



Centro de Investigación Científica de Yucatán, A.C.
Posgrado en Ciencias Biológicas

**CARACTERIZACIÓN MOLECULAR DE GENES
HOMÓLOGOS A FACTORES DE TRANSCRIPCIÓN
DE LA FAMILIA ERF EN *Carica papaya***

Tesis que presenta
MIGUEL ÁNGEL VALLEJO REYNA

En opción al título de
MAESTRO EN CIENCIAS BIOLÓGICAS
(Opción Biotecnología)

Mérida, Yucatán. Enero de 2011





RECONOCIMIENTO



Por medio de la presente, hago constar que el trabajo de tesis titulado “Caracterización molecular de genes homólogos a factores de transcripción de la familia ERF en *Carica papaya*” fue realizado en los laboratorios de la Unidad de Biotecnología del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del Dr. Santy Peraza Echeverría, dentro de la Opción Maestría en Biotecnología, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

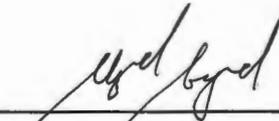
Para los efectos que sean necesarios,

Dr. Oscar Alberto Moreno Valenzuela
Director Académico
Centro de Investigación Científica de Yucatán A.C.

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 dicha información le pertenece en términos de la Ley de la Propiedad Industrial, por lo que no me reservo ningún derecho sobre ello.

Mérida, Yucatán a 10 de enero de 2011



Miguel Ángel Vallejo Reyna

AGRADECIMIENTOS:

A mi asesor Dr. Santy Peraza Echeverría por la oportunidad de desarrollar esta investigación y sobre todo por la confianza brindada a lo largo de la realización de este trabajo. Quiero agradecer el hecho de escuchar y discutir mis opiniones brindándome un trato de colega y amigo más allá de una relación entre asesor y alumno, así como el tiempo, las enseñanzas y consejos que permitieron alcanzar mi objetivo personal de ser mejor profesional del que era al inicio de este posgrado.

Al comité tutorial integrado por el Dr. Santy Peraza Echeverría, el Dr. Jorge Santamaría Fernández y el Dr. Víctor Suárez Solís por los comentarios, sugerencias y apoyo durante los exámenes tutoriales.

Al comité revisor integrado por la Dra Virginia Herrera Valencia, el Dr. Santy Peraza Echeverría, el Dr. Jorge Santamaría Fernández, el Dr. Luis Sáenz Carbonell y el Dr. Víctor Suárez Solís. Por la revisión, comentarios y sugerencias para llevar a término la escritura de la presente tesis.

A la Unidad de Biotecnología y al laboratorio de Fitopatología Molecular y Biotecnología de Cultivos Tropicales (clave BT04) por el espacio y equipos proporcionados.

Al Centro de Investigación Científica de Yucatán A.C., por las facilidades para llevar a cabo este proyecto.

Al Consejo Nacional de Ciencia y Tecnología por el financiamiento de la beca de maestría No. 228306

Al Consejo Nacional de Ciencia y Tecnología por el apoyo al proyecto de Ciencia Básica clave 83829.

A la Dra. Gabriela Fuentes Ortiz y el Tec. Eduardo Castillo Castro por las recomendaciones y consejos durante el desarrollo del trabajo.

A la Lic. Gilma Michell por el apoyo y la confianza al ingreso de este posgrado.

A los compañeros del CICY por compartir conmigo su tiempo y espacio durante el transcurso de estos dos años.

Dedicada:

A mis padres Don Chava y Doña Paz por inculcarme el estudio más allá de una obligación.

A mis hermanos Alberto, Ernesto, Joel y Max.

A Yajima por acompañarme en la vida.

Y en especial a mi hermana Claudia por todo el apoyo que me has brindado para ser el biólogo que ahora soy.

Nunca consideres el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber.

