cumpliendo un papel importante en el cuerpo humano, como prevenir el cáncer, enfermedades cardiovasculares y mejorar de la inmunidad (Zhao *et al.*, 2020; Alcaíno *et al.*, 2016; Fraser y Bramley, 2004). Los beneficios en plantas y el ser humano, hacen del estudio del metabolismo de los carotenoides y los genes implicados, indispensable para lograr su comprensión.

2.1. Biosíntesis de carotenoides en plantas

Los carotenoides se sintetizan de *novo* en las plantas, donde, el primer paso para la síntesis, es la vía MEP (metileritritol fosfato) localizada en el plastidio para la formación de geranylgeranyl difosfato sintasa (GGPP) (Figura 1.1). Esta comienza cuando se induce al gen DXS2a para condensar los substratos iniciales, gliceraldehido 3 fosfato (GADP) y piruvato para producir isopentenil pirofosfato (IPP), un precursor universal de los isoprenoides (Carballo-Uicab, 2019; Cárdenas-Conejo *et al.*, 2015). Luego, la condensación de tres IPP y una molécula de dimetilalil difosfato (DMAPP) mediante la enzima geranylgeranyl difosfato sintasa (GGDS) genera geranylgeranyl difosfato sintasa (GGPP) (Alcaíno *et al.*, 2016; Cárdenas-Conejo *et al.*, 2015; Trujillo Hernández, 2013; Cunningham y Gantt, 1998).

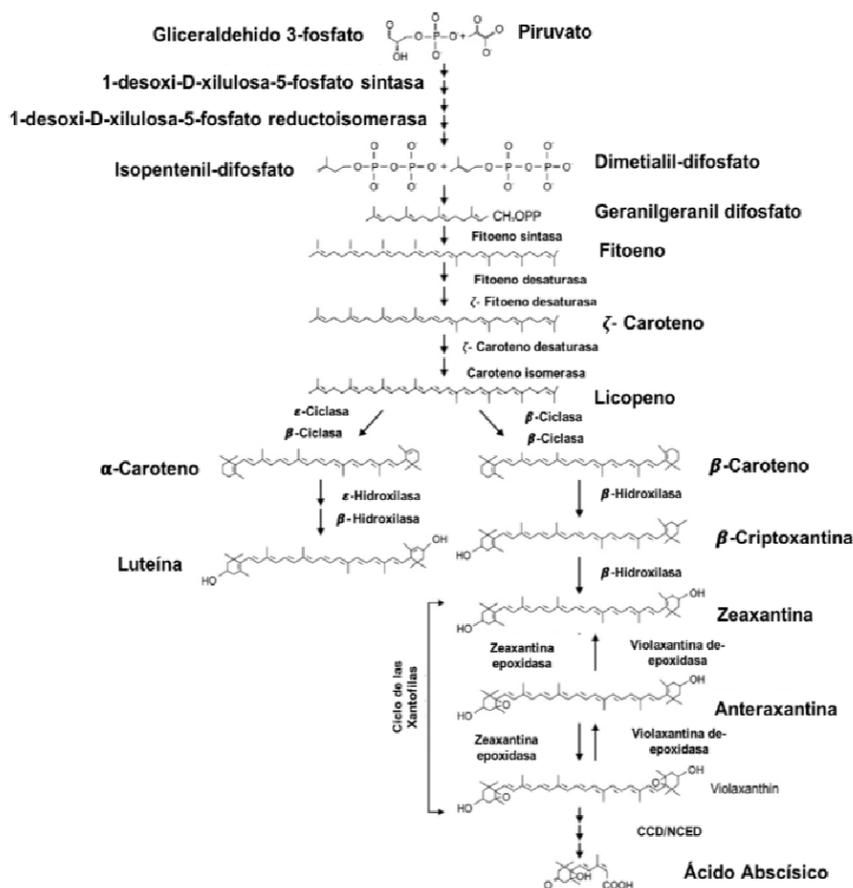


Figura 1.1. Vía metabólica de la síntesis de carotenoides en plantas superiores (Modificado de: Hannoufa y Hossain, 2012).

En el siguiente paso, se convierten dos moléculas de GGDP a diferentes productos carotenoides. La enzima fitoeno sintasa (PYS), que al actuar en GGDP genera 15*cis*-fitoeno (C₄₀), el primer carotenoide de la ruta y estructura básica de todos los carotenoides predecesores (Carballo-Uicab, 2019; Raddatz-Mota *et al.*, 2017; Alcaíno *et al.*, 2016; Cárdenas-Conejo *et al.*, 2015; Hannoufa y Hossain, 2012; Cunningham y Gantt, 1998). A partir de 15*cis*-fitoeno, una serie de cuatro desaturaciones y dos isomerizaciones son catalizadas por un conjunto de enzimas. Primero, dos fitoeno desaturasa (PDS1 y PDS2) para formar tris-*cis*-ζ-caroteno; luego, la enzima 15-*cis*-ζ-caroteno isomerasa (Z-ISO) produce di-*cis*-ζ-caroteno y una ζ-caroteno desaturasa (ZDS) forma tetra-*cis*-licopeno (Figura 1.1). La enzima caroteno *cis-trans* isomerasa (CRTISO) interviene para producir a licopeno, una estructura lineal que funge como sustrato durante su ciclación. La ciclación del final de la cadena principal de hidrocarburos de 40 carbonos es un paso crucial de ramificación para el metabolismo de los carotenoides (Song *et al.*, 2022; Maoka, 2020). Aquí interviene la familia de licopeno ciclase, compuesta por licopeno beta-ciclase (β-

LCY) con la adición de anillo β y licopeno ϵ -ciclase (ϵ -LCY) mediante la adición del anillo ϵ . Las actividades de ϵ -LCY y β -LCY determinan el flujo relativo y la composición de carotenoides de las plantas superiores; por esto, es considerado el punto de bifurcación de la vía de los carotenoides (Song *et al.*, 2022; Maoka, 2020; Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Lu y Li, 2008; Cunningham y Gantt, 1998) y central para la síntesis de bixina (Carballo-Uicab, 2019; Escobar-Turriza, 2017; Cárdenas-Conejo *et al.*, 2015; Trujillo Hernández, 2013; Fraser y Bramley 2004; Bouvier *et al.*, 2003; Jacko *et al.*, 2002).

2.2. Ruta de biosíntesis de bixina

La ruta de síntesis del pigmento bixina aún queda parcialmente incierta, donde hasta la fecha se reportan varios modelos de esta posible ruta. Los primeros fueron Jacko y colaboradores (2002), al plantear la hipótesis que está debía partir de la vía principal de los carotenoides; y que los genes participantes tendrían similitud con los implicados en la ruta sintética del ácido abscísico, como dioxigenasas, aldehído oxidasas, aldehído deshidrogenasas y metil-transferasas. En 2003, Bouvier y colaboradores, basándose en la similitud estructural entre bixina, azafrán y secuencias de enzimas homólogas de *Arabidopsis thaliana*, reportaron tres genes implicados en la biosíntesis de bixina, al lograr que *Escherichia coli* acumulara licopeno. De esta forma, el licopeno es convertido a bixina aldehído por la enzima licopeno dioxigenasa (BoLCD), luego la enzima deshidrogenasa (BoBADH) produce norbixina y, por último, la enzima norbixina metiltransferasa (BoBMT) generan bixina (Bouvier *et al.*, 2003). Sin embargo, ningún equipo de trabajo ha logrado reproducir esta ruta con éxito.

En 2015, Cárdenas-Conejo y colaboradores reportaron el transcriptoma para la especie; donde mediante análisis informático, las secuencias correspondientes al de las enzimas antes reportadas por Bouvier y colaboradores (2003) no fueron encontradas. Sin embargo, el transcriptoma proporcionó nuevas y diferentes secuencias que codifican enzimas que podrían estar implicadas en la síntesis de bixina, como carotenoides dioxigenasas (BoCCDs), aldehído deshidrogenasas (BoALDHs) y metiltransferasas (BoMTs). Con esto, los autores propusieron un modelo de dos rutas metabólicas para la síntesis de bixina, difiriendo en la localización subcelular en la cual se lleva a cabo la biosíntesis, sea en plástoglobulos de los plastidios o en el citosol. El autor denota que ambas rutas necesitan de tres pasos moleculares que involucran la expresión coordinada de genes, la vía MEP, la ruta de los carotenoides y la ruta de biosíntesis de bixina (Cárdenas-Conejo *et al.*, 2015) (Figura 1.2).

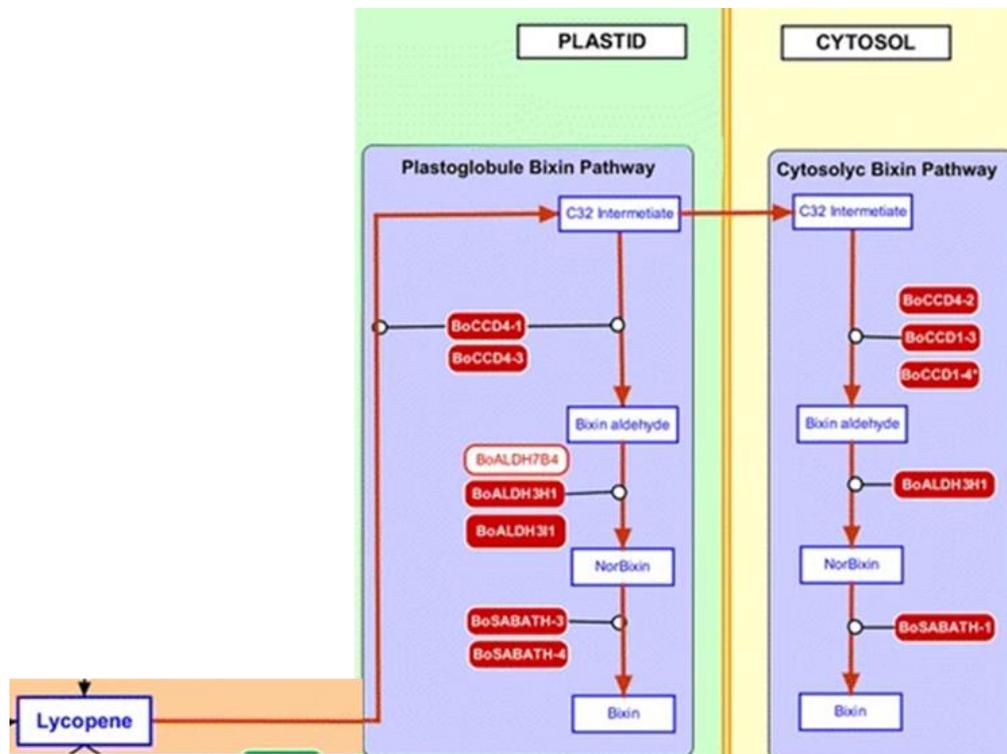


Figura 1.2. Modelo de regulación génica en la biosíntesis de bixina (Modificado de: Cárdenas-Conejo *et al.*, 2015).

En la ruta hipotética de biosíntesis de bixina, los genes licopeno ciclasas (*Boβ-LYC1*, *Boβ-LYC2*, *Boε-LYC*) se desactivan impidiendo el flujo del licopeno por la vía de los carotenoides, punto de bifurcación, provocando una disminución en la concentración de carotenoides cíclicos que aumenta la disponibilidad del licopeno para la producción de bixina (Cárdenas-Conejo *et al.*, 2015). Debido al bloqueo de los genes licopeno ciclasa, se activa una serie de genes dioxigenasa de escisión de caroteno (*BoCCD*), aldehído deshidrogenasa (*BoALDH*) y metiltransferasas (*BoSABATH*) para convertir licopeno a bixina. Dependiendo del compartimiento en donde se lleve la conversión, es el conjunto de enzimas que entran en acción. En los plastoglóbulos de los plastidios, las *BoCCD4-1*, *BoCCD4-3* obtienen bixina aldehído, las *BoADHL3H1*, *BoADHL3H1* generan norbixina y finalmente las enzimas *BoSABATH-3* y *BoSABATH-4* producen bixina. En citosol, las enzimas *BoCCD4-2*, *BoCCD1-3*, *BoCCD1-4*, *BoADHL3H1* y *BoSABATH-1* llevan a cabo el mismo proceso para obtener bixina (Cárdenas-Conejo *et al.*, 2015) (Figura 1.2). Recientemente, Carballo-Uicab y colaboradores (2019), propusieron que ambas rutas se podrían presentar simultáneamente, y una vez terminada la síntesis del pigmento, este se dirige a células de almacenamiento de carotenoides (Louro y Santiago, 2016).

2.3 Bixina

El pigmento es un apocarotenoide, un derivado de carotenoide (C_{40}) de tonalidad naranja-rojiza que se encuentra en todos sus tejidos. En bixina, se presenta principalmente en el arilo de sus semillas donde representa más del 80%; siendo *B. orellana* la única planta que acumula este tinte en una gran cantidad de forma natural (da Silva *et al.*, 2018; Avedaño-Arrazate *et al.*, 2018; Zarza García *et al.*, 2017; Chisté *et al.*, 2011; Akshatha *et al.*, 2011; Gómez-Ortíz *et al.*, 2010; Pérez *et al.*, 2003; Pineda y Calderón, 2002; Arce, 1999). Bixina (metil hidrógeno 9'-cis-6, 6' 'diapocaroteno-6, 6' dioato) es un *cis*-carotenoide de estructura básica ($C_{25}H_{30}O_4$) que consiste de un grupo carboxilo libre y otro estratificado (Carballo-Uicab, 2019; Gómez-Ortíz *et al.*, 2010), gracias a este último, es soluble en aceites y grasas, pero se degrada en presencia de luz y altas temperaturas (Carballo-Uicab, 2019; Akshatha *et al.*, 2011) (Figura 1.3).

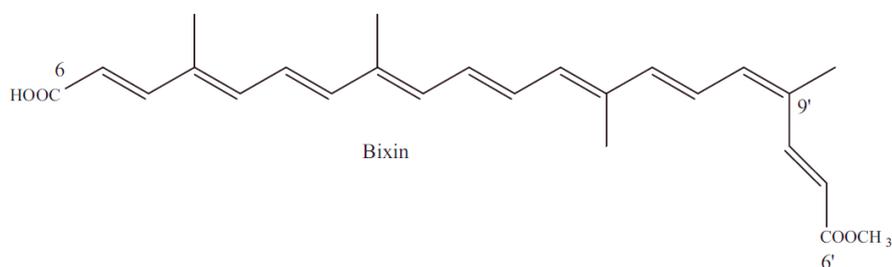


Figura 1.3. Estructura de bixina ($C_{25}H_{30}O_4$) (Chisté *et al.*, 2011).

La bixina se ha reportado con actividad antioxidante, así como con la capacidad de proteger al DNA de sustancias dañinas, además de no ser genotóxica ni hepatocarcinogénica (Agner *et al.*, 2004; Trujillo-Hernández, 2013; Escobar-Turriza, 2017). Por esto, las industrias han optado por usar este pigmento para el consumo humano, donde la secretaría de salud en México (NOM-119-SSA1-1994) y organismos mundiales como la Organización Mundial de la Salud (OMS) y la FAO (Food and Agriculture Organization) la han reconocido con nula toxicidad para su empleo en la industria de alimentos, farmacéutica y cosmética (Dzib-Cauich *et al.*, 2020; Carballo-Uicab, 2019; Pinzón López *et al.*, 2018; Avedaño-Arrazate *et al.*, 2018; Rivera Madrid *et al.*, 2016; Cárdenas Conejo *et al.*, 2015; Rodríguez *et al.*, 2014; Leal y Michelangeli, 2010; Valdez Ojeda *et al.*, 2008; Rivera Madrid *et al.*, 2006; Agner *et al.*, 2004; Pérez *et al.*, 2003).

2.4 Licopeno ciclasas

Los genes licopeno ciclasas son el punto de bifurcación de la ruta de síntesis de carotenoides, generando una diversidad de apocarotenoides (Song *et al.*, 2022; Maoka, 2020; Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Rosas-Saavedra y Stange, 2016; Nisar *et al.*, 2016; Lu y Li, 2008; Cunningham y Gantt, 1998). Estas cumplen un papel importante para el estrés abiótico al que se enfrentan los organismos durante su crecimiento y desarrollo; por lo que, la función principal de los genes es regular el estado fisiológico de la planta aumentando o disminuyendo la síntesis de los carotenoides (Zhao *et al.*, 2020). El compuesto inicial para la escisión oxidativa de los carotenos mediante los genes licopeno ciclasa es el licopeno, una molécula lineal que sufre en sus extremos una serie de ciclaciones según la enzima licopeno ciclasa que intervenga, difiriendo solo en la posición de un doble enlace en el anillo ciclohexano (Zhao *et al.*, 2020; Maoka, 2020; Nisar *et al.*, 2016; Trujillo-Hernández, 2013; Moreno *et al.*, 2013; Martín-Rodríguez, 2010; Fraser y Bramley 2004; Cunningham y Gantt, 1998). En plantas superiores se han identificado dos tipos de *LCY*, β -ciclasa y ϵ -ciclasa que compiten por el licopeno dando a carotenos con anillos β - o ϵ - (Maoka, 2020; Rosas-Saavedra y Stange, 2016).

La enzima β -*LCY* actúa en secuencia, introduce un anillo β - en un extremo produciendo γ -caroteno, luego introduce un segundo anillo en el otro extremo originando el pigmento color naranja β -caroteno, cuyo producto final de esta vía es ácido abscísico (Kössler *et al.*, 2021; Zhao *et al.*, 2020; Nisar *et al.*, 2016; Trujillo-Hernández, 2013; Martín-Rodríguez, 2010; Lu y Li, 2008; Fraser y Bramley 2004; Cunningham y Gantt, 1998) (Figura 1.4). En especies como *Arabidopsis*, arroz y maíz, β -licopeno ciclasa se presentan como copia única en los cloroplastos (Nisar *et al.*, 2016; Moreno *et al.*, 2013); sin embargo, algunas especies presentan isoformas específicas de tejido para ajustar el contenido de sus carotenoides (Rosas-Saavedra y Stange, 2016; Nisar *et al.*, 2016; Moreno *et al.*, 2013). En especies como tomate, sandía, papaya, cítrico y azafrán, presentan un gen β LCY específico de cromoplastos que se expresa en frutos y flores, caracterizándose por una acumulación de pigmentos (Rosas-Saavedra y Stange, 2016). La expresión de los genes β -*LCY*, muestran un aumento del contenido de apocarotenoides de la rama β como β -caroteno, zeaxantina y violaxantina que ayudan a las plantas a la tolerancia al estrés abiótico, fotoprotección y eficiencia fotosintética (Kössler *et al.*, 2021); y a su vez, son precursores de hormonas como estrigolactonas y ácido abscísico, reguladores de crecimiento y desarrollo de las plantas (Sun *et al.*, 2022; Kössler *et al.*, 2021). Cuando interviene la enzima ϵ -*LCY*, esta agrega el anillo ϵ - únicamente en un solo extremo del licopeno originando δ -caroteno,

extremo donde puede actuar la enzima β -LCY dando lugar a α -caroteno, también pigmento de color naranja, cuyo producto final de esta vía es luteína (Song *et al.*, 2022; Zhao *et al.*, 2020; Nisar *et al.*, 2016; Trujillo-Hernández, 2013; Martín-Rodríguez, 2010; Cunningham y Gantt, 1998) (Figura 1.1). La presencia de la luteína en las plantas aumenta la capacidad para la dispersión de calor, así como reducir el daño por la luz alta; protegiendo así a los órganos fotosintéticos, especialmente del fotosistema II. El gen ϵ -LCY se presenta generalmente como copia única; y su nivel de expresión en plantas superiores juega un papel importante en la regulación del contenido y composición de carotenoides β -caroteno/luteína, así como en la resistencia al estrés en plantas (Song *et al.*, 2022).

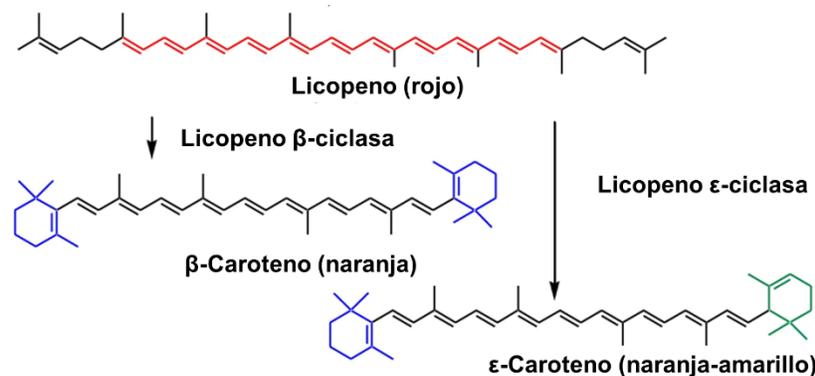


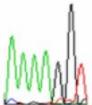
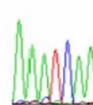
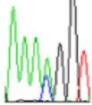
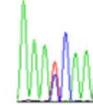
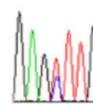
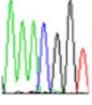
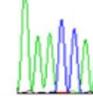
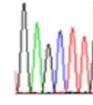
Figura 1.4. Ciclación del licopeno por las enzimas licopeno β -ciclasa (azul) y ϵ -ciclasa (verde) (Modificado de: Maoka, 2020).

En achiote, la regulación de los genes licopeno ciclasa y su sustrato licopeno son clave para la síntesis del pigmento (Escobar-Turriza, 2017; Cárdenas-Conejo *et al.*, 2015; Trujillo-Hernández, 2013; Martín-Rodríguez, 2010; Jacko *et al.*, 2002); donde, el efecto más directo de la regulación de β -ciclasa sobre el contenido de licopeno se refleja en rasgos de color en flores y frutos (Zhao *et al.*, 2020; Hao *et al.*, 2012). En este sentido, en la especie se ha reportado polimorfismo en los alelos de estos genes. En 2010, Martín aplicando la técnica polimorfismo de conformación de cadena simple (SSCP) en secuencias de los genes *pds2*, *pds3* y β -LYC, detectó un patrón de bandeo relacionado con las observaciones morfológicas reportadas por Valdez-Ojeda y colaboradores (2008). Esto representa, que la heterogeneidad de las variantes de achiote en características de flor, fruto y contenido de bixina, está relacionado la variación alélica de estos genes analizados; donde, el polimorfismo observado específicamente en el gen β -LYC1 para las

variantes rojas, estaría implicado en eventos de regulación negativa, acumulación de licopeno y, por consiguiente, mayor disponibilidad del sustrato en estas variantes para la síntesis de bixina. En esta línea, Trujillo-Hdz y colaboradores (2016) emplearon al gen licopeno beta-ciclasa (*Boβ-LCY1*), para analizar la variación alélica de un solo nucleótido (SNP) en morfotipos con características morfológicas contrastantes (color de flor, fruto, dehiscencia y contenido de bixina) en un enfoque de asociación genotipo-fenotipo; logrando identificar SNPs funcionales asociados al morfotipo, arrojando la formación de tres grupos morfológicos, e identificando al grupo con mayor producción del pigmento, morfotipo de flor rosa (Cuadro 1.1). Este estudio nos presenta al gen licopeno ciclasa asociado a rasgos de color de flor y frutos, como marcador molecular ideal para predecir producción de bixina. En 2017, Escobar-Turriza, continuó con el análisis de variaciones polimórficas en el gen de licopeno, *Boβ-LCY2* asociados a cromoplastos; logró amplificar el gen a partir del DNA genómico de 6 plantas de achiote, obteniendo 3 secuencias que contenían el ORF del gen *Boβ-LCY2*: *Boβ-LCY2-1* (1497pb) con dos cambios puntuales, *Boβ-LCY2-2* (1497pb) un solo cambio puntual y *Boβ-LCY2-3* (1948pb) con una inserción y 5 cambios puntuales. Las secuencias obtenidas presentaron ausencia de intrones dentro de la región ORF, pero se desconoce si en las regiones URT del gen exista la presencia de intrones, lo cual podría influir en la regulación y acumulación del sustrato.

Zhao y colaboradores, (2020) mencionan que el efecto de la regulación de las enzimas licopeno ciclasa se refleja directamente rasgos en color de flores y frutos; donde, la expresión de las enzimas, sea en aumento o disminución, provocará un cambio en el flujo de las vías que emplean a licopeno como sustrato, lo cual se verá reflejado en cambios en la síntesis de carotenoides y su contenido. Carballo-Uicab, 2019, caracterizó la expresión de los genes *Boβ-LCY1* y *Boε-LCY* para conocer su potencial participación en la biosíntesis del pigmento; este reportó que la expresión y acumulación de bixina incrementó durante el desarrollo de la semilla, acumulándose en los plastoglobulos (células de almacenamiento de bixina). Los genes se expresaron más en las acciones de flor blanca, con menor contenido de bixina, y se expresaron menos en la accesión rosa, con mayor contenido de bixina. Esto indica que los niveles de expresión de los genes se relacionan con la disponibilidad del sustrato licopeno para la biosíntesis de bixina en achiote, el cual se encuentra disponible por la regulación de los genes *licopeno ciclasa* (Cárdenas-Conejo *et al.*, 2015).

Cuadro 1.1. Tres morfotipos caracterizados de la especie *Bixa orellana* L. en Yucatán.

MORFOTIPO	PRODUCCIÓN DE BIXINA	SNP <i>Boβ-LCY1</i>		
		IV A473C	VII T532C	VIII T552C
 <p>G1</p>	<p>Menor</p> <p>6.25 – 12.67</p> <p>mg/gDW</p>	<p>AAAAGGT K</p> 	<p>AAATCAA I</p> 	<p>GAGTTTG F</p> 
 <p>G2</p>	<p>Mayor</p> <p>16.90 – 20.56</p> <p>mg/gDW</p>	<p>AAAAGGT K AAACGGT N</p> 	<p>AAATCAA I AAATCAA T</p> 	<p>GAGTTTG F GAGTTTG L</p> 
 <p>G3</p>	<p>Intermedia</p> <p>8.76 – 10.76</p> <p>mg/gDW</p>	<p>AAACGGT N</p> 	<p>AAATCAA T</p> 	<p>GAGTTTG L</p> 

3. GENOMA DE ACHIOTE

3.1 Variación genética

La variabilidad genética describe los cambios en las secuencias de DNA entre los individuos de una especie (Rimieri, 2017); siendo el genoma, el conjunto completo de todas estas secuencias, la información genética y heredable de un organismo (Herrera, 2007; Rimieri, 2017; Montoya *et al.*, 2014). El genoma está compuesto por secuencias de DNA de bajo número de copias que codifican proteínas (genes) y por una cantidad importante de secuencias con diferente grado de repetitividad y organización en el genoma (secuencias repetitivas) (Lee y Kim, 2014). El gen, es la unidad principal donde se origina la variación genética, estructura conformada principalmente por tres tipos de segmentos: secuencias codificantes (exones), no codificantes (intrones) y secuencias reguladoras (promotores), necesarios para expresar y producir una secuencia proteica con actividad biológica (Montoya *et al.*, 2014; Barrett *et al.*, 2012). Estos tienen un efecto directo en los individuos, que no se restringe a la secuencia codificante, las secuencias reguladoras que la flanquean, moderan en qué condiciones ambientales, células o tejidos ocurre esta expresión (Wray *et al.*, 2003); por lo que, una parte importante de la variación genética se

encuentra en estas secuencias promotoras (sitios de unión) que en conjunto con sus factores de transcripción (proteínas) generan diferentes efectos a nivel fenotipo en los organismo (Wray *et al.*, 2003).

Sin embargo, el genoma está compuesto en su mayoría por secuencias repetitivas de DNA (Meng *et al.*, 2021; Biscotti *et al.*, 2015; Lee y Kim, 2014; López-Flores y Garrido Ramos, 2012). El contenido de estos elementos, reordenamiento o duplicaciones, induce diversos efectos mutacionales que pueden tener un impacto importante en la estructura del genoma, en la expresión de genes cercanos, generando variabilidad y plasticidad genómica que se refleja en variaciones en el fenotipo (Zhang *et al.*, 2020; Lanciano y Cristofari, 2020; Boutte *et al.*, 2020; Wang *et al.*, 2016; Lee y Kim, 2014; López-Flores y Garrido Ramos, 2012). Por lo tanto, las características morfológicas que observamos hoy en día, son el resultado de una relación compleja entre las secuencias nucleotídicas, su repetitividad y regulación dentro del genoma (Montoya *et al.*, 2014). Por esto mismo, la variabilidad genética se presenta con mayor frecuencia en especies domesticadas, donde los humanos han creado razas y variedades cultivadas a través de métodos de selección o de mejoramiento genético para transmitir esas características particulares a su prole (Rimieri, 2017; Montoya *et al.*, 2014). En este sentido, disponer del genoma de una especie es contar con el conjunto completo de genes, sus secuencias de DNA, vital para vincular secuencias y regulación con su fenotipo, para así, entender la variabilidad intra e interespecífica de una especie y sus poblaciones (Montoya *et al.*, 2014; Feuillet *et al.*, 2011).

En este contexto, achote es una especie cultivada que presenta diferencias morfológicas marcadas entre sus poblaciones, principalmente en el fruto, donde se obtiene el pigmento bixina. El primer trabajo en reportar esta alta variación morfológica en las características del fruto (forma, color y pubescencia), fue el de Carvalho y colaboradores (2005) en germoplasma de achote en Brasil mediante el uso de isoenzimas. También, basados en variaciones morfológicas, incluido el contenido de bixina; Rivera-Madrid y colaboradores (2006) caracterizaron cultivares en México de la especie, mediante la polinización artificial de la especie, reportando genotipos contrastantes. Siguiendo este tópico, en 2008, Valdez-Ojeda y colaboradores, confirmaron la alta variabilidad morfológica y genética de achote empleando marcadores SRAP (polimorfismo amplificado de secuencias relacionadas), identificando un grupo de individuos con características morfológicas deseables. Estos estudios, destacaron a *B. orellana* por ser especie de polinización cruzada y altamente heterocigota (Carvalho *et al.*, 2005; Rivera-Madrid *et al.*, 2006); y esto, se refleja en características morfológicas contrastantes en fruto, color de flores y

en la producción del pigmento *bixina*, el cual es el principal carácter de interés agronómico (Avedaño-Arrazate *et al.*, 2018). Son estas diferencias morfológicas, las que se presentan como variantes o cultivares dentro las especies, donde ciertas características o rasgos deseables suelen ser seleccionadas por su interés agronómico (Poggio y Naranjo, 2004).

En Yucatán, la selección de estos rasgos fenotípicos en color de flor, fruto y contenido de bixina, se ha dado en la especie, pero sin una base científica. Estas selecciones se mantienen como variantes en tres tipos de sistemas agronómicos: solar, milpa y monocultivo, que difieren principalmente en las prácticas de manejo, como patrones de siembra y procedencia de las semillas (Pech-Hoil, *et al.*, 2017). Por esto, Pech-Hoil y colaboradores 2017, caracterizaron los sistemas de apareamiento de dos fenotipos de achiote entre los sistemas agronómicos, empleando marcadores moleculares de polimorfismo amplificado relacionado con la secuencia (SARP). Reportaron a la especie con diferentes sistemas de apareamiento entre los fenotipos y los sistemas agronómicos; solar con apareamiento mixto y milpa, monocultivo con apareamiento cruzado. Los autores mencionan que los dos sistemas producen semillas viables; donde, el cruzamiento promovería la recombinación y la autofecundación la estabilidad de rasgos deseables (Pech-Hoil, *et al.*, 2017). Sin embargo, en la actualidad, los mecanismos genéticos y moleculares que regulan los rasgos morfológicos de la especie son poco conocidos.

En la actualidad, gracias a las tecnologías de secuenciación masiva de nueva generación (NGS), se producen lecturas que difieren en longitud, número y tasa de error, que proporcionan grandes cantidades de información que permiten a los investigadores generar bases de datos para el estudio del genoma y el análisis del transcriptoma de las especies (Chan-Rodríguez y Rivera-Madrid, 2020). Illumina genera millones de secuencias de alta calidad con lecturas de hasta 300pb, esta tecnología es la más empleada para el análisis de expresión de genes con un error de <0.1%; Pacific Bioscience (PacBio) genera miles de secuencias con lecturas de hasta ~20Kb con un error de 11%; que junto a Oxford secuenciación de molécula única que produce cientos de secuencias de hasta 150kb con un error >12%, se emplean para el estudio de genomas completos (Chan-Rodríguez y Rivera-Madrid, 2020; Lu *et al.*, 2016; Rhoads y Au, 2015). Este tipo de estudios se basan en analizar las diferencias en las secuencias de nucleótidos, la búsqueda de secuencias conservadas, que no hayan sufrido grandes cambios genéticos, es decir secuencias homólogas. Estas son secuencias similares para una o varias de proteínas que presentan un mismo origen evolutivo. Las secuencias homólogas pueden ser de dos tipos: ortólogas, genes de diferentes especies que evolucionaron de un ancestro en común, cuya

divergencia es debido a la especiación, que generalmente presentan la misma función; y parálogos, genes que no comparten el último ancestro común debido a eventos de duplicación, presentes están en la misma especie y generalmente presentan funciones diferentes, son copias con posibilidad de evolucionar (Millo-Sánchez *et al.*, 2014; Koonin, 2005).

En 2015, Cárdenas-Conejo y colaboradores reportaron un estudio del transcriptoma de *Bixa orellana*; donde mediante análisis informático, las secuencias correspondientes al de las enzimas antes reportadas por Bouvier y colaboradores (2003) no fueron encontradas. Sin embargo, el transcriptoma proporcionó nuevas y diferentes secuencias que codifican enzimas que podrían estar implicadas en la síntesis de bixina, como carotenoides dioxigenasas (BoCCDs), aldehído deshidrogenasas (BoALDHs) y metiltransferasas (BoMTs). Con esto, los autores proporcionaron una base de datos de cDNAs correspondientes a los diversos genes implicados en la ruta de biosíntesis de los carotenoides, y por lo tanto de biosíntesis de bixina, incluidos licopeno ciclasa, para futuros estudios comparativos del genoma de las poblaciones de la especie.

La reciente secuenciación de una base de datos del genoma de *B. orellana* (aún no publicada) nos proporciona herramientas para caracterizar la estructura y explorar la regulación de estos genes; así como secuencias de interés para la ubicación de estos en el genoma, que en conjunto con el transcriptoma de la especie ya reportado (Cárdenas-Conejo *et al.*, 2015) nos brinde secuencias para relacionar las observaciones fenotípicas con las diferencias en sus secuencias genómicas y transcriptómicas (Chan-Rodríguez y Rivera-Madrid, 2020) y nos ayude a dilucidar la variación en la producción del pigmento bixina entre la especie.

3.2 Citogenética

Se ha reportado que diferencias o cambios en los fenotipos son el reflejo de las variaciones en los cromosomas, que se expresan en características fácilmente analizables como el número, forma y tamaño de los cromosomas (Windham *et al.*, 2020; Li *et al.*, 2017; Murray, 2012; Poggio y Naranjo, 2004) que, a su vez, son el reflejo de la variación intraespecífica del genoma (Silva *et al.*, 2018; Ceccarelli *et al.*, 2011), lo que podría afectar la regulación de genes implicados con fenotipos específicos como por ejemplo, la producción de pigmentos.

El DNA que conforma el genoma de un individuo se encuentra en diferentes niveles de condensación, que va desde la fibra desnuda (2nm) hasta los cromómeros (200 a 300 nm), estos

se superenrollan para formar las cromáticas que conforman los cromosomas (Ferguson-Smith, 2015; Herrera, 2007; Wanner *et al.*, 2005). La descripción de cada conjunto cromosómico de un individuo o especie ordenado de acuerdo a la morfología y tamaño de los cromosomas conforma el cariotipo, representando el nivel más alto de organización estructural y funcional del genoma nuclear (Carta *et al.*, 2020; Silva *et al.*, 2018; Soltis *et al.*, 2009; Herrera, 2007; García, 1985; Palomino, 1985). Aunque las características morfológicas de los cromosomas son constantes dentro de una misma especie, existe una gran variación a nivel de especie en el cariotipo y número cromosómico (Windham *et al.*, 2020; Li *et al.*, 2017; Soltis *et al.*, 2009; Brown y Bertke, 1979). Las características estructurales y cuantitativas de los cromosomas visualizadas en el cariotipo, son importantes en la investigación taxonómica y aplicada, al contribuir sobre la relación genética entre especies o poblaciones y a una mejor comprensión de la forma en que divergieron entre sí; siendo los cromosomas, parte de un sistema dinámico que está moldeando el proceso evolutivo (Li *et al.*, 2017; Ferguson-Smith, 2015; Guerra, 2008; Poggio y Naranjo, 2004; Swanson *et al.*, 1981).

Los estudios citogenéticos se han centrado en la estructura y evolución del genotipo de las diferentes especies vegetales, tanto silvestres como cultivadas (Windham *et al.*, 2020; Aiassa *et al.*, 2015; Herrera, 2007); aportando información básica como la determinación del número cromosómico, cariotipo, detección de anomalías en la división celular; donde, la estabilidad de las características cariotípicas asegura la transferencia y fidelidad del material genético a las próximas generaciones, mientras que su variación proporciona soporte genético para la diferenciación y adaptación (Carta *et al.*, 2020). Con base a estas características, es posible determinar variaciones genéticas, plasticidad a diversos entornos ecológicos, facilitando la manipulación y selección de genotipos de alto rendimiento.

Para *Bixa orellana*, existen trabajos sobre el estudio del número cromosómico y descripción de su cariotipo. La especie reporta un contenido bajo de 4C DNA de 0.83 pg (Ohri *et al.*, 2004) y 0.78 pg (Handson *et al.*, 2001); cuyo número cromosómico, según una búsqueda en Index to Plant Chromosome Number (2023) de números cromosómicos para la especie es de $2n=14,16$ cromosomas (Cuadro 1. 2). Esta serie de reportes ha generado un debate sobre si el número cromosómico de la especie ha sufrido algún evento de pérdida de un solo cromosoma (aneuploidía), pérdida de fragmentos cromosómicos (delección) (Bear, 1976) o incluso que el cromosoma 1 pudo haberse fragmentado durante el proceso de aplastamiento de tejido (Almeida *et al.*, 2006). Sin embargo, no hay estudios de descripción cariotípica para ejemplares con $2n=$

16 cromosomas para confirmar esta conjetura. En 2007, Martín-Rodríguez como trabajo de tesis de licenciatura, formó parte del equipo de trabajo de la Dra. Rivera Madrid donde estableció una metodología para el análisis del complemento cromosómico. Esta se realizó en una accesión de Yucatán identificada como peruana roja, caracterizada por flores rosas, frutos dehiscentes color verde y espinas rojas; reportando un número cromosómico $2n= 14$. No obstante, no se realizó una caracterización cromosómica ni se reportó el cariotipo de la accesión estudiada.

Cuadro 1.2. Registros de números cromosómicos para *Bixa orellana* L.

Referencia	Origen	$2n/n$
Michelangeli <i>et al.</i> , 2002	Venezuela	$2n= 16$
Hanson <i>et al.</i> , 2001	Reino Unido	$2n= 14$
Carvalho <i>et al.</i> , 1991	Brasil	$2n= 14$
Krishnan y Ayyangar, 1987	India	$2n= 14, 16$
Krishnan, 1977	India	$2n= 16$
Mukherjee, 1975	India	$n=7, 2n= 14$

No obstante, la descripción de los cromosomas se ha realizado mediante técnicas moleculares de bandeo (Ag-NOR y C) y marcadores rDNA (45S y 5S) con hibridación fluorescente en ejemplares de Brasil. Para el cariotipo somático, se describió a los pares 1 a 4 y 6 como metacéntricos y los pares 5 y 7 submetacéntricos, con la presencia de las bandas Ag-NOR en el cromosoma 1; con un cromosoma de 3.82 μm (par uno) y 1.24 μm (par 7) (Almeida *et al.*, 2006). Para el cariotipo haploide, se describió al cromosoma 1 a 6 como metacéntricos y el 7 submetacéntrico, donde el marcador 45S (negro) se presentó en el cromosoma 1 y 5S (gris) en los cromosomas 4 y 5; con un cromosoma de 3.2 μm (uno) y cromosomas entre 1.9 a 1.4 μm (dos a siete) (Lombello y Pinto-Maglio, 2014) (Figura 1.5). Estos estudios para individuos de Brasil, muestran diferencias en tamaño y morfología en los cromosomas de sus poblaciones, señalando al cromosoma 1, por presentar casi el doble del tamaño de los otros cromosomas. Sin embargo, hay poca información relacionada a los cultivares de la planta y sus características particulares.

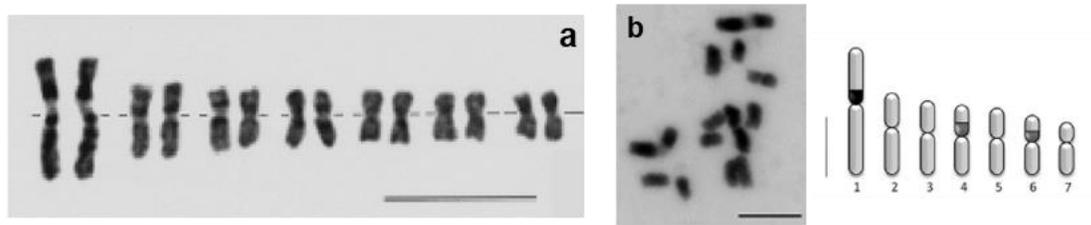
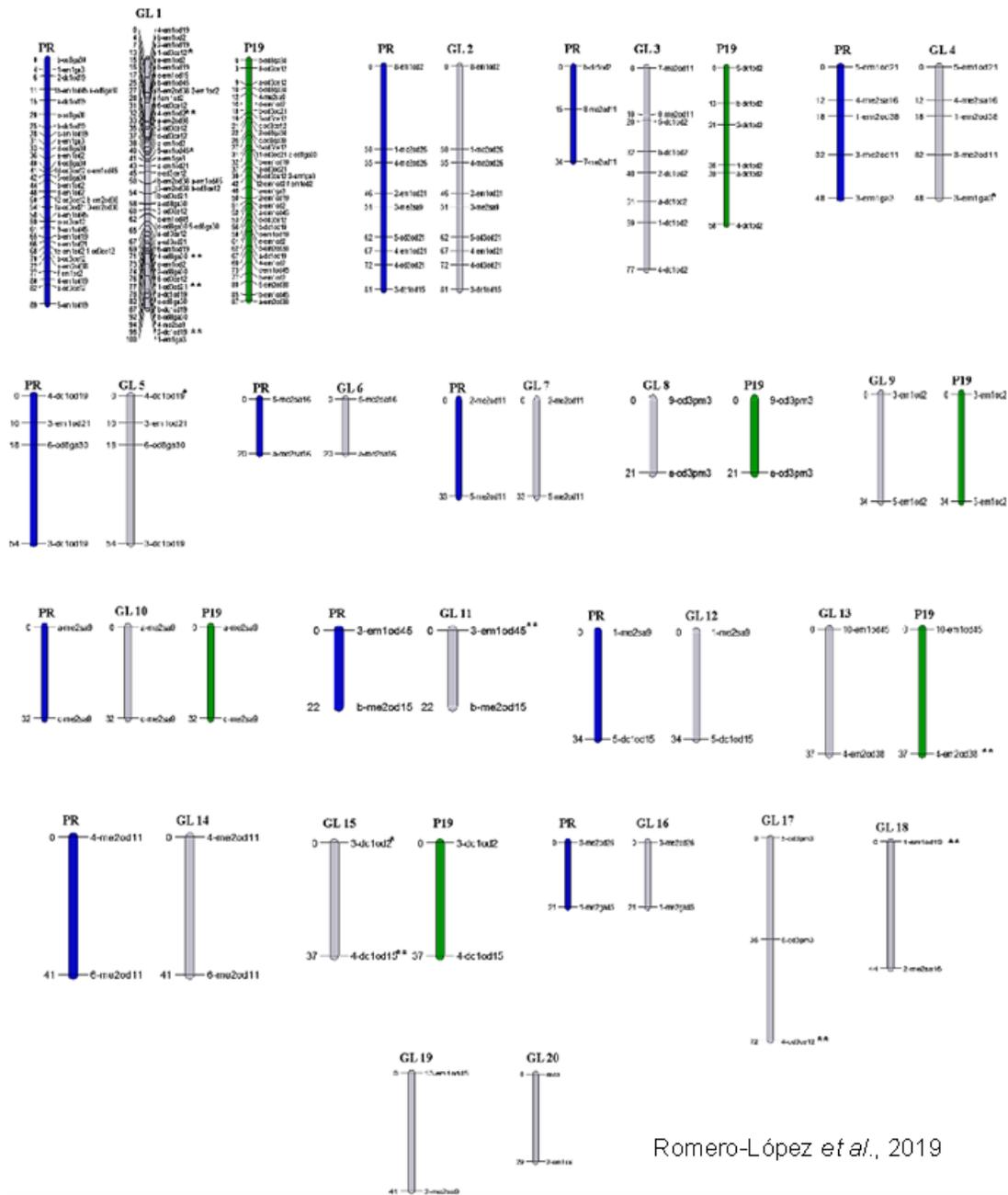


Figura 1.5. Cariotipo e ideograma de *Bixa orellana* L. a) Cariotipo somático $2n= 14$ (Tomado de: Almeida *et al.*, 2006). b) Cariotipo e ideograma haploide $n= 7$ (Tomado de: Lombello y Pinto-Maglio, 2014).

Una de las maneras para estudiar la variación fenotípica presente en las poblaciones de una especie, es la construcción de mapas genéticos, citogenéticos o físicos, proporcionando una estructura general del genoma (Chaney *et al.*, 2016). Para individuos de achiotte de Yucatán, Romero-López y colaboradores (2019) emplearon al marcador funcional β -ciclase para el desarrollo del primer mapa de ligamiento genético integrado a través de marcadores moleculares obtenidos por polimorfismo amplificado relacionado con secuencia (SRAP) y al gen funcional β -ciclase, mediante la estrategia bidireccional “pseudo-testcross”. Esta herramienta de estudio, mediante el mapeo de marcadores moleculares, proporciona un panorama de la estructura y organización del genoma, para la futura identificación de genes específicos vinculados a rasgos fenotípicos. Este cubrió una longitud de 879.054cM (78.8% del genoma) comprendiendo 103 marcadores (102 SRAP y uno de β -ciclase) incluidos en 20 grupos de enlace (GL); donde, los grupos 1 y 3 fueron lo más importantes al permitir la construcción del mapa integrado. El grupo de enlace uno, que correspondería al cromosoma 1, fue el que comprendió la mayoría de los marcadores y el GL 20 al marcador β -ciclase (Figura 1.6). No obstante, según la teoría de mapas de ligamiento, los grupos de enlace deberían ser igual al número cromosómico haploide (Venkat *et al.*, 2014); por lo tanto, para achiotte, no se puede determinar la ubicación del marcador β -ciclase.

Por otra parte, es importante señalar que la información cromosómica de la especie se limita a individuos de Brasil no caracterizados morfológicamente y, por consiguiente, es importante el estudio cromosómico de las poblaciones caracterizadas, así como describir el cariotipo de la especie. Con lo anterior, se infiere que las técnicas citogenéticas serían idóneas para confirmar el número cromosómico de la especie, así como generar un panorama general de su genoma que nos proporcione más información para dilucidar las diferencias en producción de bixina en los tres morfotipos de achiotte para futuros trabajos de mejoramiento genético.



Romero-López *et al.*, 2019

Figura 1.6. Mapas genéticos parentales e integrados de *Bixa orellana*. El mapa integrado GL (gris) contiene 103 marcadores incluidos en 20 grupos de enlace. El mapa P19 (verde) que corresponde al progenitor femenino, contiene 52 marcadores en siete GL y el mapa PR (azul) que corresponde al progenitor masculino, contiene 70 marcadores en 12 GL.

RECAPITULACIÓN

De lo presentado anteriormente, se observa que *Bixa orellana* es una especie que se caracteriza por presentar un fenotipo contrastante en cuanto a las características en flor, fruto y en el contenido de bixina. La bixina es un pigmento derivado de la ruta de biosíntesis de los carotenoides, de tonalidad naranja-rojiza presente en las semillas. La bixina se emplea para elaborar colorantes en la industria de alimentos, farmacéutica y cosmética. La variabilidad fenotípica del achiote se relaciona con el hecho de que es una especie que presenta polinización cruzada, lo que lo hace altamente heterocigoto que a su vez genera las diferencias mencionadas en la calidad y cantidad del pigmento que produce. Se ha reportado que los genes licopeno ciclasa cumplen un papel importante al regular la síntesis y acumulación de carotenos en las plantas durante su crecimiento y desarrollo, con el fin de mantener su estado fisiológico en condiciones de estrés. En achiote, estudios que involucran a los genes licopeno ciclasa, han mostrado una relación entre SNPs y características morfológicas, determinando tres grupos contrastantes e identificando al morfotipo con mayor contenido de bixina. En este contexto, el análisis de los genes licopeno ciclasa de la base de datos del genoma (aún no publicado) y de los transcriptomas de morfotipos de achiote, nos proporcionaría información para entender la variación intraespecífica en la especie.

HIPÓTESIS

Bixa orellana presenta diferentes contenidos de bixina entre sus morfotipos. Dado que los genes licopeno ciclasa se encuentran entre los principales reguladores de la síntesis de estos pigmentos carotenoides, las diferencias en producción de bixina podría estar asociado a diferencias en la estructura, regulación y ubicación genómica de los genes licopeno ciclasa en la especie.

OBJETIVO GENERAL

Caracterizar los genes *licopeno ciclasa* (*BoLCY*) del genoma de *Bixa orellana* L.

OBJETIVOS ESPECÍFICOS

1. Caracterizar la estructura y regulación de los genes *licopeno ciclasa* (*Boβ-LCY1*, *Boβ-LCY2*, *Boε-LCY*) del genoma de *Bixa orellana* L.
2. Analizar los genes *licopeno ciclasa* (*Boβ-LCY1*, *Boβ-LCY2*, *Boε-LCY*) del transcriptoma de tres morfotipos de *Bixa orellana* L. con diferente producción de bixina.
3. Determinar las características estructurales y cuantitativas de los cromosomas en un morfotipo de *Bixa orellana* L. con alta producción de bixina.

JUSTIFICACIÓN

Bixina es el pigmento principal del achiote, empleado en la industria alimentaria, farmacéutica y cosmética. En México, existen morfotipos con diferente contenido de bixina, el cual, ha sido asociado con el polimorfismo de los genes licopeno beta ciclasa; punto de bifurcación de la ruta de biosíntesis de los carotenoides. Estos fenotipos son reflejo de la variación intraespecífica del genoma; siendo los cromosomas la representación del nivel más alto y el gen, la unidad fundamental de organización estructural y funcional del DNA nuclear. En la actualidad, las técnicas de secuenciación masiva nos permiten detectar cambios nucleotídicos, así como analizar secuencias que regulan la expresión y variabilidad de los genes; que, en conjunto con técnicas citogenéticas que nos permitan visualizar los cromosomas, nos brindarían un panorama general del genoma de la especie. Con esto, el estudio de la ubicación, estructura y probable regulación de los genes licopeno ciclasa del genoma de achiote es valioso, con el fin de elucidar la relación entre estos genes y la producción diferencial de bixina en la especie, para la futura selección de genotipos y programas de mejoramiento genético.

ESTRATEGIA EXPERIMENTAL

Con el fin de cumplir los objetivos planteados en este trabajo, se describe la siguiente estrategia experimental. Para los análisis *in silico*, se emplearán dos estrategias bioinformáticas para el estudio de los genes licopeno ciclasa de Achote. Primero, para extraer e identificar las secuencias codificantes y regulatorias, a fin de visualizar su estructura. Segundo, para identificar y analizar su expresión en tres morfotipos contrastantes. En cuanto al análisis citogenético, se empleará para examinar las características cromosómicas (morfología y número cromosómico), con el propósito de caracterizar el cariotipo y construir el ideograma de la especie (Fig. 1.7.)

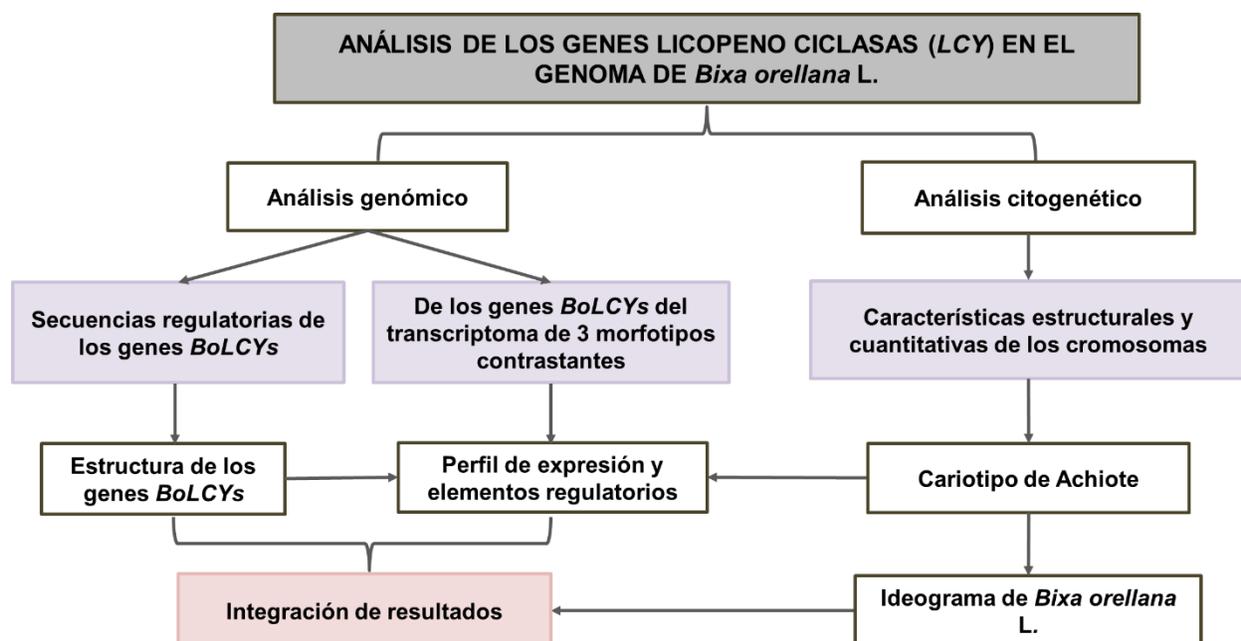


Figura 1.7. Estrategia experimental para el análisis de los genes licopeno ciclasa en *Bixa orellana* L.

CAPÍTULO II

ANÁLISIS DE LOS GENES LICOPENO CICLASA (*Boβ-LCY1*, *Boβ-LCY2*, *Boε-LCY*) DEL GENOMA DE *Bixa orellana* L.

2.1. INTRODUCCIÓN

Los genes licopeno ciclasas (β - y ϵ -) son un punto clave en la ruta de síntesis de biosíntesis de los carotenoides (Song *et al.*, 2022; Maoka, 2020; Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Lu y Li, 2008), cuya función principal es regular el estado fisiológico de la planta aumentando o disminuyendo la síntesis de los carotenoides (Zhao *et al.*, 2020). Estos pigmentos son primordiales para procesos como la fotosíntesis en la absorción de luz (380 a 500nm), fotoprotección, antioxidantes, componentes estructurales de la membrana, y precursores de fitohormonas como ácido abscísico y estrigolactonas, involucradas en procesos de desarrollo y defensa a diversos factores de estrés abióticos y bióticos (Kössler *et al.*, 2021; Zhao *et al.*, 2020; Alcaíno *et al.*, 2016; Fraser y Bramley, 2004).

En plantas superiores se han identificado dos tipos de *LCY*, β -licopeno ciclasa y ϵ -licopeno ciclasa (Maoka, 2020; Rosas-Saavedra y Stange, 2016). En especies como *Arabidopsis*, arroz y maíz, β -*LCY* se presentan como copia única en los cloroplastos (Nisar *et al.*, 2016; Moreno *et al.*, 2013); sin embargo, algunas especies presentan isoformas específicas de tejido para ajustar el contenido de sus carotenoides (Rosas-Saavedra y Stange, 2016; Nisar *et al.*, 2016; Moreno *et al.*, 2013). Este caso se presenta en especies como tomate, sandía, papaya, cítrico y azafrán, que presentan un gen β -*LCY* específico de cromoplastos que se expresa en frutos y flores, caracterizándose por una acumulación de pigmentos (Rosas-Saavedra y Stange, 2016). En cuanto a ϵ -*LCY*, se presenta generalmente como copia única en los órganos fotosintéticos, ayudando a reducir el daño por la luz (Song *et al.*, 2022).

Se ha reportado que el efecto más directo de la regulación de los genes licopeno ciclasas en las plantas, se refleja directamente en rasgos del color de flores y frutos (Zhao *et al.*, 2020). Este efecto se ha observado en *Bixa orellana*, especie que posee una diversidad de fenotipos en flores y frutos, debido a los pigmentos carotenoides presentes en sus semillas (Valdez-Ojeda *et al.*, 2010; Rivera-Madrid *et al.*, 2006); los cuales se emplean en la elaboración de colorantes en la industria de alimentos, farmacéutica y cosmética. Por esto, se ha empleado a los genes licopeno ciclasa como objeto de estudio para direccionar la búsqueda de genotipos con caracteres deseables para su aprovechamiento. Producto de esto, en 2015, Cárdenas-Conejo y colaboradores, identificaron dos genes licopeno beta ciclasa (*Boβ-LCY1* y *Boβ-LCY2*) y una

épsilon ciclasa (*Bo ϵ -LCY*) en el transcriptoma de *novo* de la especie. Luego, en 2016, Trujillo-Hdz y colaboradores, detectaron variaciones alélicas de un solo nucleótido (SNP) en el gen *Bo β -LCY1*, lo que les permitió identificar tres grupos morfológicos y al grupo con mayor producción de bixina.

No obstante, el efecto de los genes sobre el fenotipo no se restringe a las secuencias codificantes. Las regiones no codificantes (no traducidas) también están involucradas en la regulación de la expresión génica (Koul *et al.*, 2018; Barrett *et al.* 2012); estas moderan cuando, donde, nivel y tejido donde se expresarán (Wray *et al.*, 2003). A su vez, las secuencias repetitivas también influyen en la variabilidad genética de las especies; sus diversos efectos mutacionales mediante contenido, reordenamiento o duplicación, pueden tener un impacto importante en la estructura del genoma y en la expresión de genes cercanos, que se refleja a su vez en diferencias en el fenotipo (Zhang *et al.*, 2020; Lanciano y Cristofari, 2020; Boutte *et al.*, 2020; Wang *et al.*, 2016; Lee y Kim, 2014; López-Flores y Garrido Ramos, 2012). Con esto, el estudio de la estructura y elementos regulatorios de los genes licopeno ciclasa de una reciente secuenciación del genoma (aún no publicado), nos brinda la oportunidad de analizar la estructura de los genes licopeno ciclasa y su potencial regulación. Por lo tanto, este apartado se centrará en desarrollar la metodología para identificar las secuencias codificantes de los genes de licopeno ciclasa de una base de datos del genoma del morfotipo B (flor rosa) obtenidos con secuenciación de molécula única (aún no publicado), así como el análisis de las secuencias regulatorias de la región río arriba.

2.2. MATERIALES Y MÉTODOS

2.2.1. Secuencias de los genes *BoLCY* y sus regiones conservadas

Se empleó una base de datos de secuenciación de molécula única que consta de 961 secuencias, 39,440 contigs con longitudes de 9,966,608 a 1,055pb correspondiente al morfotipo B (flor rosa), obtenidas por el equipo de laboratorio 27 de la UBI (sin publicar), así como el cDNA de los genes *LCY* de achote obtenidos de GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) como referencia. Estas se alinearon mediante el programa BLASTx (Johnson *et al.*, 2008). con el fin de identificar secuencias con coberturas >90% y la búsqueda de posibles ortólogos. El rango de las secuencias resultantes se extrajo del contig, se analizaron en ORFinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) para detectar la región codificante, se tradujo a proteína con Six frame translation (https://www.bioline.com/media/calculator/01_13.html) para obtener las

secuencias de aminoácidos, estas se alinearon en la plataforma CDD (Lu *et al.*, 2020) con referencia en BioEdit (Hall T. A., 1999) junto con la secuencia proteica reportada para cada gen, así, obtener las secuencias codificantes y no codificantes para cada gen *LCY* de achote. Las secuencias codificantes se alinearon con las secuencias reportadas de las regiones conservadas y motivos característicos de los genes licopeno ciclasa: región conservada β -LCYs, sitio de unión de dinucleótidos, motivo ciclasa I, motivo ciclasa II, motivo β -LCYs y dominio de actividad catalítica β -LCY (Moreno *et al.*, 2013; Alquézar *et al.*, 2009).

2.2.2. Análisis de elementos regulatorios *cis* de los genes *BoLCY*

Para caracterizar los elementos que comprenden la región regulatoria de los genes *LCY*; se realizó una búsqueda a 3,000 pb río arriba para cada gen *LCY* con el programa PLACE database (Plant Cis-acting Regulatory DNA Elements) (Higo *et al.*, 1999), para identificar las regiones regulatorias *cis* y sus asociados factores de transcripción (TFs) con respuesta a estímulos/función. Con estos, se generó una base de datos de presencia/ausencia para los tres genes, clasificándose en las categorías: fotosíntesis, reproducción, sistema oxidativo, defensa a factores ambientales, señalización, ciclo celular, metabolismo de carbohidratos, respuesta a enfermedades, desconocidos y transcripción. Estas se visualizaron utilizando el entorno de programación R (RStudio Team, 2015) empleando el paquete “ggplots” (Warnes *et al.*, 2016). Posteriormente, los motivos de los elementos *cis* seleccionados se identificaron en los contigs, esto mediante el programa Ugene 48.1 (Okonechnikov *et al.*, 2012). Posteriormente, en el programa se construyó la estructura de los genes *LCY*: secuencia codificante, no codificante y región promotora, con sus respectivas secuencias regulatorias.

2.2.3. Análisis de miRNAs de los genes *BoLCY*

Para la búsqueda de microRNAs, se analizó 1,000 pb río arriba con el programa miRBase (Kozomara *et al.*, 2019), con la opción miRNAs maduros y BLASTN para los tres genes *LCY* de achote. Las secuencias con E-value < 2.8 se seleccionaron, se confirmó su presencia en la secuencia de la región promotora de cada gen mediante alineamiento en la plataforma CDD (Lu *et al.*, 2020) con referencia en BioEdit (Hall T. A., 1999) y posteriormente se añadieron al Ugene 48.1 (Okonechnikov *et al.*, 2012) para visualizar su ubicación y organización en los genes. Para verificar su relación con los factores de transcripción, se analizó las secuencias miRNAs obtenidas en la base de datos PLACE para obtener sus secuencias *cis* regulatorias y sus TFs asociadas.

2.2.4. Secuencias repetitivas de DNA de los genes *BoLCY*

Para la búsqueda de regiones repetitivas, se analizó a 2,000pb río arriba para cada gen *LCY* en la base de datos Gypsy Database 2.0 (Llorens *et al.*, 2011). Estas secuencias se alinearon en la plataforma CDD (Lu *et al.*, 2020) con referencia en BioEdit (Hall T. A., 1999) y se analizaron en MEME suite 5.4.1 (Bailey *et al.*, 2015) para obtener secuencias repetitivas conservadas. Los motivos obtenidos, se analizó su abundancia y presencia en la región río arriba de los genes *LCY*. Los datos obtenidos se ubicaron en el contig de cada gen *LCY*, se agregaron a la estructura ya obtenida de los genes *LCY* para visualizarlos, mediante el programa Ugene 48.1 (Okonechnikov *et al.*, 2012). Este tipo de análisis nos brinda información para diseñar sondas que nos permitirá conocer la ubicación y organización de los genes *LCY* en los cromosomas de la especie.

2.3. RESULTADOS

2.3.1. Secuencias de los genes *LCY* de achote y sus regiones conservadas

El análisis BLASTx a la base de datos secuenciación de molécula única que consta de 39,440 contigs, logró identificar a los genes en tres contigs que contienen secuencias de >95% de cobertura: *Boβ-LCY1* en contig 113, *Boβ-LCY2* en contig 74 y *Boε-LCY* en contig 1,290, así como determinar el rango de su presencia en el contig (Cuadro 2.1). Estas secuencias se analizaron en el programa ORFinder para los marcos de lectura y en Six frame translation para obtener las secuencias de aminoácidos correspondientes; estos se alinearon en Bioedit con la proteína ya reportada para cada gen, y se determinó las secuencias codificantes completas para *β-LCY1* y *β-LCY2* y no codificantes para *ε-LCY* (Cuadro 2.2).

Cuadro 2.1. Secuencias de los genes *BoLCY* obtenidos del genoma de *Bixa orellana*, morfotipo B.

Gen	Contig	Longitud (bp)	Inicio región río arriba	Fin región río arriba	Inicio secuencia del gen	Fin secuencia del gen
<i>Boβ-LCY1</i>	000113	9,966,608	9,083,139	9,086,139	9,086,140	9,087,625
<i>Boβ-LCY2</i>	000074	6,753,459	2,495,907	2,498,907	2,498,908	2,500,402
<i>Boε-LCY</i>	001296	6,033,754	2,742,197	2,745,197	2,745,198	2,750,071

El gen *Boβ-LCY1* presentó una región codificante de 1,512 nts correspondiente a una proteína

de 504 aa; el gen *Boβ-LCY2* presentó una región codificante de 1,494 nts correspondiente a una proteína de 498 aa; y el gen *ε-LCY* presentó 11 exones con una región codificante de 1,604 nts correspondientes a una proteína de 527 aa (Cuadro 2.2). Con esto, se corroboraron las secuencias codificantes de los genes *LCY* reportados anteriormente mediante el análisis de transcriptoma para la especie (Cárdenas-Conejo *et al.*, 2015). En la secuencia del gen *Boβ-LCY1* se realizó la búsqueda de los SNPs reportados por Trujillo-Hdz y colaboradores (2016), al no encontrar ortólogos para el gen, solo se encontró la secuencia nucleotídica correspondiente al aminoácido y el nucleótido que presenta el cambio (Figura 2.1).

El alineamiento múltiple de las regiones conservadas y motivos característicos de los licopeno ciclasa: región conservada β -LCYs, sitio de unión a dinucleótidos, motivo específico LCYs, motivo ciclasa I, motivo ciclasa II, motivo β -LCYs y dominio de actividad catalítica β -LCY de *Bixa orellana*, *Daucus carota* y *Citrus sinensis* para el gen β LCY1 y, *Bixa orellana* y *Citrus sinensis* para el gen β LCY2; mostró completas y parciales las secuencias de las regiones conservadas y motivos característicos de los licopeno ciclasa en achiote (Figura 2.2). Los tres genes presentaron regiones altamente conservadas en sitio de unión de dinucleótidos, motivo ciclasa I y motivo ciclasa II. Sin embargo, los beta licopeno son las que mostraron mayor similitud de secuencia, principalmente β LCY1 que se mostró conservados en casi todas las secuencias características con *Daucus carota* (Moreno *et al.*, 2013) y *Citrus sinensis* (Alquézar *et al.*, 2009).

No obstante, β LCY2 también mostró una alta similitud en sitio de unión de dinucleótidos, motivo ciclasa I y dominio de actividad catalítica β -LCY pero con algunos cambios en aminoácidos en la región conservada β -LCYs, motivo específico LCY, motivo ciclasa II y motivo β -LCYs con respecto a *Citrus sinensis* (Alquézar *et al.*, 2009). Sin embargo, el fenotipo no se restringe a las secuencias codificantes; las regiones no codificantes (no traducidas) son regiones importantes involucradas en la regulación, estas moderan cuando, donde, nivel y tejido de la expresión (Koul *et al.*, 2018; Barrett *et al.* 2012; Wray *et al.*, 2003). Por esto, se realizó una serie de análisis para la región río arriba de los genes *LCY* para determinar los elementos que la conforman.

Cuadro 2.2. Estructura de los genes *LCY* de *B. orellana*

Gen	ORF (pb)	aa	Exon	Longitud (pb)	Inicio	Fin	Intron	Longitud (pb)	Inicio	Fin	Referencia (id)	ORF (pb)	No. Amino ácidos
<i>Boβ-LCY1</i>	1512	504	1	1512	9,086,140	9,087,625	NA				KT359004.1	1512	504
<i>Boβ-LCY2</i>	1495	498	1	1495	2,498,908	2,500,402	NA				KT359005.1	1494	498
			1	286	2,745,198	2,745,483	1	196	2,745,484	2,745,679			
			2	173	2,745,680	2,745,852	2	271	2,745,853	2,746,123			
			3	139	2,746,124	2,746,262	3	235	2,746,263	2,746,497			
			4	112	2,746,498	2,746,609	4	84	2,746,610	2,746,693			
			5	106	2,746,694	2,746,799	5	452	2,746,800	2,747,251			
<i>Boε-LCY</i>	1581	527	6	142	2,747,252	2,747,393	6	81	2,747,394	2,747,474	KT359006.1	1448	270
			7	109	2,747,475	2,747,583	7	842	2,747,584	2,748,425			
			8	100	2,748,426	2,748,525	8	84	2,748,526	2,748,609			
			9	127	2,748,610	2,748,736	9	665	2,748,737	2,749,401			
			10	136	2,749,402	2,749,537	10	371	2,749,538	2,749,908			
			11	163	2,749,909	2,750,071							

Figura 2.1. Alineamiento de secuencias de nucleótidos y aminoácidos de los genes *Boβ-LCY1*, *Boβ-LCY2* y *ε-LCY*.

Boβ-LCY1

```

βLCY1      M D T L L K T H N K L E F L P Q L H G F S E K V S I L N S S K L Q S K E L K L G S S R P C L R K G R N S C L K A N S S T L T E L V P E
nt6140     atggacacgctacttaagacgcacaacaagcttgagttcttccacaacttcattgggttttcggagaaagtgagtattttgaactcctccaagcttcaaaagcaaggagcttaagcttggcttagcagccttgcctgagaaaggtaggaatagttgccttaaggccaattctagcactctcacggaacttgttccoga

βLCY1      I K K E N L E F E L P L Y D P S K G L V V D L A V V G G G P A G L A V A Q Q V S E A G L S V C T I D P S P K L I W P N N Y G V W V D
nt...     aaacaagaaaggagaatcttgagtttgagcttctctctatgaccatcaaaaagccttgggtggatctagctgtttaggaggtggccctgcaggtctcgcctgttgcagcaggttccagagcagggcttgcctgacagattgaccccttcccaaatgattggcctaataattatgggtttgggtggatg

βLCY1      E F E A M D L L D C L D T T W S G A V V Y I N D Q T K K D L D R P Y G R V N R R Q L K S K M L Q K C I S N G V K F H Q A K V I K V I H
nt...     aatttgaggctatggatttgcctgattgcttgcaccacctggagtggtgctgttttacatataatgaccaacaagaaggatcttagatagaccttatggagagtaataagaagcagctgaagtcctcaaatgctgcagaaatgcataatcaatgggttcaaatttcatcaggcttaaggttaaggttatccat

βLCY1      E Q S K S L L I C N D G V T V Q A T L V L D A T G F S R C L V Q Y D K P Y N P G Y Q V A Y G I L A E V E E H P F D V D K M V F M D W R
nt...     gagcagtcctcaaatctctattgattgcaatgattggtttacaggttcaggctaccttgggttctgatgcaacaggggtttctaggtgcttgcagtagataagccctataatcctggttaccaagtggttatggaatcttagccaggtagaagaacatccattgtagtggaataagatgggttttatggattggag

βLCY1      D S H L T N N K E L R E R N S K I P T F L Y A M P F S S E R I F L E E T S L V A R P G V Q M K D I Q E R M V A R L R H L G I K V K S
nt...     agattcacacctcaccacaataagagctgagggagagaataatagataccaactttctctatgcaatgccctttcatctgagagaatattcttgaagagacttcccttggctgcagctgagtagacaatgaaggacatccaggaagaatgggtggtaggttaagacacttaggcataaaagtgaagta

βLCY1      I E E D E H C V I P M G G P L P V L P Q R V V G I G G T A G M V H P Q L G Y M V A R T L A A A P I V A N S I V Q C L G S E R S L S A D
nt...     ttgaagagtagagcattggtgattccgagggcgacccttctctgcttccccaaagagttytggaaatcgggtgaaacagctgggatggtgacctcaactgggttagcatgtagcagctccaattggtgcaactccattgctcagtccttggtaagagagcctatcagcagat

βLCY1      V L S E V V W K D L W P I E R R R Q R E F F C F G M D I L L K L D L Q A T R R F F D A F F N L E P R Y W H G F L S S R L F L P D L L F
nt...     gttgtgctgaagtagtttgaaagatttatggccatagaagagagagagaggaatctctgttttggcatggatattctgctgaagctagattacaagctacaagaagattttttagtcatttttaacttggaaacctgtagattggcatggattcttgcactctcctccagacctttatt

βLCY1      L G L S L F S N A S N T S R I E I M A K G T L P L V K M I N N L I Q D S D
nt7652     tttggcctttctctattctctaatgcttctaatctcttagatagagattatggccaagggagccttcttggtaagatgatcaacaacttaatacaggatagcagat
    
```

Boβ-LCY2

```

βLCY2      M G T L V K L C S P A F T T R N K F F G S P S V T T P F T S A K T Q I K K T S S K T H Q K V R V S K F G N F L E L K P E S Q R E F L D
nt498908   atggaaactctgctcaagttatgttcaccggctttacaaccaggaacaatttttggttcaccctcagtcaccactcaattacttctgcaaaaaccagatcaagaagacctcaagcaagacctcagaagaatcagagtcagcaaatgggaacttctggaatgaaagccagagtcacagcgggagttcttggga

βLCY2      Y D L P L F D P S H E R P R F D V I V I G T G P A G L R L A E Q V S G R G I K V C C V D P S P L S M W S N N Y G V W V D E F K S L G
nt...     ttatgatctgcttgtttgatccatcccatgagcggcctcgtttgatgtagctctcggcaccggcccgaggtctccgtctcggcagcaagtttctggcctgggaatcaaggtatgttgcctgcagctccctcaccacttccatgtggtctaaacattatggttctgggtagatgagttcaagagcttggggc

βLCY2      L D D C L D K T W P M T V H I D D K K T K Y L D R P Y C R V S R K K L K S R L M E S C T S N G V K F H K A K V W K V R H E E F E S I
nt...     tctgatgattttggacaagacatggccgatgactacggtccatagatgataaaaaacaaatctctggaccgacctgactgtgctgactagaaagaagctcaagcttagatataatggagctctgcagctcgaacggagttaaagttccataaagccaaggttggaaagtgaagaccgaagaattgagtcata

βLCY2      I E C D D G S E F K A S L I V D A S G F A S P F I K Y D K P R N D G Y Q L A H G I L A E V E S H P F D L D K M V L M D W R D S H L A N
nt...     atcaggtgtagcagtagggagtagttcaagccagcttgatcgtcagatgagagtgatctgcaagcccttccataaaatagcagaaagcctggaagctgggtatcaactcggccatggcattctggctgaagtgaagaagcctcttccgattggacaagatggttctatggattggagagattcccatgtggcaaa

βLCY2      E P D L R E R N A K L P T F L Y V M P F D S N L V F L E E T S L V G R P V L P Y M E V K Q R M V A R L R H L G I I V K R V I E E E K
nt...     tgagcctgacttacgcgaacgaacgcaagctcccaactttctctatgtaatgccattcgattcgaatttgggtttcttggaaagacttctctggtcggctgcagcgggtctcggctacatggaggtgaacagagagtaggtgagcaggctgaggcatttagaatcatagtagaagagatgatagaagaagaaaaat

βLCY2      C L I P M G G P L P K L P Q N V L A M G G N S G L V H P S T G Y T V Y R T M A M A P V V A D A I A E C L G S T R M M R G S L L H R R V
nt...     gcctgattcccaatggcggcctttccaaaactccctcagaatgattggcagtagggaggaatcttggcctgtaaaaatccgctgacaggtgtaccgagcagtagggcctggttggcggagctattgagcagaggtttgggttcgacgagtagatgagagggagcctttacatcggagaggt

βLCY2      W N G L W P M E R R V E R Q L Y C F G M E T L L N L D L A G T R R F M G E F F Q L D P Y Q W K G F L S Q S L S T P E I V A L L S T M F
nt...     tggaaacgactgtggcagtaggagagagagtagtagcggcaattatactgttttgggatggagactttgctgaatctgatttgcgggaccagggcagcttccagggcaattctccagctggatccttcaatggaagagcttctctcgaagcctcgcagcccgaatttggctctattgagtagcagatgtt

βLCY2      I K A S V P C K I D L A T K C P L P M A R T L G K I A L D A I
nt500402   cataaaagcatcgggtccatgcaagatcagatttggctaccaagtgcctttaccgatggctcgcagctgggaaaaattgcaactgacgccatc
    
```

Boε-LCY

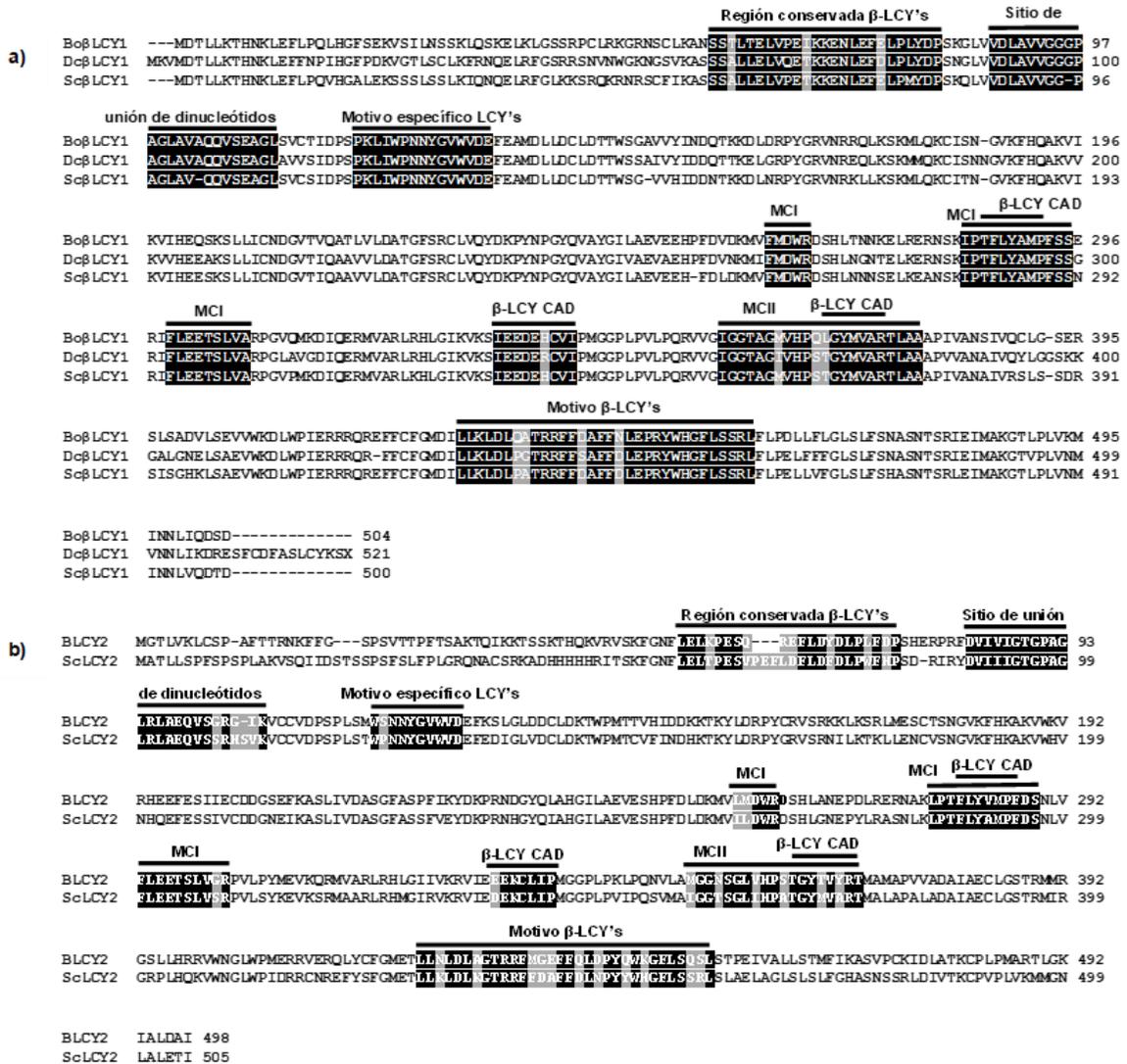


Figura 2.2. Alineamiento de regiones características de la beta licopeno ciclasa. a) β-LCY1, *Bixa orellana*, *Daucus carota* (Moreno *et al.*, 2013) y *Citrus sinensis* (Alquézar *et al.*, 2009). b) β-LCY2, *Bixa orellana* y *Citrus sinensis* (Alquézar *et al.*, 2009). Regiones características de los genes LCY: región conservada β-LCYs, sitio de unión de dinucleótidos, motivo ciclasa I, II (CDMI, CDMII), motivo β-LCYs y dominio de actividad catalítica β-LCY (Hugueyney *et al.*, 1995; Cunningham *et al.*, 1996). Los aminoácidos con fondo negro y letras blancas son idénticos en todas las secuencias, las letras blancas con fondo gris son aminoácidos diferentes.

2.3.2. Factores de transcripción putativos de los genes *BoLCY*

Con base a las secuencias codificantes obtenidas para los genes del genoma, se analizó los elementos regulatorios *cis* presentes en su región promotora, así como sus factores de transcripción putativos asociados (Anexo 1). Se observaron dos grupos principales de TFs; el primer grupo se divide en dos subgrupos: el subgrupo I con las categorías de fotosíntesis y reproducción, y el subgrupo II con las categorías de sistema oxidativo y defensa frente a factores ambientales (Fig. 2.3) (Anexo 2). El primer subgrupo de TFs cubre la categoría de fotosíntesis con los motivos CACTFTPPCA1 (5'-CACT-3'), DOFCOREZM (5'-AAAG-3'), y ROOTMOTIFTAPOX1 (5'-ATATT-3') (Figura. 2.4a). En cuanto a la reproducción, los motivos CAATBOX1 (5'-CAAT-3'), POLLEN1LELAT52 (5'-AGAAA-3') y GTGANTG10 (5'-GTGA-3') tuvieron una mayor presencia (Fig. 2.4b). El subgrupo II comprende las categorías de sistema oxidativo y estrés ambiental. En el caso del sistema oxidativo, los motivos GT1CONSENSUS (5'-GRWAAW-3'), GATABOX (5'-GATA-3') y EBOXBNNAPA (5'-CANNTG-3') están regulados en respuesta a la luz (Fig. 2c). Para la categoría de estrés ambiental, sólo los motivos MYCCONSENSUSAT (5'-CANNTG-3') emergieron como un TF abundante (Fig.2.4d).