Albert Einstein

CONTENIDO

RESUMEN	1
ABSTRACT	3
INTRODUCCIÓN	5
CAPÍTULO I	7
1.1 ANTECEDENTES GENERALES	7
1.1.1 Origen e historia de la papaya	7
1.1.2 Biología y clasificación taxonómica de papaya	8
1.1.3 La papaya como cultivo agrícola	10
1.1.4 Producción de papaya a nivel mundial	14
1.1.5 Producción de papaya en México	15
1.1.6 Problemáticas de la Producción	17
1.1.7 Biotecnología aplicada en papaya	18
1.1.8 Mecanismos moleculares de resistencia de las plantas en respuesta al ataque de patógenos.	20
1.1.9 Rutas de señalización en los mecanismos de defensa de plantas	23
1.1.10 Factores de transcripción y su participación en la resistencia a patógenos	25
1.1.11 Sobreexpresión de Factores de transcripción ERF en plantas transgénicas y su efecto en la resistencia contra patógenos.	34
1.2 HIPÓTESIS	38
1.3 OBJETIVOS	38
1.3.1 General	38
1.3.2 Particulares	38
1.4 ESTRATEGIA EXPERIMENTAL	39
BIBLIOGRAFÍA	40
CAPÍTULO II	49
CARACTERIZACIÓN <i>IN SILICO</i> DE GENES HOMÓLOGOS A FACTORES DE TRANSCRIPCIÓN DE LA FAMILIA ERF EN EL GENOMA SECUENCIADO DE LA PAPAYA TRANSGÉNICA "SUNUP"	49
2.1 INTRODUCCIÓN	49
2.2 MATERIALES Y MÉTODOS	52
2.2.1 Identificación de secuencias homólogas a FT-ERF en papaya.	52
2.2.2 Obtención de los marcos de lectura abierta y medición del porcentaje de identidad	52
2.2.3 Alineamiento múltiple de secuencias	53
2.2.4 Modelaje por homología de la estructura tridimensional de las secuencias tipo ERF de papaya	53
2.2.5 Análisis filogenético.	54
2.2.6 Predicción de elementos regulatorios en cis de las regiones promotoras putativas de las secuencias tipo ERF de papaya.	55
2.3 RESULTADOS	56
2.3.1 Identificación de secuencias homólogas a FT-ERF en papaya.	56
2.3.2 Marcos de lectura abierta y porcentaje de identidad	59

2.3.4 Alineamiento múltiple clustal	65
2.3.5 Estructura tridimensional de las Proteínas CpERF1-CpERF5	71
2.3.6 Análisis filogenético.	71
2.3.7 Elementos Regulatorios en cis predichos en las regiones promotoras de los genes CpERF1-CpERF5	74
2.4 DISCUSIÓN	75
2.4.1 CpERF1: un gen homólogo a Pti4 posiblemente implicado en la interacción con una proteína R del tipo PTO	76
2.4.2 El gen CpERF2 presenta homología con TSI1	77
2.4.3 El gen CpERF3 homólogo al grupo VII y su posible interacción con proteínas cinasas en la generación de resistencia a estrés biótico y abiótico	78
2.4.4 CpERF4 el homólogo a ERF1 de <i>Arabidopsis</i> podría estar implicado en la ruta de señalización del etileno.	79
2.4.5 CpERF5 y su homólogo JcERF son miembros de la subfamilia DREB que podrían conferir resistencia a estrés biótico.	81
2.4.6 Posible regulación de la expresión de los factores CpERF1-5	82
2.5 CONCLUSIONES	84
2.5 BIBLIOGRAFÍA	85
CAPÍTULO III	93
ANÁLISIS DE LA EXPRESIÓN DE LOS GENES CpERF1-CpERF5 Y MAPEO DE LOS EXTREMOS 5' Y 3' DE SUS CORRESPONDIENTES ADNc EN PAPAYA MARADOL	93
3.1 INTRODUCCIÓN	93
3.2 MATERIALES Y MÉTODOS	95
3.2.1 Material vegetal	95
3.2.2 Extracción de ácido ribonucleico	95
3.2.3 Cuantificación de ácidos nucleicos mediante espectrofotometría	96
3.2.4 Síntesis de ADN complementario	96
3.2.5 Transcripción Reversa y Reacción en Cadena de la Polimerasa (RT-PCR)	97
3.2.6 Mapeo de los extremos terminales del ADNc mediante la técnica de amplificación de los extremos terminales del ADNc (RACE)	97
3.2.7 Clonación de los extremos 5' y 3' terminales del ADNc en el vector pGEMT-easy	99
3.2.8 Secuenciación.	101
3.2.9 Análisis Bioinformático	102
3.3 RESULTADOS	103
3.3.1 Aislamiento de ARN total.	103
3.3.2 Expresión Basal de los genes CpERF1-CpERF5 en diferentes tejidos de una planta de papaya Maradol.	105
3.3.3 Amplificación de los extremos 5' y 3' de los ADNc de CpERF1-CpERF5	106
3.3.4 Clonación de los productos del RACE de CPERF1-CPERF5	107
3.3.5 Mapeo de los extremos 5' y 3' de los ADNc de CpERF1-CpERF5	109
3.4 DISCUSIÓN	115
3.4.1 Expresión basal y espacial de los genes CpERF1-CpERF5	115