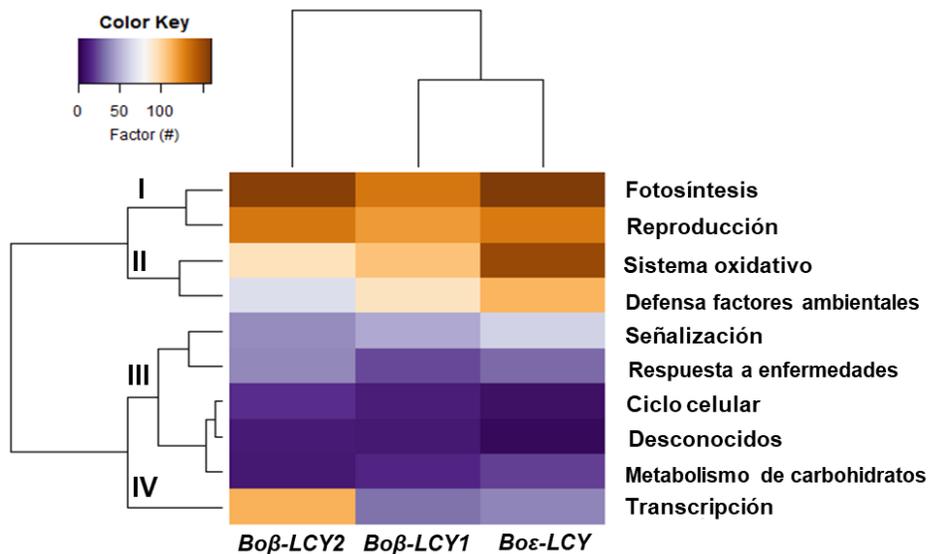


Figura 2.3. Mapa de calor de categorías de los elementos *cis* regulatorios de los genes *LCY* de *Bixa orellana*.

El segundo grupo también cuenta con dos subgrupos, el subgrupo III cuenta con cinco categorías con TFs implicados en: señalización, respuesta a enfermedades, ciclo celular, desconocidos y metabolismo de carbohidratos. El subgrupo IV cuenta con una categoría y un factor con mayor repetición, transcripción: POLASIG3 (5'-AATAAT-3') con un 1% implicado en la activación, expresión e iniciación de la transcripción (Higo *et al.*, 1999). Las secuencias de los elementos regulatorios *cis* identificados, junto con las secuencias codificantes para cada gen, se ingresaron en el programa UGENE para visualizar la estructura de los genes *BoLCY* de achote (Figura 2.5).

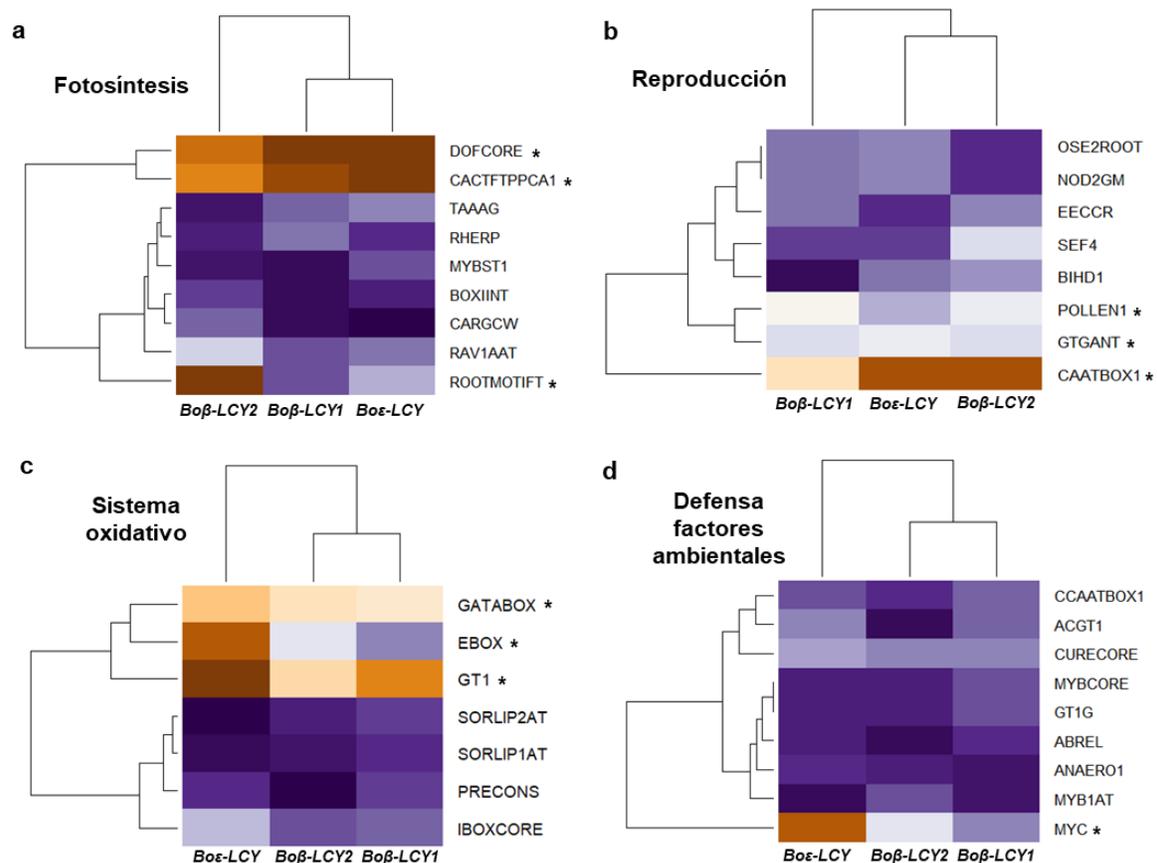


Figura 2.4. Mapa de calor de las categorías principales de elementos *cis* y sus factores de transcripción putativos asociados: a) Fotosíntesis (DOFCOREZM, CACTFTPPCA1, ROOTMOTIFTAPOX1); b) Reproducción (CAATBOX1, POLLEN1LELAT52, GTGANTG10); c) Sistema oxidativo (GT1CONSENSUS, GATABOX, EBOXBNNAPA) y d) Defensa a factores ambientales (MYCCONSENSUSAT).

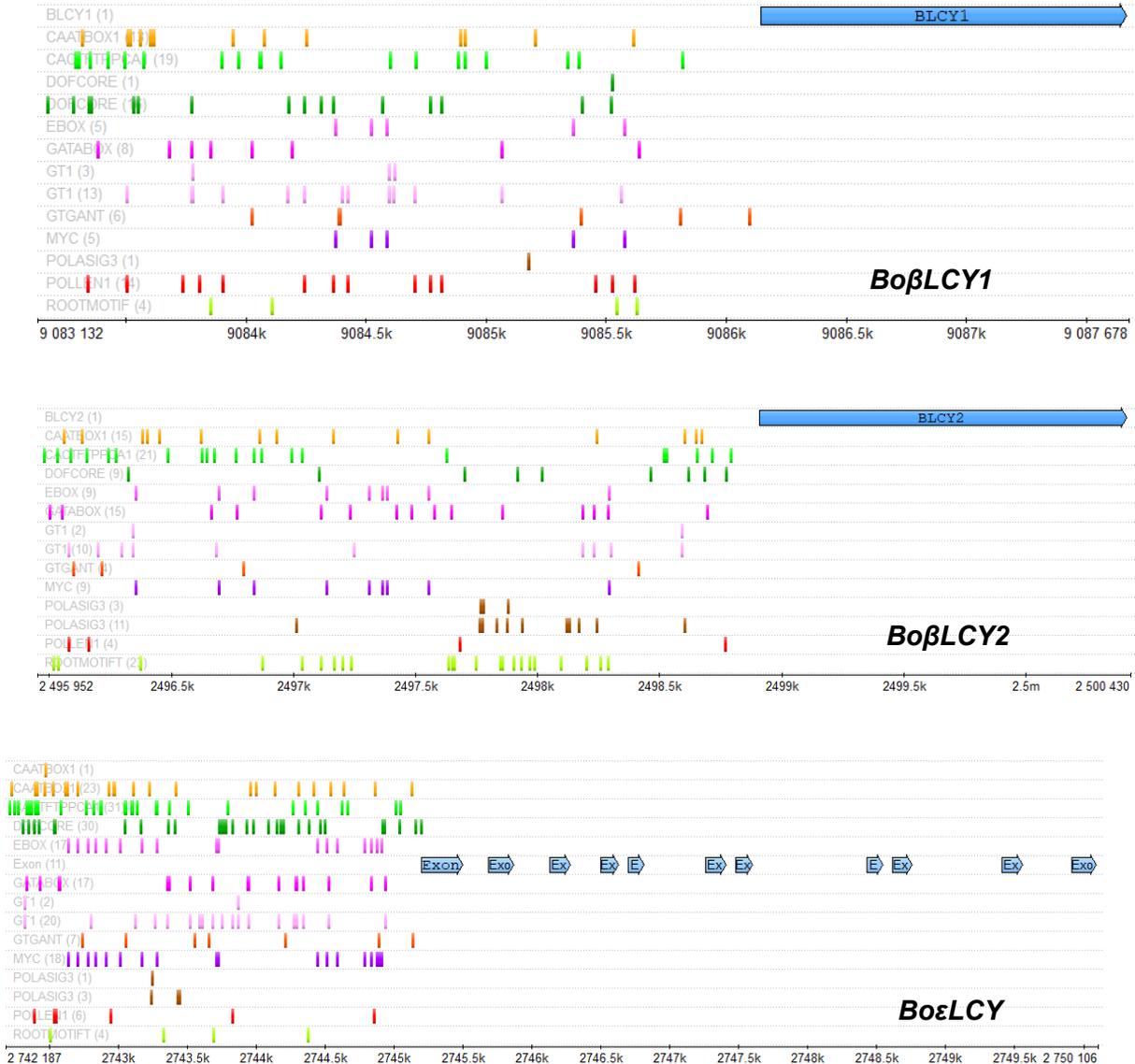


Figura 2.5. Estructura de los genes *BoLCY* en Ugene. En color azul, secuencias codificantes; en color verde factores putativos de fotosíntesis; rojo reproducción; lila sistema oxidativo; morado defensa a factores ambientales y color café transcripción. 3,000pb.

Los TF de la categoría fotosíntesis (verde) (CACTFTPPCA1, DOFCOREZM y ROOTMOTIFTAPOX1) se distribuyeron homogéneamente en los tres genes *BoLCY*, al igual que CAATBOX1 de la categoría reproducción (amarillo-naranja), pero solo en el gen *Boε-LCY*. Se encontraron 6 hotzones a lo largo de los 1500pb río arriba de los genes *Bo-LCY*. Para la región más cercana al de inicio de la transcripción (250-320pb) los TF presentes fueron de la categoría fotosíntesis (DOFCOREZM) para *Boβ-LCY1*, reproducción (CAATBOX1) y sistema oxidativo

(GATABOX/EBOXBNNAPA) para *Boβ-LCY2* y *BoεLCY* respectivamente, y defensa a factores ambientales (MYCCONSENSUSAT) solo para *Boε-LCY*. Para la región la región cercana a los 1000pb (930pb) los TF fueron de la categoría reproducción (CAATBOX1) para *Boβ-LCY1*; y fotosíntesis (ROOTMOTIFTAPOX1/DOFCOREZM) *Boβ-LCY2* y *BoεLCY* respectivamente, y sistema oxidativo (GT1CONSENSUS) solo para *Boε-LCY*. La hotzone con el mayor número de TF (650-770pb), con las categorías reproducción (POLLEN1LELAT52) y fotosíntesis (DOFCOREZM) para *Boβ-LCY1*; fotosíntesis (ROOTMOTIFTAPOX1), reproducción (CAATBOX1), sistema oxidativo (GT1CONSENSUS, GATABOX) y Transcripción (POLASIG3) para *Boβ-LCY2* y reproducción (CAATBOX1), sistema oxidativo (GT1CONSENSUS, GATABOX, EBOXBNNAPA), fotosíntesis (CACTFTPPCA1) y defensa a factores ambientales (MYCCONSENSUSAT) para *Boε-LCY* (Figura 2.6).

En cuanto a los genes *LCY*; *Boε-LCY* presenta el mayor número de secuencias de sitios de unión de TF y repeticiones en las categorías: fotosíntesis, reproducción, sistema oxidativo y defensa a factores ambientales; al contrario del gen *Boβ-LCY1*, que tiene el menor número de sitios en las mismas categorías para sus TF. *Boβ-LCY2* es quien muestra diferentes categorías de TF para sus sitios unión: fotosíntesis, sistema oxidativo y transcripción. En general, los *BoLCYs* muestran diferentes sitios de unión reguladores muchos de los cuales están orientados a la producción y acumulación de carotenoides en diferentes tejidos (fotosintéticos y no fotosintéticos) para cumplir funciones esenciales en la planta.

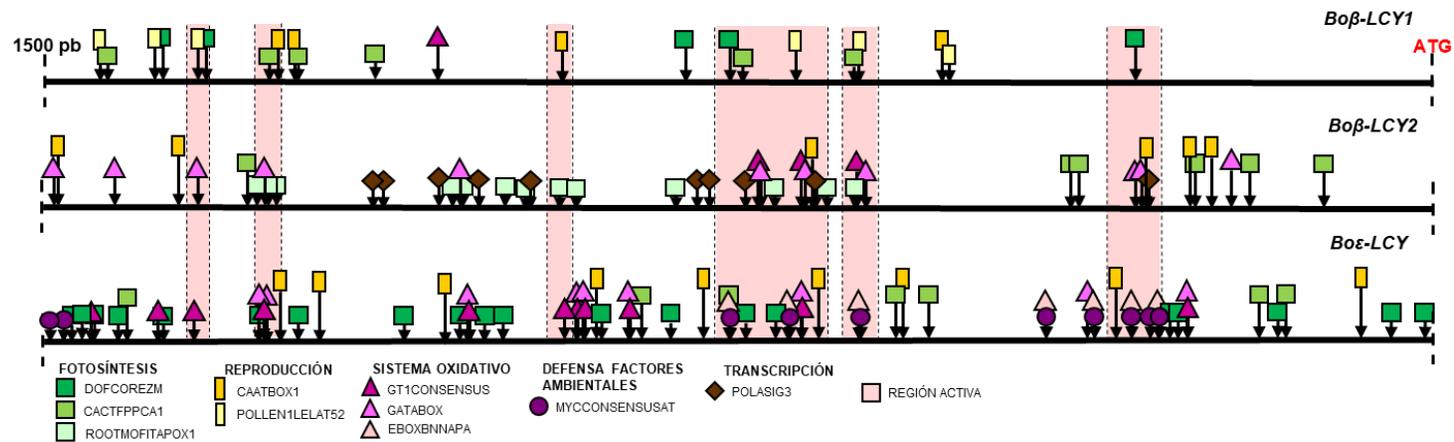


Figura 2.6. Sitios de unión de los factores de transcripción putativos para los genes *BoLCY*. 1,500 pb río arriba.

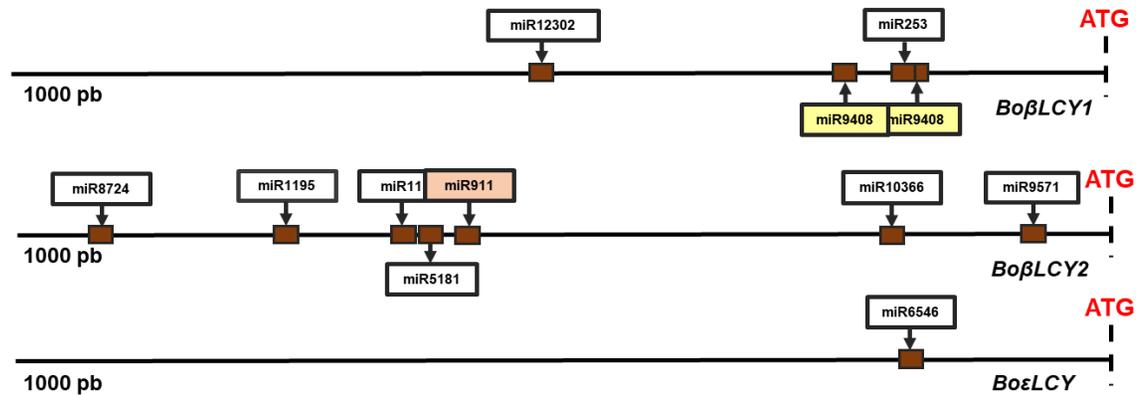


Figura 2.7. Secuencias miRNAs identificadas para los genes *BoLCY*, 1,000pb río arriba. (Amarillo; naranja, miRNAs que pertenecen a la misma familia).

2.3.3. miRNAs de los genes *BoLCY*

Los genes *LCY* presentaron 11 familias de miRNAs, las cuales se clasifican como no conservados; los cuales, suelen ser específicos de la especie o intervienen en procesos únicos para la planta (Sun *et al.*, 2012). Para *Boβ-LCY1*, presentó tres familias: miR253, miR9408 (con dos presencias) y miR12302; ubicadas entre los 490 a 820 pb río arriba (Figura 2.7). *BoβLCY2* presentó 7 familias: miR991 (con dos presencias), miR1195, miR11, miR5181c, miR8724, miR9571a y miR10366a, los cuales se encontraron dispersos en toda la región promotora (Figura 2.7). En cuanto a *εLCY*, solo presentó una secuencia: miR6546, ubicado en la posición 803 pb río arriba (Figura 2.7). Se sabe que la mayoría de los objetivos de los miRNAs son mRNAs que codifica para factores de transcripción, siendo en conjunto, un mecanismo de regulación recíproca para la expresión génica (Mejia-Guerra *et al.*, 2012). Por esto, se verificó que secuencias miRNAs presentes en la región río arriba están asociados a factores de transcripción, mediante su análisis en la base de datos PLACE. De las 11 familias, solo tres miRNAs tuvieron homólogos en plantas: miR9408 en *Brassica aleracea* y *B. rapa* involucrado en el desarrollo del polen (Lukasik *et al.*, 2013); miR5181 en *Brachypodium distachyon* implicado crecimiento de hojas durante la sequía (Bertolini *et al.*, 2013) y miR8724 en *Gossypium raimondii* involucrado en el alargamiento de las células de fibra de algodón (Xue *et al.*, 2013). Con esto, podría suponerse que los miRNAs asociados a los genes *LCY* de achiote, podrían estar inmersos en procesos de desarrollo y respuesta a estrés específicos de la planta.

Para el gen *Boβ-LCY1*, el miRNA 12,302 presentó secuencias para elementos reguladores *cis* asociadas a los factores de transcripción: POLLEN1LELAT52, ROOTMOTIFTAPOX1 y GATABOX. En el gen *Boβ-LCY2*, el miRNA 10,366a presentó secuencia asociada a CACTFTPPCA1 y miRNA 5,181c con MYCCONSENSUSAT y EBOXBNNAPA. *Boε-LCY* no presentó secuencias regulatorias *cis*. Los dominios de las secuencias miRNAs identificadas se vinculan a factores de transcripción con funciones importantes para la planta, como POLLEN1LELAT52 en la reproducción, específicamente en la maduración del polen; CACTFTPPCA1 y ROOTMOTIFTAPOX1 en la fotosíntesis; EBOXBNNAPA para sistema oxidativo; y MYCCONSENSUSAT para defensa a factores ambientales.

2.3.4. Secuencias repetitivas de DNA de los genes *BoLCY*

El análisis de secuencias repetitivas para los genes *LCY* de achiote, identificaron tres motivos de retrotransposones de repetición de terminal larga (LTR) asociados a los genes *LCY*. El motivo 1

(azul) se presentó en *Boβ-LCY1* y *Boε-LCY*, correspondiente al LTR Ogre perteneciente a la familia Ty3/Gypsy; el motivo 2 (naranja) y 3 (rojo) se presentaron únicamente en el gen *Boε-LCY*, correspondiente a los LTR Del y Diaspora de la familia Ty3/Gypsy; siendo *Boβ-LCY2* el gen que no presentó motivos, pero se identificó al LTR V12 (verde) correspondiente a la familia Ty1/Copia (Figura 2.8). Estos motivos serían correspondientes a dominios LTR, por lo que se confirmó su presencia en la región río arriba de los *BoLCY*. También, se identificaron secuencias de dominios de proteínas asociados a los motivos LTR; *Boβ-LCY1* presentó al dominio GAG y *Boβ-LCY2* a RNaseH (Figura 2.8).

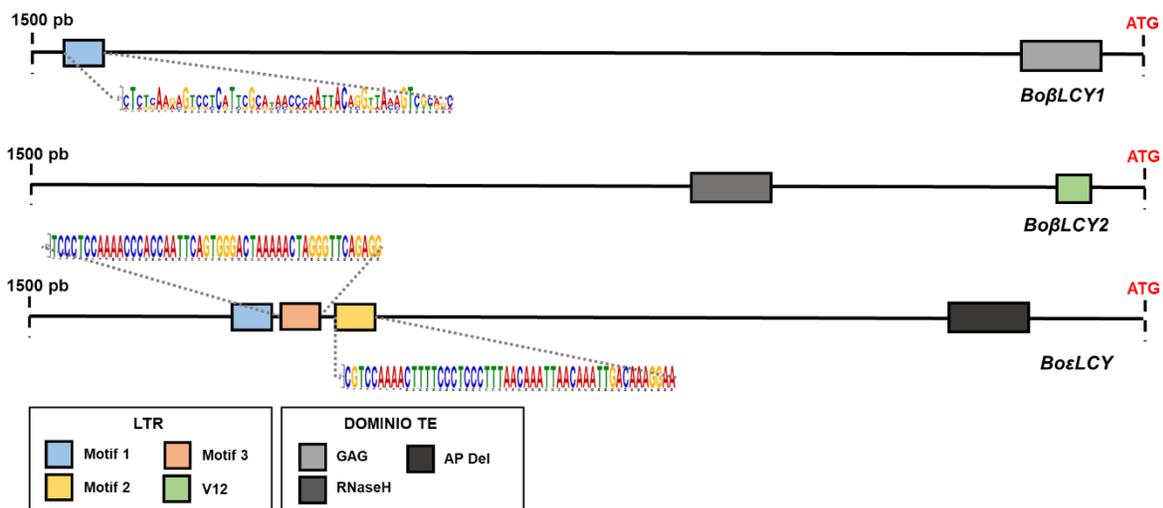


Figura 2.8. Motivos LTR y dominios relacionados a proteína de los genes *BoLCY*.

Boε-LCY que muestra tres motivos asociados a LTR de la familia Ty3/Gypsy (Ogre, Del y Diaspora), es el gen que se presenta, según nuestro análisis de factores de transcripción, el mayor potencial de regulación de los genes *LCY* (Figura 2.9). Esto podría ser resultado del mecanismo de integración de los extremos cohesivos de los fragmentos retrotransponibles; lo cual podría explicar en parte, el bajo número de TF presentes en *Boβ-LCY1* con solo un motivo LTR (Ogre) pero de categorías similares a épsilon ciclasa, involucrados en funciones de captación de luz, fotooxidación y reproducción. *Boβ-LCY2* no presentó motivos, sin embargo, se identificó un LTR de la familia Ty1/Copia (V12), el cual estaría vinculado a diferentes factores de transcripción, pero siempre inmersos en procesos de fotosíntesis y fotoprotección, vitales para la regulación de la planta durante sus diferentes etapas de desarrollo.

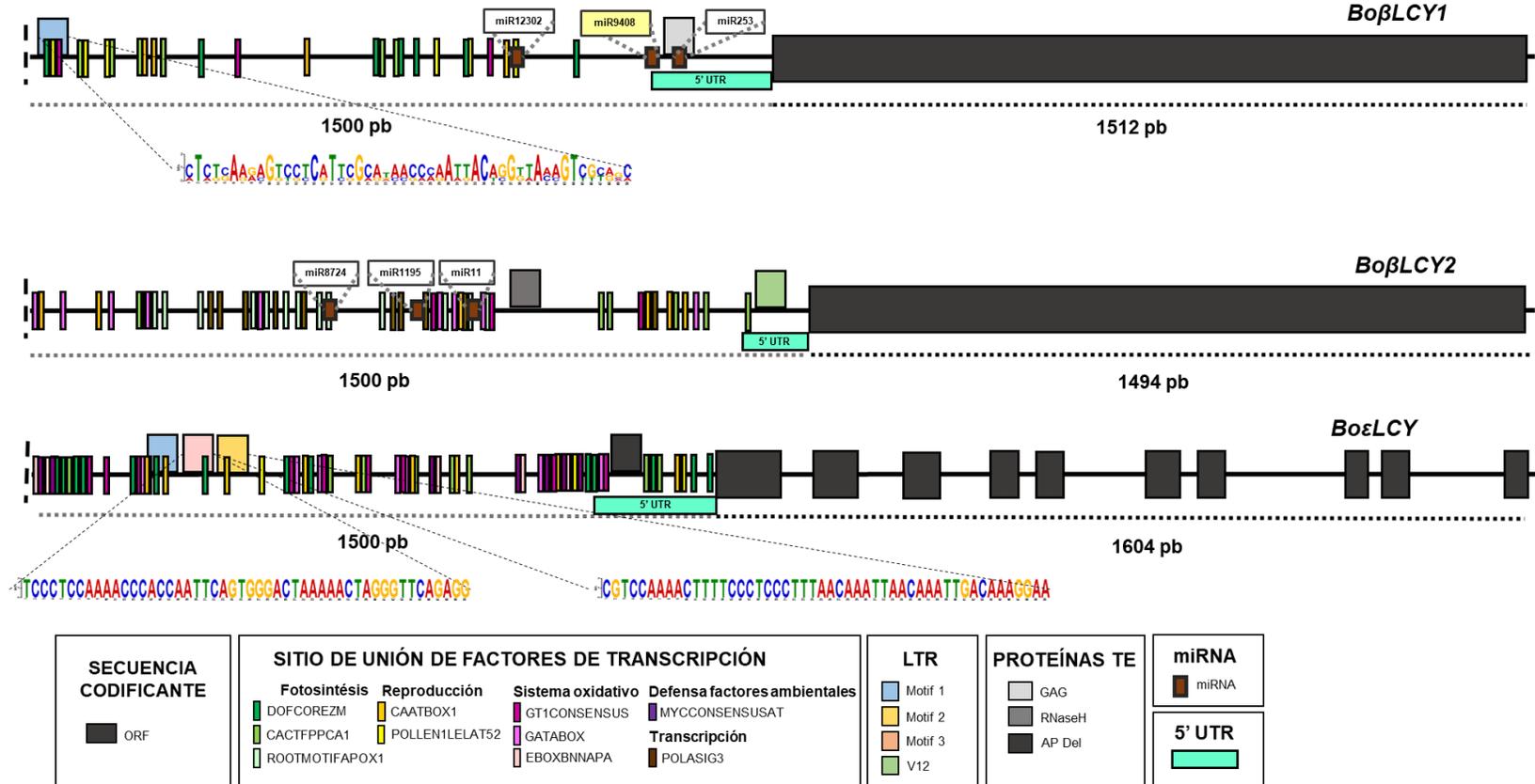


Figura 2.9. Estructura de los genes *LCY* de achote. *Boβ-LCY1*, ORF de 1512bp. En región usptream, principales TFs putativos: fotosíntesis (DOFCOREZM, CACTFTPPCA1) y reproducción (CAATBOX1); motivo 1 LTR y dominio de proteína GAG; miRNAs asociados a TFs: miR12302, miR9408 y miR253. *Boβ-LCY2*, ORF de 1495bp. En región río arriba, principales Tfs putativos: fotosíntesis (ROOTMOTIFAPOX1) y reproducción (CAATBOX1); V12 LTR y dominio de proteína RNaseH; miRNAs asociados a TFs: miR8724, miR1195 y miR11. *Boε-LCY* con 11 exones y 10 intrones, ORF de 1581 bp. En region río arriba, principales Tfs putativos: fotosíntesis (DOFCOREZM, CACTFTPPCA1) y sistema oxidativo (GT1CONSENSUS, EBOXBNNAPA); motivo 1, 2 y 3 LTR y dominio de proteína AP Del.

2.4. DISCUSIÓN

2.4.1 Caracterización de los genes *LCY* de Achiote

El análisis de la base de datos de secuenciación del genoma, confirmó las secuencias codificantes de los genes *LCY* de achiote ya reportados por Cárdenas-Conejo y colaboradores (2015). *Boβ-LCY1* presentó 1,512pb con una proteína de 504aa y una identidad de 100% con *B. orellana* (KT359006.1), 84% con *Theobroma cacao* (LT594792.1) y 82% con *Vitis vinífera* (AM432668.1) y *Citrus sinensis* (MG492005.1). Los cambios puntuales reportados por Trujillo-Hdz y colaboradores (2016) en la secuencia de *Boβ-LCY1* que determinaron tres grupos con diferencias en producción de bixina no se observaron; sin embargo, esto puede deberse a que nosotros solo trabajamos con una accesión rosa. *Boβ-LCY2* tuvo 1494pb y una proteína de 498aa e identidad de 100% con *B. orellana* (KT359005.1), y 74% con *Actinidia chinensis* (FJ427508.1) y *Actinidia deliciosa* (FJ427509). Ambos *Boβ-LCY* se caracterizan por presentar una sola secuencia codificante. *Boε-LCY* presentó 1,604pb y 527aa, con identidad de 99% pero una cobertura de 28% con la accesión reportada por Trujillo-Hdz y colaboradores (2016) (KT359006.1). *Boε-LCY*, a diferencia de los genes beta, tiene una estructura conformada por 11 exones de entre 286 a 100pb y 10 intrones de entre 843 a 78pb distribuidas en una longitud de 4,871pb. El análisis de identidad mostró 90 % con *Citrus sinensis* (AY533827.1), 87% con *Solanum lycopersicum* (CP023768.1) y 84% con *Theobroma cacao* (LT594796.1), pero con coberturas mínimas de 1, 5 y 7% respectivamente, correspondientes, posiblemente a solo un exón. Por lo que las secuencias de β - y ϵ - de achiote, cuentan con el número de aminoácidos aproximado de 400aa y con una similitud significativa reportada para las ciclasas de plantas (Rosas-Saavedra y Stange, 2016).

En cuanto a las regiones características de los genes licopeno ciclasa; los genes presentaron regiones altamente conservadas en: sitio de unión de dinucleótidos, una firma presente en todo caroteno ciclasa conocidas (Cunningham *et al.*, 1996), motivo ciclasa I y motivo ciclasa II, dominios esenciales para la actividad catalítica (Alquézar *et al.*, 2009). Las β LCYs fueron las que presentaron las secuencias mejor conservadas; principalmente *Boβ-LCY1* que se mostró todas las secuencias características de los licopeno ciclasa conservadas con *Daucus carota* (Moreno *et al.*, 2013) y *Citrus sinensis* (Alquézar *et al.*, 2009), denotando su importancia en la síntesis y regulación de carotenoides en los cloroplastos. No obstante, *Boβ-LCY2* también mostró una alta similitud en las secuencias de sitio de unión de dinucleótidos, motivo ciclasa I y dominio de actividad catalítica β -LCY, pero con algunos cambios de aminoácidos en la región conservada β -

LCYs, motivo específico LCY, motivo ciclasa II y motivo β -LCYs con respecto a *Citrus sinensis* (Alqu  zar *et al.*, 2009). Esto podr  a deberse a que β -LCY2 se asocia a cromoplastos y en algunas especies, el gen se presenta como isoforma, lo cual podr  a tener cambios requeridos para la acumulaci  n de pigmentos en tejidos no fotosint  ticos. No obstante, se observa que los tres genes LCY tienen grandes similitudes en sus secuencias de amino  cidos, donde, Rosas-Saavedra y Stange (2016) mencionan que podr  a deberse a que se originaron a partir de un ancestro com  n, quiz  a a trav  s de un proceso de duplicaci  n.

2.4.2. Elementos regulatorios *cis* de los genes *BoLCY*

En el presente estudio, se identificaron dos grupos importantes de factores de transcripci  n asociados principalmente a funciones de: I. fotos  ntesis (DOFCOREZM y CACTFTPPCA1) y reproducci  n (CAATBOX1, POLLEN1LELAT52 y GTGANTG10) y II. sistema oxidativo (GATABOX y GT1CONSENSUS) y defensa a factores ambientales (MYCCONSensusAT). Todas ellas con funciones esenciales en las plantas, llevadas a cabo por los carotenoides sintetizados *de novo* en los pl  stidos de los organismos fotosint  ticos y no fotosint  ticos (Sun *et al.*, 2022; Maoka, 2020; Nisar *et al.*, 2016; Rosas-Saavedra y Stange, 2016). Para la categor  a fotos  ntesis, sus factores est  n relacionados con la expresi  n de brotes, ra  ces y espec  ficamente al mes  filo de las hojas, donde los carotenoides captan y absorben la luz en rangos donde la clorofila no es funcional durante la fotos  ntesis (380 a 500nm) (Lohani *et al.*, 2021; Aguilar-Espinosa *et al.*, 2020; Gowik *et al.*, 2004; Yanagisawa 2002; Higo *et al.*, 1999). DOFCOREZM fue de especial atenci  n, este cuenta con funciones clave en diversos tejidos como fotos  ntesis, floraci  n, germinaci  n y desarrollo de semillas reportado para *Arabidopsis*, almacenamiento en las semillas en tomate (Swinnen *et al.*, 2016) y tolerancia al estr  s abiotico en *Glycine max*, *Medicago truncatula*, *Oryza sativa*, *Phaseolus vulgaris* y *Populus trichocarpa* (Benson *et al.*, 2014). CACTFTPPCA1 componente del m  dulo de expresi  n de mes  filo, es parte del promotor del gen fosfoenolpiruvato carboxilasa C4 en la *Flaveria trinervia*; sin embargo, en las plantas C3 se asume tendr  a una funci  n diferente, y en c  tricos como un factor de respuesta al hambre asociado al gen *LCYB1* (Lu *et al.*, 2021). En cuanto a reproducci  n, cuenta con los TF de mayor repetici  n a: CAATBOX1 (5'-CAAT-3'), POLLEN1LELAT52 (5'-AGAAA-3') y GTGANTG10 (5'-GTGA-3') (Figura 2.4b) relacionados con la expresi  n de polen y semilla. CAATBOX1 es promotor responsable de la actividad en tejido de semilla del ch  charo (Li *et al.*, 2014; Sharma *et al.*, 2011) y POLLEN1LELAT52 es uno de dos elementos reguladores que activan de manera espec  fica al polen del gen *LAT52* en *Lycopersicon esculentum* (Li *et al.*, 2014). En achiote, el efecto indirecto de la protecci  n al estr  s foto-oxidativo por la acumulaci  n

de carotenoides, proporcionan a las flores, frutos y semillas tonalidades entre amarillas y rojizas, importantes para la reproducción de la planta al atraer polinizadores, garantizando la reproducción de la planta (Dzib-Cauich *et al.*, 2020; Rogers *et al.*, 2001; Higo *et al.*, 1999; Bate *et al.*, 1998).

El subgrupo II comprende a las categorías sistema oxidativo y defensa a factores ambientales. Para sistema oxidativo, los TF con mayores repeticiones fueron: GT1CONSENSUS (5'-GRWAAW-3'), GATABOX (5'-GATA-3') y EBOXBNNAPA (5'-CANNTG-3') (Figura 2.4c) regulados como respuesta a la luz, donde EBOXBNNAPA también se le asocia con el almacenamiento en semillas (Li *et al.*, 2014). En tejidos fotosintéticos la señal más importante para la regulación de la expresión de genes estructurales de los carotenoides es la luz (Sun *et al.*, 2022). Estos ayudan a la membrana al defender al aparato fotosintético de la fotooxidación, acumulándose en sus tejidos; en achote, esto le confiere a sus frutos y semillas colores amarillos rojizos característicos (Dzib-Cauich *et al.*, 2020; Muthamilarasan *et al.*, 2014; Teakle *et al.*, 2002; Reyes *et al.*, 2004; Zhou *et al.*, 1999; Higo *et al.*, 1999). En la categoría defensa a factores ambientales, se presentó un solo factor de transcripción con el mayor número de repeticiones: MYCCONSUSAT (5'-CANNTG-3') (Figura 2.4d), como respuesta al estrés biótico y abiótico. Este es un sitio de reconocimiento MYC involucrado a respuestas de defensa abióticos como ataques a patógenos, heridas (Deb *et al.*, 2015) y bióticos como deshidratación y frío (Li *et al.*, 2014). Este es inducido por ABA, un apocarotenoide importante en la respuesta al estrés ambiental, fitohormona resultado de la escisión de los carotenoides la cual, también está relacionada con procesos de desarrollo y crecimiento en la planta (Rivera-Madrid & Ramamoorthy, 2020; Higo *et al.*, 1999).

No obstante, los genes LCY presentaron diferente número de repeticiones de TFs. *Boβ-LCY1* presentó la menor repetición de TFs, indicando un menor nivel de regulación génica. Este se expresa en los cloroplastos, acumulando carotenoides en las membranas de los tilacoides y complejos de recolección de luz (Trujillo-Hernández, 2016; Rosas-Saavedra y Stange, 2016; Ruiz-Sola y Rodríguez-Concepción 2012), actuando como refuerzo en la absorción de luz en rangos donde la clorofila no es funcional durante la fotosíntesis (380 a 500nm) (Yanagisawa, 2002; Gowik *et al.*, 2004; Aguilar-Espinosa *et al.*, 2020; Lohani *et al.*, 2021). En tejidos fotosintéticos y raíces, como la zanahoria (Llorente *et al.* 2017), es la señal más importante para la regulación de la expresión de genes estructurales de los carotenoides es la luz (Sun *et al.*, 2022). Sin embargo, para los órganos reproductores, las señales de desarrollo parecieran ser el

determinante principal de la regulación de los *LCY* (Sun y Li, 2020; Sun *et al.*, 2022), como se ha reportado en *C. papaya* y *Actinidia spp.* (Zhou *et al.* 2019; Ampomah-Dwamena *et al.* 2009). En tomate, durante la maduración del fruto, los cloroplastos sufren una transición a cromoplastos (tejidos no fotosintéticos) (Sun *et al.*, 2022; Escobar-Turriza, 2017; Bian *et al.*, 2011), sitio principal de acumulación de diversos y grandes cantidades de carotenoides (Sun *et al.*, 2022). En achiote, *Boβ-LCY2* presentó un mayor número de repeticiones de TFs que *Boβ-LCY1*, el cual se expresa en tejidos no fotosintéticos, cromoplastos (Vishnevetsky *et al.*, 1999; Escobar-Turriza, 2017), sitio principal de acumulación de diversos y grandes cantidades de carotenoides (Sun *et al.*, 2022). Este gen ha reportado mutaciones funcionales y no funcionales que favorecen la diversidad y cantidad de acumulación de carotenoides en especies como tomate (Sun *et al.*, 2022), así como papaya (Zhou *et al.* 2019) y sandía (Zhang *et al.* 2020). Por último, *Boε-LCY* presenta el mayor número de TFs en las categorías: fotosíntesis, reproducción, sistema oxidativo y defensa frente a factores ambientales. *ε-LCY* tiene un papel regulador muy importante ya que interviene junto con *β-LCY* genera una serie de productos diferentes para su vía (Zhao *et al.*, 2020; Trujillo-Hernández, 2013; Martín-Rodríguez, 2010), como se ha demostrado en maíz (Harjes *et al.*, 2008); por lo que sería lógico se encuentre altamente regulado en la ruta de los carotenoides.

2.4.3. miRNAs de los genes *BoLCY*

Los microARN (miARN) son una clase amplia de ARN pequeños no codificantes (aproximadamente 22 nt) que regulan la expresión génica a nivel postranscripcional a través de la complementariedad de secuencias, involucrados en numerosos procesos biológicos en eucariotas (Owusu *et al.*, 2021; Zuo *et al.*, 2020; Song *et al.*, 2019; Ciu *et al.*, 2017; Sun *et al.*, 2012). Recientemente, se ha demostrado su participación activa en la regulación del metabolismo secundario, como lo son los carotenoides, vitales en funciones de defensa, reproducción, crecimiento y desarrollo de las plantas (Owusu *et al.*, 2021; Bulgakov y Avramenko, 2015). El análisis de miRNAs asociados a los *BoLCY*, reportó 11 familias de miRNAs reconocidas como no conservados. En las plantas, los miRNA se han identificado en dos grupos: los conservados, se producen en diferentes familias (Owusu *et al.*, 2021; Ciu *et al.*, 2017) y regulan la expresión de procesos básicos del desarrollo (Owusu *et al.*, 2021; Song *et al.*, 2019; Dalmay, 2010); las no conservadas, suelen ser específicos de la especie, caracterizados por una baja expresión o presentar copias únicas (Owusu *et al.*, 2021; Sun *et al.*, 2012), los cuales se ha propuesto podrían involucrarse en el desarrollo de rasgos específicos para ciertos taxos (Dalmay, 2010). Para achiote, tres de sus miRNAs asociados a los genes *LCY* tuvieron homólogos en

plantas: miR9408 en *Brassica aleracea* y *Brassica rapa* involucrado en el desarrollo del polen (Lukasik *et al.*, 2013); miR5181 en *Brachypodium distachyon* implicado crecimiento de hojas durante la sequía (Bertolini *et al.*, 2013) y miR8724 en *Gossypium raimondii* involucrado en el alargamiento de las células de fibra de algodón (Xue *et al.*, 2013). Esta información nos da una idea de los posibles procesos de regulación en los que estén involucrados los miRNAs de achiote, más específicamente, en la producción y regulación de los carotenoides. Para *Boβ-LCY1* asociado a cloroplastos con función de captación de luz (Lohani *et al.*, 2021; Aguilar-Espinosa *et al.*, 2020; Trujillo-Hernández, 2016), la familia miR9408 podría estar involucrado en la regulación negativa del desarrollo del polen. Para *Boβ-LCY2*, gen asociado a cromoplastos y acumulación de grandes cantidades de carotenoides (Sun *et al.*, 2022; Escobar-Turriza, 2017; Bian *et al.*, 2011), miR5181 podría implicarse a procesos en las hojas como respuesta a estrés abiótico (luz) y miR8724 en la acumulación de carotenoides en tejidos no fotosintéticos, ambos en una posible regulación positiva.

En especies de importancia económica como tomate, miRNAs conservados y no conservados participan en la regulación de la maduración del fruto (Gao *et al.*, 2015; Dalmay, 2010). Según el modelo de regulación de la maduración de los frutos del tomate propuesto por Gao y colaboradores (2015), el factor de transcripción RIN (inhibidor de maduración) trabaja en conjunto con miRNAs a nivel postranscripcional para que los genes relacionados con la maduración estén activados a niveles necesarios, manifestando un mecanismo de regulación recíproca para la expresión génica (Mejia-Guerra *et al.*, 2012). Por esto, se verificó si los miRNAs presentes podrían estar asociados a los factores de transcripción identificados y por ende, regulando la expresión de los *BoLCY*. Se obtuvo que solo *Boβ-LCY1* y *Boβ-LCY2* se relacionaron a los principales grupos de TF involucrados a fotosíntesis (DOFCOREZM), reproducción (POLLEN1LELAT52) y sistema oxidativo (GT1CONSENSUS), consistente con las posibles funciones que los miRNAs podrían estar regulando en la producción y acumulación de los carotenoides en achiote, como se ha mencionado anteriormente. Mientras que *Boε-LCY*, su único miRNA no se presentó relacionado con TF. Esto sugiere una posible regulación recíproca entre los TFs y miRNAs de los genes *Boβ-LCY* relacionados a funciones vitales en la especie.

2.4.4. Secuencias repetitivas de DNA de los genes *BoLCY*

Se investigó la variabilidad de los *BoLCY* como consecuencia de la aparición de LTRs. Estos son una clase de elementos transponibles que se mueven mediante un intermediario de ARN, se

transcriben, luego, inversamente se transcribe a un DNA complementario que se ubica en una nueva posición en el genoma; lo que le confiera una naturaleza replicativa, siendo propensos a mutaciones, amplificarse y moverse dentro del genoma (Butte *et al.*, 2020; Wang *et al.*, 2016; Lee y Kim, 2014; López-Flores y Garrido Ramos, 2012; Wicker *et al.*, 2007). Se identificaron tres motivos LTR en la región río arriba. *Boβ-LCY1* presentó un motivo correspondiente al LTR Ogre y un dominio de proteína GAG (proteína de replicación). *Boε-LCY* presentó tres motivos LTR: Ogre, Diaspora y Del con el dominio de proteína AP Del (proteína de envoltura), pertenecientes a la familia Ty3/Gypsy. Por su parte, *Boβ-LCY2* no presentó motivos relacionados a los *BoLCY*; pero se identificó al LTR V12 con un dominio de proteína RNaseH (proteína para la transcripción) correspondiente a la familia Ty1/Copia. Los dominios de proteína generalmente se encuentran en la región interna de los elementos retrotransponibles que codifican para diversos productos proteicos necesarios para la replicación y el mecanismo de transposición (Lee y Kim, 2014). Estos se comprenden por dos ORF: GAG, corresponde a una secuencia que codifica para una proteína de cápside y POL, secuencia que codifica una serie de proteínas involucradas en la transcripción de los retroelementos: proteasa (PR), integrasa (INT), transcriptasa inversa (RT) y RNaseH (RH) (Lee y Kim, 2014; Llorens *et al.*, 2011).

Diversos elementos transponibles son remanentes de eventos de transposiciones pasadas que ocurrieron hace millones de años pero que aún pueden tener efectos en el genoma (Lanciano y Cristofari, 2020). Se ha reportado que los elementos transponibles pueden contener secuencias que actúan en *cis*, al fragmentarse e integrarse en los genes o cerca de ellos generan efectos en los niveles de expresión, principalmente con genes cercanos regulados al alza en respuesta a diversos tipos de estrés abiótico, biótico y de carácter reproductivo (Boutte *et al.*, 2020; Wang *et al.*, 2016). *Boε-LCY* que presentó tres motivos LTR, fue el gen que presentó el mayor número de factores de transcripción en su región río arriba. Se asume que este proceso de integración pudo influir en el elevado número de TFs en *Boε-LCY* en comparación de los genes *Boβ-LCY1* y *Boβ-LCY2*, e igualmente estaría influyendo en su nivel de regulación. involucrados en funciones de captación de luz, fotooxidación y reproducción vitales para la regulación de la planta durante sus diferentes etapas de desarrollo.

CAPÍTULO III

METAANÁLISIS DE LOS GENES LICOPENO CICLASA (*Boβ-LCY1*, *Boβ-LCY2*, *Boε-LCY*) DEL TRANSCRIPTOMA DE TRES MORFOTIPOS DE *Bixa orellana* L.

3.1. INTRODUCCIÓN

Los genes licopeno ciclasa (β - y ϵ -) son claves en la ruta de síntesis de carotenoides, al ser el punto de bifurcación de la vía (Song *et al.*, 2022; Maoka, 2020; Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Lu y Li, 2008) y central para la síntesis de bixina en *B. orellana* (Carballo-Uicab, 2019; Escobar-Turriza, 2017; Cárdenas-Conejo *et al.*, 2015; Trujillo Hernández, 2013; Fraser y Bramley 2004; Bouvier *et al.*, 2003; Jacko *et al.*, 2002). Estos genes cumplen un papel importante en el estrés abiótico al que se enfrentan los organismos durante su crecimiento y desarrollo; por lo que, su función principal es regular el estado fisiológico de las plantas aumentando o disminuyendo la síntesis de los carotenoides (Song *et al.*, 2022; Maoka, 2020; Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Lu y Li, 2008).

Estudios sobre los efectos de los licopeno ciclasa, confirman el impacto que estos genes tienen principalmente en la coloración de flores y frutos de las plantas, resultado de la acumulación de diversos pigmentos derivados del licopeno (Hao *et al.*, 2012; Zhao *et al.*, 2020). En *Lycopersicon esculentum*, la supresión del gen *LCY* aumentó el contenido de licopeno (Wan *et al.*, 2007); en sandía un SNP del gen *LCYB* determina el color sea amarillo o rojo de la pulpa de la fruta (Bang *et al.*, 2007); en *Citrus sinensis*, la expresión diferencial del gen β *LCY* y la regulación negativa de ϵ -*LCY* caracteriza la diferente coloración de la pulpa en dos de sus variantes (Alquezar *et al.*, 2008); en Kiwi, se identificó al nivel de expresión del gen *LCYB* asociado a la acumulación total de carotenoides y β -caroteno (Ampomah-Dwamena *et al.*, 2009); en *Solanum habrochaite*, la caracterización del promotor del gen licopeno ciclasa (*ShCYC-B*), mostró que su expresión se mantuvo similar en las etapas verdes pero se regulaba a la positiva en las etapas de maduración (Dalal *et al.*, 2010); en papaya, la expresión del gen *lcy-β2* aumenta durante la maduración del fruto y se correlaciona directamente con el color del fruto, su mutación inactiva su actividad lo que genera la diferencia en producción de carotenoides y coloración en sus frutos (Devitt *et al.*, 2010); en *Cucumis mello* L., la expresión diferencial en dos genotipos, sugiere que la acumulación de β -caroteno está relacionada con el nivel de transcripción de *CmLcyb1* en la fruta (Hao *et al.*, 2012). En achiote, al reportarse el transcriptoma de la especie (Cárdenas-Conejo *et al.*, 2015), se identificaron dos licopeno beta ciclasa y una épsilon ciclasa, donde, según el

modelo de biosíntesis de bixina, los genes se desactivan reprimiendo el flujo del licopeno por la vía principal de los carotenoides, provocando una disminución en la concentración de carotenoides cíclicos que aumenta la disponibilidad del licopeno para la producción de bixina (Cárdenas-Conejo *et al.*, 2015), cumpliendo una función reguladora (Zhao *et al.*, 2020; Raddatz-Mota *et al.*, 2017; Lu y Li, 2008). Recientemente, el trabajo de Trujillo-Hdz y colaboradores (2016) reportaron SNPs del gen *Boβ-LCY1* está asociada a un grupo morfológico (flor rosa), caracterizado con una mayor producción de bixina.

En achote, al reportarse el transcriptoma de la especie (Cárdenas-Conejo *et al.*, 2015), se identificaron dos licopeno beta ciclasa y una épsilon ciclasa; donde, según el modelo de biosíntesis de bixina, los genes se desactivan reprimiendo el flujo del licopeno por la vía principal de los carotenoides, provocando una disminución en la concentración de carotenoides cíclicos que aumenta la disponibilidad del licopeno para la producción de bixina (Cárdenas-Conejo *et al.*, 2015). Luego, para el gen *Boβ-LCY1*, Trujillo-Hdz y colaboradores (2016) reportaron que la variación alélica de un solo nucleótido (SNP) permitió identificar tres grupos morfológicos: G1, flor blanca, frutos verdes, indehisciente; G2, flor rosa, fruto verde con espinas rojas, dehiscente; y G3, flor purpura, frutos rojos, indehisciente; así como identificar al grupo con mayor contenido de bixina, G2 de flor rosa, dejando al G1 de flor blanca con el contenido más bajo. Recientemente, Cárdenas-Conejo y colaboradores (2023) realizaron una base de datos transcriptómica de los morfotipos identificados por Trujillo-Hdz y colaboradores (2016) para diferentes etapas de desarrollo de la semilla: S1 (0-7d), S3 (14-21d), S5 (28-42d) y hoja, con el fin de identificar a los genes relacionados con la biosíntesis de carotenoides, principalmente bixina. Aquí, los autores reportan para los genes licopeno ciclasa una regulación al alza, pero una correlación negativa con la producción de bixina durante el desarrollo de las semillas en todas las accesiones. Por lo tanto, en este capítulo se presenta la metodología para analizar las secuencias codificantes de los genes licopeno ciclasa, así como sus perfiles de expresión de la base de datos del transcriptoma de los morfotipos A (flor blanca), B (flor rosa) y C (flor púrpura) en tres etapas de desarrollo S1 (0-7d), S3 (14-21d), S5 (28-42d) y hoja; con el fin de confirmar las secuencias obtenidas del genoma y brindar información para dilucidar el papel de los genes en la producción diferencial de bixina en la especie.

3.2. MATERIALES Y MÉTODOS

3.2.1 Identificación de los genes *LCY* del transcriptoma de tres morfotipos de *Achiote*

Para identificar los genes *BoLCY* obtenidos de la base de datos del genoma de *achiote*; se analizaron las secuencias codificantes *LCY* genómicas mediante Blastn suite contra la de datos del transcriptoma de tres tejidos (hoja joven, semilla inmadura y semilla madura) de tres morfotipos de *Bixa orellana*, correspondiente al Bioproyecto: Transcriptome sequencing of *Bixa orellana* L. to study the gene expression profile during seed development (NCBI: PRJNA895001) (Cárdenas-Conejo *et al.*, 2023). Se analizó el porcentaje de identidad, cobertura para cada gen *BoLCY*, y la posible presencia de polimorfismos y secuencias homólogas.

3.2.2. Análisis de miRNAs de los genes *BoLCY*

Para identificar la presencia de los microRNAs identificados en la región promotora de los genes *BoLCY* del genoma; se empleó de referencia la secuenciación del transcriptoma de tres tejidos (hoja joven, semilla inmadura y semilla madura) de tres morfotipos de *Bixa orellana* (NCBI: PRJNA895001) (Cárdenas-Conejo *et al.*, 2023) y las secuencias maduras de los miRNAs genómicos identificados miRBase database (Kozomara *et al.*, 2019), así como las secuencias de los miRNAs de tomate (Arazi y Khedia *et al.*, 2022) y 9 familias conservadas de las plantas terrestres (Axtell y Meyers, 2018), para confirmar su presencia en la especie.

3.2.3. Análisis de expresión de los genes *LCY* del transcriptoma de tres morfotipos de *B. orellana*

Se empleó una base de datos de 36 bibliotecas de ADNc indexadas (3 réplicas biológicas) obtenida con la plataforma Illumina, correspondiente a los morfotipos A (flor blanca), B (flor rosa) y C (flor púrpura) con tres etapas de maduración de la semilla: S1 (0-7d), S3 (14-21d), S5 (28-42d) y hoja (NCBI: PRJNA895001) (Cárdenas-Conejo *et al.*, 2022). Se identificaron los contigs 40528 para β *LCY1*, 16795 para β *LCY2* y 18699, 40767 para ϵ *LCY*. Los niveles de expresión relativos de los genes *BoLCY* en sus diferentes etapas de desarrollo de la semilla, se calcularon usando TPM (Transcripciones por millón) y se visualizaron en el entorno de programación de R (RStudio Team, 2015) usando el paquete “ggplots” (Warnes *et al.*, 2016).

3.3. RESULTADOS

3.3.1. Secuencias de los genes *LCY* del transcriptoma de tres morfotipos de *Achiote*

El análisis BLAST de las secuencias genómicas de los *BoLCY* de achote en la base de datos de secuenciación del transcriptoma de la especie (accesión: PRJNA895001) (Cárdenas-Conejo *et al.*, 2023). Se confirmó la presencia de los genes *Boβ-LCY1*, *Boβ-LCY2* y *Boε-LCY* con un 99.80%, 100% y 98.5% de identidad respectivamente; con una cobertura de 100% para los genes β-ciclasa y un 64 y 24% para los contigs (18,699 y 40,767) del gen ε-ciclasa (Figura 3.1) (Anexo 3). No se detectaron polimorfismos ni secuencias homólogas.

Boβ-LCY1

```

BoLCY1 Gen ATGGACACGCTACTTAAGACGCACAACAAGCTTGAGTTCCTGCCACAACCTCATGGGTTTTCCGGAGAAAGTGAGTATTTTGAACCTCCCAAGCTTCAA 100
BoLCY1 tma ATGGACACGCTACTTAAGACGCACAACAAGCTTGAGTTCCTGCCACAACCTCATGGGTTTTCCGGAGAAAGTGAGTATTTTGAACCTCCCAAGCTTCAA 100

BoLCY1 Gen GCAAGGAGCTTAAGCTTGGTTCAGCAGGCCTTGCTGAGAAAAGGTAGGAATAGTTGCCTTAAGGCCAATTCTAGCACTCTCACGGAACCTTGTTC 200
BoLCY1 tma GCAAGGAGCTTAAGCTTGGTTCAGCAGGCCTTGCTGAGAAAAGGTAGGAATAGTTGCCTTAAGGCCAATTCTAGCACTCTCACGGAACCTTGTTC 200

BoLCY1 Gen AATCAAGAAGGAGAATCTTGAGTTTGAGCTTCCTCTCTATGACCCATCAAAGGCCCTTGTTGGTGGATCTAGCTGTTGTAGGAGGTGGCCCTGCAGGTCTC 300
BoLCY1 tma AATCAAGAAGGAGAATCTTGAGTTTGAGCTTCCTCTCTATGACCCATCAAAGGCCCTTGTTGGTGGATCTAGCTGTTGTAGGAGGTGGCCCTGCAGGTCTC 300

BoLCY1 Gen GCTGTTGCCAGCAGGTTTCAGAGGCAGGCTTTCGGTTCGCACGATTGACCCCTCTCCCAAATGATTTGGCCTAATAAATATGGTGTGGTGGATG 400
BoLCY1 tma GCTGTTGCCAGCAGGTTTCAGAGGCAGGCTTTCGGTTCGCACGATTGACCCCTCTCCCAAATGATTTGGCCTAATAAATATGGTGTGGTGGATG 400

BoLCY1 Gen AATTTGAGGCTATGGATTTGCTTGATTGCCTTGACACCACCTGGAGTGGTGTGTTGTTTACATTAATGACCAAACAAGAAGGATCTTGATAGACCTTA 500
BoLCY1 tma AATTTGAGGCTATGGATTTGCTTGATTGCCTTGACACCACCTGGAGTGGTGTGTTGTTTACATTAATGACCAAACAAGAAGGATCTTGATAGACCTTA 500

BoLCY1 Gen TGGGAGAGTTAATAGAAGGCAGCTGAAGTCCAAAATGCTGCAGAAATGCATATCAAATGGTGTCAAATTTTCATCAGGCTAAGGTAATTAAGGTTATCCAT 600
BoLCY1 tma TGGGAGAGTTAATAGAAGGCAGCTGAAGTCCAAAATGCTGCAGAAATGCATATCAAATGGTGTCAAATTTTCATCAGGCTAAGGTAATTAAGGTTATCCAT 600

BoLCY1 Gen GAGCAGTCCAAATCTCTATTGATTGCAATGATGGTGTACGGTTCAGGCTACCTTGGTCTCTGATGCAACAGGGTTTTCTAGGTGCCTTGTGCAGTATG 700
BoLCY1 tma GAGCAGTCCAAATCTCTATTGATTGCAATGATGGTGTACGGTTCAGGCTACCTTGGTCTCTGATGCAACAGGGTTTTCTAGGTGCCTTGTGCAGTATG 700

BoLCY1 Gen ATAAGCCCTATAATCCTGGTTACCAAGTGGCTTATGGAATTTAGCCGAGGTAGAAGAACATCCATTGTATGTGGATAAGATGGTTTTTATGGATTGGAG 800
BoLCY1 tma ATAAGCCCTATAATCCTGGTTACCAAGTGGCTTATGGAATTTAGCCGAGGTAGAAGAACATCCATTGTATGTGGATAAGATGGTTTTTATGGATTGGAG 800

BoLCY1 Gen AGATTACACCTCACCAACAATAAGGAGCTGAGGGAGAGAAATAGTAAGATACCAACATTTCTCTATGCAATGCCCTTTTCATCTGAGAGAATATTTCTT 900
BoLCY1 tma AGATTACACCTCACCAACAATAAGGAGCTGAGGGAGAGAAATAGTAAGATACCAACATTTCTCTATGCAATGCCCTTTTCATCTGAGAGAATATTTCTT 900

BoLCY1 Gen GAAGAGACTTCCCTTGTGCTCGACCTGGAGTACAAATGAAGGACATCCAGGAAAGAAATGGTGGCTAGGTTAAGACACTTAGGCATAAAAAGTAAAAGTA 1000
BoLCY1 tma GAAGAGACTTCCCTTGTGCTCGACCTGGAGTACAAATGAAGGACATCCAGGAAAGAAATGGTGGCTAGGTTAAGACACTTAGGCATAAAAAGTAAAAGTA 1000

BoLCY1 Gen TTGAAGAGGATGAGCATTGTGTGATTCCGATGGGCGGACCCCTTCCCTGTGCTTCCCCAAAGAGTTGTGGGAATCGGTGGAACAGCTGGATGGTGCACCC 1100
BoLCY1 tma TTGAAGAGGATGAGCATTGTGTGATTCCGATGGGCGGACCCCTTCCCTGTGCTTCCCCAAAGAGTTGTGGGAATCGGTGGAACAGCTGGATGGTGCACCC 1100

BoLCY1 Gen TCAACTGGGTTACATGGTAGCGAGGACCCTAGCTGCAGCTCCAATTTGTTGCCAATCCATTGTCAGTGCCTTGGTTCTGAAAGGAGCCTATCAGCAGA 1199
BoLCY1 tma TCAACTGGGTTACATGGTAGCGAGGACCCTAGCTGCAGCTCCAATTTGTTGCCAATCCATTGTCAGTGCCTTGGTTCTGAAAGGAGCCTATCAGCAGA 1199

BoLCY1 Gen TGTGTTGTCTGAAGTAGTTTGGAAAGATTTATGGCCATAGAAAAGGAGAGACAGAGGAAATCTTCTGTTTTGGCATGGATATTTCTGCTGAAGCTAGAT 1299
BoLCY1 tma TGTGTTGTCTGAAGTAGTTTGGAAAGATTTATGGCCATAGAAAAGGAGAGACAGAGGAAATCTTCTGTTTTGGCATGGATATTTCTGCTGAAGCTAGAT 1299

BoLCY1 Gen TTACAAGCTACAAGAAGATTTTTTGTATGCATTTTTTAATCTGGAACCTCGATATTGGCATGGATTCTTGTCTATCTCGGTTGTTCTTCCAGACCTTTTAT 1399
BoLCY1 tma TTACAAGCTACAAGAAGATTTTTTGTATGCATTTTTTAATCTGGAACCTCGATATTGGCATGGATTCTTGTCTATCTCGGTTGTTCTTCCAGACCTTTTAT 1399

BoLCY1 Gen TTTTGGCCCTTCTCTATCTCTAATGCTTCTAATACTTCTAGGATAGAGATTATGGCCAAAGGAAACGCTTCCCTTGGTTAAGATGATCAACAACCTTAAT 1499
BoLCY1 tma TTTTGGCCCTTCTCTATCTCTAATGCTTCTAATACTTCTAGGATAGAGATTATGGCCAAAGGAAACGCTTCCCTTGGTTAAGATGATCAACAACCTTAAT 1499

BoLCY1 Gen ACAGGATAGCGAT 1512
BoLCY1 tma ACAGGATAGCGAT 1512

```

Bo β -LCY2

BoLCY2 gen AGATCAAGAAGACCTCAAGCAAGACTCATCAGAAAGTCAGAGTCAGCAAATTTGGGAACCTTCTGGAATTGAAGCCAGAGTCACAGCGGGAGTTCTTGGA 200
BoLCY2 tma AGATCAAGAAGACCTCAAGCAAGACTCATCAGAAAGTCAGAGTCAGCAAATTTGGGAACCTTCTGGAATTGAAGCCAGAGTCACAGCGGGAGTTCTTGGA 200

BoLCY2 gen TTATGATCTGCCTTTGTTTGTATCCATCCCATGAGCGGCCTCGCTTTGATGTGATCGTCATCGGCACCGGGCCCGCAGGTCTCCGTCTCGCCGAGCAAGTT 300
BoLCY2 tma TTATGATCTGCCTTTGTTTGTATCCATCCCATGAGCGGCCTCGCTTTGATGTGATCGTCATCGGCACCGGGCCCGCAGGTCTCCGTCTCGCCGAGCAAGTT 300

BoLCY2 gen TCTGGCCGTGGAATCAAGGTATGTTGCGTCGATCCCTCACCACCTTCCATGTGGTCTAACAAATTATGGTGTCTGGGTAGATGAGTTCAAGACTTGGGGC 400
BoLCY2 tma TCTGGCCGTGGAATCAAGGTATGTTGCGTCGATCCCTCACCACCTTCCATGTGGTCTAACAAATTATGGTGTCTGGGTAGATGAGTTCAAGACTTGGGGC 400

BoLCY2 gen TCGATGATTGTTTGGACAAGACATGGCCGATGACTACGGTCCATATAGATGATAAAAAACCAAATATCTGGACCAGCCGACTGTCTGTCTAGTAGAAA 500
BoLCY2 tma TCGATGATTGTTTGGACAAGACATGGCCGATGACTACGGTCCATATAGATGATAAAAAACCAAATATCTGGACCAGCCGACTGTCTGTCTAGTAGAAA 500

BoLCY2 gen GAAGCTCAAGTCTAGATTAATGGAGTCCTGCACGTGCAACGGAGTTAAGTCCATAAAGCCAAAGTTTGGAAAGTGAACACGAGAAATTTGAGTCCATA 600
BoLCY2 tma GAAGCTCAAGTCTAGATTAATGGAGTCCTGCACGTGCAACGGAGTTAAGTCCATAAAGCCAAAGTTTGGAAAGTGAACACGAGAAATTTGAGTCCATA 600

BoLCY2 gen ATCGAGTGTGACGATGGGAGTGAGTTCAAGCCAGCTTGATCGTCGATGCGAGTGGATTGCAAGCCCTTTCATAAAAAACGATAAGCCCTCGGAACGATG 700
BoLCY2 tma ATCGAGTGTGACGATGGGAGTGAGTTCAAGCCAGCTTGATCGTCGATGCGAGTGGATTGCAAGCCCTTTCATAAAAAACGATAAGCCCTCGGAACGATG 700

BoLCY2 gen GGTATCAACTCGCCCATGGCATTCTGGCTGAAGTAGAAAGCCATCCTTTCGATTTGGACAAGATGGTTCTTATGGATTGGAGAGATTTCCCATTTGGCAAA 800
BoLCY2 tma GGTATCAACTCGCCCATGGCATTCTGGCTGAAGTAGAAAGCCATCCTTTCGATTTGGACAAGATGGTTCTTATGGATTGGAGAGATTTCCCATTTGGCAAA 800

BoLCY2 gen TGAGCCTGACTTACGCGAACGAAACGCAAAGCTCCCAACTTTTCTCTATGTAATGCCATTTCGATTTCGAATTTGGTGTCTTCTGGAAGAACTTCTCTGGTC 900
BoLCY2 tma TGAGCCTGACTTACGCGAACGAAACGCAAAGCTCCCAACTTTTCTCTATGTAATGCCATTTCGATTTCGAATTTGGTGTCTTCTGGAAGAACTTCTCTGGTC 900

BoLCY2 gen GGTGACCCGGTTCTGCCGTACATGGAGGTGAAACAGAGGATGGTGGCAGGCTGAGGCATTTAGGAATCATACTGAAAGAGTATAGAAAGAAAAAT 1000
BoLCY2 tma GGTGACCCGGTTCTGCCGTACATGGAGGTGAAACAGAGGATGGTGGCAGGCTGAGGCATTTAGGAATCATACTGAAAGAGTATAGAAAGAAAAAT 1000

BoLCY2 gen GCCTGATCCCAATGGGGGGCCCTTTCGCAAACTCCCTCAGAATGATTTGGCGATGGGAGGAAATTCGGGCTGGTCCATCCGTCGACTGGGTACACGGT 1100
BoLCY2 tma GCCTGATCCCAATGGGGGGCCCTTTCGCAAACTCCCTCAGAATGATTTGGCGATGGGAGGAAATTCGGGCTGGTCCATCCGTCGACTGGGTACACGGT 1100

BoLCY2 gen GTACCCGACGATGGCGATGGCCCTGTTGTGGCGGATGCTATTGCGGAGTGTGGGTTTCGACGAGGATGATGAGAGGGAGCCTTTTACATCGGAGAGTT 1200
BoLCY2 tma GTACCCGACGATGGCGATGGCCCTGTTGTGGCGGATGCTATTGCGGAGTGTGGGTTTCGACGAGGATGATGAGAGGGAGCCTTTTACATCGGAGAGTT 1200

BoLCY2 gen TGGAAACGGACTGTGGCCGATGGAGAGGAGAGTGGAGCGGCAATTATACTGTTTGGGATGGAGACTTTGCTGAATCTCGATTGGCCGGGACCAGGCGAT 1300
BoLCY2 tma TGGAAACGGACTGTGGCCGATGGAGAGGAGAGTGGAGCGGCAATTATACTGTTTGGGATGGAGACTTTGCTGAATCTCGATTGGCCGGGACCAGGCGAT 1300

BoLCY2 gen TCATGGGCGAATTTCTCCAGCTGGATCCTTATCAATGGAAGGCTTTCTCTCGAAAAGCCTCTCGACGCCGAAATTTGTGCTCTATTGAGTACGATGTT 1400
BoLCY2 tma TCATGGGCGAATTTCTCCAGCTGGATCCTTATCAATGGAAGGCTTTCTCTCGAAAAGCCTCTCGACGCCGAAATTTGTGCTCTATTGAGTACGATGTT 1400

BoLCY2 gen CATAAAAGCATCGGTTCCATGCAAGATCGATTTGGCTACCAAGTGTCTTTACCCGATGGCTCGCACGCTGGGAAAAATTCGACTTGACGCCATCTGA 1497
BoLCY2 tma CATAAAAGCATCGGTTCCATGCAAGATCGATTTGGCTACCAAGTGTCTTTACCCGATGGCTCGCACGCTGGGAAAAATTCGACTTGACGCCATCTGA 1497

Bo ϵ -LCY

```

ELCY gen ATGATGGAGTGTATTGGAGCACGGAACCTCGGGGCAATGGCGGTCTCCACCTTCTCTACTTGGAACTCTGGAGAAGGACGCTGAGAACCAGTCAAGCGC 100
ELCY tma -----TGGAACTCTGGAGAAGGACGCTGAGAACCAGTCAAGCGC 40

ELCY gen TTTCCAACCATCGATGTTTCTCGTTCAGAGTGAGAGCCAGCGGTGCGGGGAGCGAGAGCTGTGTGGCGGTTAAAGGAGGTTTGTCTGATGAAGAAGATTA 200
ELCY tma TTTCCAACCATCGATGTTTCTCGTTCAGAGTGAGAGCCAGCGGTGCGGGGAGCGAGAGCTGTGTGGCGGTTAAAGGAGGTTTGTCTGATGAAGAAGATTA 140

ELCY gen TTACCACCTATATCGACTGGAGATAAGATACTGGACTTGGTGGTGGTGGATGGATGGTCCAGCTGGTCTTGCTCTTGCTGCAGAGTCCGCTAAATGGAT 299
ELCY tma CTTACCACCTATATCGACTGGAGATAAGATACTGGACTTGGTGGTGGTGGATGGATGGTCCAGCTGGTCTTGCTCTTGCTGCAGAGTCCGCTAAATGGAT 240

ELCY gen TTGAACGTTGGTTAATTTGGAGCTTGGACTTGAAGTTGTATTGAACATGTTTGGGGGATACCATCGTTACCTAGACAATAATGACCCAATTTGTGTGG 399
ELCY tma TTGAACGTTGGTTAATTTGGAGCTTGGACTTGAAGTTGTATTGAACATGTTTGGGGGATACCATCGTTACCTAGACAATAATGACCCAATTTGTGTGG 336

ELCY gen AGTCAGTGTTCGTATCTAAACTCAAAGGTGGAAGCATTATGGAAGCTACCAACGGTCATAATCTTGTAGCTTGTGAACAGAACATCGTTGTTCCCTG 499
ELCY tma -----CGTATCTAAACTCAAAGGTGGAAGCATTATGGAAGCTACCAACGGTCATAATCTTGTAGCTTGTGAACAGAACATCGTTGTTCCCTG 424

ELCY gen CCGGCTTGCTACTGTTGCTTCAGGAGCAGCTTCAGGAACTGTTGCAATATGAAGTGGGAGTCCCTAGGGTCTCTGTTCAA-----CAGC 586
ELCY tma CCGGCTTGCTACTGTTGCTTCAGGAGCAGC-----A-----GC 457

ELCY gen CTACGGCATTGAGGTTGAGGTGAGAAACAATCCTTATGATCCTAGCCTAATGGTTTTTCATGGATTACAGAGACTATGTTAAACACCAAGCTCCGTGTTTA 686
ELCY tma CTACGGCATTGAGGTTGAGGTGAGAAACAATCCTTATGATCCTAGCCTAATGGTTTTTCATGGATTACAGAGACTATGTTAAACACCAAGCTCCGTGTTTA 557

ELCY gen GAATGTCAATATCCAACATTTCTTTATGTAATGCCCATTTTCATCAACAAGAGTTTTTTTTGAGGAACTTGTGTTGGCTTCAAAGATGCCATGCCTTTTG 786
ELCY tma GAATGTCAATATCCAACATTTCTTTATGTAATGCCCATTTTCATCAACAAGAGTTTTTTTTGAGGAACTTGTGTTGGCTTCAAAGATGCCATGCCTTTTG 657

ELCY gen AGTTGCTAAAGGAAAAGCTCATGTTGAGGTTAGATACCATGGGAATCCGGATTCTTAAGACTTACGAAGAGGAATGGTCTTATATCCCAGTCGGTGGGTC 886
ELCY tma AGTTGCTAAAGGAAAAGCTCATGTTGAGGTTAGATACCATGGGAATCCGGATTCTTAAGACTTACGAAGAGGAATGGTCTTATATCCCAGTCGGTGGGTC 757

ELCY gen CTTACCAAAACACAGAGCAAAAAACCTTGCATTTGGTGTCTGCTAGCATGGTGCATCCAGCTACAGGTAAGACTAAAACAGTGGTGGATATTGAC 986
ELCY tma CTTACCAAAACACAGAGCAAAAAACCTTGCATTTGGTGTCTGCTAGCATGGTGCATCCAGCTACAGGT----- 827

ELCY gen TTAAGTGGTTCAGACTTATGGTGTCTACTGATTGTTAAACCTTTATTCAGTTGTCAGATCATTGTCTGAAGCTCCTAAATATGCTTCTGTAATTGCAAA 1086
ELCY tma -----TATTCAGTTGTCAGATCATTGTCTGAAGCTCCTAAATATGCTTCTGTAATTGCAAA 884

ELCY gen ATATTGAACAAGGGTCCCTTCTAAAGGCATGCTTTCTGATGAAAGAAATAATGAGAATATCTCAATGCAAGCTTGGAACTCTATGGCCACAAGAAAGGA 1186
ELCY tma ATATTGAACAAGGGTCCCTTCTAAAGGCATGCTTTCTGATGAAAGAAATAATGAGAATATCTCAATGCAAGCTTGGAACTCTATGGCCACAAGAAAGGA 984

ELCY gen AGAGGCAAAGGGCATTCTTTCTGTTGGATTGGCACTAATTTTGAACCTGGATATTGAAGGCATTAGAACATCTTTTCCACTTTTTCCGCTTACCAG 1286
ELCY tma AGAGGCAAAGGGCATTCTTTCTGTTGGATTGGCACTAATTTTGAACCTGGATATTGAAGGCATTAGAACATCTTTTCCACTTTTTCCGCTTACCAG 1084

ELCY gen CTGGATGTGGCAAGGATTTCTTGGCTTACACTTTCCCTCGGCTGACCTCATGCTGTTGCTTTTTATATGTTTGTGTTTGTAGCTCCAAATGATATGAGAAT 1386
ELCY tma CTGGATGTGGCAAGGATTTCTTGGCTTACACTTTCCCTCGGCTGACCTCATGCTGTTGCTTTTTATATGTTTGTGTTTGTAGCTCCAAATGATATGAGAAT 1184

ELCY gen TCCCTTGTTCAGGCATCTAGTTTCAGACCCAACCTGGTGAACATGATAAAGACATATCTCACAATA 1452
ELCY tma TCCCTTGTTCAGGCATCTAGTTTCAGACCCAACCTGGTGAACATGATAAAGACATATCTCACAATA 1250

```

Figura 3.1. Alineamiento de secuencias de nucleótidos obtenidas de las bases de datos del genoma y transcriptoma de los genes *Bo β -LCY1*, *Bo β -LCY2* y *ϵ -LCY*.

3.2.2. miRNAs de los genes *BoLCY*

Los miARNs tienen funciones importantes en la regulación de la expresión génica, involucrados en numerosos procesos biológicos en las plantas (Owusu *et al.*, 2021; Gao *et al.*, 2015). Para achiote, se confirmó la presencia de los miRNAs identificados en la región promotora de los genes *BoLCY*: miR253, miR408, miR12302, miR991, miR1195, miR11, miR5181, miR8724, miR9571, miR10366 y miR6546, en la base de datos de transcriptoma de los morfotipos. Las secuencias de los 12 miRNAs identificados de la base de datos de secuenciación de genoma, observados en la base de datos de los transcriptomas de la especie, presentaron una identidad del 100%. De igual manera, se identificó que el contig 16,795 donde se encuentra la secuencia del gen *Bo β LCY2*, es la que presentan los transcritos de 4 miRNAs identificados en el genoma: miR991-3p, miR11-5p, miR5181-3p y miR8724 (Cuadro 3.1).

Cuadro 3.1. microRNAs transcritos de los genes *BoLCY* de achote

microRNA	LONGITUD pb	% IDENTIDAD	% QUERY COVER	HITS
miR253	21	100	100	11
miR408	23	100	100	40
miR408-3p	22	100	100	48
miR12302-3p	22	100	90	73
*miR991-3p	22	100	100	103
miR1195	23	100	91	31
*miR11-5p	20	100	100	96
*miR5181-3p	21	100	94	35
*miR8724	23	100	95	30
miR9571	21	100	100	135
miR10366	22	100	100	74
miR6546-5p	24	100	87	8

*miRNAs identificados en el contig 16,725 (*Boβ-LCY2*).

El gen *Boβ-LCY2* fue el que presentó el mayor número de secuencias miRNAs en su región promotora (miR991, miR1195, miR11, miR5181, miR8724, miR9571, miR10366); donde, solo 4 miRNAs se presentaron en la base de datos del transcriptoma (miR991-3p, miR11-5p, miR5181-3p y miR8724). De igual manera, se analizó si los miRNAs vinculados a los genes *LCY* identificados en el transcriptoma podrían estar relacionados con genes estructurales de la ruta de los carotenoides: DXS2, PYS2 y PSD2. Sin embargo, los contigs identificados que contienen a los genes DXS2 (15,072 y 45,501), PYS2 (807 y 3,892) y PDS2 (13,471; 26,642; 23,617; 42,007 y 31,880), no se vincularon con los miRNAs de los genes *BoLCY*.

Recientemente, se documentó 39 familias conservadas de miARNs de plantas, donde, nueve de estas (miR156, miR160, miR166, miR171, miR319, miR390, miR477, miR529 y miR535) se consideran las mejor conservadas por presentarse en la mayoría o en todos los linajes de plantas terrestres (Arazi y Khedia *et al.*, 2022; Axtell y Meyers, 2018). Una especie modelo donde se han realizado la identificación de miRNAs ha sido el tomate, donde miRNAs conservados como no conservados tienen un papel esencial en el desarrollo y maduración de sus frutos (Arazi y Khedia *et al.*, 2022; Gao *et al.*, 2015). Por esta razón, fue de interés el analizar la presencia de los miRNAs conservados reportados en tomate (156, 157, 159, 160, 162, 164, 165, 166, 167, 168, 169a, 171a, 172, 390a, 394, 396a, 397, 399, 403 y 408) (Gao *et al.*, 2015) en la base de datos del transcriptoma de achote. Esto arrojó la presencia de todos los miRNAs conservados en tomate y entre ellas, 5 miRNAs que pertenecen a las familias mejor conservadas en las plantas

terrestres (miR156, miR160, miR166, miR171 y miR390) con una identidad del 100% (Cuadro 3.2). Por lo que estos miRNAs conservados presentes en los transcriptomas de achiote, podrían estar llevando estas funciones en el desarrollo de la planta.

Cuadro 3.2 miRNAs conservados de tomate en transcriptomas de Achiote

microRNA	LONGITUD pb	% IDENTIDAD	% QUERY COVER	HITS
*miR156	20	100	100	131
miR157	20	100	100	131
miR159	21	100	100	31
*miR160	21	100	100	23
miR162	21	100	94	94
miR164	21	100	90	52
miR165	21	100	94	43
*miR166	21	100	100	61
miR167	21	100	100	70
miR168	21	100	95	14
miR169a	21	100	100	44
*miR171a	21	100	100	33
miR172	20	100	95	224
miR319	21	100	94	50
*miR390a	21	100	90	39
miR394a	20	100	100	19
miR396a	21	100	100	127
mir397	21	100	100	32
miR399	21	100	95	16
miR403	21	100	94	32
miR408	21	100	90	11

* miRNAs reportados como mejor conservadas en las plantas terrestres.

3.2.3. Expresión de los genes *LCY* del transcriptoma de tres morfotipos de *B. orellana*

El análisis de expresión relativa de los genes *Boβ-LCY1*, *Boβ-LCY2*, *Boε-LCY1* identificados en los contigs 40,528; 16,795 y 18,699-40,767 respectivamente en la base de datos de los tres morfotipos de achiote, se visualizaron en el entorno de programación R (RStudio Team, 2015) en mapas de calor, con el fin de dilucidar el papel que juegan en la biosíntesis y acumulación de carotenoides, como bixina, en *Bixa orellana*.

Para el gen *Boβ-LCY1*, la expresión relativa se observó con valores similares para los tres morfotipos en hoja y semilla, siendo la etapa S5 (28-42d) con el menor nivel de expresión (Figura 3.2). *Boβ-LCY2* presentó una alta expresión en hoja en los morfotipos flor rosa (B) y flor púrpura (C), y baja en morfotipo flor blanca (A) (Figura 3.2), esto estaría aunado al trabajo en conjunto con *Boβ-LCY1* para la producción de productos carotenoides de la rama β . En las semillas, el gen específico de cromoplastos también presentó una alta expresión, aunque de manera diferencial en los morfotipos, principalmente en el morfotipo flor rosa (B), reportada con mayor contenido de bixina (Trujillo-Hdz *et al.*, 2016). En cuanto a *Boε-LCY*, presentó un nivel de expresión alto para hoja para los morfotipos flor rosa (B) y flor púrpura (C), pero débil en semilla para todos los morfotipos (Figura 3.2).

3.4. DISCUSIÓN

3.4.1. Secuencias de los genes *LCY* del transcriptoma de tres morfotipos de Achiote

Las secuencias codificantes de los genes licopeno ciclasa de achiote se identificaron gracias a la secuenciación del transcriptoma *de novo* del morfotipo N4 (flor rosa) de achiote, reportada con el mayor contenido de bixina (Cárdenas-Conejo *et al.*, 2015). La secuenciación del genoma de la especie (aún no publicada), nos permitió confirmar la presencia y extraer las secuencias genómicas de los genes *LCY* de achiote. Los genes β -*LCYs*, se presentan con un solo ORF, mientras que *Boε-LCY* se presentó con 11 exones y 10 intrones. La reciente secuenciación del transcriptoma de los morfotipos caracterizados de la especie (flor rosa, flor blanca y flor púrpura) reportadas con diferentes contenidos de bixina (accesión: PRJNA895001) (Cárdenas-Conejo *et al.*, 2023), nos brindó la oportunidad de extraer las secuencias para compararlas con las del genoma, donde, los genes *Boβ-LCY1*, *Boβ-LCY2* y *Boε-LCY* mostraron una identidad de 98.5%, 99.80% y 100% respectivamente, con una cobertura de 100%. No obstante, no se presentaron secuencias homologas o los polimorfismos antes reportados para el gen *Boβ-LCY1* en el morfotipo flor rosa (Trujillo-Hdz *et al.*, 2016).