3.4.2 Mapeo de los extremos terminales del ADNc de CpERF1- CpERF5	116
3.5 CONCLUSIÓN	120
3.6 BIBLIOGRAFIA	121
CAPITULO IV	125
4.1 DISCUSIÓN GENERAL	125
4.2 CONCLUSIÓN	129
4.3 PERSPECTIVAS	130
4.4 BIBLIOGRAFIA	132

ÍNDICE DE FIGURAS

Figura 1.1 Árbol de papaya	9
Figura 1.2 Plantas de papaya en una plantación comercial.....	10
Figura 1.3 Valor de exportación de papaya a nivel mundial.....	15
Figura 1.4 Porcentaje de producción de papaya en la República Mexicana	16
Figura 1.5 Producción y exportación de papaya en México	17
Figura 1.6 Modelo de evolución de la resistencia a microorganismos en plantas	22
Figura 1.7 Co-evolución entre los genes R y elicitores.	23
Figura 1.8 Interacción entre rutas de señalización reguladas por fitohormonas durante la respuesta de las plantas al ataque de patógenos.....	25
Figura 1.9 Estructura tridimensional del dominio AP2/ERF.....	29
Figura 1.10 Representación esquemática de la interacción entre el dominio AP2/ERF y el elemento “GCC box”	30
Figura 1.11 . Modelo de evolución del dominio de unión a ADN AP2/ERF en plantas.....	31
Figura 1.12 Árbol filogenético de los miembros de la familia ERF de <i>Arabidopsis</i> <i>thaliana</i>	33
Figura 1.13 Sobreexpresión de factores de transcripción tipo ERF en plantas	37
Figura 2.1 Marco de lectura abierta de la secuencia CpERF1.....	60
Figura 2.2 Marco de lectura abierta de la secuencia CpERF2.....	61
Figura 2.3 Marco de lectura abierta de la secuencia CpERF3.....	62
Figura 2.4 Marco de lectura abierta de la secuencia CpERF4.....	63
Figura 2.5 Marco de lectura abierta de la secuencia CpERF5.....	64
Figura 2.6 Alineamiento clustal X del grupo IXa.	66
Figura 2.7 Alineamiento clustal X del grupo VI.....	67
Figura 2.8 Alineamiento clustal X del grupo VII	69
Figura 2.9 Alineamiento clustal X del grupo IXc	69
Figura 2.10 Alineamiento clustal X del grupo I.....	70

Figura 2.11 Análisis de la estructura tridimensional de los dominios AP2/ERF.....	72
Figura 2.12 Análisis filogenético realizado de las secuencias tipo ERF.....	73
Figura 2.13 Elementos regulatorios de los genes CpERF1-CpERF5.....	74
Figura 3.1 Estructura típica de un ARN mensajero	98
Figura 3.2 Mapa del vector de clonación pGEM [®] -T Easy.....	99
Figura 3.3 Electroforesis de ARN total	104
Figura 3.4 Amplificación por medio de RT-PCR de los genes CpERF1-CpERF5.....	105
Figura 3.5 Productos del 5' RACE de los factores de transcripción ERF de papaya	106
Figura 3.6 Productos del 3' RACE de los factores de transcripción ERF de papaya	107
Figura 3.7 Digestión del vector pGEM-T Easy con los productos del 5'RACE	108
Figura 3.8 Digestión del vector pGEM-T Easy con los productos del 3'RACE	108
Figura 3.9 Alineamiento del ORF de CpERF1 predicho <i>in silico</i> de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol.	110
Figura 3.10 Alineamiento del ORF de CpERF2 predicho <i>in silico</i> de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol.	111
Figura 3.11 Alineamiento del ORF de CpERF3 predicho <i>in silico</i> de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol.	112
Figura 3.12 Alineamiento del ORF de CpERF4 predicho <i>in silico</i> de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol.	113
Figura 3.13 Alineamiento del ORF de CpERF5 predicho <i>in silico</i> de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol.	113