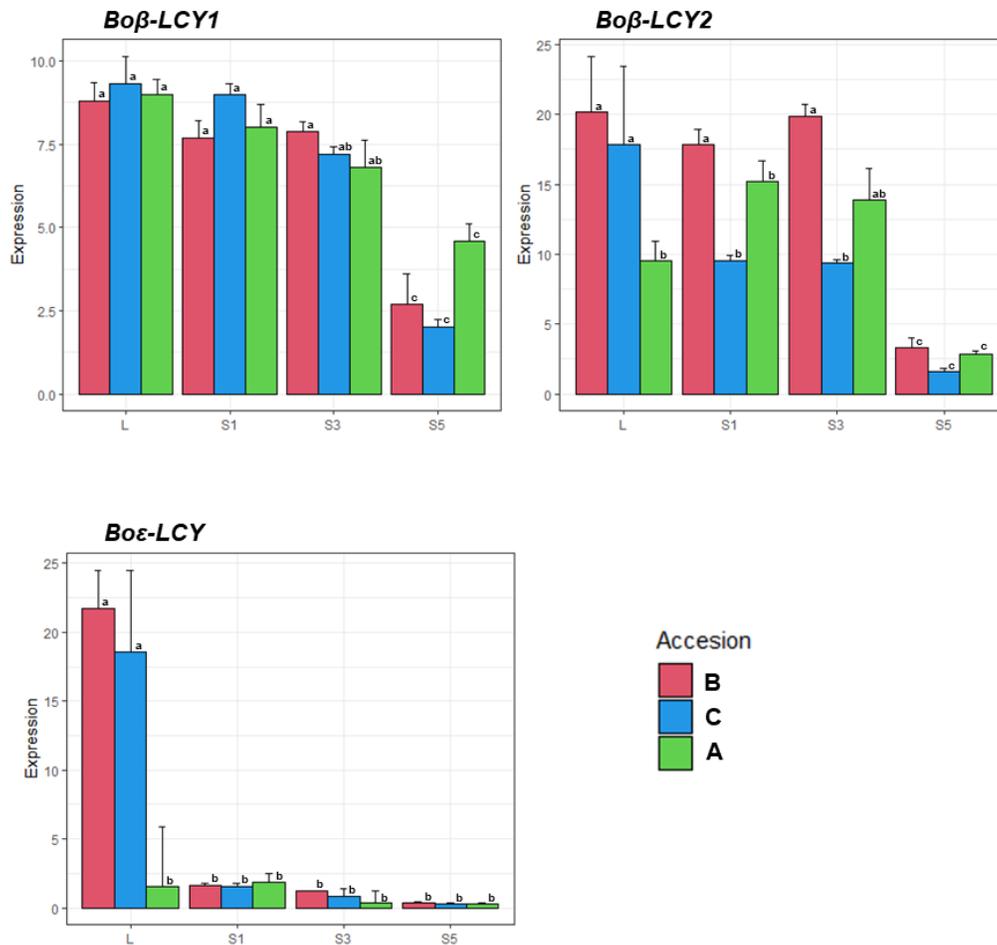


Figura 3.2. Expresión relativa de los genes *BoLCY* del transcriptoma de tres morfotipos de *Bixa orellana*. A (flor blanca), B (flor rosa) y C (flor púrpura) para tres etapas de desarrollo de semilla S1 (0-7d), S3 (14-21d), S5 (28-42d) y hoja. Se analizaron tres réplicas en cada muestra. Las letras indican grupos significativos de Tukey.

3.4.2. miRNAs de los genes *BoLCY*

Se identificó la presencia de las secuencias miRNAs del genoma en la base de datos de transcriptomas de los morfotipos de achote. Estos correspondieron a miRNAs no conservados (miR253, miR408, miR12302, miR991, miR1195, miR11, miR5181, miR8724, miR9571, miR10366 y miR6546); los cuales, suelen ser específicos de la especie o intervienen en procesos únicos para la planta (Sun *et al.*, 2012). No obstante, solo cuatro miRNAs: miR991-3p, miR11-5p, miR5181-3p y miR8724 se vincularon con un gen, *Boβ-LCY2*. Dos de estos miRNAs, identificados como no conservados, tienen reportado sus funciones en otras especies: miR5181

en *Brachypodium distachyon* implicado en el crecimiento de las hojas durante la sequía (Bertolini *et al.*, 2013) y miR8724 en *Gossypium raimondii* involucrado en el alargamiento de las células de fibra de algodón (Xue *et al.*, 2013), los cuales, podrían estar regulando procesos específicos en el desarrollo de la planta, como en el agrandamiento de las células que acumulan carotenoides y el crecimiento de hojas durante eventos de estrés, como la fotooxidación. Esto podría suponer que de los genes *BoβLCY*, solo *BoβLCY2* contaría con miRNAs con potencial para reprimir su expresión o la de genes cercanos, regulando procesos específicos en el desarrollo de la planta. Sin embargo, para LCYb no es único reporte de vinculación con miRNAs no conservados. En naranja (*Citrus sinensis*), el miRNA1857 y su interacción con el gen licopeno beta ciclasa podría estar regulando la acumulación al alta de licopeno en su mutante de rasgo pulpa roja (Xu *et al.*, 2010); esto debido al deterioro del flujo descendente del licopeno por la vía de los carotenoides. Sin embargo, para achiote, el transcrito del miRNA1857 no se encontró en la base de datos de transcriptoma de la especie. Esto podría deberse a que los miRNAs no conservados al ser específicas de la especie, difieren su expresión y/o función según el órgano o especie donde se presenten (Owusu *et al.*, 2021; Song *et al.*, 2019; Xu *et al.*, 2010).

En cuanto a los miRNAs conservados, en las plantas regulan la expresión de procesos básicos del desarrollo (Arazi y Khedia *et al.*, 2022; Owusu *et al.*, 2021; Song *et al.*, 2019; Gao *et al.*, 2015; Dalmay, 2010). Estos se han aislado principalmente en *Arabidopsis*, los cuales se consideran evolutivamente conservados en todas las especies de plantas (Xu *et al.*, 2010). Mas recientemente, se ha reportado 9 familias (miR156, miR160, miR166, miR171, miR319, miR390, miR477, miR529 y miR535) que se consideran las mejor conservadas por presentarse en la mayoría o en todos los linajes de plantas terrestres (Arazi y Khedia *et al.*, 2022; Axtell y Meyers, 2018). En un sistema modelo como tomate, sus miRNAs conservados se han reportado con un papel esencial en el desarrollo y maduración de sus frutos (Arazi y Khedia *et al.*, 2022; Gao *et al.*, 2015). En achiote, se identificaron secuencias de todas las familias conservadas reportadas en tomate y entre ellas, 5 miRNAs que pertenecen a las familias mejor conservadas en las plantas terrestres (miR156, miR160, miR166, miR171 y miR390). miR156 presentó el mayor número de hits para la base de datos del transcriptoma seguido de miR166 y miR172. En tomate, miR156 presenta funciones en el desarrollo de la flor y tiempo de floración, miR166 en desarrollo de tallo, hoja y ovulo de la flor (Arazi y Khedia *et al.*, 2022) y miR172 que regula a la familia de factores de transcripción APETALA (AP2) involucrado en el tiempo de floración y morfología de la flor (Chung *et al.*, 2020). De igual manera, miR156, esta reportado vinculado al flujo de los carotenoides, su sobreexpresión en *Brassica napus* aumenta la luteína y β -caroteno en sus

semillas y raíces (Wei *et al.*, 2012). Por lo que su regulación podría tener un efecto importante en los contenidos finales de carotenoides de la especie.

Las especies comprenden un conjunto de miRNAs tanto conservados como no conservados (Xu *et al.*, 2010); las conservadas cumpliendo funciones esenciales y conservadas en las plantas en procesos de desarrollo en los tejidos, así como no conservadas con procesos específicas de la especie, pero también importantes en la regulación (Arazi y Khedia *et al.*, 2022; Owusu *et al.*, 2021; Song *et al.*, 2019). Para *Bixa orellana*, se observa que cuenta con la presencia de miRNAs conservados con procesos reportados con funciones en el desarrollo de los tejidos en crecimiento, así como miRNAs no conservados que podrían estar regulando procesos específicos en la planta involucrados en la biosíntesis y acumulación de los carotenoides como bixina.

3.4.3. Expresión de los genes *LCY* del transcriptoma de tres morfotipos de *B. orellana*

Los genes licopeno ciclasa regulan el flujo de los carotenoides mediante la disponibilidad del sustrato licopeno (Zhao *et al.*, 2020). En especies como *Arabidopsis*, tomate y arroz, los niveles de expresión de estos genes afectan la síntesis de carotenoides cíclicos, determinando la relación de sus diferentes productos (Rosas-Saavedra y Stange, 2016).

3.4.3.1. *βLCY1*

El análisis de expresión relativa de los genes *LCYs* identificados en la base de datos de los transcriptomas de achote, mostró al gen *βLCY1* con una expresión con valores similares para los tres morfotipos en hoja y semilla (Figura 3.3). El gen *βLCY1* se expresa en cloroplastos, donde forman parte de las membranas tilacoidales y plato-glóbulos que secuestran y acumulan los carotenoides sintetizados para la captación y dispersión de luz (Maoka *et al.*, 2020). En plantas modelo como *Arabidopsis*, *Oryza sativa* y *Zea mays*, solo se presenta una copia del gen, expresándose preferentemente en órganos verdes (hojas) (Moreno *et al.*, 2013). En tabaco, la introducción y expresión de *βLCY1* influyó en la biomasa vegetal y aumento de los niveles de ABA y GA, resultando en un mejor rendimiento de la planta (Kössler *et al.*, 2021). Para achote, *Boβ-LCY1* ha reportado la presencia de SNPs que generan un polimorfismo funcional en el morfotipo B (flor rosa), el cual correlaciona con una mayor producción de pigmento en la especie (Trujillo-Hdz *et al.*, 2016). Por lo que, los polimorfismos presentes en los genes ciclasa podrían ayudar a la planta a regular su expresión y ser de beneficio para la planta.

3.4.3.2. β LCY2

No obstante, la sobreacumulación de pigmentos carotenoides no es un fenotipo común en las plantas (Zhang *et al.*, 2020). Por esto, algunas plantas emplean isoformas de los genes licopeno ciclasa específicos de tejido, como el cromoplasto (β LCY2), para acumular grandes cantidades de pigmentos en la planta (Sun *et al.*, 2022). Estos se expresan en los frutos y flores de especies como tomate, papaya y azafrán modulando los niveles de síntesis y acumulación de apocarotenoides según los requerimientos de la planta (Zhang *et al.*, 2020; Kössler *et al.*, 2021; Rosas-Saavedra y Stange, 2016). En achiote, $Bo\beta$ LCY2 presentó una alta expresión en hoja en los morfotipos B (flores rosas) y C (púrpura) y baja en el morfotipo A (flor blanca), esto estaría aunado al trabajo en conjunto con β LCY1 para la producción de β -caroteno, pigmento estructural de la fotosíntesis. En semillas, $Bo\beta$ LCY2 también presentó la mayor expresión, aunque de manera diferencial en los morfotipos, principalmente en el morfotipo B (flor rosa) (Figura 3.3), reportada con mayor contenido de bixina (Trujillo-Hdz *et al.*, 2016). Este mismo patrón se observó en níspero para el gen beta específico de cromoplastos ($CYC-\beta$); los niveles de expresión fueron diferenciales entre los frutos de cascará color amarillo y blanco, alcanzando el mayor nivel de expresión en la etapa de maduración (Hong *et al.*, 2019). En la especie, esta acumulación de pigmentos carotenoides, tiene como efecto indirecto fenotipos diversos en las flores, frutos y semillas con tonalidades entre amarillas y rojizas, importantes para la reproducción de la planta al atraer polinizadores (Maoka *et al.*, 2020; Dzib-Cauich *et al.*, 2020; Rogers *et al.*, 2001; Higo *et al.*, 1999; Bate *et al.*, 1998). En este sentido, los genes β LCYs trabajan en las hojas para producir β -caroteno, pigmento estructural de los fotosistemas, en las anteras para capturar luz y en los centros de reacción para dispersarse durante la fotosíntesis. En las semillas, su expresión en tejidos no fotosintéticos, permitiría la acumulación de diversos pigmentos apocarotenoides en flores y frutos.

3.4.3.3. ϵ LCY

En cuanto a $Bo\epsilon$ -LCY, presentó un nivel importante de expresión en hoja para los morfotipos flor rosa (B) y flor púrpura (C), pero débil en semilla para todos los morfotipos (Figura 3.3). La expresión de ϵ -LCY junto a las β -LCYs favorece la rama de productos α , como luteína; pigmento carotenoide que forma parte de las anteras para la captación de luz, esencial para la fotoprotección del fotosistema II en las plantas (Song *et al.*, 2022). Esto se ha observado en mutantes del gen en especies como canola, papa y tabaco, donde, la regulación a la baja y/o silenciamiento del gen propicia la producción de productos carotenoides de la rama β en hojas

(Song *et al.*, 2022; Ke *et al.*, 2019). En cuanto a la expresión débil en las semillas, se ha documentado que su regulación a la baja está aunado al proceso de maduración del fruto, como se ha reportado en tomate (Fu *et al.*, 2019; Rosas-Saavedra y Stange, 2016); lo que parece propiciar la acumulación de apocarotenoides en tejidos no fotosintéticos (cromoplastos). Por lo tanto, se pensaría que *Boε-LCY* llevaría el mismo patrón para la maduración de su fruto y la acumulación de productos carotenoides de la rama β . Con esto, los genes *LCY* muestran una expresión diferencial en achote, orientadas principalmente a la síntesis de carotenoides esenciales en los procesos de fotosíntesis, y en tejidos no fotosintéticos a la acumulación de apocarotenoides durante la maduración del fruto.

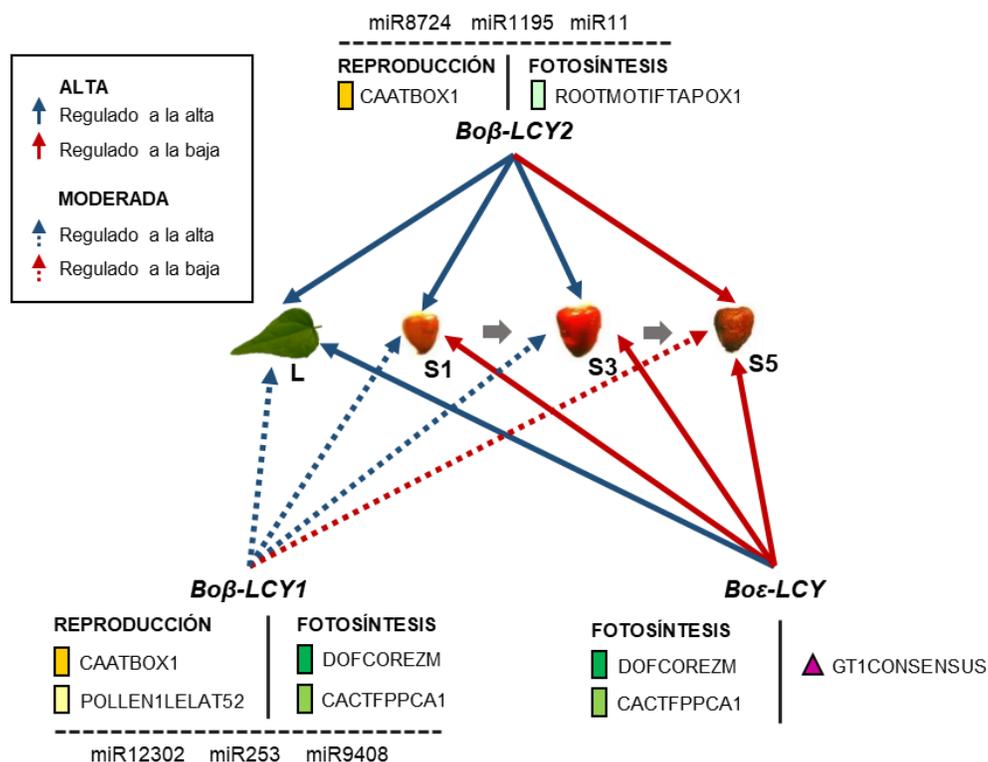


Figura 3.3. Modelo de expresión de las *BoLCY* de *B. orellana* en tejidos fotosintéticos (hojas) y no fotosintéticos (estadios de desarrollo de la semilla). Se muestran los principales TFs (más abundantes) y sus miRNAs asociados en cada gen *BoLCY*. Los estadios de desarrollo de las semillas son: S1 (0-7 d), S3 (14-21 d), S5 (28-42 d). Las líneas rojas representan una regulación al alza y las azules una regulación a la baja. Las líneas punteadas indican una regulación más moderada.

CAPÍTULO IV

ANÁLISIS CITOGÉNÉTICO DE *Bixa orellana* L.

4.1. INTRODUCCIÓN

Los estudios citogenéticos se centran en el estudio de la estructura y evolución de los cromosomas de las diferentes especies (Windham *et al.*, 2020; Aiassa *et al.*, 2015; Herrera, 2007). Estos aportan información básica sobre el número cromosómico, ploidía, detección de anomalías en la división celular; siendo la estabilidad de estas características importantes para la transferencia y fidelidad del material genético a las próximas generaciones, mientras que su variación proporciona soporte genético para la diferenciación y adaptación (Carta *et al.*, 2020). En las especies, las características cromosómicas suelen ser constantes, sin embargo, en las plantas puede existir una gran variación en la morfología y/o número cromosómico dentro de una especie (Windham *et al.*, 2020; Li *et al.*, 2017; Soltis *et al.*, 2009; Brown y Bertke, 1979). Las características estructurales y cuantitativas de los cromosomas son visualizadas en el cariotipo, el cual, es la descripción de cada conjunto cromosómico de un individuo o especie ordenado de acuerdo a la morfología y tamaño, representando el nivel más alto de organización estructural y funcional del genoma nuclear (Carta *et al.*, 2020; Silva *et al.*, 2017; Soltis *et al.*, 2009; Herrera, 2007; García, 1985). Este tipo de estudios son importantes en la investigación taxonómica y aplicada, al contribuir sobre la relación genética entre especies o poblaciones y una mejor comprensión de la forma en que divergieron entre sí; siendo los cromosomas, parte de un sistema dinámico que está moldeando el proceso evolutivo (Li *et al.*, 2017; Ferguson-Smith, 2015; Guerra, 2008; Poggio y Naranjo, 2004; Swanson *et al.*, 1981).

Para *Bixa orellana*, una búsqueda en Index to Plant Chromosome Number (2023), una base de datos de números cromosómicos de especies vegetales, registra $2n= 14$ y 16 cromosomas diploides para individuos de India, Brasil, Reino Unido y Venezuela (Mukherjee, 1975, Krishnan, 1977, Krishnan y Ayyangar, 1987, Carvalheira *et al.*, 1991, Hanson *et al.*, 2001), generando una discrepancia entre diversos autores sobre el número cromosómico de la especie. Con esto, Michelangeli y colaboradores (2002, 2006) analizaron el complemento cromosómico de células madre del grano de polen, para determinar la carga cromosómica haploide de genotipos de Venezuela, reportando un número haploide de $n= 8$ cromosomas. Para Brasil, el estudio de sus accesiones, lograron reportar el número cromosómico diploide de $2n= 14$ y haploide de $n= 7$, así como describir el cariotipo de la especie (Almeida *et al.*, 2006; Lombello y Pinto-Maglio, 2004). En las plantas, estas variaciones en número, forma o tamaño de los cromosomas pueden

observarse en cambios en el fenotipo de la especie (Windham *et al.*, 2020; Li *et al.*, 2017; Murray, 2012), que se presentan como variantes o cultivares con rasgos particulares (Poggio y Naranjo, 2004).

Para las poblaciones de Achiote de Yucatán, se ha reportado una alta variabilidad morfológica en color de fruto, flores y producción de *bixina*, principal carácter de interés agronómico (Rivera-Madrid *et al.*, 2006; Valdez-Ojeda *et al.*, 2008; Avedaño-Arrazate *et al.*, 2018). Por esto mismo, se caracterizaron tres grupos de accesiones con morfologías contrastantes: G1 de flores blancas, frutos indehiscentes color verde y menor contenido de bixina; G2 de flores rosas, frutos dehiscentes, color verde y espinas rojas con mayor contenido de bixina; y G3 de flores purpura, fruto dehiscentes color rojo y contenido intermedio de bixina; identificando al grupo G2 (flor rosa) con el mayor contenido de bixina, característica deseable para su aprovechamiento (Trujillo-Hdz *et al.*, 2016). Por esto, fue de interés analizar el complemento cromosómico de la accesión “Peruana roja”; caracterizada por flores rosas y alto contenido de bixina, reportando un complemento cromosómico diploide $2n=14$ (Martín-Rodríguez, 2007). Estos fenotipos son reflejo de la variación intraespecífica del genoma; siendo los cromosomas la representación del nivel más alto y el gen, la unidad fundamental de organización estructural y funcional del DNA nuclear. La ubicación de los genes responsables de la diferenciación de estos fenotipos en los cromosomas, particularmente de aquellos involucrados en la producción de pigmentos, como es el caso de los genes licopeno ciclasa, apoyaría junto con el estudio de la estructura y regulación de estos el establecimiento de criterios agronómicos para la futura selección de genotipos y programas de mejoramiento genético en esta planta. Sin embargo, las variaciones en número, forma o tamaño de los cromosomas de acuerdo con el fenotipo de esta especie de planta dificultan este análisis. Por lo tanto, el presente capítulo aborda la implementación y análisis de las características cuantitativas y estructurales de los cromosomas, para determinar el cariotipo de las poblaciones de *Bixa orellana* en la península de Yucatán, así como construir el ideograma con el fin de tener un panorama general del genoma de la especie, y específicamente de las accesiones con fenotipos altamente productores de bixina.

4.2. MATERIALES Y MÉTODOS

4.2.1 Material vegetal

Se emplearon semillas del morfotipo B (accesión N4) de *Bixa orellana*, previamente caracterizadas en el trabajo de Trujillo-Hdz y colaboradores (2016) de acuerdo a sus

características morfológicas: plantas con flores color rosa, frutos color verde con espinas rojas dehiscentes y mayor contenido de bixina. El material fue proporcionado por el laboratorio 27 de la Unidad de Bioquímica y Biología Molecular de Plantas del CICY, correspondiente a la cosecha de 2 de marzo 2019 ubicado en Dzityá, Mérida, Yucatán. Para obtener tejido meristemático para la técnica citogenética; las semillas se hidrataron en agua entre 10-15 días hasta romper la testa. Luego, se pusieron en charolas de germinación en cuarto de cultivo en condiciones de luz y temperatura estable (25°C) hasta germinar.

4.2.2. Citogenética de achiote

El análisis citogenético se realizó con modificaciones a la técnica reportada para la especie por Martín (2007). La colecta de raíces se realizó entre las 8:30 a 10:30am, estas se pretrataron con 8-hidroxiquinoleína 0.002M durante 5h, luego se lavaron con agua y fueron fijadas en metanol-ácido acético 3:1 a 4°C por 24h hasta su uso. Para la hidrólisis, las raíces se pusieron en HCl 1N a 60°C por 15min; seguido, las raíces fueron teñidas con 50µl de solución carbón-fucsina por 20 min en oscuridad a temperatura ambiente. Los meristemos teñidos se seleccionaron y se colocan en portaobjetos agregando una gota de aceto-orceína 1.8% realizando un ligero aplastamiento, para luego pasar por flama para adherir el tejido. La observación de las preparaciones cromosómicas se realizó en el microscopio Carl Zeiss Axioplan en objetivo 40X con una apertura numérica (A.N) de 0.17. Las imágenes se capturaron en la cámara AxioCamICc5 en objetivo 40X con A.N. 0.17 y 100X con A.N. 0.17 mediante el software AxioVision SEG4 Rel. 4.8. Las fotografías en metafase seleccionadas, se procesaron en el programa Fiji (Schindelin *et al.*, 2012) para unificarlas a escala de grises, e identificar la morfología de los cromosomas y realizar el conteo cromosómico.

4.2.3. Caracterización de los cromosomas de *Bixa orellana*

Las imágenes se cargaron en el software KaryoMeasure (Mahmoudi y Mirzaghaderi, 2023) para realizar el análisis del cariotipo y posteriormente, construir el ideograma de la especie. Para el cariotipo, se midieron los siguientes parámetros: longitud del brazo largo (L), longitud del brazo corto (S), longitud del cromosoma (CL); con lo cual, se calculó la relación del brazo (AR) y se clasificó a los cromosomas según la posición del centrómero (Levan *et al.*, 1964). Después, los parámetros del cariotipo obtenidos se emplearon para construir el ideograma de la especie, mediante la opción “generación de ideograma”.

4.3. RESULTADOS

4.3.1 Citogenética de Achiote

Se modificó la técnica establecida por Martín (2007) para obtener células completas y cromosomas dispersos. La colecta de raíces se hizo en material recién germinado (4 a 5 días luego de ser puestas en cajas de germinación), la cual se realizó entre 8:30 a 10:30am, momento donde se observó la división celular y mayor presencia de células con cromosomas en metafase. El material se fijó por 5h dando un mejor resultado para conservar la morfología de los cromosomas. Seguido, los meristemos se sometieron a una hidrólisis ácida con HCl 1N por 15min, donde las células mantuvieron su forma. La tinción de fucsina básica junto con la contratinción de aceto-orceína 1.8% previo al aplastamiento fue efectiva para la visualización de las fases mitóticas y la identificación de los cromosomas en las células (Figura 4.1).

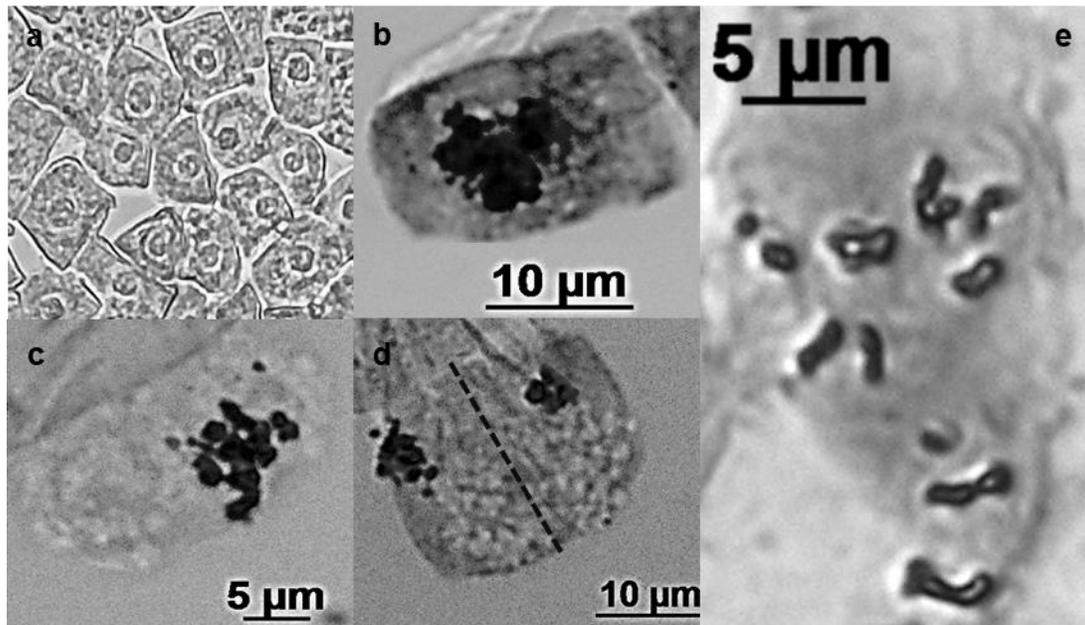


Figura 4.1 Fases mitóticas de achiote, morfotipo B (flor rosa). a) Interfase, b) Profase, c) Metafase y d) Telofase 40X, A.N. 0.17; e) Metafase 100X A.N. 0.17.

4.3.2 Cariotipo e ideograma de *Bixa orellana*

El análisis citogenético de la especie mostró un complemento cromosómico de $2n= 14$ (Figura 4.2). El análisis de los cromosomas realizado con el programa KaryoMeasure (Mahmoudi y Mirzaghaderi, 2023) identificó a los pares de cromosomas 1 a 6 como metacéntricos (m) y el par

7 como submetacéntrico (sm), según los parámetros de longitud de sus brazos (L y S), su relación (AR) y la posición del centrómero, según la clasificación de Levan y colaboradores (1964) (Tabla 4.1). En cuanto al tamaño; los cromosomas metafásicos de la especie variaron en de 3.2 μm (par 1) que presentó casi el doble de longitud que los demás cromosomas, a 2.2 μm (par 2) a .90 μm (par 7). Las características de los cromosomas observados en este trabajo para achiote, específicamente el morfotipo B (flor rosa), entran en el rango de tamaño cromosómico antes reportado por Almeida y colaboradores (2006) y Lombello y Pinto-Maglio (2004) para la especie (Tabla 4.1).

Cuadro 4.1. Longitud de los pares de cromosomas de *B. orellana*. Longitud del brazo largo (L), longitud del brazo corto (S), longitud del cromosoma (CL), relación del brazo (AR) y clasificación según la ubicación del centrómero (Levan *et al.*, 1964).

Par	L (μm)	S (μm)	CL (μm)	AR	Clasificación
1	1.95	1.24	3.19	1.57	m
2	1.26	0.95	2.21	1.32	m
3	0.97	0.85	1.82	1.14	m
4	1	0.75	1.75	1.32	m
5	0.85	0.78	1.64	1.08	m
6	0.63	0.48	1.12	1.30	m
7	0.58	0.33	0.91	1.72	sm

4.4. DISCUSIÓN

4.4.1 Citogenética de Achiote

La investigación citogenética estudia la organización del genoma a nivel cromosómico, siendo uno de sus objetivos básicos el determinar el número y morfología de los cromosomas (Schwarzacher *et al.*, 2023). Un paso preliminar para esto, es obtener buenas preparaciones cromosómicas, lo cual, implica un gran esfuerzo y en ocasiones es necesario evaluar y modificar los procedimientos (Schwarzacher *et al.*, 2023; Maravilla *et al.*, 2023). Para achiote, la metodología establecida por Martín (2007) en una población Yucateca para el análisis cromosómico, se mostró demasiado violento con el tejido meristemático lo que dificultó la observación y análisis de cromosomas. Por esto, se modificó el protocolo en tiempos de colecta,

pretratamiento y tinción, con el fin de obtener material que nos permita identificar los cromosomas de la especie.

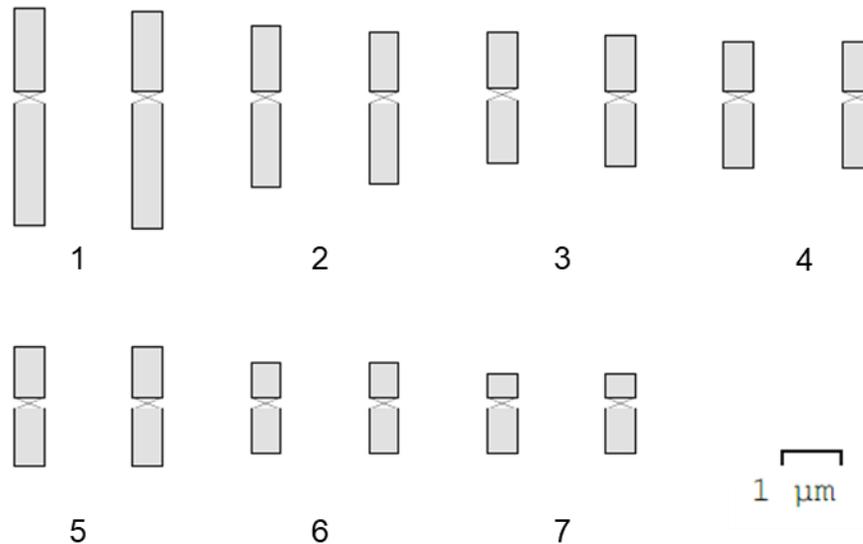


Figura 4.2. Ideograma somático de *Bixa orellana*, morfolo B (flor rosa). Barra 1 μm

La colecta de material se realizó de entre 8:30 a 10:30am donde se observó un mayor número de células en profase y metafase, a 5h de pretratamiento, para mejorar el empaquetamiento del material genético. Es en estas fases que durante el proceso de mitosis las copias duplicadas de los cromosomas presentan la mayor condensación (se vuelven visibles) y se separan más entre sí (Schwarzacher *et al.*, 2023; Swanson *et al.*, 1981). Seguido fue la hidrólisis, esta ayuda a eliminar tejido, como la pared celular, que impida visualizar los cromosomas (Marcela-Delgado *et al.*, 2010). No obstante, la hidrólisis más la digestión enzimática presentada por Martín (2007) se mostró agresiva con el tejido; por lo que se optó por implementar una hidrólisis ácida con HCl1N por 15min a 60°C, dando mejores resultados en la integridad del tejido meristemático. Por último, para la tinción se añadió aceto-orceína al 1.8% como contratinción previo al aplastamiento (Marcela-Delgado *et al.*, 2010; Rodríguez y Bueno, 2006), lo que resultó efectivo para la visualización, diferenciación y conteo de los cromosomas. No obstante, los estudios citogenéticos no solo ayudan a determinar el número de cromosomas, sus protocolos permiten analizar su forma y tamaño (Schwarzacher *et al.*, 2023; Mahmouding y Mirzaghaderi, 2023). Estas características tienen implicaciones significativas para entender la organización y comportamiento de los cromosomas, y es una primera etapa para interpretar genotipos, eventos de hibridación y evolutivos de las especies (Schwarzacher *et al.*, 2023; Carta *et al.*, 2020). Por lo

que, los estudios citogenéticos contribuyen a una mejor comprensión de los cromosomas y la organización del genoma (Schwarzacher *et al.*, 2023; Carta *et al.*, 2020; Li *et al.*, 2017; Ferguson-Smith, 2015).

4.4.2 Cariotipo e ideograma de *Bixa orellana* L.

El análisis cromosómico realizado en una accesión de achiote (morfortipo B, flor rosa), reportó $2n= 14$ cromosomas y concuerda con lo reportado para accesiones de Brasil (Lombello y Pinto-Maglio, 2014; Almeida *et al.*, 2006) y para una población del estado de Yucatán (Martín, 2007). Sin embargo, una serie de reportes $n= 8$ cromosomas que datan desde 1954 (Baer, 1976; Simmonds, 1954) y en accesiones de Venezuela e India (Michelangeli *et al.*, 2006, 2002; Krishnan y Ayyangar, 1987; Krishnan), han mantenido el debate sobre los eventos que pudo haber sufrido la especie, aneuploidía, delección (Bear, 1976) o incluso fragmentación de cromosomas (Almeida *et al.*, 2006) durante la evolución de *Bixa*. Aunque en las plantas la variación en las características cromosómicas (morfología y/o número cromosómico) dentro de una especie es algo bastante común (Windham *et al.*, 2020; Li *et al.*, 2017; Soltis *et al.*, 2009; Zhao-Yang *et al.*, 2002); estos cambios suelen tener consecuencias en el fenotipo (Windham *et al.*, 2020; Li *et al.*, 2017; Murray, 2012), que luego se observan como variantes o cultivares con rasgos particulares (Poggio y Naranjo, 2004). Por lo que, estudios de las características cuantitativas (número) y estructurales (tamaño y forma) de los cromosomas, que comprende el cariotipo; nos proporcionaría un panorama general de la estructura, organización y funcionalidad del genoma de una especie (Schwarzacher *et al.*, 2023; Carta *et al.*, 2020; Silva *et al.*, 2017; Guerra, 2008).

B. orellana presentó mayormente cromosomas pequeños, sobresaliendo el par uno con 3.1 μm , de los seis pares de cromosomas restantes que oscilaron entre 2.2 a 0.90 μm . Estos tamaños entraron en el rango presentado para accesiones de Brasil (3.2 a 1.2 μm) (Lombello y Pinto-Maglio, *et al.*, 2014; Almeida *et al.*, 2006) para los pares 1 a 6; donde solo el par siete se presentó un poco más pequeño de lo reportado. Esta variación cariomorfológica intraespecífica, principalmente en tamaño de los cromosomas también se ha reportado en la especie *Spiraea japónica* (Zhao-Yang *et al.*, 2002). En cuanto a la clasificación de los cromosomas, se identificó a los pares 1 a 6 como metacéntricos y al par 7 como submetacéntrico (6m+1sm) según el sistema de terminología cromosómica (Levan *et al.*, 1964). Esto concuerda el reporte de Lombello y Pinto-Maglio, (*et al.*, 2014) pero difiere de lo presentado por Almeida (*et al.*, 2006)

con una fórmula cariotípica $5m+2sm$.

Las variaciones en las características cariotípicas reportadas para de la especie, podrían estar aunadas a las diferentes poblaciones con características contrastantes de la especie (Valdez-Ojeda *et al.*, 2008; Rivera-Madrid *et al.*, 2006), principalmente en el fruto, de donde se extrae bixina, pigmento de importancia económica (Rivera-Madrid *et al.*, 2006; Avedaño-Arrazate *et al.*, 2018). No obstante, la mayoría de los análisis citogenéticos reportados de la especie, no cuentan con una caracterización o descripción morfológica de la accesión. Para poblaciones de Yucatán, se han caracterizado tres grupos morfológicos (flor blanca, flor rosa y flor purpura), identificando al grupo con mayor producción del pigmento (flor rosa). Por lo que el presente estudio, sería el primer reporte de análisis cariotípico para una accesión caracterizada morfológicamente, presentada en un ideograma, cuyo principal interés y característica es una alta producción de bixina en comparación con sus otras poblaciones. Esto aunado a técnicas de hibridación *in situ* de secuencias repetitivas y de genes involucrados con la producción de bixina, como los genes licopeno ciclasa, nos ayudaría a explicar la variación intraespecífica; y/o de genes de interés, nos aportarían información para futuros trabajos de mejoramiento genético.

CAPÍTULO V

DISCUSIÓN, CONCLUSIONES Y PERSPECTIVAS

5.1. DISCUSIÓN

Bixa orellana es una especie que se caracteriza por la tonalidad naranja-rojiza de sus semillas debido al compuesto bixina (Avedaño-Arrazate *et al.*, 2018; Pérez *et al.*, 2003; Arce, 1999), este cuanta con una importancia económica al emplearse en la elaboración de colorantes para la industria de alimentos, farmacéutica y cosmética (Cárdenas-Conejo *et al.*, 2023). No obstante, la especie presenta una alta heterocigocidad, generando una gran variación en la calidad y producción del pigmento, lo que afectan su producción a mayor escala y plantaciones con caracteres deseables y homogéneas (Rivera-Madrid *et al.*, 2006). En 2015, Cárdenas-Conejo y colaboradores propusieron un modelo de biosíntesis de bixina, donde se menciona a los genes licopeno ciclasa (*Boβ-LYC1*, *Boβ-LYC2*, *Boε-LYC*) como clave en la producción de bixina, al desactivarse impido el flujo de licopeno por la vía de los carotenoides, aumentando su disponibilidad para la producción de bixina. En 2016, Trujillo-Hdz y colaboradores (2016), lograron caracterizar tres morfotipos de achiote con diferente producción de bixina mediante la variación alélica de un solo nucleótido (SNP) del gen *Boβ-LCY1*, así como determinar al grupo con mayor producción de bixina. Por esta razón, el presente estudio se centra en el análisis de la estructura de los genes *LCY*, elementos regulatorios, así como la organización del genoma a nivel cromosómico, como un conjunto de factores que nos brindaría información para dilucidar la variación específica de la especie y determinar su relación con su producción de bixina.

5.1.1. Estructura de los genes *BoLCY*

Se identificaron tres *LCY* en el genoma de achiote que corresponden con las reportadas en el transcriptoma de *novo* para la especie (Cárdenas-Conejo *et al.*, 2015). Los genes *Boβ-LCY* se caracterizan por presentar una única secuencia codificante, mientras que *Boε-LCY* mostró 11 secuencias codificantes. En general, todos los genes *BoLCY* tienen similitudes significativas con otros ortólogos reportados en plantas (Rosas-Saavedra y Stange 2016). *Boβ-LCY* presentó regiones altamente conservadas, particularmente el sitio de unión de dinucleótidos, una firma presente en todo caroteno ciclasa conocidas (Cunningham *et al.*, 1996). También se identificaron el motivo ciclasa I y el motivo ciclasa II, dominios esenciales para su actividad catalítica (Moreno *et al.*, 2013; Alquézar *et al.*, 2009). *Boβ-LCY2* mostró algunos cambios de aminoácidos en la región conservada β-LCYs, en el motivo específico *LCY*, motivo ciclasa II y motivo β-LCYs en

comparación con *C. sinensis* (Alqu zar et al., 2009). Esto podr a deberse al hecho de que β -*LCY2* generalmente est a presente en los cromoplastos, como una isoforma que podr a contener algunos cambios necesarios para la acumulaci n de pigmento en tejidos no fotosint ticos (Sun et al., 2022). Para *Bo β -LCY1*, se ha reportado que hay SNPs relacionados con una alta acumulaci n de pigmento en las diferentes accesiones (Trujillo-Hdz et al., 2016). Sin embargo, ninguna de estas variaciones de nucle tidos se observ  en el genoma del morfotipo B (flor rosa), probablemente como consecuencia de analizar un solo genoma y no una poblaci n de plantas (Du et al., 2019).

5.1.2. Regulaci n de los genes *BoLCY*

El an lisis de la regi n aguas arriba (3000 pb) de los tres *BoLCY* identific  los sitios putativos de uni n para los TF que podr an estar regulando su expresi n g nica. La identificaci n de los TF es importante en la v a de los carotenoides, ya que muchos de ellos podr an ser importantes para la regulaci n de la transcripci n, las se ales ambientales, el desarrollo y el metabolismo de los carotenoides (Sun y Li, 2020) mediante la activaci n o supresi n de genes estructurales (Sun et al., 2022). Se identificaron dos principales grupos de TF asociados a funciones de: I. fotos ntesis (DOFCOREZM y CACTFTPPCA1) y reproducci n (CAATBOX1, POLLEN1LELAT52 y GTGANTG10) y II. sistema oxidativo (GATABOX y GT1CONSENSUS) y defensa a factores ambientales (MYCCONSENSUSAT). En la categor a de fotos ntesis; los motivos CACTFTPPCA1 (5'-CACT-3'), DOFCOREZM (5'-AAAG-3') y ROOTMOTIFTAPOX1 (5'-ATATT-3') est n relacionados con la expresi n en tejidos verdes (brotes, ra ces y mes filo de las hojas) (Yanagisawa, 2000; Higo et al., 1999), involucrados en la captura y absorci n de luz en rangos donde la clorofila no es funcional durante la fotos ntesis (380–500 nm) (Lohani et al., 2021; Maoka, 2020; Aguilar-Espinosa et al., 2022; Gowik et al., 2004). En cuanto a la categor a reproducci n, los motivos CAATBOX1 (5'-CAAT-3'), POLLEN1LELAT52 (5'-AGAAA-3') y GTGANTG10 (5'-GTGA-3') tambi n tuvieron una alta presencia. CAATBOX1 es un elemento regulador de la actividad en el tejido de la semilla de guisante (Li et al., 2014; Sharma et al., 2011), y POLLEN1LELAT52 es uno de los dos elementos reguladores activadores espec ficos del polen del gen LAT52 en *L. esculentum* (Li et al., 2014; Filichkin et al., 2004). En tejidos como semillas, flores y frutos, los carotenoides se acumulan en grandes cantidades, proporcionando tonalidades atractivas a los polinizadores y asegurando la reproducci n de las plantas (Dzib-Cauch et al., 2020; Rogers et al., 2001; Higo et al., 1999; Bate et al., 1998). Respecto a las categor as del subgrupo II; el sistema oxidativo tuvo los motivos GT1CONSENSUS (5'-

GRWAAW-3'), GATABOX (5'-GATA-3') y EBOXBNNAPA (5'-CANNTG-3') que se regulan en respuesta a la luz. En los tejidos fotosintéticos, la luz es la señal más importante para la regulación de la expresión de genes estructurales de carotenoides (Sun *et al.*, 2022). Esta activa la síntesis y acumulación de pigmentos carotenoides que ayudan a la membrana a defender el aparato fotosintético de la fotooxidación (Maoka, 2020; Dzib-Cauich *et al.*, 2020; Muthamilarasan *et al.*, 2014; Reyes *et al.*, 2004; Higo *et al.*, 1999). Para la categoría de defensa a factores ambientales, los motivos MYCCONSENSUSAT (5'-CANNTG-3') estuvieron ampliamente representados. Este es un sitio de reconocimiento MYC involucrado en respuestas de defensa biótica como ataques de patógenos y heridas (Deb *et al.*, 2015), y respuestas de defensa abiótica como deshidratación y frío (Li *et al.*, 2014). Estas respuestas son inducidas por ABA, un apocarotenoide importante en la respuesta al estrés ambiental relacionado con los procesos de desarrollo y crecimiento de la planta (Ramamoorthy *et al.*, 2020; Higo *et al.*, 1999). En general, todos los TF putativos identificados desempeñan funciones esenciales para las plantas y son llevadas a cabo por carotenoides sintetizados de *novo* en los plastidios de los tejidos fotosintéticos y no fotosintéticos (Sun *et al.*, 2022; Maoka, 2020; Nisar *et al.*, 2015; Rosas-Saavedra y Stange, 2016).

Por otra parte, los *BoLCY* se observaron con diferencias en la presencia de factores de transcripción putativos. *Boβ-LCY1* presentó la presencia más baja, lo que indica un nivel más específico de regulación genética. La menor aparición de TF también podría sugerir un patrón de expresión más estable y constitutivo en los cloroplastos para una acumulación eficiente de carotenoides y absorción de luz. En la raíz de zanahoria, la luz es la señal más importante para la regulación de la expresión de genes estructurales de carotenoides (Llorente *et al.*, 2017; Sun *et al.*, 2022). Con respecto a los órganos reproductivos, las señales de desarrollo parecían ser el principal determinante de la regulación de los genes (Sun y Li 2020; Sun *et al.*, 2022), como se ha informado en *C. papaya* y *Actinidia spp.* (Zhou *et al.*, 2019; Ampomah-Dwamena *et al.*, 2009) y tomate (Bian *et al.*, 2011). Es en este último que, durante la maduración del fruto, los cloroplastos sufren una transición a cromoplastos (orgánulos no fotosintéticos), principal sitio de acumulación de diversas y grandes cantidades de carotenoides (Sun *et al.*, 2022; Bian *et al.*, 2011). En Achiote, se informa que *Boβ-LCY2* está asociado con cromoplastos (Vishnevetsky *et al.*, 1999). Este gen mostró un mayor número de TF con respecto a *Boβ-LCY1*. De acuerdo con la diversidad y número de TF que ocurren en *Boβ-LCY2*, es posible que su gen desempeñe un papel regulador en la biosíntesis y acumulación de carotenoides en *B. orellana*. Finalmente, *Boε-LCY* presenta el mayor número de TF en las categorías de fotosíntesis, reproducción, sistema

oxidativo y defensa a factores ambientales. Este tiene un papel regulador muy importante al intervenir junto con β -LCY para generar una gran diversidad de compuestos (Zhao *et al.*, 2020), especialmente carotenoides, como se ha reportado en el maíz (Harjes *et al.*, 2008).

5.1.3. microRNAs de los genes *BoLCY*

Los microRNAs tienen una función reguladora en numerosos procesos biológicos en las plantas (Owusu *et al.*, 2021; Gao *et al.*, 2015) y activa en el metabolismo secundario, como los carotenoides (Owusu *et al.*, 2021; Bulgakov y Avramenko, 2015). Para *B. orellana*, se confirmó en la base de datos del transcriptoma, la presencia de 11 familias de miRNAs no conservados, identificados en la región promotora de los genes *BoLCY*: miR253, miR408, miR12302, miR991, miR1195, miR11, miR5181, miR8724, miR9571, miR10366 y miR6546. Los miRNAs no conservados, suelen ser específicos de la especie o presentar una baja expresión (Owusu *et al.*, 2021; Sun *et al.*, 2012), involucrándose en el desarrollo de rasgos específicos (Dalmay, 2010). Solo tres miRNAs identificados reportaron función en otros sistemas: miR9408 en *Brassica aleracea* y *B. rapa* en desarrollo del polen (Lukasik *et al.*, 2013) asociado a *Bo β -LCY1*; miR5181 en *Brachypodium distachyon* implicado crecimiento de hojas durante la sequía (Bertolini *et al.*, 2013) y miR8724 en *Gossypium raimondii* involucrado en el alargamiento de las células de fibra de algodón (Xue *et al.*, 2013) asociados a *Bo β -LCY2*. Para *Bo β -LCY1* presente en cloroplastos y con función de captación de luz (Lohani *et al.*, 2021; Aguilar-Espinosa *et al.*, 2020; Trujillo-Hernández, 2016), la familia miR9408 podría estar involucrado en la regulación negativa del desarrollo del polen. Para *Bo β -LCY2*, asociado a cromoplastos y acumulación de grandes cantidades de carotenoides (Sun *et al.*, 2022; Escobar-Turriza, 2017; Bian *et al.*, 2011); miR5181 podría implicarse a procesos en las hojas como respuesta a estrés abiótico (luz) y miR8724 en la acumulación de carotenoides en tejidos no fotosintéticos, ambos en una posible regulación positiva.

Por otra parte, los miRNAs conservados se presentan en diferentes familias (Owusu *et al.*, 2021; Ciu *et al.*, 2017) y regulan la expresión de procesos básicos del desarrollo (Owusu *et al.*, 2021; Song *et al.*, 2019; Dalmay, 2010). Se han documentado 9 familias mejor conservadas de miARNs por presentarse en la mayoría o en todos los linajes de plantas terrestres (miR156, miR160, miR166, miR171, miR319, miR390, miR477, miR529 y miR535) (Arazi y Khedia *et al.*, 2022; Axtell y Meyers, 2018). Una especie donde se ha realizado el estudio de miRNAs tanto conservados como no conservados, es el tomate, especie modelo y de gran importancia

económica, ya que tienen un papel esencial en el desarrollo y maduración de sus frutos (Arazi y Khedia *et al.*, 2022; Gao *et al.*, 2015; Dalmay, 2010). Para achiote, se observó la presencia de todos los miRNAs conservados en tomate y entre ellas, 5 miRNAs de las familias mejor conservadas en las plantas terrestres (miR156, miR160, miR166, miR171 y miR390). Con esto, en este estudio se identificó miRNAs no conservados asociados a los genes β -LCY posiblemente involucrados en procesos específicos del metabolismo de los carotenoides, y miRNAs conservados inmersos en procesos esenciales como el desarrollo y crecimiento de la planta.

5.1.4. Secuencias repetitivas de DNA de los genes *BoLCY*

Los elementos repetitivos de DNA tienen un impacto importante en la organización, complejidad, función y estabilidad del genoma. Su capacidad de replicación y movimiento de sus copias resulta en cambios en el tamaño y la estructural cromosómica (Boutte *et al.*, 2020; Wang *et al.*, 2016; Lee y Kim, 2014) que, a su vez, pueden inducir una diversidad de mutaciones que pueden afectar la expresión de genes cercanos (Zhang *et al.*, 2020b). En las plantas con flores, los elementos LTR son los más abundantes como resultado de la dinámica entre la inserción de nuevas copias y la eliminación por eventos de recombinación (Suguiyama *et al.*, 2019). Se investigó la variabilidad de los *BoLCY* como consecuencia de la presencia de secuencias LTR.

Se detectaron diversos motivos LTR y proteínas asociadas. La mayoría de estos se presentaron en el gen *Bo ϵ -LCY* (motivos 1, 2 y 3, y dominio proteico AP Del), mientras que el gen *Bo β -LCY1* tiene los motivos M1 y el dominio proteico GAG. Todos los motivos pertenecían a los LTR (clase I) pertenecientes a la familia Ty3/Gypsy. Finalmente, *Bo β -LCY2* no se detectó ningún motivo, pero muestra un LTR V12 (familia Ty1/Copia) y un fragmento de proteína RNasaH. Estos motivos LTR corresponden a las secuencias cohesivas de las LTR, cuya función es integrarse en los sitios objetivo de una nueva ubicación en el genoma (Lee y Kim, 2014) y se encuentran ampliamente en las plantas (Suguiyama *et al.*, 2019). El dominio GAG y las regiones RNasaH identificadas generalmente se encuentran en la región interna de elementos retrotransponibles que codifican diversos productos proteicos necesarios para los mecanismos de replicación y transposición (Lee y Kim, 2014). Otros motivos identificados en el estudio actual incluyeron regiones relacionadas con la proteína GAG, correspondientes a una proteína de la cápside y POL, involucradas en la transcripción de retroelementos (Lee y Kim, 2014; Llorens *et al.*, 2011). Los elementos transponibles pueden contener secuencias que actúan en cis y que, al

fragmentarse e integrarse en genes o cerca de ellos, generan efectos en los niveles de expresión, principalmente con genes cercanos regulados positivamente en respuesta a varios tipos de estrés abiótico, biótico y reproductivo, significativos en función del genoma (Boutte *et al.*, 2020; Wang *et al.*, 2016). En el actual estudio, *Boε-LCY* fue el gen que exhibió la mayor cantidad de motivos LTR y dominios proteicos, lo que concuerda con la gran cantidad de TF identificados, indicando que este gen desempeña un papel regulador en la producción de pigmentos para tejidos fotosintéticos y en los cromoplastos.

5.1.5. Expresión de los genes *LCY* de tres morfotipos de *Bixa orellana* L.

Los genes *LCY* regulan el flujo de los carotenoides a través de la disponibilidad de licopeno (Zhao *et al.*, 2020), sustrato considerado precursor de la bixina (Cárdenas-Conejo *et al.*, 2023; Carballo-Uicab *et al.*, 2019). En especies como *Arabidopsis*, tomate y arroz, los niveles de expresión de estos genes afectan la síntesis de carotenoides cíclicos, determinando la proporción de sus diferentes productos (Rosas-Saavedra y Stange 2016). Los genes *BoLCY* se identificaron como parte de un análisis de transcriptoma en las semillas de tres morfotipos contrastantes (A, flor blanca; B, flor rosa y C, flor purpura) (Cárdenas-Conejo *et al.*, 2023). Por lo que el presente estudio muestra un análisis *in silico* de los *BoLCY* en las semillas y su comparación con las hojas para cada morfotipo. *Boβ-LCY1* mostró un comportamiento de expresión bajo en comparación con *Boβ-LCY2* en todos los tejidos, disminuyendo en las semillas maduras en todas los morfotipos. *Boβ-LCY1* suele expresarse en cloroplastos, formando parte de las membranas tilacoides y plastoglobulos (Maoka *et al.*, 2020), donde los carotenoides capturan y absorben luz en rangos donde la clorofila no es funcional durante la fotosíntesis (380-500 nm) (Lohani *et al.*, 2021; Maoka, 2020; Aguilar-Espinosa *et al.*, 2020). *Boβ-LCY2* mostró una expresión diferencial en hojas, pero alta en los morfotipos B (flor rosa) y C (flor purpura) y baja en la A (flor blanca). Esto estaría relacionado con la función de *β-LCY1* de producir *β*-caroteno, un pigmento estructural de la fotosíntesis en las hojas. Algunas plantas ajustan su contenido de carotenoides empleando *β-LCY2* específicos de tejidos. Estos son principalmente de cromoplastos, se expresan en frutos y flores de especies como tomate, papaya y azafrán, modulando los niveles de síntesis de apocarotenoides según los requerimientos de la planta (Zhang *et al.*, 2020; Kässler *et al.*, 2021; Rosas-Saavedra y Strange, 2016). Un patrón similar se observó en el níspero (*Eriobotrya japonica*) para el gen beta específico de cromoplasto (*CYC-β*), donde los niveles de expresión fueron diferenciales entre frutos de cáscara amarilla y blanca, alcanzando su pico en la etapa de maduración (Hong *et al.*, 2019). Con esto, las diferencias en la expresión de *BoLCY*

entre accesiones con diferentes rasgos florales podrían estar involucradas con un efecto indirecto sobre el color de las flores, un rasgo esencial para la reproducción de las plantas al atraer polinizadores (Maoka, 2020; Dzib-Cauich *et al.*, 2020; Rogers *et al.*, 2001; Higo *et al.*, 1999; Bate *et al.*, 1998). En este sentido, los β -LCY se expresan en las hojas para producir β -caroteno, un pigmento estructural de los fotosistemas. Curiosamente, los $Bo\beta$ -LCY, principalmente β -LCY2, se expresan durante las etapas iniciales de desarrollo de la semilla, lo que podría contribuir a la acumulación de compuestos de caroteno cíclicos en estas etapas de la semilla.

$Bo\epsilon$ -LCY mostró una regulación positiva en las hojas del morfotipo B (flor rosada) y C (flor purpura), pero débil durante el desarrollo de la semilla. La expresión de ϵ -LCY junto con β -LCY produce luteína, un pigmento carotenoides que forma parte de las anteras para la captación de luz y es esencial para la fotoprotección del fotosistema II en las plantas (Song *et al.*, 2022). Estudios sobre mutantes de ϵ -LCY en especies como canola, papa y tabaco, demostraron que la regulación negativa y/o el silenciamiento de ϵ -LCY favorece la producción de productos carotenoides de la rama β en las hojas, cuyas funciones son esenciales para el desarrollo (Song *et al.*, 2022; Ke *et al.*, 2019). Se ha documentado una expresión débil de ϵ -LCY en semillas durante la maduración del fruto de tomate (Fu *et al.*, 2019; Rosas-Saavedra y Stange, 2016). Por tanto, la regulación negativa de ϵ -LCY parece propiciar la acumulación de apocarotenoides en tejidos no fotosintéticos. Finalmente, los genes LCY muestran una expresión diferencial durante las etapas de desarrollo, orientados a la síntesis y acumulación de carotenos, esenciales para procesos de fotosíntesis, fotoprotección y reproducción de la planta.

5.1.6. Citogenética de *Bixa orellana* L.

Como parte inicial de caracterizar la ubicación cromosómica de los genes Licopeno ciclasa para ahondar en su regulación y posible involucramiento con el morfotipo productivo de *Bixa orellana*, se implementó el análisis citogenético de esta planta. Los estudios citogenéticos se centran en determinar el número, forma y tamaño de los cromosomas (cariotipo) dentro del núcleo celular durante la división celular, momento donde los cromosomas se condensan y se vuelven visibles (Schwarzacher *et al.*, 2023; Mahmounding y Mirzaghaderi, 2023). Estos se han aplicado desde el siglo XIX, siendo la microscopía la mejor técnica para determinar sus características (Schwarzacher *et al.*, 2023). Para achiote, se modificó la metodología presentada por Martín (2007) para una población Yucateca en tiempos de colecta, pretratamiento y tinción. Esto mejoró la visualización de los cromosomas, para determinar el complemento cromosómico de la especie

(Maravilla *et al.*, 2023) y la región del centrómero, punto de referencia para la identificación morfológica (Levan *et al.*, 1964).

Las observaciones cromosómicas realizadas para el morfotipo B (flor rosa) de achiote, mostraron $2n= 14$ cromosomas, concuerdando con lo reportado para *B. orellana* (Lombello y Pinto-Maglio, 2014; Almeida *et al.*, 2006; Martín, 2007). Se observaron cromosomas pequeños de entre 2.2 a 0.90 μm para los pares 2 a 7, sobresaliendo el par 1 con un gran tamaño (3.1 μm). Estos datos coinciden con los reportes para las accesiones de Brasil de entre 3.2 a 1.2 μm (Lombello y Pinto-Maglio, *et al.*, 2014; Almeida *et al.*, 2006), donde solo el par 7 presenta tamaños más pequeños en este estudio. En cuanto a su forma, se identificó a los pares 1 a 6 como metacéntricos y al par 7 como submetacéntrico, según la clasificación de cromosomas basada en la ubicación del centrómero (Levan *et al.*, 1964). La fórmula cariotípica de la especie reportada en este estudio (6m+1sm) concuerda con lo reportado por Lombello y Pinto-Maglio, (*et al.*, 2014) pero difiere de lo presentado por Almeida (5m+2sm) (*et al.*, 2006).

En las especies, los caracteres cromosómicos suelen ser constantes; sin embargo, en las plantas puede existir variación en la morfología y/o número cromosómico dentro de una especie (Windham *et al.*, 2020; Li *et al.*, 2017; Soltis *et al.*, 2009; Brown y Bertke, 1979). En consecuencia, estos cambios pueden verse reflejados en el fenotipo de la planta (Windham *et al.*, 2020; Li *et al.*, 2017; Murray, 2012) que se presentan como variantes o cultivares con rasgos particulares (Poggio y Naranjo, 2004). Por lo que estas variaciones intraespecíficas en tamaño y forma de los cromosomas podrían estar asociadas a las diferencias morfológicas observadas en la especie (Valdez-Ojeda *et al.*, 2008; Rivera-Madrid *et al.*, 2006). Sin embargo, no se cuentan con análisis citogenéticos que vinculen la variación fenotípica y cariomorfológica de la especie para confirmar esta conjetura. Con esto, el presente trabajo, sería un aporte de relación entre el análisis cariotípico para una accesión caracterizada por una alta producción del pigmento bixina.

5.2. CONCLUSIONES

El objetivo de este trabajo fue caracterizar la estructura y la potencial regulación de los genes licopeno ciclasa (*Bo β -LCY1*, *Bo β -LCY2*, *Bo ϵ -LCY*) del genoma de *Bixa orellana* L, con el fin de aportar información sobre la relación de estos y las diferencias en contenido de bixina en la especie. Los resultados mostraron que las diferencias en contenido del pigmento bixina entre los morfotipos contrastantes de Achiote, estaría vinculado a diferencias en la estructura de las

secuencias codificantes y elementos regulatorios de los genes *LCY*, así como posibles variaciones fenotípicas en los cromosomas de la especie. Para esto, se realizó el análisis de secuencias de los genes *LCY* extraídas del genoma, el metaanálisis de su expresión en tres morfotipos de la especie, así como la caracterización a nivel cromosómico la organización estructural del genoma de la especie.

Para el primer objetivo, las secuencias codificantes de *BoLCY* reportadas en el transcriptoma de la especie, se confirmaron con este enfoque genómico. Para los genes *Boβ-LCY1* y *Boβ-LCY2* presentan un solo exón codificante, mientras que *Boε-LCY* muestra 11 exones, lo que sugiere una mayor complejidad estructural. Los genes exhibieron motivos característicos de los *LCY*, como el sitio de unión de dinucleótidos, una firma presente en todo caroteno ciclasa conocidos, y los motivos ciclasa I y II, que son dominios esenciales para la actividad catalítica. *Boβ-LCY1* y *Boβ-LCY2* exhibieron una gran similitud entre sus secuencias, pero con algunos cambios de aminoácidos en sus motivos conservados. En cuanto a su potencial regulación, los principales TF identificados se asociaron a funciones de fotosíntesis, reproducción y del sistema oxidativo, particularmente en *Boε-LCY*. También se analizaron los elementos repetitivos de ADN que flanquean las regiones con mayor abundancia de TFs; estos se relacionaron con LTR de la familia Ty3/Gypsy, principalmente en *Boε-LCY*, sugiriendo un papel regulador. De igual manera, se confirmó la presencia de 11 familias miRNAs no conservadas que podrían estar regulando procesos específicos y miRNAs conservados reportados en tomate y en las familias mejor conservadas en las plantas terrestres, cumpliendo funciones esenciales y conservadas en las plantas.

Respecto al segundo objetivo, el análisis de la expresión de tres morfotipos (A, flor blanca; B, flor rosa y C, flor purpura) en tres estados de desarrollo de la semilla (S1, 0-7d; S3 14-21d y S5, 28-42d), mostró al gen *Boε-LCY* fuertemente regulado a la baja durante el desarrollo de la semilla, mientras que los *Boβ-LCY* están regulados al alza en las etapas iniciales de desarrollo de la semilla (S1, S3) y en las hojas. En general, todos los *BoLCY* mostraron una expresión principalmente orientada a la síntesis de carotenoides esenciales en los procesos de fotosíntesis, mientras que los *Boβ-LCY* también parecen participar en las etapas iniciales de acumulación de carotenoides de las semillas inmaduras de *B. orellana*.

Referente al tercer objetivo, y como fase inicial de caracterizar la ubicación cromosómica de los

genes Licopeno ciclasa se implementó el análisis citogenético en el morfotipo productivo de *Bixa orellana*. A nivel cromosómico, el genoma de *Bixa orellana* se organizó en 14 cromosomas somáticos. Estos se caracterizaron por presentar tamaños pequeños (3.1 a 0.90 μm), donde el par uno resaltó al presentar un gran tamaño en consideración con los demás pares (2.2 a 0.90 μm). De igual manera, según los parámetros cariotípicos, se identificaron a los pares 1 a 6 como metacéntricos y al par 7 como submetacéntrico; reportando una fórmula cariotípica para el morfotipo B (flor rosa) de achiote de $6m+1sm$.

5.3. PERSPECTIVAS

El presente trabajo aborda a los genes licopeno ciclasa del genoma de *Bixa orellana* como centro de estudio. Los resultados aquí obtenidos, resaltan la importancia de seguir explorando las tecnologías de secuenciación masiva para el análisis de la organización del genoma de la especie (mapa físico). Esto nos proporcionaría una base para el mapeo de secuencias informativas, como las secuencias repetitivas (variabilidad en la especie) o genes de interés (LCY). Por otra parte, la implementación de técnicas citogenéticas moleculares, como la hibridación fluorescente *in situ* en los cromosomas de la especie; nos corroboraría los resultados obtenidos por el mapeo físico. Por último, el conjunto de análisis bioinformáticos y citogenéticos serían herramientas valiosas para caracterizar los genomas de los morfotipos y su relación con su producción de bixina, para la futura selección de genotipos y programas de mejoramiento genético.

BIBLIOGRAFÍA

- Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell*, *15*(1), 63–78. <https://doi.org/10.1105/tpc.006130>
- Acevedo Gasman, F., E. Huerta Ocampo, S. Lorenzo Alonso & Ortiz García S. (2009). La bioseguridad en México y los organismos genéticamente modificados: cómo enfrentar un nuevo desafío. *En: Dirzo, R., R. González e I.J. March. Capital Natural de México, vol. II: Estado de conservación y tendencias de cambio.* Conabio. México.
- Agner, A. R., Barbisan, L. F., Scolastici, C., & Salvadori, D. M. F. (2004). Absence of carcinogenic and anticarcinogenic effects of annatto in the rat liver medium-term assay. *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, *42*(10), 1687–1693. <https://doi.org/10.1016/j.fct.2004.06.005>
- Aguilar-Espinosa M, Carballo-Uicab VM, Rivera-Madrid R. (2020). Biology and chemistry of apocarotenoids. *Biology, chemistry, and applications of apocarotenoids (1st Edition)*. CRC Press, Boca Raton, pp 1–9
- Aiassa D., Bosh B., Gentile N., Mañas F. y Gorla N. (2015). Citogenética: Teoría y Práctica, Manual. *Colección Manuales*. Córdoba. CEPYD. E-Book.
- Akshatha, V., Giridhar, P., & Ravishankar, G. A. (2011). Morphological diversity in *Bixa orellana* L. and variations in annatto pigment yield. *The Journal of Horticultural Science & Biotechnology*, *86*(4), 319–324. <https://doi.org/10.1080/14620316.2011.11512767>
- Alcaíno, J., Baeza, M., & Cifuentes, V. (2016). Carotenoid Distribution in Nature. In *Subcellular Biochemistry* (pp. 3–33). Springer International Publishing.
- Aliphath, M. M. (2009). Huertos y cacaotales Mayas: Un análisis agroecosistémico. *En XXII Simposio de Investigaciones Arqueológicas en Guatemala, 2008* (editado por J.P. Laporte, B. Arroyo y H. Mejía), pp.267-275. Museo Nacional de Arqueología y Etnología, Guatemala (versión digital).

-
- Almeida P. M., Carvalho C. R. y Clarindo W. R. (2006). Classical and Molecular Cytogenetic Tools to Resolve the Bixa Karyotypes. *Cytologia*, 71(4): 391-398.
- Alquezar B, Rodrigo MJ, Zacarías L. (2008) Regulation of carotenoid biosynthesis during fruit maturation in the red-fleshed orange mutant cara cara. *Phytochemistry* 69(10):1997–2007. <https://doi.org/10.1016/j.phytochem.2008.04.020>
- Alquézar B, Zacarías L, Rodrigo MJ. (2009) Molecular and functional characterization of a novel chromoplast-specific lycopene beta-cyclase from Citrus and its relation to lycopene accumulation. *J Exp Bot* 60(6):1783–1797. <https://doi.org/10.1093/jxb/erp048>
- Ampomah-Dwamena C, McGhie T, Wibisono R, Montefiori M, Hel- lens RP, Allan AC. (2009). The kiwifruit lycopene beta-cyclase plays a significant role in carotenoid accumulation in fruit. *J Exp Bot* 60(13):3765–3779. <https://doi.org/10.1093/jxb/erp218>
- Arazi, T., & Khedia, J. (2022). Tomato MicroRNAs and their functions. *International Journal of Molecular Sciences*, 23(19), 11979. <https://doi.org/10.3390/ijms231911979>
- Arce J. (1999). El achiote *Bixa orellana* L. cultivo promisorio para el trópico. *EARTH Press*, Costa Rica.
- Avendaño Arrazate C.H., Pinzón-López L.L., Mendoza-López A., Campos Rojas E., Correa Navarro P.J., Godoy Hernández G., Mijangos Cortés J.O. y Rivera Madrid R. (2018). Rescate y conservación del achiote (*Bixa Orellana* L.) en México. *Agro productividad*.
- Axtell, M. J., & Meyers, B. C. (2018). Revisiting criteria for plant MicroRNA annotation in the era of big data. *The Plant Cell*, 30(2), 272–284. <https://doi.org/10.1105/tpc.17.00851>
- Bailey TL, Johnson J, Grant CE, Noble WS (2015). The MEME suite. *Nucleic Acids Res* 43(W1) W39–W49.
- Bang H, Kim S, Leskovar D, King S (2007). Development of a codominant CAPS marker for allelic selection between canary yellow and red watermelon based on SNP in lycopene β -cyclase (LCYB) gene. *Mol Breed New Strateg Plant Improv* 20(1):63– 72. <https://doi.org/10.1007/s11032-006-9076-4>

-
- Bate N, Twell D (1998). Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant Mol Biol* 37(5):859–869. <https://doi.org/10.1023/a:1006095023050>
- Barrett, L. W., Fletcher, S., & Wilton, S. D. (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences: CMLS*, 69(21), 3613–3634. <https://doi.org/10.1007/s00018-012-0990-9>
- Bear, D. (1976). Systematic of the genus *Bixa* and geography of the cultivated annatto tree. PhD Thesis. University of California, Los Angeles. 525 p.
- Benson B. V., Nepall M. P., Davison L. G. Davison, Duwadi and P. B., & Moore B. G. (2014). In silico analysis of cis-regulatory elements of disease resistance genes across six plant species. *Proceedings of the South Dakota Academy of Science*, 93: 133.
- Bertolini, E., Verelst, W., Horner, D. S., Gianfranceschi, L., Piccolo, V., Inzé, D., Pè, M. E., & Mica, E. (2013). Addressing the Role of microRNAs in Reprogramming Leaf Growth during Drought Stress in *Brachypodium distachyon*. *Molecular Plant*, 6(2), 423–443. <https://doi.org/10.1093/mp/sss160>
- Bian W, Barsan C, Egea I, Purgatto E, Chervin C, Zouine M, Latché A, Bouzayen M, Pech J-C (2011). Metabolic and molecular events occurring during chromoplast biogenesis. *J Bot* 2011:1– 13. <https://doi.org/10.1155/2011/289859>
- Biscotti, M. A., Olmo, E., & Heslop-Harrison, J. S. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 23(3), 415–420. <https://doi.org/10.1007/s10577-015-9499-z>
- Bittencourt C, Felicissimo MP, Pireaux JJ, Houssiau L (2005). Study of annatto from *Bixa orellana* seeds: an application of time-of-flight secondary ion mass spectrometry. *Spectrosc Eur* 17:16–22
- Boutte J, Maillet L, Chaussepied T, Letort S, Aury J-M, Belser C, Boideau F, Brunet A, Coriton O, Deniot G, Falentin C, Huteau V, Lodé-Taburel M, Morice J, Trotoux G, Chèvre A-M, Rousseau-Gueutin M, de Ferreira J (2020) Genome size variation and comparative genomics

-
- reveal intraspecific diversity in *Brassica rapa*. *Front Plant Sci* 11:577536. <https://doi.org/10.3389/fpls.2020.577536>
- Bouvier, F., Dogbo, O., & Camara, B. (2003). Biosynthesis of the food and cosmetic plant pigment bixin (annatto). *Science (New York, N.Y.)*, 300(5628), 2089–2091. <https://doi.org/10.1126/science.1085162>
- Brown W. V. & Bertke E. M. (1979). *Citología*. Ediciones Omega, España. pp. 334, 360-370.
- Bulgakov, V. P., & Avramenko, T. V. (2015). New opportunities for the regulation of secondary metabolism in plants: focus on microRNAs. *Biotechnology Letters*, 37(9), 1719–1727. <https://doi.org/10.1007/s10529-015-1863-8>
- Carballo-Uicab VM, Cárdenas-Conejo Y, Vallejo-Cardona AA, Aguilar-Espinosa M, Rodríguez-Campos J, Serrano-Posada H, Narváez-Zapata JA, Vázquez-Flota F, Rivera-Madrid R (2019). Isolation and functional characterization of two dioxygenases putatively involved in bixin biosynthesis in annatto (*Bixa orellana* L). *PeerJ* 7(e7064) e7064. <https://doi.org/10.7717/peerj.7064>
- Carballo-Uicab V. M (2019). Análisis y caracterización de genes potencialmente involucrados en la biosíntesis de bixina en *Bixa orellana* L. Tesis de Doctorado. Centro de Investigación Científica de Yucatán, A C. México. pp. 111
- Cárdenas-Conejo Y, Carballo-Uicab V, Lieberman M, Aguilar-Espinosa M, Comai L, Rivera-Madrid R (2015). *De novo* transcriptome sequencing in *Bixa orellana* to identify genes involved in methylerythritol phosphate, carotenoid and bixin biosynthesis. *BMC Genomics* 16(1):877. <https://doi.org/10.1186/s12864-015-2065-4>
- Cárdenas-Conejo Y, Narváez-Zapata JA, Carballo-Uicab VM, Aguilar-Espinosa M, Us-Camas R, Escobar-Turriza P, Comai L, Rivera-Madrid R (2023). Gene expression profile during seed development of *Bixa orellana* accessions varying in bixin pigment. *Front Plant Sci* 14:1066509. <https://doi.org/10.3389/fpls.2023.1066509>
- Carta, A., Bedini, G., & Peruzzi, L. (2020). A deep dive into the ancestral chromosome number and genome size of flowering plants. *The New Phytologist*, 228(3), 1097–1106. <https://doi.org/10.1111/nph.16668>

-
- Carvalho, G. M. G., Guerra, M., Santos, G. A. dos, Andrade, V. C. de, & Farias, M. C. A. de. (1991). Citogenética de angiospermas coletadas em Pernambuco: IV. *Acta botanica Brasílica*, 5(2), 37–51. <https://doi.org/10.1590/s0102-33061991000200003>
- Carvalho, J. F. R. P. de, Robinson, I. P., & Alfenas, A. C. (2005). Isozymic variability in a Brazilian collection of annatto (*Bixa orellana* L.). *Pesquisa Agropecuária Brasileira*, 40(7), 653–660. <https://doi.org/10.1590/s0100-204x2005000700005>
- Chaney, L., Sharp, A. R., Evans, C. R., & Udall, J. A. (2016). Genome mapping in plant comparative genomics. *Trends in Plant Science*, 21(9), 770–780. <https://doi.org/10.1016/j.tplants.2016.05.004>
- Chan-Rodríguez D. & R. Rivera-Madrid (2020). Unravelling Apocarotenoid Biosynthesis Pathways: Transcriptomic Tools for Gene Discovery. *Biology, chemistry, and applications of apocarotenoids (1st Edition)*. CRC Press, Boca Raton, pp 59-72
- Chisté, R. C., Yamashita, F., Gozzo, F. C., & Mercadante, A. Z. (2011). Simultaneous extraction and analysis by high performance liquid chromatography coupled to diode array and mass spectrometric detectors of bixin and phenolic compounds from annatto seeds. *Journal of Chromatography A*, 1218(1), 57–63. <https://doi.org/10.1016/j.chroma.2010.10.094>
- Chung, M.-Y., Nath, U. K., Vrebalov, J., Gapper, N., Lee, J. M., Lee, D.-J., Kim, C. K., & Giovannoni, J. (2020). Ectopic expression of miRNA172 in tomato (*Solanum lycopersicum*) reveals novel function in fruit development through regulation of an AP2 transcription factor. *BMC Plant Biology*, 20(1). <https://doi.org/10.1186/s12870-020-02489-y>
- Cui, J., You, C., & Chen, X. (2017). The evolution of microRNAs in plants. *Current Opinion in Plant Biology*, 35, 61–67. <https://doi.org/10.1016/j.pbi.2016.11.006>
- Clement, C., De Cristo-Araújo, M., Coppens D'Eeckenbrugge, G., Alves Pereira, A., & Picanço-Rodrigues, D. (2010). Origin and domestication of native amazonian crops. *Diversity*, 2(1), 72–106. <https://doi.org/10.3390/d2010072>

-
- Cunningham FX, Gantt E (1998). Genes and enzymes of carotenoid biosynthesis in plants. *Annu Rev Plant Physiol Plant Mol Biol*, 49(1):557–583. <https://doi.org/10.1146/annurev.arplant.49.1.557>
- Cunningham Jr FX, Pogson B, Sun Z, McDonald KA, DellaPenna D, Gantt E (1996). Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell*, 8(9):1613–1626. <https://doi.org/10.1105/tpc.8.9.1613>
- da Silva, R. A., Souza, G., Lemos, L. S. L., Lopes, U. V., Patrocínio, N. G. R. B., Alves, R. M., Marcellino, L. H., Clement, D., Micheli, F., & Gramacho, K. P. (2017). Genome size, cytogenetic data and transferability of EST-SSRs markers in wild and cultivated species of the genus *Theobroma* L. (Byttnerioideae, Malvaceae). *PloS One*, 12(2), e0170799. <https://doi.org/10.1371/journal.pone.0170799>
- Dalal M, Chinnusamy V, Bansal KC (2010) Isolation and functional characterization of lycopene beta-cyclase (CYC-B) promoter from *Solanum habrochaites*. *BMC Plant Biol* 10(1):61. <https://doi.org/10.1186/1471-2229-10-61>
- Dalmay, T. (2010). Short RNAs in tomato. *Journal of Integrative Plant Biology*, 52(4), 388–392. <https://doi.org/10.1111/j.1744-7909.2009.00871.x>
- Deb A, Kundu S (2015) Deciphering Cis-regulatory element mediated combinatorial regulation in rice under blast infected condition. *PLoS ONE* 10(9) e0137295. <https://doi.org/10.1371/journal.pone.0137295>
- Dequigiovanni, G., Ramos, S. L. F., Alves-Pereira, A., Fabri, E. G., Picanço-Rodrigues, D., Clement, C. R., Gepts, P., & Veasey, E. A. (2018). Highly structured genetic diversity of *Bixa orellana* var. *urucurana*, the wild ancestor of annatto, in Brazilian Amazonia. *PloS One*, 13(6), e0198593. <https://doi.org/10.1371/journal.pone.0198593>
- Devitt LC, Fanning K, Dietzgen RG, Holton TA (2010) Isolation and functional characterization of a lycopene beta-cyclase gene that controls fruit colour of papaya (*Carica papaya* L.). *J Exp Bot* 61(1):33–39. <https://doi.org/10.1093/jxb/erp284>

-
- Du H, Yang J, Chen B, Zhang X, Zhang J, Yang K, Wen C (2019) Target sequencing reveals genetic diversity, population structure, core-SNP markers, and fruit shape-associated loci in pepper varieties. *BMC Plant Biol* 19(1):1–16. <https://doi.org/10.1186/s12870-019-2122-2>
- Dzib-Cauich J, Us-Camas R, Rivera-Madrid R (2020) Natural sources of apocarotenoids and their applications. *Biology, chemistry, and applications of apocarotenoids*, 1st edn. CRC Press, Florida, pp 11–39
- Escobar Turriza P. J. (2017). Identificación de marcadores moleculares para el mejoramiento genético del Achiote (*Bixa orellana* L.). Tesis de Maestría. Centro de Investigación Científica de Yucatán, A C. México. pp. 80
- Ferguson-Smith, M. A. (2015). History and evolution of cytogenetics. *Molecular Cytogenetics*, 8(1). <https://doi.org/10.1186/s13039-015-0125-8>
- Feuillet, C., Leach, J. E., Rogers, J., Schnable, P. S., & Eversole, K. (2011). Crop genome sequencing: lessons and rationales. *Trends in Plant Science*, 16(2), 77–88. <https://doi.org/10.1016/j.tplants.2010.10.005>
- Filichkin SA, Leonard JM, Monteros A, Liu P-P, Nonogaki H (2004) A novel endo-beta-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. *Plant Physiol* 134(3):1080–1087. <https://doi.org/10.1104/pp.103.035998>
- Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. *Prog Lipid Res* 43(3):228–265. <https://doi.org/10.1016/j.plipres.2003.10.002>
- Fu X, Cheng S, Feng C, Kang M, Huang B, Jiang Y, Duan X, Grierson D, Yang Z (2019) Lycopene cyclases determine high α -/ β -carotene ratio and increased carotenoids in bananas ripening at high temperatures. *Food Chem* 283:131–140. <https://doi.org/10.1016/j.foodchem.2018.12.121>
- Gao, C., Ju, Z., Cao, D., Zhai, B., Qin, G., Zhu, H., Fu, D., Luo, Y., & Zhu, B. (2015). MicroRNA profiling analysis throughout tomato fruit development and ripening reveals potential regulatory role of RIN on microRNAs accumulation. *Plant Biotechnology Journal*, 13(3), 370–382. <https://doi.org/10.1111/pbi.12297>

-
- García V. A. (1985). Sistemas Robertsonianos: su papel en evolución cromosómica en plantas superiores. Memorias del seminario sobre la investigación genética básica en el conocimiento y evaluación de los recursos genéticos. Sociedad Mexicana de Fitogenética. pp. 41.
- Giuliano, G., Rosati, C., & Bramley, P. M. (2003). To dye or not to dye: biochemistry of annatto unveiled. *Trends in Biotechnology*, 21(12), 513–516. <https://doi.org/10.1016/j.tibtech.2003.10.001>
- Gomes Pacheco, T., de Santana Lopes, A., Monteiro Viana, G. D., Nascimento da Silva, O., Morais da Silva, G., do Nascimento Vieira, L., Guerra, M. P., Nodari, R. O., Maltempi de Souza, E., de Oliveira Pedrosa, F., Otoni, W. C., & Rogalski, M. (2019). Genetic, evolutionary and phylogenetic aspects of the plastome of annatto (*Bixa orellana* L.), the Amazonian commercial species of natural dyes. *Planta*, 249(2), 563–582. <https://doi.org/10.1007/s00425-018-3023-6>
- Gómez-Ortíz, N. M., Vázquez-Maldonado, I. A., Pérez-Espadas, A. R., Mena-Rejón, G. J., Azamar-Barrios, J. A., & Oskam, G. (2010). Dye-sensitized solar cells with natural dyes extracted from achiote seeds. *Solar Energy Materials and Solar Cells: An International Journal Devoted to Photovoltaic, Photothermal, and Photochemical Solar Energy Conversion*, 94(1), 40–44. <https://doi.org/10.1016/j.solmat.2009.05.013>
- Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P (2004) Cis-regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene[W]. *Plant Cell* 16(5):1077–1090. <https://doi.org/10.1105/tpc.019729>
- Guerra, M. (2008). Chromosome numbers in plant cytotaxonomy: concepts and implications. *Cytogenetic and Genome Research*, 120(3–4), 339–350. <https://doi.org/10.1159/000121083>
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp Ser* 41(41):95–98

-
- Hannoufa, A., & Hossain, Z. (2012). Regulation of carotenoid accumulation in plants. *Biocatalysis and Agricultural Biotechnology*, 1(3), 198–202.
<https://doi.org/10.1016/j.bcab.2012.03.004>
- Hanson, L. (2001). First nuclear DNA C-values for 25 angiosperm families. *Annals of Botany*, 87(2), 251–258. <https://doi.org/10.1006/anbo.2000.1325>
- Hao H, Ma L, Cong H, Li Q, Yu X (2012) Isolation and characterization of a muskmelon cDNA encoding lycopene beta-cyclase. *Gene* 503(1):147–151.
<https://doi.org/10.1016/j.gene.2012.04.059>
- Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319(5861):330–333.
<https://doi.org/10.1126/science.1150255>
- Herrera, J. C. (2007). La citogenética molecular y su aplicación en el estudio de los genomas vegetales. *Agronomía Colombiana*, 25(1), 26–35.
<https://www.redalyc.org/articulo.oa?id=180316240004>
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27(1):297–300.
<https://doi.org/10.1093/nar/27.1.297>
- Hong M, Chi Z-H, Wang Y-Q, Tang Y-M, Deng Q-X, He M-Y, Wang R-K, He Y-Z (2019) Expression of a chromoplast-specific lycopene β -cyclase gene (CYC-B) is implicated in carotenoid accumulation and coloration in the loquat. *Biomolecules* 9(12):874.
<https://doi.org/10.3390/biom9120874>
- Imelfort, M., Batley, J., Grimmond, S., & Edwards, D. (2009). Genome sequencing approaches and successes. In *Plant Genomics* (pp. 345–358). Humana Press.
- Index to plant chromosome numbers. (1979-2011). P. Goldblatt & D. E. Johnson, eds. Missouri Botanical Garden, St. Louis. <http://legacy.tropicos.org/project/ipcn>

-
- Jako, C., Coutu, C., Roewer, I., Reed, D. W., Pelcher, L. E., & Covello, P. S. (2002). Probing carotenoid biosynthesis in developing seed coats of *Bixa orellana* (Bixaceae) through expressed sequence tag analysis. *Plant Science: An International Journal of Experimental Plant Biology*, 163(1), 141–145. [https://doi.org/10.1016/s0168-9452\(02\)00083-3](https://doi.org/10.1016/s0168-9452(02)00083-3)
- Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Mad-den TL (2008) NCBI BLAST: a better web interface. *Nucleic Acids Res* 36(suppl2) W5–W9
- Ke Q, Kang L, Kim HS, Xie T, Liu C, Ji CY, Kim SH, Park WS, AhnM-J, Wang S, Li H, Deng X, Kwak S-S (2019) Down-regulation of lycopene ϵ -cyclase expression in transgenic sweet potato plants increases the carotenoid content and tolerance to abiotic stress. *Plant Sci Int J Exp Plant Biol* 281:52–60. <https://doi.org/10.1016/j.plantsci.2019.01.002>
- Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics*, 39(1), 309–338. <https://doi.org/10.1146/annurev.genet.39.073003.114725>
- Kössler S, Armarego-Marriott T, Tarkowská D, Turečková V, Agrawal S, Mi J, de Souza LP, Schöttler MA, Schadach A, Fröhlich A, Bock R, Al-Babili S, Ruf S, Sampathkumar A, Moreno JC (2021). Lycopene β -cyclase expression influences plant physiology, development, and metabolism in Tobacco plants. *J Exp Bot* 72(7):2544–2569. <https://doi.org/10.1093/jxb/erab029>
- Koul, A., Sharma, D., Kaul, S., & Dhar, M. K. (2019). Identification and in silico characterization of cis-acting elements of genes involved in carotenoid biosynthesis in tomato. *3 Biotech*, 9(7). <https://doi.org/10.1007/s13205-019-1798-1>
- Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Research*, 47(D1), D155–D162. <https://doi.org/10.1093/nar/gky1141>
- Krishnan, N. & K. R. Ayyangar. (1987). Taxonomical relationships of *Bixa* L. *Proc. Indian Sci. Congr. Assoc.* 74 (3, VI): 105.
- Krishnan, N. (1977). Cytotaxonomical studies on Bixaceae and Samydaceae from south India with collected evidences from palynology, anatomy and biochemistry. Ph.D. Thesis, Annamalai University. 156 pp.