ÍNDICE DE CUADROS

Cuadro 1.1 Factores de transcripción asociados con la defensa a patógenos en <i>Arabidopsis thaliana</i> (Eulgem, 2005)	27
Cuadro 1.2 FT-ERF que han generado resistencia a patógenos cuando son sobreexpresados en plantas de <i>Arabidopsis</i> o tabaco.....	36
Cuadro 2.1 Resultados del análisis TBLASTN para Pti4.....	57
Cuadro 2.2 Resultados del análisis TBLASTN para TSI1	57
Cuadro 2.3 Resultados del análisis TBLASTN para GmERF3.....	58
Cuadro 2.4 Resultados del análisis TBLASTN para ERF1.....	58
Cuadro 2.5 Resultados del análisis TBLASTN para JcERF	58
Cuadro 2.6 Porcentajes de identidad entre secuencias tipo ERF de papaya y las secuencias Pti4, TSI1, GmERF3, ERF1 y JcERF.....	65
Cuadro 3.1 Oligos específicos para la amplificación de fragmentos de ADNc de los genes de papaya homólogos a factores de transcripción tipo ERF.	97
Cuadro 3.2 Oligos específicos para la PCR anidada 5' RACE (Antisentido 5' RACE) y 3' RACE (Sentido 3' RACE) de CpERF1-CpERF5.....	99
Cuadro 3.3 Rendimiento y grado de pureza de los extractos de ARN total de diferentes tejidos de papaya	104

RESUMEN

La papaya es una planta tropical de alto valor nutricional, siendo México el primer exportador mundial de esta fruta. No obstante, al igual que otros cultivos tropicales, la papaya es susceptible al ataque de patógenos, tal es el caso de *Colletotrichum gloeosporioides*, agente causal de antracnosis y principal causa de pérdidas a nivel post-cosecha. La disponibilidad pública de la secuencia del genoma de papaya ha abierto la posibilidad de acelerar el descubrimiento de genes involucrados en la resistencia a patógenos y al consecuente mejoramiento genético de este cultivo. Entre los genes envueltos en el sistema inmune de plantas, diversos factores de transcripción (FT) tienen un papel importante en la regulación positiva de la expresión los genes de defensa conocidos como genes relacionados con patogénesis (PR). Los factores de respuesta a etileno (ERF por sus siglas en inglés) son FT presentes en diferentes especies de plantas que regulan de manera positiva la expresión de genes de defensa en respuesta a la acumulación de diferentes fitohormonas relacionadas con la resistencia a enfermedades como son el ácido salicílico, ácido jasmónico y etileno. Los FT-ERF cuentan con dominio de unión a ADN llamado AP2/ERF, el cuál es único de plantas. Interesantemente, la sobreexpresión de los genes ERF en *Arabidopsis* y tabaco ha mostrado un incremento en la resistencia contra una amplia gama de patógenos. El objetivo del presente trabajo fue caracterizar la estructura de genes homólogos a la familia ERF en el genoma de papaya y evaluar sus patrones de expresión. Los resultados permitieron identificar y caracterizar cinco secuencias que presentan homología con secuencias ERF previamente caracterizadas en otras especies de plantas. Estas secuencias nombradas CpERF1-CpERF5 y su expresión basal fueron analizadas en cuatro diferentes tejidos de plantas adultas comerciales de papaya Maradol. Finalmente, los extremos 5' y 3' de los ADNc de estas cinco secuencias tipo ERF de papaya fueron mapeados y comparados con sus correspondientes secuencias *in silico*. Las secuencias CpERF aisladas en este estudio son secuencias de gran valor que podrían ayudar en conferir resistencia a papaya.