-
- Leal F. y C. Michelangeli de Clavijo. (2010). Acerca de la historia, taxonomía, botánica y usos de *Bixa Orellana* L. *Revista Unellez Ciencia y Tecnología*. 28: 78-86.
- Lee S-I, Kim N-S (2014) Transposable elements and genome size variations in plants. *Genomics Inf* 12(3):87–97. <https://doi.org/10.5808/GI.2014.12.3.87>
- Levan, A., Fredga, K., & Sandberg, A. A. (2009). Nomenclature for centromeric position on chromosomes. *Hereditas*, 52(2), 201–220. <https://doi.org/10.1111/j.1601-5223.1964.tb01953.x>
- Li J., Yuan J., Li M. (2014). Characterization of putative cis-regulatory elements in genes preferentially expressed in *Arabidopsis* male meiocytes. *BioMed Res Int*. <https://doi.org/10.1155/2014/708364>
- Li, S.-F., Su, T., Cheng, G.-Q., Wang, B.-X., Li, X., Deng, C.-L., & Gao, W.-J. (2017). Chromosome evolution in connection with repetitive sequences and epigenetics in plants. *Genes*, 8(10), 290. <https://doi.org/10.3390/genes8100290>
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods (San Diego Calif)* 25(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
- Llorens C, Futami R, Covelli L, Domínguez-Escribá L, Viu JM, Tamarit D, Aguilar-Rodríguez J, Vicente-Ripolles M, Fuster G, Bernet GP, Maumus F, Munoz-Pomer A, Sempere JM, Latorre A, Moya A (2011). The Gypsy database (GyDB) of mobile genetic elements: release 2.0. *Nucleic Acids Res* 39(Database issue). D70-74. <https://doi.org/10.1093/nar/gkq1061>
- Llorente B, Martinez-Garcia JF, Stange C, Rodriguez-Concepcion M (2017). Illuminating colors: regulation of carotenoid biosynthesis and accumulation by light. *Curr Opin Plant Biol* 37:49–55. <https://doi.org/10.1016/j.pbi.2017.03.011>
- Lohani N, Babaei S, Singh MB, Bhalla PL (2021) Genome-wide in silico identification and comparative analysis of Dof gene family in *Brassica napus*. *Plants* 10(4):709. <https://doi.org/10.3390/plants10040709>

-
- Lombello, R. A., & Pinto-maglio, C. A. F. (2014). Cytogenetics and Reproductive Biology of *Bixa orellana* L. (Bixaceae). *Cytologia*, 79(3), 379–386. <https://doi.org/10.1508/cytologia.79.379>
- López-Flores, I., & Garrido-Ramos, M. A. (2012). The repetitive DNA content of eukaryotic genomes. In *Genome Dynamics* (pp. 1–28). S. Karger AG.
- Lourido Pérez H. C. y Martínez Sánchez G. (2010). La *Bixa orellana* L. en el tratamiento de afecciones estomatológicas, un tema aún por estudiar. *Revista Cubana de Farmacia*. 44(2): 231-244.
- Lu, H., Giordano, F., & Ning, Z. (2016). Oxford secuenciación de molécula única MinION sequencing and genome assembly. *Genomics, Proteomics & Bioinformatics*, 14(5), 265–279. <https://doi.org/10.1016/j.gpb.2016.05.004>
- Lu S, Li L (2008) Carotenoid metabolism: biosynthesis, regulation, and beyond. *J Integr Plant Biol* 50(7):778–785. <https://doi.org/10.1111/j.1744-7909.2008.00708.x>
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Marchler-Bauer A (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 48(D1). D265–D268
- Lukasik, A., Pietrykowska, H., Paczek, L., Szweykowska-Kulinska, Z., & Zielenkiewicz, P. (2013). High-throughput sequencing identification of novel and conserved miRNAs in the Brassica oleracea leaves. *BMC Genomics*, 14(1). <https://doi.org/10.1186/1471-2164-14-801>
- Mahmoudi, S., & Mirzaghaderi, G. (2023a). Tools for Drawing Informative Idiograms. In *Methods in Molecular Biology* (pp. 515–527). Springer US.
- Maravilla, A. J., Rosato, M., & Rosselló, J. A. (2023). Preparation of mitotic chromosomes with the squash technique. In *Methods in Molecular Biology* (pp. 141–149). Springer US.
- Maoka T (2020) Carotenoids as natural functional pigments. *J Nat Med* 74(1):1–16. <https://doi.org/10.1007/s11418-019-01364-x>
- Marcela Delgado, L., Uribe Lastra, M., & Marulanda Angel, M. L. (2010). Estandarización de la técnica citogenética “squash” para conteo de cromosomas mitóticos en *Rubus glaucus*

- Martín-Rodríguez M. E. (2010). Uso de Genes Involucrados en la Síntesis de Carotenoides como Marcadores moleculares de Pigmentos en Achiote. Tesis de Maestría. Centro de Investigación Científica de Yucatán, A. C. pp. 80
- Mejía-Guerra, M. K., Pomeranz, M., Morohashi, K., & Grotewold, E. (2012a). From plant gene regulatory grids to network dynamics. *Biochimica et Biophysica Acta. Gene Regulatory Mechanisms*, 1819(5), 454–465. <https://doi.org/10.1016/j.bbagrm.2012.02.016>
- Meng, Y., Su, W., Ma, Y., Liu, L., Gu, X., Wu, D., Shu, X., Lai, Q., Tang, Y., Wu, L., & Wang, Y. (2021c). Assessment of genetic diversity and variety identification based on developed retrotransposon-based insertion polymorphism (RBIP) markers in sweet potato (*Ipomoea batatas* (L.) Lam.). *Scientific Reports*, 11(1), 1–12. <https://doi.org/10.1038/s41598-021-95876-w>
- Mercadante AZ, Steck A, Pfander H (1997) Isolation and identification of new apocarotenoids from annatto (*Bixa orellana*) seeds. *J Agric Food Chem* 45(4):1050–1054. <https://doi.org/10.1021/jf960412k>
- Mercadante AZ, Steck A, Rodriguez-Amaya D, Pfander H, Britton G (1996) Isolation of methyl 9'Z-apo-6'-lycopenoate from *Bixa orellana*. *Phytochemistry* 41(4):1201–1203. [https://doi.org/10.1016/0031-9422\(95\)00784-9](https://doi.org/10.1016/0031-9422(95)00784-9)
- Michelangeli C., Maureen Medina A., Artioli P. y Mata J. (2002). Microsporogénesis y microgametogénesis de onoto (*Bixa Orellana* L.). *Acta Científica Venezolana*, 53:171-175.
- Michelangeli de Clavijo C. C., Medina M. A. M., Artioli G. P.I. y Mata J. (2006). Estudios citogenéticos de estructuras embriogénicas y plántulas regeneradas de onoto. *Agronomía Tropical*, 56(1): 129-137.
- Millo Sánchez, R., Galpert Cañizares, D., Casa Cardoso, G., Grau Ábalo, R., Arco García, L., García Lorenzo, M. M., & Fernández Marin, M. Á. (2014). Agregación de medidas de similitud para la detección de ortólogos: validación con medidas basadas en la teoría de

conjuntos aproximados. *Computación y Sistemas*, 18(1), 19–35.
<https://www.redalyc.org/articulo.oa?id=61530484003>

- Montoya, C., Ávila, K., Reyes, P., Navia, M., & Romero, H. M. (2014). Secuenciación del genoma de las especies vegetales: implicaciones y perspectivas. *Palmas*, 35(3), 11–22.
- Moreira, P. A., Lins, J., Dequigiovanni, G., Veasey, E. A., & Clement, C. R. (2015). The Domestication of Annatto (*Bixa orellana*) from *Bixa urucurana* in Amazonia. *Economic Botany*, 69(2), 127–135. <https://doi.org/10.1007/s12231-015-9304-0>
- Moreno JC, Pizarro L, Fuentes P, Handford M, Cifuentes V, Stange C (2013) Levels of lycopene β -cyclase 1 modulate carotenoid gene expression and accumulation in *Daucus carota*. PLoS ONE 8(3)e58144. <https://doi.org/10.1371/journal.pone.0058144>
- Mukherjee, P. (1975). Cytotaxonomical studies on *Bixa* and *Flacourtia*. *Bull. Bot. Soc. Bengal*, 29: 25–27.
- Murray, B. G. (2013). Karyotype Variation and Evolution in Gymnosperms. In *Plant Genome Diversity Volume 2* (pp. 231–243). Springer Vienna.
- Muthamilarasan M, Khandelwal R, Yadav CB, Bonthala VS, Khan Y, Prasad M (2014) Identification and molecular characterization of MYB transcription factor superfamily in C4 model plant foxtail millet (*Setaria italica* L.). PLoS ONE 9(10)e109920
- Nisar N, Li L, Lu S, Khin NC, Pogson BJ (2015) Carotenoid metabolism in plants. *Mol Plant* 8(1):68–82. <https://doi.org/10.1016/j.molp.2014.12.007>
- Ohri, D., Bhargava, A., & Chatterjee, A. (2004). Nuclear DNA amounts in 112 species of tropical hardwoods - new estimates. *Plant Biology (Stuttgart, Germany)*, 6(5), 555–561. <https://doi.org/10.1055/s-2004-821235>
- Okonechnikov, K., Golosova, O., Fursov, M., & the UGENE team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics (Oxford, England)*, 28(8), 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>

-
- Owusu Adjei, M., Zhou, X., Mao, M., Rafique, F., & Ma, J. (2021). MicroRNAs roles in plants secondary metabolism. *Plant Signaling & Behavior*, 16(7), 1915590. <https://doi.org/10.1080/15592324.2021.1915590>
- Palomino H. G. (1985). Los estudios citogenéticos como apoyo al conocimiento de los recursos genéticos. Memorias del seminario sobre la investigación genética básica en el conocimiento y evaluación de los recursos genéticos. Sociedad Mexicana de Fitogenética. pp. 82.
- Pech-Hoil R., Ferrer M. M., Aguilar-Espinosa M., Valdez-Ojeda R., Garza-Caligaris L. E. & Rivera-Madrid R (2017). Variation in the mating system of *Bixa orellana* L. (achiote) under three different agronomic systems. *Scientia Horticulturae*, 223: 31–37.
- Perez S., Cuen M., Becerra R. 2003. El achiote. *Biodiversitas*, 46: 7-11.
- Pineda, J. E., & Calderón, L. (2003). Planta piloto para obtener colorante de la semilla del achiote (*Bixa orellana*). *Revista universidad EAFIT*, 39(131), 8–22. <https://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/899>
- Pinzón López L. L., Avendaño Arrazate C. H., Mijangos Cortés J. O., Gody Hernández G.C., Cituk Chan D. E., Correa Navarro P. J. y Campos Rojas E. (2018). Situación del Recurso Achiote en México (*Bixa Orellana* L.). Red de achiote. SINAREFI.
- Poggio. L. & Naranjo C. A (2004). Citogenética, en V. Echenique, C. Rubinstein & L. Mroginski (eds.). *Biotechnología y Mejoramiento Vegetal*. INTA. Buenos Aires. pp. 69-79.
- Raddatz-Mota D, Pérez-Flores LJ, Carrari F, Mendoza-Espinoza JA, de León-Sánchez FD, Pinzón-López LL, Godoy-Hernández G, Rivera-Cabrera F (2017) Achiote (*Bixa orellana* L.): a natural source of pigment and vitamin E. *J Food Sci Technol* 54(6):1729–1741. <https://doi.org/10.1007/s13197-017-2579-7>
- Ramamoorthy S, Madrid RR, Doss CGP (2020) Biology, chemistry, and applications of apocarotenoids. In: Ramamoorthy S, Rivera-Madrid R, Doss CGP (Eds.), 1st Edition. CRC Press

-
- Reyes JC, Muro-Pastor MI, Florencio FJ (2004) The GATA family of transcription factors in arabidopsis and rice. *Plant Physiol* 134(4):1718–1732. <https://doi.org/10.1104/pp.103.037788>
- Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics*, 13(5), 278–289. <https://doi.org/10.1016/j.gpb.2015.08.002>
- Rimieri P (2017). La diversidad genética y la variabilidad genética: dos conceptos diferentes asociados al germoplasma y al mejoramiento genético vegetal. *Journal of Basic and Applied Genetics*, 28(2): 7-13
- Rivera-Madrid R, Escobedo-GM RM, Balam-Galera E, Vera-Ku M, Harries H (2006) Preliminary studies toward genetic improvement of annatto (*Bixa orellana* L.). *Sci Hort* 109(2):165–172. <https://doi.org/10.1016/j.scienta.2006.03.011>
- Rivera-Madrid, R., Aguilar-Espinosa, M., Cárdenas-Conejo, Y., & Garza-Caligaris, L. E. (2016). Carotenoid Derivates in Achioté (*Bixa orellana*) Seeds: Synthesis and Health Promoting Properties. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01406>
- Rivera-Madrid R & Ramamoorthy S. (2020). Apocarotenoid Molecules Are Continuously Being Discovered in All Taxa. *Biology, chemistry, and applications of apocarotenoids (1st Edition)*. CRC Press, Boca Raton, pp 41-58
- Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140
- Rodríguez-Ávila NL, Narváez-Zapata JA, Aguilar-Espinosa ML, Rivera-Madrid R (2009) Full-length gene enrichment by using an optimized RNA isolation protocol in *Bixa orellana* recalcitrant tissues. *Mol Biotechnol* 42(1):84–90. <https://doi.org/10.1007/s12033-008-9138-4>
- Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D (2001) Functional analysis of cis-regulatory elements within the promoter of the Tobacco late pollen gene g10. *Plant Mol Biol* 45(5):577–585. <https://doi.org/10.1023/a:1010695226241>

-
- Romero-López, N., Luna-Martínez, F., Gallegos-Brito, C., Escobar-Turriza, P., Aguilar-Espinosa, M., Valdez-Ojeda, R., Simpson, J., & Rivera-Madrid, R. (2019). An integrated genetic linkage map of *Bixa orellana* L. *Tree Genetics & Genomes*, 15(4). <https://doi.org/10.1007/s11295-019-1371-4>
- Rosas-Saavedra C, Stange C (2016) Biosynthesis of carotenoids in plants: enzymes and color. *Subcell Biochem* 79:35–69. https://doi.org/10.1007/978-3-319-39126-7_2
- RStudio Team. (2015). RStudio: Integrated Development Environment for R. Boston, MA. <http://www.rstudio.com/>
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. doi:10.1038/nmeth.2019
- Schwarzacher, T., Liu, Q., & Heslop-Harrison, J. S. (2023). Plant Cytogenetics: From Chromosomes to Cytogenomics. In *Methods in Molecular Biology* (pp. 3–21). Springer US.
- Sharma N, Russell SD, Bhalla PL, Singh MB (2011) Putative cis- regulatory elements in genes highly expressed in rice sperm cells. *BMC Res Notes* 4(1):319. <https://doi.org/10.1186/1756-0500-4-319>
- Silva, J. C., Carvalho, C. R., & Clarindo, W. R. (2018). Updating the maize karyotype by chromosome DNA sizing. *PloS One*, 13(1), e0190428. <https://doi.org/10.1371/journal.pone.0190428>
- SIRE. 2003. *Bixa orellana* L. Paquetes tecnológicos. Sistema de Información para la reforestación. Comisión Nacional Forestal.
- Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., Sankoff, D., de Pamphilis, C. W., Wall, P. K., & Soltis, P. S. (2009). Polyploidy and angiosperm diversification. *American Journal of Botany*, 96(1), 336–348. <https://doi.org/10.3732/ajb.0800079>

-
- Song, X., Li, Y., Cao, X., & Qi, Y. (2019). MicroRNAs and their regulatory roles in plant–environment interactions. *Annual Review of Plant Biology*, 70(1), 489–525. <https://doi.org/10.1146/annurev-arplant-050718-100334>
- Song W, Wei F, Gao S, Dong C, Hao J, Jin L, Li F, Wei P, Guo J, Wang R (2022) Functional characterization and comparison of lycopene epsilon-cyclase genes in *Nicotiana tabacum*. *BMC Plant Biol* 22(1):252. <https://doi.org/10.1186/s12870-022-03634-5>
- Sugiyama VF, Vasconcelos LAB, Rossi MM, Biondo C, de Setta N (2019) The population genetic structure approach adds new insights into the evolution of plant LTR retrotransposon lineages. *PLoS ONE* 14(5) e0214542. <https://doi.org/10.1371/journal.pone.0214542>
- Sun, G. (2012). MicroRNAs and their diverse functions in plants. *Plant Molecular Biology*, 80(1), 17–36. <https://doi.org/10.1007/s11103-011-9817-6>
- Sun T, Li L (2020) Toward the golden era: the status in uncovering the regulatory control of carotenoid accumulation in plants. *Plant Sci Int J Exp Plant Biol* 290(110331):110331. <https://doi.org/10.1016/j.plantsci.2019.110331>
- Sun T, Rao S, Zhou X, Li L (2022) Plant carotenoids: recent advances and future perspectives. *Mol Hortic*. <https://doi.org/10.1186/s43897-022-00023-2>
- Swanson C. P. 1981. *Cytogenetics: the chromosome in division, inheritance, and evolution*. Prentice-Hall. Englewood Cliffs, NJ. pp. 1,103.
- Swinnen, G., Goossens, A., & Pauwels, L. (2016a). Lessons from domestication: Targeting Cis - regulatory elements for crop improvement. *Trends in Plant Science*, 21(6), 506–515. <https://doi.org/10.1016/j.tplants.2016.01.014>
- Teakle, G. R., Manfield, I. W., Graham, J. F., & Gilmartin, P. M. (2002). Arabidopsis thaliana GATA factors: organisation, expression and DNA-binding characteristics. *Plant Molecular Biology*, 50(1), 43–56. <https://doi.org/10.1023/a:1016062325584>
- Tamayo-García R, Narváez-Zapata JA, Ku-González A, Aguilar-Espinosa M, Gutiérrez-Pacheco LC, Rivera-Madrid R (2022) Gene expression analysis during the fruit development in

-
- dehiscent and indehiscent *Bixa orellana* L. accessions. *Physiology and Molecular Biology of Plants*, 28(4), 709–718. <https://link.springer.com/article/10.1007/s.2298-022-01180-w>
- Teixeira da Silva JA, Dobránszki J, Rivera-Madrid R (2018) The biotechnology (genetic transformation and molecular biology) of *Bixa orellana* L. (achiote). *Planta* 248(2):267–277. <https://doi.org/10.1007/s00425-018-2909-7>
- Trujillo Hernández J. A. (2013). Caracterización de la variación alélica en genes involucradas en la biosíntesis de carotenoides en *Bixa orellana* L. Tesis de Maestría. Centro de Investigación Científica de Yucatán, A C. México. pp. 112
- Trujillo-Hdz JA, Cárdenas-Conejo Y, Turriza PE, Aguilar-Espinosa M, Carballo-Uicab V, Garza-Caligaris LE, Comai L, Rivera-Madrid R (2016) Functional polymorphism in lycopene beta-cyclase genes as a molecular marker to predict bixin production in *Bixa orellana* L. (achiote). *Mol Breed New Strateg Plant Improv*. <https://doi.org/10.1007/s11032-016-0555-y>
- Valdez-Ojeda R, Quiros CF, de Lourdes Aguilar-Espinosa M, Rivera-Madrid R (2010) Outcrossing rates in annatto determined by sequence-related amplified polymorphism. *Agron J* 102(5):1340–1345. <https://doi.org/10.2134/agronj2009.0510>
- Venkat, S. K., Bommisetty, P., Patil, M. S., Reddy, L., & Chennareddy, A. (2014). The genetic linkage maps of *Anthurium* species based on RAPD, ISSR and SRAP markers. *Scientia Horticulturae*, 178, 132–137. <https://doi.org/10.1016/j.scienta.2014.08.017>
- Vishnevetsky M, Ovadis M, Vainstein A (1999) Carotenoid sequestration in plants: the role of carotenoid-associated proteins. *Trends Plant Sci* 4(6):232–235. [https://doi.org/10.1016/s1360-1385\(99\)01414-4](https://doi.org/10.1016/s1360-1385(99)01414-4)
- Wan Q, Zhang X-G, Song M (2007) Fruit-specific RNAi-mediated lcy gene enhances content of lycopene in tomatoes silencing. *Sheng Wu gong Cheng xue bao [Chin J Biotechnol]* 23(3):429–433. [https://doi.org/10.1016/s1872-2075\(07\)60034-x](https://doi.org/10.1016/s1872-2075(07)60034-x)
- Wanner, G., Schroeder-Reiter, E., & Formanek, H. (2005). 3D Analysis of chromosome architecture: advantages and limitations with SEM. *Cytogenetic and Genome Research*, 109(1–3), 70–78. <https://doi.org/10.1159/000082384>

-
- Wang K, Huang G, Zhu Y (2016) Transposable elements play an important role during cotton genome evolution and fiber cell development. *Sci China Life Sci* 59(2):112–121. [https://doi.org/ 10.1007/s11427-015-4928-y](https://doi.org/10.1007/s11427-015-4928-y)
- Warnes MGR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A (2016) Package ‘gplots’. various R programming tools for plotting data. https://www.researchgate.net/profile/Arni-Magnusson-2/publication/303186599_gplots_Various_R_programming_tools_for_plotting_data/data/5a8fd3b80f7e9ba4296a11a1/gplots.pdf
- Wei, S., Gruber, M. Y., Yu, B., Gao, M.-J., Khachatourians, G. G., Hegedus, D. D., Parkin, I. A. P., & Hannoufa, A. (2012). Arabidopsis mutant sk156 reveals complex regulation of SPL15 in a miR156-controlled gene network. *BMC Plant Biology*, 12(1). <https://doi.org/10.1186/1471-2229-12-169>
- Windham, M. D., Pryer, K. M., Poindexter, D. B., Li, F.-W., Rothfels, C. J., & Beck, J. B. (2020). A step-by-step protocol for meiotic chromosome counts in flowering plants: A powerful and economical technique revisited. *Applications in Plant Sciences*, 8(4). <https://doi.org/10.1002/aps3.11342>
- Wray, G. A. (2003). The evolution of transcriptional regulation in eukaryotes. *Molecular Biology and Evolution*, 20(9), 1377–1419. <https://doi.org/10.1093/molbev/msg140>
- Xu, Q., Liu, Y., Zhu, A., Wu, X., Ye, J., Yu, K., Guo, W., & Deng, X. (2010). Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type. *BMC Genomics*, 11(1), 246. <https://doi.org/10.1186/1471-2164-11-246>
- Xue, W., Wang, Z., Du, M., Liu, Y., & Liu, J.-Y. (2013). Genome-wide analysis of small RNAs reveals eight fiber elongation-related and 257 novel microRNAs in elongating cotton fiber cells. *BMC Genomics*, 14(1). <https://doi.org/10.1186/1471-2164-14-629>
- Yanagisawa S (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize: roles of dof proteins in maize. *Plant J Cell Mol Biol* 21(3):281–288. <https://doi.org/10.1046/j.1365-313x.2000.00685.x>
- Zarza-García, A. L., Sauri-Duch, E., Raddatz-Mota, D., Cuevas-Glory, L. F., Pinzón-López, L. L., Rivera-Cabrera, F., & Mendoza-Espinoza, J. (2017). Pharmacological, phytochemical and

-
- morphological study of three Mayan accessions of *Bixa orellana* L. leaves. *Emirates Journal of Food and Agriculture*, 163-169.
- Zhang J, Sun H, Guo S, Ren Y, Li M, Wang J, Zhang H, Gong G, Xu Y (2020) Decreased protein abundance of lycopene β -cyclase contributes to red flesh in domesticated watermelon. *Plant Physiol* 183(3):1171–1183. <https://doi.org/10.1104/pp.19.01409>
- Zhang, X., Zhao, M., McCarty, D. R., & Lisch, D. (2020b). Transposable elements employ distinct integration strategies with respect to transcriptional landscapes in eukaryotic genomes. *Nucleic Acids Research*, 48(12), 6685–6698. <https://doi.org/10.1093/nar/gkaa370>
- Zhao Z, Liu Z, Mao X (2020) Biotechnological advances in lycopene β -cyclases. *J Agric Food Chem* 68(43):11895–11907. <https://doi.org/10.1021/acs.jafc.0c04814>
- Zhao-Yang, Z., Hang, S., & Zhi-Jian, G. (2002). Karyomorphological study of the *Spiraea japonica* complex (Rosaceae). *Brittonia*, 54(3), 168–174. [https://doi.org/10.1663/0007-196x\(2002\)054\[0168:ksotsj\]2.0.co;2](https://doi.org/10.1663/0007-196x(2002)054[0168:ksotsj]2.0.co;2)
- Zhou D, Shen Y, Zhou P, Fatima M, Lin J, Yue J, Zhang X, Chen L-Y, Ming R (2019) Papaya CpbHLH1/2 regulate carotenoid biosynthesis-related genes during papaya fruit ripening. *Hortic Res* 6(1):80. <https://doi.org/10.1038/s41438-019-0162-2>
- Zuo, J., Grierson, D., Courtney, L. T., Wang, Y., Gao, L., Zhao, X., Zhu, B., Luo, Y., Wang, Q., & Giovannoni, J. J. (2020). Relationships between genome methylation, levels of non-coding RNAs, mRNAs and metabolites in ripening tomato fruit. *The Plant Journal: For Cell and Molecular Biology*, 103(3), 980–994. <https://doi.org/10.1111/tpj.14778>



RESEARCH ARTICLE



Gene structure and potential regulation of the lycopene cyclase genes in *Bixa orellana* L.

G. Rivero-Manzanilla¹ · J. A. Narváez-Zapata² · M. Aguilar-Espinosa¹ · V. M. Carballo-Uicab¹ · R. Rivera-Madrid¹ Received: 20 July 2023 / Revised: 9 October 2023 / Accepted: 31 October 2023
© Prof. H.S. Srivastava Foundation for Science and Society 2023

Abstract

Lycopene cyclases (LCYs) are a key branching point in regulating the carotenoid biosynthesis pathway in plants. *Bixa orellana* L. is characterized by the presence in its seed of bixin, an apocarotenoid of significant importance in the food, pharmaceutical, and cosmetic industries. Gene analysis provides the opportunity to investigate the *LCY* gene structure in plant species and its relationship with the synthesis of carotenoids. Coding sequences of the *LCY* genes were retrieved from a *B. orellana* genome DNA. *Boβ-LCY1* and *Boβ-LCY2* genes exhibit 100% of identity to their respective cDNA accessions, and exhibit a single coding region of 1512 bp (504 aa) and 1495 bp (498 aa), respectively. In contrast, *Boe-LCY* gene shows a coding region of 1581 bp (527 aa) with 10 introns of diverse lengths. Putative Transcription Factors (TFs) binding sites were upstream (3000 bp) identified for each *LCY* gene. TFs cover two groups, one with the categories of photosynthesis, reproduction, and oxidative processes that are frequent. The second one with the categories of defense, cell cycle, signaling, and carbohydrate metabolism, which are poorly represented. Besides, repetitive DNA elements showed motifs and proteins related to LTR from the Ty3/Gypsy family, were associated with the TFs regions. In general, TFs vary in the different *BoLCY* genes, being more abundant in the *Boe-LCY* gene. *LCY* expression analyzed from a transcriptome database, and validated by RT-qPCR, shows an upregulation of the three *LCYs*, mainly oriented to the synthesis of essential carotenoids in photosynthetic tissues (leaves), as well as an upregulation of the *Boβ-LCY2* gene in the non-photosynthetic tissues (firsts seed development stages) related to the bixin accumulation.

Keywords Lycopene cyclase · Carotenoids · Bixin · *Bixa orellana* · Transcriptional factors

Introduction

The lycopene cyclase genes (*β-LCY* and *ε-LCY*) are key points in the carotenoid biosynthesis pathway (Song et al. 2022; Maoka 2020; Zhao et al. 2020; Raddatz-Mota et al. 2017; Lu and Li 2008). They play an important role in organisms during growth and development regulating the

physiological state of plants either increasing or decreasing carotenoid synthesis (Song et al. 2022; Maoka 2020; Zhao et al. 2020; Raddatz-Mota et al. 2017; Lu and Li 2008). These pigments are secondary metabolites consisting of a 40-carbon hydrocarbon main chain (Alcaño et al. 2016; Rosas-Saavedra and Stange 2016; Lu and Li 2008,) and their cycling is a crucial step for pathway bifurcation and generation of a diversity of apocarotenoid products (Song et al. 2022; Maoka 2020). Cycling involves the addition of β- and ε-rings, which determines the carotenoids composition in higher plants (Song et al. 2022; Maoka 2020; Zhao et al. 2020; Raddatz-Mota et al. 2017; Lu and Li 2008; Cunningham and Gantt 1998). In plants, carotenoids participate in primordial processes such as photosynthesis, photoprotection, antioxidants, membrane structural components, and precursors of phytohormones such as abscisic acid and strigolactones, which are involved in developmental processes and defense against abiotic and biotic stress factors (Kössler et al. 2021; Zhao et al. 2020; Alcaño et al. 2016; Fraser

✉ J. A. Narváez-Zapata
jnarvaez@ipn.mx

✉ R. Rivera-Madrid
renata@cicy.mx

¹ Unidad de Bioquímica y Biología Molecular de Plantas, Centro de Investigación Científica de Yucatán, A.C. Calle 43 # 130, Chuburná de Hidalgo, 97205 Mérida, Yucatán, Mexico

² Instituto Politécnico Nacional - Centro de Biotecnología Genómica, Blvd Del Maestro esq. Elias Piña, 88710 Reynosa, Tamaulipas, Mexico

Anexo 2. Ubicación de las secuencias cis de la región río arriba de los genes *BoLCY*, 3,000 pb.

CATEGORY	FACTOR	SEQUENCE	PUTATIVE CIS-ACTING ELEMENTS IN UPSTREAM REGION		
			<i>BoβLCY1</i>	<i>BoβLCY2</i>	<i>BoεLCY</i>
FOTOSÍNTESIS	DOFCOREZM	AAAG	31 (+), 139 (+), 201 (+), 213 (+), 343 (-), 353 (-), 389 (+), 406 (+), 440 (-), 531 (-), 517 (-), 631 (+), 727 (-), 838 (-), 984 (-), 1033 (+), 1097 (+), 1138 (-), 1167 (+), 1221 (+), 1425 (+), (+) 1623, (+) 1670, (-) 1730, 1739 (-), 1771 (-), 1855 (-), 1961 (-), 1968 (-), 2213 (-), 2256 (+), 2268 (-), 2336 (-), 2340 (-), 2375 (+), 2379 (+), 2556 (-), 2562 (-), 2570 (-), 2612 (-), 2678 (-), 2685 (-), 2723 (-), 2782 (-)	279 (-), 313 (-), 334 (-), 354 (-), 370 (-), 411(+), 497 (-), 694 (-), 716 (-), 723 (-), 905 (-), 916 (-), 1015 (-), 1195 (+), 1234 (-), 1375 (-), 1615 (-), 1789 (+), 2007 (+), 2108 (+), 2513 (-), 2520 (-), 2527 (-), 2551 (+), 2618 (-), 2708 (+), 2772 (+), 2790 (-), 2810 (-), 2859 (+), 2907 (-), 2953 (-), 2963 (-), 2986 (-)	77 (-), 88 (-), 95 (-), 99 (-), 103 (-), 108 (+), 150 (+), 190 (+), 226 (+), 336 (+), 343 (+), 385 (-), 398 (-), 509 (-), 596 (-), 694 (+), 753 (-), 847 (+), 895 (-), 902 (-), 964 (+), 1162 (+), 1171 (-), 1211 (+), 1307 (-), 1372 (+), 1377 (-), 1448 (-), 1529 (+), 1544 (+), 1554 (+), 1579 (+), 1627 (+), 1668 (+), 1698 (-), 1706 (-), 1712 (-), 1730 (+), 1773 (+), 1886 (+), 1946 (+), 1973 (+), 1996 (+), 2107 (+), 2181 (+), 2239 (-), 2262 (+), 2295 (+), 2640 (-), 2716 (+), 2725 (-), 2833 (+), 2843 (-), 2955 (+), 2975 (-), 2979 (-), 2991 (+)
	CACTFTPPCA1	YACT	26 (-), 144 (+), 162 (+), 206 (+), 283 (+), 293 (-), 296 (-), 307 (-), 327 (-), 351 (+), 428 (+), 544 (-), 582 (-), 754 (+), 776 (-), 817 (-), 881 (-), 911 (+), 916 (+), 1000 (+), 1175 (-), 1294 (-), 1455 (+), 1563 (+), 1737 (+), 1753 (-), 1769 (+), 1853 (+), 1959 (+), 1992 (-), 2071 (-), 2187 (-), 2194 (+), 2240 (+), 2422 (-), 2456 (-), 2674 (+)	13 (-), 68 (+), 93 (-), 121 (+), 175 (+), 243 (+), 308 (-), 332 (+), 361 (+), 252 (-), 575 (+), 616 (-), 714 (+), 735 (+), 763 (+), 852 (+), 926 (+), 957 (+), 989 (-), 1038 (+), 1126 (+), 1150 (-), 1717 (+), 2009 (-), 2451 (-), 2492 (-), 2607 (+), 2616 (+), 2700 (-), 2741 (+), 2802 (+), 2880 (+)	12 (+), 44 (+), 55 (-), 70 (+), 136 (+), 154 (-), 172 (+), 198 (+), 216 (+), 315 (-), 383 (+), 394 (-), 563 (+), 618 (+), 666 (+), 672 (+), 678 (+), 843 (+), 849 (-), 856 (+), 893 (+), 900 (+), 933 (+), 1071 (+), 1081 (+), 1148 (-), 1169 (+), 1305 (+), 1452 (-), 1581 (-), 1592 (+), 1631 (-), 1748 (-), 1804 (-), 2062 (+), 2151 (+), 2191 (-), 2244 (+), 2422 (+), 2459 (+), 2687 (-), 2811 (+), 2841 (+)
	ROOTMOTIFTAPOX1	ATATT	367 (-), 683 (-), 707 (+), 964 (+), 975 (-), 2398 (+), 2483 (+)	105 (+), 126 (-), 127 (+), 427 (-), 463 (+), 963 (+), 1049 (-), 1202 (+), 1252 (+), 1255 (-), 1289 (+), 1324 (+), 1487 (-), 1725 (+), 1738 (+), 1743 (-), 1746 (+), 1835 (+), 1842 (-), 1858 (-), 1936 (+), 1991 (+), 2018 (-), 2021 (+), 2058 (+), 2076 (+), 2057 (-), 2058 (+), 2075 (-), 2076 (+), 2094 (-), 2123 (+), 2123 (+), 2147 (-), 2182 (-), 2183 (+), 2287 (-), 2288 (+), 2343 (-), 2346 (+), 2377 (+)	28 (-), 306 (+), 604 (-), 945 (-), 1128 (-), 1129 (+), 1186 (-), 1216 (-), 1488 (-), 1489 (+), 1614 (-), 2169 (-), 2170 (+)
REPRODUCCIÓN	POLLEN1LELAT52	AGAAA	198 (+), 358 (+), 501 (-), 590 (+), 760 (+), 1099 (+), 1218 (+), 1278 (+), 1556 (+), 1621 (+), 1667 (+), 2214 (-), 2310 (+), 2333 (-), 2337 (-), 2377 (+), 2474 (+), 2648 (-), 2783 (-), 2808 (-)	169 (+), 250 (+), 351 (-), 378 (-), 650 (-), 695 (-), 725 (-), 771 (+), 900 (-), 918 (-), 1189 (-), 1612 (-), 1709 (-), 2619 (-), 2857 (+), 2904 (-), 2919 (-), 2987 (-), 2987 (-)	78 (-), 96 (-), 100 (-), 188 (+), 334 (+), 345 (+), 742 (+), 754 (-), 878 (-), 1623 (+), 2642 (-), 2649 (+), 2976 (-)
	GTGANTG10	GTGA	394 (+), 882 (+), 957 (-), 1003 (-), 1240 (+), 1245 (+), 1315 (-), 1364 (-), 1443 (-), 1768 (-), 1864 (-), 2218 (-), 2252 (+), 2452 (+), 2664 (+), 2770 (-), 2953 (+)	114 (-), 187 (+), 242 (-), 305 (+), 360 (-), 561 (-), 574 (-), 851 (-), 885 (+), 925 (-), 1125 (-), 1311 (-), 1505 (-), 1570 (-), 2503 (+), 2735 (-), 2841 (-)	197 (-), 539 (+), 825 (-), 842 (-), 850 (+), 855 (-), 905 (+), 1070 (-), 1080 (-), 1346 (-), 1351 (+), 1453 (+), 1497 (-), 2006 (+), 2688 (+), 2773 (-), 2820 (-), 2870 (-), 2929 (+)
	CAATBOX1	CAAT	44 (+), 174 (+), 366 (+), 374 (+), 415 (+), 458 (+), 470 (+), 800 (+), 931 (+), 1024 (-), 1109 (+), 1418 (-), 1748 (+), 1765 (+), 1810 (-), 1817 (-), 1917 (-), 2010 (-), 2061 (+), 2249 (-), 2345 (-), 2467 (+), 2794 (-), 2964 (-)	22 (-), 107 (-), 129 (-), 148 (+), 165 (-), 195 (-), 223 (+), 302 (-), 347 (-), 388 (-), 457 (-), 470 (+), 488 (+), 534 (-), 542 (+), 711 (+), 950 (+), 1019 (+), 1079 (-), 1163 (-), 1251 (+), 1296 (-), 1513 (+), 1540 (-), 1564 (-), 1643 (+), 1645 (-), 1817 (-), 2125 (-), 2164 (-), 2331 (+), 2365 (-), 2691 (+), 2738 (+), 2761 (+), 2949 (-), 2972 (-)	27 (+), 194 (+), 213 (+), 265 (+), 269 (+), 274 (-), 325 (+), 413 (+), 429 (+), 503 (+), 505 (-), 552 (-), 731 (+), 764 (+), 774 (+), 908 (+), 1025 (+), 1055 (-), 1215 (+), 1223 (-), 1329 (-), 1390 (-), 1470 (-), 149 (-), 1756 (+), 1798 (+), 1880 (-), 1934 (+), 2085 (-), 2102 (+), 2217 (+), 2340 (+), 2368 (-), 2430 (+), 2660 (+), 2904 (-), 2922 (+)
SISTEMA OXIDATIVO	GATABOX	GATA	83 (-), 130 (-), 238 (+), 371 (-), 537 (+), 628 (+), 706 (+), 835 (-), 878 (+), 898 (-), 981 (-), 1020 (-), 1046 (+), 1492 (-), 1755 (-), 1921 (+), 1994 (-), 1998 (-), 2493 (+), 2580 (-), 2762 (-), 2889 (-), 2993 (-)	33 (-), 140 (+), 358 (-), 417 (-), 754 (+), 791 (+), 838 (-), 858 (+), 893 (-), 923 (-), 1201 (+), 1261 (-), 1323 (+), 1448 (-), 1509 (+), 1574 (+), 1665 (+), 1681 (-), 1735 (+), 1944 (+), 2274 (+), 2320 (+), 2376 (+), 2782 (+)	30 (-), 134 (+), 229 (+), 262 (-), 368 (+), 377 (+), 533 (-), 914 (-), 947 (-), 1049 (-), 1155 (-), 1165 (+), 1315, (+) 1485 (+), 1701 (-), 1734 (+), 1740 (+), 1959 (+), 2081 (+), 2088 (+), 2137 (+), 2164 (-), 2321 (+), 2628 (+), 2736 (+), 2751 (-), 2878 (-)
	EBOXBNNAPA	CANN TG	1230 (-), 1230 (+), 1379 (+), 1379 (-), 1444 (-), 1444 (+), 2219 (-), 2219 (+), 2429 (-), 2429 (+)	445 (-), 445 (+), 783 (-), 783 (+), 926 (-), 926 (+), 1225 (-), 1225 (+), 1398 (-), 1398 (+), 1451 (-), 1451 (+), 1471 (-), 1471 (+), 1643 (-), 1643 (+), 2382 (-), 2382 (+)	437 (-), 437 (+), 503 (-), 503 (+), 577 (-), 577 (+), 637 (-), 637 (+), 712 (-), 712 (+), 812 (-), 812 (+), 970 (-), 970 (+), 1081 (-), 1081 (+), 1507 (-), 1507 (+) 1522 (-), 1522 (+), 2244 (-), 2244 (+), 2307 (-), 2307 (+), 2383 (-), 2383 (+), 2583 (-), 2583 (+), 2634 (-), 2634 (+), 2674 (-), 2674 (+), 2695 (-), 2695 (+), 2703 (-), 2703 (+)
	GT1CONSENSUS	GRWAAW	243 (-), 359 (+), 526 (-), 628 (+), 634 (+), 635 (+), 761 (+), 1027 (+), 1100 (+), 1257 (+), 1279 (+), 1449 (+), 1450 (+), 1471 (+), 1472 (+), 1490 (-), 1557 (+), 1740 (-), 1741 (-), 1891 (-), 1921 (+), 2331 (-), 2404 (-)	84 (-), 170 (+), 287 (+), 356 (-), 384 (+), 432 (+), 433 (+), 648 (-), 718 (-), 772 (-), 836 (-), 891 (-), 1327 (-), 1335 (+), 1412 (-), 1558 (-), 1616 (-), 1707 (-), 2274 (+), 2320 (+), 2387 (+), 2676 (+), 2680 (+), 2917 (-)	122 (+), 123 (+), 260 (-), 531 (-), 601 (+), 920 (+), 1047 (-), 1065 (+), 1155 (+), 1315 (+), 1386 (+), 1409 (+), 1437 (-), 1485 (+), 1551 (+), 1624 (+), 1664 (+), 1665 (+), 1699 (-), 1740 (+), 1853 (-), 1959 (+), 2067 (+), 2081 (+), 2088