ABSTRACT

Papaya is a tropical plant of high nutritional value and Mexico is the leading exporter of this tropical fruit. However, like other tropical crops, papaya is susceptible to the attack by pathogens, such as the fungus *Colletotrichum gloeosporioides*, causal agent of anthracnose disease, which causes serious losses in the post-harvesting stage. The public availability of the papaya genome sequence has opened the possibility of accelerating the discovery of genes involved in disease resistance and also the molecular breeding of this crop. Among the genes involved in plant immunity, several transcription factors (TF) play a major role in up-regulating the expression of defense genes known as pathogenesis-related genes (PR). The Ethylene Response Factors (ERF) are TF present in different plant species that up-regulate the expression of PR genes in response to the accumulation of plant hormones involved in disease resistance such as salicylic acid, jasmonic acid or ethylene. The ERF transcription factors possess a DNA binding domain known as AP2/ERF, which is unique in plants. Interestingly, the overexpression of ERF genes in *Arabidopsis* and tobacco has been shown to increase resistance against a broad range of pathogens. The aim of this study was to characterize the structure of ERF gene homologues in the papaya genome sequence and evaluate their expression patterns. We identified five sequences that showed homology to previously characterized ERF sequences from other plants. These sequences were named as CpERF1-CpERF5 and their basal expression was observed in four different tissues from adult commercial plants of papaya Maradol. Finally, the 5' and 3' ends of the five papaya ERF cDNA sequences were mapped and compared to their corresponding *in silico* sequences. The CpERF sequences isolated in this study represent a valuable sequence resource that could assist in the development of disease resistance in papaya.

INTRODUCCIÓN

La papaya (*Carica papaya* Linnaeus) es una planta arborescente tropical originaria de Mesoamérica y ampliamente cultivada en las zonas tropicales del continente americano, así como en África y Asia. De manera silvestre, la papaya se puede encontrar en México en los estados de Veracruz y Yucatán (Claridades Agropecuarias, 1999). Es una especie que tiene múltiples usos, desde ornamental, artesanal, forrajero, cosmético, insecticida hasta medicinal, sin embargo su principal uso es la comercialización y consumo del fruto fresco. La papaya es una fruta con gran cantidad de vitamina A, vitamina C, potasio, folato, niacina, tiamina, riboflavina, hierro, calcio y fibra, por lo cual es altamente recomendado su consumo (Claridades Agropecuarias, 1999; Vázquez-Yanes *et al.*, 1999).

La papaya al igual que otros frutos tropicales, es susceptible al ataque de distintos patógenos que comprometen la producción. Tal es el caso del hongo *Colletotrichum gloeosporioides*, agente causal de la enfermedad antracnosis que ocasiona manchas necróticas en el fruto ocasionando pudrición del mismo a nivel post cosecha, provocando así graves problemas en su comercialización (Zavala-Leon *et al.*, 2005). Por lo anterior, el cultivo de papaya generalmente suele estar acompañado por un uso intensivo de pesticidas; de igual manera se requiere tratamiento frecuente de fertilizantes químicos, a fin que el fruto se desarrolle de acuerdo a los requerimientos del mercado (Plan Rector Sistema Nacional Papaya, 2006). Para lo anterior, el uso de la tecnología de ADN recombinante es una herramienta valiosa que nos puede ayudar a encontrar soluciones a los daños causados por estrés biótico, principalmente la antracnosis que causa grandes pérdidas económicas en este fruto (Massieu-Trigo, 2009).

La secuenciación de genomas completos como el de *Arabidopsis* ofrecen una gran ayuda en la identificación de genes que participen en las rutas bioquímicas del sistema inmune de plantas y que puedan ser de utilidad para el desarrollo biotecnológico de diferentes variedades. El genoma de *Carica papaya* Linnaeus var. SunUp ha sido recientemente secuenciado (Ming *et al.*, 2008), y su disponibilidad al público facilita la búsqueda de genes involucrados en la resistencia a patógenos.

Por su parte, los factores de transcripción (FT) son proteínas que actúan como reguladores maestros de la transcripción de genes, por lo que son excelentes candidatos para utilizarlos en la regulación de mecanismos que le permitan a la planta resistir el ataque de patógenos (Century *et al.*, 2008).