Anexo 3. Matriz de elementos *cis* por categorías de los genes *BoLCY*

CATEGORIA	FACTOR	CLAVE	BoβLCY1	BoβLCY2	BoεLCY
METABOLISMO DE CARBOHIDRATOS	ACGTABOX	ACGTABOX	0	0	2
	AMYBOX1	AMYBOX1	2	1	3
	AMYBOX2	AMYBOX2	0	1	1
	CGACGOSAMY3	CGACGAMY3	4	0	2
	MYBGAHV	MYBGAHV	0	1	3
	PYRIMIDINEBOXOSRAMY1A	PYRIMIDINEBOX	3	2	2
	SP8BFIBSP8BIB	SP8BF	1	0	1
	SREATMSD	SER	1	0	4
	TATCCAOSAMY	TATCCA	1	2	1
	TATCCAYMOTIFOSRAMY3D	TATCCAMOTIF	0	1	1
CICLO CELULAR	WBOXHVISO1	WBOXISO1	4	3	2
	CELLCYCLESC	CELLCYC	1	0	0
	E2FCONSENSUS	E2F	0	2	2
	MARABOX1	MARABOX1	1	2	0
	MARTBOX	MARTBOX	7	9	1
	MARARS	MARARS	0	1	0
	MYBCOREATCYCB1	MYBCOREAT	4	2	4
	QELEMENTZMZM13	QELEMENT	0	2	1
TRANSCRIPCIÓN	INRNTPSADB	INRNT	8	13	1
	POLASIG1	POLASIG1	7	15	6
	POLASIG2	POLASIG2	2	13	6
	POLASIG3	POLASIG3	4	24	8
	TATABOX1	TATABOX1	0	1	0
	TATABOX2	TATABOX2	0	9	4
	TATABOX3	TATABOX3	0	11	1
	TATABOX4	TATABOX4	2	6	2
	TATABOX5	TATABOX5	8	15	8
	TATABOXOSPAL	TATABOX	1	2	2
DESCONOCIDOS	TATAPVTRNALEU	TATA	1	5	0
	P1BS	P1BS	0	2	0
	PREATPRODH	PREAT	2	2	2
	S1FSORPL21	S1F	0	0	1
	SITEIIATCYTC	SITEII	5	4	0
	SURECOREATSULTR11	SURECORE	4	3	2
RESPUESTA A ENFERMEDADES	SPHCOREZMC1	SPHCORE	0	1	0
	ABREATCONSENSUS	ABREAT	1	0	1
	ABRELATERD1	ABREL	5	2	4
	ABREOSRAB21	ABREOS	3	0	0
	ABRERATCAL	ABRER	3	0	4
	ACGTABREMOTIFA2OSEM	ACGT2	1	0	2
	ACGTATERD1	ACGT1	8	2	10
	ANAERO1CONSENSUS	ANAERO1	3	4	5
	ANAERO2CONSENSUS	ANAERO2	2	1	2
	ANAERO3CONSENSUS	ANAERO3	0	0	1
	CBFHV	CBFHV	1	0	1
	CCAATBOX1	CCAATBOX1	8	5	7
	CERREGLUBOX1PSLEGA	CERREGLUBOX1	1	0	0
	CERREGLUBOX2PSLEGA	CERREGLUBOX2	1	0	0
	CURECORECR	CURECORE	10	10	12
	DPBFCOREDCDC3	DPBFCORE	1	2	5
	DRECRTCOREAT	DRECR	1	0	0
	DRE2COREZMRAB17	DRE2CORE	0	0	1
	ELRECOREPCRP1	ELRECORE	2	1	1
	EMBP1TAEM	EMBP1	1	0	1
	GT1GMSCAM4	GT1G	7	4	4
	IRO2OS	IRO2	1	0	1
	LTRE1HVBLT49	LTRE1	2	1	0
	LTRECOREATCOR15	LTRECORE	3	0	2
	MYB1AT	MYB1AT	3	7	2
	MYB2AT	MYB2AT	2	0	0
	MYB2CONSENSUSAT	MYB2	2	2	2
	MYBCORE	MYBCORE	7	4	4
	MYBPLANT	MYBPLANT	1	1	0
	MYCCONSENSUSAT	MYC	10	18	36
	MYCATERD1	MYCAT	0	0	2
	MYCATRD22	MYCATR	0	0	2
	RBCSCONSENSUS	RBCS	2	0	1
	RYREPEATBNNAPA	RYREPEATB	2	3	0
DEFENSA FACTORES AMBIENTALES	GCCCORE	GCCCORE	1	0	1
	SEBFCONSSTPR10A	SEBFCONS	0	2	0
	WBOXATNPR1	WBOXNPR1	3	12	6

	WBOXPCWRKY1	WBOX	0	0	1	
	WBOXNTCHN48	WBOXCHN	1	1	1	
	WBOXNTERF3	WBOXERF3	7	5	5	
	WRKY71OS	WRKY7	12	19	17	
SEÑALIZACIÓN	AGCBOXNPLB	AGCBOX	1	0	0	
	ARFAT	ARFAT	0	1	1	
	ARR1AT	ARR1AT	34	22	43	
	ASF1MOTIFCAMV	ASF1MOTIF	3	3	3	
	AUXREPSIAA4	AUXRE	1	0	0	
	CAREOSREP1	CAREREP1	1	3	0	
	CATATGGMSAUR	CATATG	2	2	4	
	CPBCSPOR	CPBC	1	0	2	
	D2GMAUX28	AUX28	0	2	0	
	ERELEE4	ERELEE4	0	1	2	
	GARE1OSREP1	GAREREP1	2	0	0	
	EVENINGAT	EVENINGAT	0	0	1	
	GAREAT	GAREAT	2	2	3	
	GADOWNAT	GADOWN	0	0	1	
	GGTCCCATGMSAUR	GMSAUR	1	0	0	
	LECPLEACS2	LECP	0	3	2	
	T/GBOXATPIN2	T/GBOX	0	1	0	
	10PEHVPSBD	PEHV	1	2	0	
	SISTEMA OXIDATIVO	BOXCPSAS1	BOXC1	1	0	0
		BOXIIPCCHS	BOXIIPC	1	0	1
CACGTGMOTIF		CACGTG	4	0	4	
CIACADIANLELHC		CIACADIAN	1	0	5	
EBOXBNNAPA		EBOXB	10	18	36	
GATABOX		GATABOX	23	24	27	
GBOXLERBCS		GBOXL	0	0	1	
GT1CONSENSUS		GT1	32	25	42	
GT1CORE		GT1CORE	0	2	3	
HDZIP2ATATHB2		HDZIP2	0	2	4	
IBOX		IBOX	2	1	0	
IBOXCORE		IBOXCORE	8	7	14	
IBOXCORENT		IBOXCORENT	1	1	0	
PALBOXAPC		BOXA	1	0	2	
PRECONSCRHSP70A		PRECONS	6	1	5	
REALPHALGLHCB21		REALPHA	1	4	0	
REBETALGLHCB2		REBETA	1	0	0	
SORLIP1AT		SORLIP1AT	5	3	2	
SORLIP2AT		SORLIP2AT	6	4	1	
SORLIP5AT		SORLIP5AT	1	0	0	
SV40COREENHAN		SV40CORE	1	0	1	
TBOXATGAPB		TBOXAT	1	1	3	
UPRMOTIFIAT		UPRMOTIF	1	0	1	
REPRODUCCIÓN		2SSEEDPROTBANAPA	2SSEED	2	1	0
		300CORE	300CORE	2	0	0
		300ELEMENT	300ELM	2	1	2
		AACACOREOSGLUB1	ACACORE	1	0	1
	ACGTCBOX	ACGTCBOX	2	0	0	
	ATHB1ATCONSENSUS	ATHB1AT	2	0	0	
	ATHB5ATCORE	ATHB5AT	0	0	2	
	BIHD1OS	BIHD1	2	11	9	
	CANBNNAPA	CANBN	3	1	1	
	ECCRCAH1	ECCR	9	10	5	
	GCN4OSGLUB1	GCN4	0	0	2	
	GTGANTG10	GTGANT	17	17	19	
	NAPINMOTIFBN	NAPIN	0	0	1	
	NODCON1GM	NOD1GM	3	0	4	
	NODCON2GM	NOD2GM	9	5	10	
	OSE1ROOTNODULE	OSE1ROOT	3	0	4	
	OSE2ROOTNODULE	OSE2ROOT	9	5	10	
	POLLEN1LELAT52	POLLEN1	21	19	13	
	POLLEN2LELAT52	POLLEN2	0	1	0	
	SEF1MOTIF	SEF1MOTIF	1	3	2	
	SEF3MOTIFGM	SEF3	1	4	2	
	SEF4MOTIFGM7S	SEF4	6	17	6	
	TGACGTMAMY	TGACG	2	0	1	
	CAATBOX1	CAATBOX1	24	37	37	
	FOTOSÍNTESIS	BOXIINTPATPB	BOXIINT	2	6	4
		BOXLCOREDCPAL	BOXLCORE	1	0	0
		BS1EGCCR	BS1	1	0	0
CACTFTPPCA1		CACTFTPPCA1	38	32	43	
CARGCW8GAT		CARGCW	2	8	1	
	CARGATCONSENSUS	CARGAT	0	2	0	

	DOFCOREZM	DOFCORE	45	34	60
	GBOX10NT	GBOX10	1	0	0
	HEXMOTIFTAH3H4	HEXMOTIFT	2	0	0
	HEXAMERATH4	HEXAME	0	0	1
	L1BOXATPDF1	L1BOX	0	1	1
	MYBPZM	MYBPZM	2	0	1
	MYBST1	MYBST1	2	3	7
	NTBBF1ARROLB	NTBBF1	1	0	4
	RAV1AAT	RAV1AAT	7	16	9
	RAV1BAT	RAV1BAT	1	1	0
	RHERPATEXPA7	RHERP	9	4	5
	ROOTMOTIFTAPOX1	ROOTMOTIFT	7	40	13
	RYREPEATGMGY2	RYREPEATG	1	1	0
	RYREPEATLEGUMINBOX	RYREPEATL	1	1	0
	S1FBOXSORPS1L21	S1FBOX	1	2	3
	TAAAGSTKST1	TAAAG	8	3	10
	UP1ATMSD	UP1	0	1	0
	WUSATAG	WUSAG	0	2	0

Anexo 4. Secuencias hits para los genes *BoLCY* del transcriptoma de achote

BoβLCY1

# 984 hits	Subject acc.ver	%Identidad	Alignment length	Mismatches	E. value	Bit score
1	SRA:SRR2131178.13957915.1	100	79	0	8.32E-33	147
2	SRA:SRR2131178.13925248.2	100	79	0	8.32E-33	147
3	SRA:SRR2131178.13925248.1	100	79	0	8.32E-33	147
4	SRA:SRR2131178.13913099.1	100	79	0	8.32E-33	147
5	SRA:SRR2131178.13908946.1	100	79	0	8.32E-33	147
6	SRA:SRR2131178.13850445.2	100	79	0	8.32E-33	147
7	SRA:SRR2131178.13767401.1	100	79	0	8.32E-33	147
8	SRA:SRR2131178.13752811.1	100	79	0	8.32E-33	147
9	SRA:SRR2131178.13710156.1	100	79	0	8.32E-33	147
10	SRA:SRR2131178.13703771.1	100	79	0	8.32E-33	147
11	SRA:SRR2131178.13700172.1	100	79	0	8.32E-33	147
12	SRA:SRR2131178.13695216.1	100	79	0	8.32E-33	147
13	SRA:SRR2131178.13650786.1	100	79	0	8.32E-33	147
14	SRA:SRR2131178.13624135.1	100	79	0	8.32E-33	147
15	SRA:SRR2131178.13622373.1	100	79	0	8.32E-33	147
16	SRA:SRR2131178.13604516.1	100	79	0	8.32E-33	147
17	SRA:SRR2131178.13603887.1	100	79	0	8.32E-33	147
18	SRA:SRR2131178.13601420.1	100	79	0	8.32E-33	147
19	SRA:SRR2131178.13598388.1	100	79	0	8.32E-33	147
20	SRA:SRR2131178.13581370.1	100	79	0	8.32E-33	147
21	SRA:SRR2131178.13560482.1	100	79	0	8.32E-33	147
22	SRA:SRR2131178.13525486.2	100	79	0	8.32E-33	147
23	SRA:SRR2131178.13525486.1	100	79	0	8.32E-33	147
24	SRA:SRR2131178.13524818.1	100	79	0	8.32E-33	147
25	SRA:SRR2131178.13513821.1	100	79	0	8.32E-33	147
26	SRA:SRR2131178.13484533.1	100	79	0	8.32E-33	147

27	SRA:SRR2131178.13473482.1	100	79	0	8.32E-33	147
28	SRA:SRR2131178.13434139.1	100	79	0	8.32E-33	147
26	SRA:SRR2131178.13400550.1	100	79	0	8.32E-33	147
27	SRA:SRR2131178.13395422.1	100	79	0	8.32E-33	147
28	SRA:SRR2131178.13378315.1	100	79	0	8.32E-33	147
29	SRA:SRR2131178.13328605.1	100	79	0	8.32E-33	147
30	SRA:SRR2131178.13323936.1	100	79	0	8.32E-33	147
31	SRA:SRR2131178.13306686.1	100	79	0	8.32E-33	147
32	SRA:SRR2131178.13306398.1	100	79	0	8.32E-33	147
33	SRA:SRR2131178.13300410.1	100	79	0	8.32E-33	147
34	SRA:SRR2131178.13263361.1	100	79	0	8.32E-33	147
35	SRA:SRR2131178.13261396.1	100	79	0	8.32E-33	147
36	SRA:SRR2131178.13259779.1	100	79	0	8.32E-33	147
37	SRA:SRR2131178.13259275.1	100	79	0	8.32E-33	147
38	SRA:SRR2131178.13227831.2	100	79	0	8.32E-33	147
39	SRA:SRR2131178.13224946.1	100	79	0	8.32E-33	147
40	SRA:SRR2131178.13212814.1	100	79	0	8.32E-33	147
41	SRA:SRR2131178.13191318.1	100	79	0	8.32E-33	147
42	SRA:SRR2131178.13154727.1	100	79	0	8.32E-33	147
43	SRA:SRR2131178.13141155.1	100	79	0	8.32E-33	147
44	SRA:SRR2131178.13123977.1	100	79	0	8.32E-33	147
45	SRA:SRR2131178.13101669.1	100	79	0	8.32E-33	147
46	SRA:SRR2131178.13041279.1	100	79	0	8.32E-33	147
47	SRA:SRR2131178.13028211.1	100	79	0	8.32E-33	147
48	SRA:SRR2131178.12994011.1	100	79	0	8.32E-33	147
49	SRA:SRR2131178.12979155.1	100	79	0	8.32E-33	147
50	SRA:SRR2131178.12961855.1	100	79	0	8.32E-33	147
51	SRA:SRR2131178.12956405.1	100	79	0	8.32E-33	147
52	SRA:SRR2131178.12952786.1	100	79	0	8.32E-33	147
53	SRA:SRR2131178.12920835.1	100	79	0	8.32E-33	147
54	SRA:SRR2131178.12908183.1	100	79	0	8.32E-33	147
55	SRA:SRR2131178.12885432.2	100	79	0	8.32E-33	147
56	SRA:SRR2131178.12885432.1	100	79	0	8.32E-33	147
57	SRA:SRR2131178.12874344.1	100	79	0	8.32E-33	147
58	SRA:SRR2131178.12857469.1	100	79	0	8.32E-33	147
59	SRA:SRR2131178.12836881.1	100	79	0	8.32E-33	147
60	SRA:SRR2131178.12815754.1	100	79	0	8.32E-33	147
61	SRA:SRR2131178.12800794.1	100	79	0	8.32E-33	147
62	SRA:SRR2131178.12777780.1	100	79	0	8.32E-33	147
63	SRA:SRR2131178.12769859.1	100	79	0	8.32E-33	147
64	SRA:SRR2131178.12766781.1	100	79	0	8.32E-33	147
65	SRA:SRR2131178.12746083.1	100	79	0	8.32E-33	147

66	SRA:SRR2131178.12712890.2	100	79	0	8.32E-33	147
67	SRA:SRR2131178.12712890.1	100	79	0	8.32E-33	147
68	SRA:SRR2131178.12699974.2	100	79	0	8.32E-33	147
69	SRA:SRR2131178.12681454.1	100	79	0	8.32E-33	147
70	SRA:SRR2131178.12677065.2	100	79	0	8.32E-33	147
71	SRA:SRR2131178.12677065.1	100	79	0	8.32E-33	147
72	SRA:SRR2131178.12675292.1	100	79	0	8.32E-33	147
73	SRA:SRR2131178.12660501.1	100	79	0	8.32E-33	147
74	SRA:SRR2131178.12660169.2	100	79	0	8.32E-33	147
75	SRA:SRR2131178.12660169.1	100	79	0	8.32E-33	147
76	SRA:SRR2131178.12641405.1	100	79	0	8.32E-33	147
77	SRA:SRR2131178.12623768.1	100	79	0	8.32E-33	147
78	SRA:SRR2131178.12614108.1	100	79	0	8.32E-33	147
79	SRA:SRR2131178.12603529.1	100	79	0	8.32E-33	147
80	SRA:SRR2131178.12586543.1	100	79	0	8.32E-33	147
81	SRA:SRR2131178.12571905.1	100	79	0	8.32E-33	147
82	SRA:SRR2131178.12568894.1	100	79	0	8.32E-33	147
83	SRA:SRR2131178.12560302.1	100	79	0	8.32E-33	147
84	SRA:SRR2131178.12532586.1	100	79	0	8.32E-33	147
85	SRA:SRR2131178.12532009.1	100	79	0	8.32E-33	147
86	SRA:SRR2131178.12491933.1	100	79	0	8.32E-33	147
87	SRA:SRR2131178.12483036.1	100	79	0	8.32E-33	147
88	SRA:SRR2131178.12415895.1	100	79	0	8.32E-33	147
89	SRA:SRR2131178.12410394.1	100	79	0	8.32E-33	147
90	SRA:SRR2131178.12402207.1	100	79	0	8.32E-33	147
91	SRA:SRR2131178.12383913.1	100	79	0	8.32E-33	147
92	SRA:SRR2131178.12374368.1	100	79	0	8.32E-33	147
93	SRA:SRR2131178.12363045.1	100	79	0	8.32E-33	147
94	SRA:SRR2131178.12360815.1	100	79	0	8.32E-33	147
95	SRA:SRR2131178.12356967.1	100	79	0	8.32E-33	147
96	SRA:SRR2131178.12341208.1	100	79	0	8.32E-33	147
97	SRA:SRR2131178.12335021.1	100	79	0	8.32E-33	147
98	SRA:SRR2131178.12323875.1	100	79	0	8.32E-33	147
99	SRA:SRR2131178.12294655.1	100	79	0	8.32E-33	147
100	SRA:SRR2131178.12279400.1	100	79	0	8.32E-33	147

Bo β LCY2

1,217 hits	Subject acc. Ver.	%Identidad	Aligment length	Mismatches	E. value	Bit score
1	SRA:SRR2131178.13952397.1	100	79	0	8.24E-33	147
2	SRA:SRR2131178.13865068.1	100	79	0	8.24E-33	147
3	SRA:SRR2131178.13861404.2	100	79	0	8.24E-33	147

4	SRA:SRR2131178.13861404.1	100	79	0	8.24E-33	147
5	SRA:SRR2131178.13853652.1	100	79	0	8.24E-33	147
6	SRA:SRR2131178.13849470.1	100	79	0	8.24E-33	147
7	SRA:SRR2131178.13840015.1	100	79	0	8.24E-33	147
8	SRA:SRR2131178.13829167.2	100	79	0	8.24E-33	147
9	SRA:SRR2131178.13827517.2	100	79	0	8.24E-33	147
10	SRA:SRR2131178.13805824.2	100	79	0	8.24E-33	147
11	SRA:SRR2131178.13790189.1	100	79	0	8.24E-33	147
12	SRA:SRR2131178.13752605.1	100	79	0	8.24E-33	147
13	SRA:SRR2131178.13673514.1	100	79	0	8.24E-33	147
14	SRA:SRR2131178.13665100.1	100	79	0	8.24E-33	147
15	SRA:SRR2131178.13582101.1	100	79	0	8.24E-33	147
16	SRA:SRR2131178.13561608.1	100	79	0	8.24E-33	147
17	SRA:SRR2131178.13555066.1	100	79	0	8.24E-33	147
18	SRA:SRR2131178.13553951.1	100	79	0	8.24E-33	147
19	SRA:SRR2131178.13548595.1	100	79	0	8.24E-33	147
20	SRA:SRR2131178.13547744.1	100	79	0	8.24E-33	147
21	SRA:SRR2131178.13484029.1	100	79	0	8.24E-33	147
22	SRA:SRR2131178.13476167.1	100	79	0	8.24E-33	147
23	SRA:SRR2131178.13433876.1	100	79	0	8.24E-33	147
24	SRA:SRR2131178.13426837.1	100	79	0	8.24E-33	147
25	SRA:SRR2131178.13422501.1	100	79	0	8.24E-33	147
26	SRA:SRR2131178.13399747.1	100	79	0	8.24E-33	147
27	SRA:SRR2131178.13380993.1	100	79	0	8.24E-33	147
28	SRA:SRR2131178.13353478.1	100	79	0	8.24E-33	147
26	SRA:SRR2131178.13330480.1	100	79	0	8.24E-33	147
27	SRA:SRR2131178.13271680.1	100	79	0	8.24E-33	147
28	SRA:SRR2131178.13251626.1	100	79	0	8.24E-33	147
29	SRA:SRR2131178.13244894.1	100	79	0	8.24E-33	147
30	SRA:SRR2131178.13237670.1	100	79	0	8.24E-33	147
31	SRA:SRR2131178.13222453.1	100	79	0	8.24E-33	147
32	SRA:SRR2131178.13202902.1	100	79	0	8.24E-33	147
33	SRA:SRR2131178.13200420.1	100	79	0	8.24E-33	147
34	SRA:SRR2131178.13187166.2	100	79	0	8.24E-33	147
35	SRA:SRR2131178.13179079.1	100	79	0	8.24E-33	147
36	SRA:SRR2131178.13145788.1	100	79	0	8.24E-33	147
37	SRA:SRR2131178.13084341.1	100	79	0	8.24E-33	147
38	SRA:SRR2131178.13080876.1	100	79	0	8.24E-33	147
39	SRA:SRR2131178.13065536.1	100	79	0	8.24E-33	147
40	SRA:SRR2131178.13028509.1	100	79	0	8.24E-33	147
41	SRA:SRR2131178.13019984.1	100	79	0	8.24E-33	147
42	SRA:SRR2131178.12991208.1	100	79	0	8.24E-33	147

43	SRA:SRR2131178.12971797.1	100	79	0	8.24E-33	147
44	SRA:SRR2131178.12959477.1	100	79	0	8.24E-33	147
45	SRA:SRR2131178.12951361.1	100	79	0	8.24E-33	147
46	SRA:SRR2131178.12951355.1	100	79	0	8.24E-33	147
47	SRA:SRR2131178.12930596.1	100	79	0	8.24E-33	147
48	SRA:SRR2131178.12915657.2	100	79	0	8.24E-33	147
49	SRA:SRR2131178.12890434.1	100	79	0	8.24E-33	147
50	SRA:SRR2131178.12881048.1	100	79	0	8.24E-33	147
51	SRA:SRR2131178.12847887.1	100	79	0	8.24E-33	147
52	SRA:SRR2131178.12843145.1	100	79	0	8.24E-33	147
53	SRA:SRR2131178.12823383.1	100	79	0	8.24E-33	147
54	SRA:SRR2131178.12817219.1	100	79	0	8.24E-33	147
55	SRA:SRR2131178.12790641.1	100	79	0	8.24E-33	147
56	SRA:SRR2131178.12763360.1	100	79	0	8.24E-33	147
57	SRA:SRR2131178.12738373.1	100	79	0	8.24E-33	147
58	SRA:SRR2131178.12735576.1	100	79	0	8.24E-33	147
59	SRA:SRR2131178.12717621.1	100	79	0	8.24E-33	147
60	SRA:SRR2131178.12700070.1	100	79	0	8.24E-33	147
61	SRA:SRR2131178.12693125.1	100	79	0	8.24E-33	147
62	SRA:SRR2131178.12690745.1	100	79	0	8.24E-33	147
63	SRA:SRR2131178.12580433.1	100	79	0	8.24E-33	147
64	SRA:SRR2131178.12568883.1	100	79	0	8.24E-33	147
65	SRA:SRR2131178.12523249.1	100	79	0	8.24E-33	147
66	SRA:SRR2131178.12500831.1	100	79	0	8.24E-33	147
67	SRA:SRR2131178.12406371.1	100	79	0	8.24E-33	147
68	SRA:SRR2131178.12385175.1	100	79	0	8.24E-33	147
69	SRA:SRR2131178.12384294.2	100	79	0	8.24E-33	147
70	SRA:SRR2131178.12381295.1	100	79	0	8.24E-33	147
71	SRA:SRR2131178.12372794.1	100	79	0	8.24E-33	147
72	SRA:SRR2131178.12342476.1	100	79	0	8.24E-33	147
73	SRA:SRR2131178.12341910.1	100	79	0	8.24E-33	147
74	SRA:SRR2131178.12321177.1	100	79	0	8.24E-33	147
75	SRA:SRR2131178.12306285.1	100	79	0	8.24E-33	147
76	SRA:SRR2131178.12296784.1	100	79	0	8.24E-33	147
77	SRA:SRR2131178.12287686.1	100	79	0	8.24E-33	147
78	SRA:SRR2131178.12284813.1	100	79	0	8.24E-33	147
79	SRA:SRR2131178.12272518.1	100	79	0	8.24E-33	147
80	SRA:SRR2131178.12272427.1	100	79	0	8.24E-33	147
81	SRA:SRR2131178.12222909.1	100	79	0	8.24E-33	147
82	SRA:SRR2131178.12207386.1	100	79	0	8.24E-33	147
83	SRA:SRR2131178.12156312.1	100	79	0	8.24E-33	147
84	SRA:SRR2131178.12077722.1	100	79	0	8.24E-33	147

85	SRA:SRR2131178.12072815.1	100	79	0	8.24E-33	147
86	SRA:SRR2131178.12042779.1	100	79	0	8.24E-33	147
87	SRA:SRR2131178.12040880.1	100	79	0	8.24E-33	147
88	SRA:SRR2131178.12034203.1	100	79	0	8.24E-33	147
89	SRA:SRR2131178.12015814.1	100	79	0	8.24E-33	147
90	SRA:SRR2131178.12009322.1	100	79	0	8.24E-33	147
91	SRA:SRR2131178.12004244.1	100	79	0	8.24E-33	147
92	SRA:SRR2131178.12000856.1	100	79	0	8.24E-33	147
93	SRA:SRR2131178.11981265.1	100	79	0	8.24E-33	147
94	SRA:SRR2131178.11953564.1	100	79	0	8.24E-33	147
95	SRA:SRR2131178.11939725.1	100	79	0	8.24E-33	147
96	SRA:SRR2131178.11933379.1	100	79	0	8.24E-33	147
97	SRA:SRR2131178.11931573.1	100	79	0	8.24E-33	147
98	SRA:SRR2131178.11924289.1	100	79	0	8.24E-33	147
99	SRA:SRR2131178.11911044.1	100	79	0	8.24E-33	147
100	SRA:SRR2131178.11907480.1	100	79	0	8.24E-33	147

BoeLCY

1,217 hits	Subject acc. Ver.	%Identidad	Aligment length	Mismatches	E. value	Bit score
1	SRA:SRR2131178.13943249.1	100	79	0	7.99E-33	147
2	SRA:SRR2131178.13892796.1	100	79	0	7.99E-33	147
3	SRA:SRR2131178.13843431.1	100	79	0	7.99E-33	147
4	SRA:SRR2131178.13804079.1	100	79	0	7.99E-33	147
5	SRA:SRR2131178.13773286.1	100	79	0	7.99E-33	147
6	SRA:SRR2131178.13672113.1	100	79	0	7.99E-33	147
7	SRA:SRR2131178.13651790.1	100	79	0	7.99E-33	147
8	SRA:SRR2131178.13608531.1	100	79	0	7.99E-33	147
9	SRA:SRR2131178.13562619.2	100	79	0	7.99E-33	147
10	SRA:SRR2131178.13562619.1	100	79	0	7.99E-33	147
11	SRA:SRR2131178.13385727.1	100	79	0	7.99E-33	147
12	SRA:SRR2131178.13345746.1	100	79	0	7.99E-33	147
13	SRA:SRR2131178.13309825.1	100	79	0	7.99E-33	147
14	SRA:SRR2131178.13248388.1	100	79	0	7.99E-33	147
15	SRA:SRR2131178.13101985.1	100	79	0	7.99E-33	147
16	SRA:SRR2131178.13098726.1	100	79	0	7.99E-33	147
17	SRA:SRR2131178.13053724.1	100	79	0	7.99E-33	147
18	SRA:SRR2131178.12869575.1	100	79	0	7.99E-33	147
19	SRA:SRR2131178.12833682.1	100	79	0	7.99E-33	147
20	SRA:SRR2131178.12667549.1	100	79	0	7.99E-33	147
21	SRA:SRR2131178.12613597.2	100	79	0	7.99E-33	147

22	SRA:SRR2131178.12450115.1	100	79	0	7.99E-33	147
23	SRA:SRR2131178.12218820.1	100	79	0	7.99E-33	147
24	SRA:SRR2131178.12076766.1	100	79	0	7.99E-33	147
25	SRA:SRR2131178.12015441.1	100	79	0	7.99E-33	147
26	SRA:SRR2131178.11985970.1	100	79	0	7.99E-33	147
27	SRA:SRR2131178.11969739.2	100	79	0	7.99E-33	147
28	SRA:SRR2131178.11969739.1	100	79	0	7.99E-33	147
26	SRA:SRR2131178.11759711.1	100	79	0	7.99E-33	147
27	SRA:SRR2131178.11600578.1	100	79	0	7.99E-33	147
28	SRA:SRR2131178.11536728.1	100	79	0	7.99E-33	147
29	SRA:SRR2131178.11480137.1	100	79	0	7.99E-33	147
30	SRA:SRR2131178.11395278.1	100	79	0	7.99E-33	147
31	SRA:SRR2131178.11382669.1	100	79	0	7.99E-33	147
32	SRA:SRR2131178.11348499.1	100	79	0	7.99E-33	147
33	SRA:SRR2131178.11253632.1	100	79	0	7.99E-33	147
34	SRA:SRR2131178.11178279.1	100	79	0	7.99E-33	147
35	SRA:SRR2131178.11144343.1	100	79	0	7.99E-33	147
36	SRA:SRR2131178.11134165.1	100	79	0	7.99E-33	147
37	SRA:SRR2131178.11121046.1	100	79	0	7.99E-33	147
38	SRA:SRR2131178.11106912.1	100	79	0	7.99E-33	147
39	SRA:SRR2131178.11103806.1	100	79	0	7.99E-33	147
40	SRA:SRR2131178.11042593.1	100	79	0	7.99E-33	147
41	SRA:SRR2131178.10906488.2	100	79	0	7.99E-33	147
42	SRA:SRR2131178.10890926.1	100	79	0	7.99E-33	147
43	SRA:SRR2131178.10875400.1	100	79	0	7.99E-33	147
44	SRA:SRR2131178.10836810.2	100	79	0	7.99E-33	147
45	SRA:SRR2131178.10814602.1	100	79	0	7.99E-33	147
46	SRA:SRR2131178.10672368.1	100	79	0	7.99E-33	147
47	SRA:SRR2131178.10635520.1	100	79	0	7.99E-33	147
48	SRA:SRR2131178.10551937.1	100	79	0	7.99E-33	147
49	SRA:SRR2131178.10457036.1	100	79	0	7.99E-33	147
50	SRA:SRR2131178.10449330.1	100	79	0	7.99E-33	147
51	SRA:SRR2131178.10048750.1	100	79	0	7.99E-33	147
52	SRA:SRR2131178.9949707.1	100	79	0	7.99E-33	147
53	SRA:SRR2131178.9878487.1	100	79	0	7.99E-33	147
54	SRA:SRR2131178.9845119.1	100	79	0	7.99E-33	147
55	SRA:SRR2131178.9817975.2	100	79	0	7.99E-33	147
56	SRA:SRR2131178.9817975.1	100	79	0	7.99E-33	147
57	SRA:SRR2131178.9787886.1	100	79	0	7.99E-33	147
58	SRA:SRR2131178.9782517.1	100	79	0	7.99E-33	147
59	SRA:SRR2131178.9732612.1	100	79	0	7.99E-33	147
60	SRA:SRR2131178.9697332.1	100	79	0	7.99E-33	147

61	SRA:SRR2131178.9614710.1	100	79	0	7.99E-33	147
62	SRA:SRR2131178.9509982.2	100	79	0	7.99E-33	147
63	SRA:SRR2131178.8427700.1	100	79	0	7.99E-33	147
64	SRA:SRR2131178.8312021.2	100	79	0	7.99E-33	147
65	SRA:SRR2131178.7731733.2	100	79	0	7.99E-33	147
66	SRA:SRR2131178.6185319.1	100	79	0	7.99E-33	147
67	SRA:SRR2131178.6108457.1	100	79	0	7.99E-33	147
68	SRA:SRR2131178.5204159.1	100	79	0	7.99E-33	147
69	SRA:SRR2131178.4609751.1	100	79	0	7.99E-33	147
70	SRA:SRR2131178.3019632.1	100	79	0	7.99E-33	147
71	SRA:SRR2131178.13651790.2	100	77	0	1.03E-31	143
72	SRA:SRR2131178.13565680.1	100	76	0	3.72E-31	141
73	SRA:SRR2131178.12906667.1	98.734	79	1	3.72E-31	141
74	SRA:SRR2131178.11467222.1	98.734	79	1	3.72E-31	141
75	SRA:SRR2131178.11252309.1	98.734	79	1	3.72E-31	141
76	SRA:SRR2131178.11249520.1	98.734	79	1	3.72E-31	141
77	SRA:SRR2131178.10815731.2	100	76	0	3.72E-31	141
78	SRA:SRR2131178.9663592.2	100	76	0	3.72E-31	141
79	SRA:SRR2131178.9663592.1	98.734	79	1	3.72E-31	141
80	SRA:SRR2131178.9632537.1	98.734	79	1	3.72E-31	141
81	SRA:SRR2131178.8312021.1	98.734	79	1	3.72E-31	141
82	SRA:SRR2131178.7731733.1	98.734	79	1	3.72E-31	141
83	SRA:SRR2131178.13943249.2	98.734	79	0	1.34E-30	139
84	SRA:SRR2131178.13672113.2	98.734	79	0	1.34E-30	139
85	SRA:SRR2131178.13289411.2	98.734	79	0	1.34E-30	139
86	SRA:SRR2131178.11480137.2	98.734	79	0	1.34E-30	139
87	SRA:SRR2131178.11106912.2	98.734	79	0	1.34E-30	139
88	SRA:SRR2131178.9846501.2	98.734	79	0	1.34E-30	139
89	SRA:SRR2131178.9788032.1	100	75	0	1.34E-30	139
90	SRA:SRR2131178.13257665.1	98.701	77	1	4.81E-30	137
91	SRA:SRR2131178.13077344.1	100	74	0	4.81E-30	137
92	SRA:SRR2131178.12818258.2	100	74	0	4.81E-30	137
93	SRA:SRR2131178.12818258.1	98.701	77	1	4.81E-30	137
94	SRA:SRR2131178.11702629.1	100	74	0	4.81E-30	137
95	SRA:SRR2131178.11252309.2	98.718	78	0	4.81E-30	137
96	SRA:SRR2131178.10815731.1	100	74	0	4.81E-30	137
97	SRA:SRR2131178.9676645.2	100	74	0	4.81E-30	137
98	SRA:SRR2131178.13776618.1	100	73	0	1.73E-29	135
99	SRA:SRR2131178.13082718.2	98.684	76	1	1.73E-29	135
100	SRA:SRR2131178.9846501.1	97.468	79	2	1.73E-29	135