Se han descrito varios factores de transcripción en *Arabidopsis* y otras plantas, dentro de esta gama se encuentran los Factores de Respuesta a Etileno (ERF por sus siglas en inglés), que son factores de transcripción altamente conservados, presentes en plantas de diferentes taxa filogenéticos. Dentro de las características de este tipo de factores de transcripción está la habilidad de activarse bajo la presencia de factores de estrés, tanto bióticos como abióticos promoviendo de esta forma la activación de la expresión de distintos genes, incluyendo aquellos que tienen que ver con la respuesta a patogénesis (genes PR). Se han estudiado y caracterizado estos factores en diversas plantas y se ha comprobado que algunos de ellos al ser sobreexpresados en plantas modelo activan genes PR que confieren tolerancia a diversos patógenos (Gutterson y Reuber, 2004).

En el presente estudio se utilizaron programas bioinformáticos para identificar genes homólogos a factores de transcripción de la familia ERF en el genoma de la papaya transgénica SunUp (Ming *et al.* 2008). Una vez identificados dichos genes, se aislaron y evaluaron sus patrones de expresión en individuos de una plantación comercial de papaya Maradol, también se mapearon y caracterizaron los extremos 5' y 3' de sus ADN complementarios (ADNc).

CAPÍTULO I

1.1 ANTECEDENTES GENERALES

1.1.1 Origen e Historia de la papaya

El fruto de la papaya (*Carica papaya* Linnaeus) se obtiene del árbol conocido como papayo, originario de las zonas tropicales de México y Centroamérica. Se cultiva en terrenos de muy distinta naturaleza, pero es fundamental que éstos sean ricos en materia orgánica y que contengan una humedad abundante. La historia de la papaya aparece por primera vez documentada por el cronista español Oviedo, director de “*Mines in Hispaniola*” en las Antillas Españolas, quien la encontró en las costas de Panamá (Oviedo, 1851). Ésta fue cultivada rápidamente en los trópicos, ya que su distribución, indudablemente fue auxiliada por la abundancia de semillas de amplia viabilidad. Este fruto se ha adaptado en diversas regiones tropicales, particularmente en áreas con suelos fértiles y lluvia abundante. La historia de la dispersión de la papaya, relatada por Oviedo describe como Alfonso de Valverde tomó semillas de papaya de las costas de Panamá hacia Darien y luego a Santo Domingo y otras islas de las indias Occidentales (Oviedo, 1851). Los españoles le dieron el nombre de Papaya y llevaron la planta hacia las Filipinas donde se expandió hacia Malasia y finalmente a la India en 1598 (Schery, 1956). Para 1600 aproximadamente, la papaya se cultivaba en regiones cálidas de Sur y Centro América, Sur de México, las Antillas, Bahamas y Bermudas. En ese siglo la semilla de papaya fue transportada de la India a Nápoles, Italia. A Hawai llegó entre 1800 y 1820. Con el tiempo la papaya se estableció en Uganda en 1874 y su distribución se dispersó a través de la mayor parte de los países tropicales y subtropicales (Teixeira da Silva, *et al* 2007). La semilla de papaya fue llevada a Florida hasta 1900, probablemente de las Bahamas. La variedad “*Solo*” se cultiva en Hawai desde 1911, originaria de Barbados y Jamaica. En 1978 se introdujeron a México las primeras semillas de la variedad *Maradol*, a través de la CONAFRUT, en Xalapa, Ver. En la actualidad el papayo se cultiva en forma comercial no solo en las regiones de América, sino también en África, Asia, Australia, Filipinas, y en los Estados Unidos (Hawai y Florida) (Claridades Agropecuarias, 1999; Teixeira da Silva, *et al.*, 2007).

1.1.2 Biología y clasificación taxonómica de papaya

Papaya es una planta arborescente perennifolia de rápido crecimiento, de 2 a 10 m de altura con un diámetro a la altura del pecho de 6 a 30 cm con un olor acre distintivo. Su copa es abierta y redondeada, con hojas grandes de pecíolo largo, de 0.7 a 1 m, con la lámina palmeada de 7 a 9 lóbulos, y éstos a su vez en lóbulos más pequeños, ligeramente gruesos y carnosos; las hojas superiores son erectas y extendidas y las inferiores son colgantes. El tronco es erguido, cilíndrico, hueco excepto en los nudos, más grueso en su base; sin ramas y con las características cicatrices que dejan las hojas al caer. Crecimiento monopódico cuando joven, exhibe una fuerte dominancia apical solamente se ramifica cuando se daña o se corta el meristemo apical y al madurar. La corteza es lisa, verde grisácea, con manchas pardas, oscuras, o bien raramente pardo pálidas, de forma irregular, lenticelas pequeñas o ausentes, cicatrices semicirculares a todo lo largo del tronco. Exudado blanco. Las flores son pistiladas, estaminadas y bisexuales, con el cáliz tubular de 8 a 10 mm de largo, verdoso; corola tubular de 10 a 20 mm de largo, blancuzca o amarilla pálida. Flores femeninas solitarias, 5 ó 6 juntas en la base de una hoja; masculinas en panículas delgadas con 15 a 20 flores o llegando a tener hasta 100 florecillas por inflorescencia. Las flores femeninas son mucho más grandes que las masculinas. Los frutos apiñados alrededor del tronco. Bayas elipsoides a esféricas, tornándose de verdes a anaranjadas en la madurez, pulpa blanda, jugo lechoso. El fruto silvestre mide de 4 a 6 cm de largo y de 3 a 4.5 cm de ancho. Cada fruto conteniendo de 200 a 400 semillas. El fruto cultivado mide de 10 a 50 cm de largo, dependiendo del cultivo. Las semillas de 3.7 a 4.5 mm de largo por 2 a 2.8 mm de ancho y 2 a 2.5 mm de grueso, esféricas, cubiertas por una capa mucilaginosa (sarcotesta); endotesta pardo negruzca y arrugada. Tiene un sistema radical pivotante. La planta puede ser dioica (más comúnmente en la papaya silvestre), monoica, hermafrodita o polígama. Ocurren cambios en la expresión sexual debido a diferentes condiciones ecológicas y otras variables. El sexo de la planta no se puede determinar sino hasta la floración (Vázquez-Yanes *et al.*, 1999).

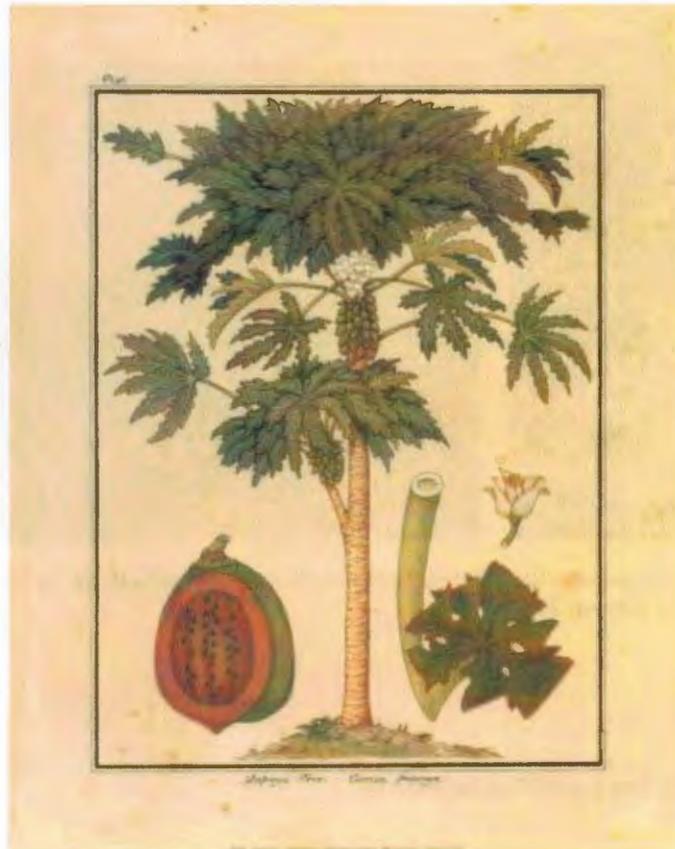


Figura 1.1 Árbol de Papaya (<http://www.keyposters.com/poster/141433.html>)

La papaya pertenece a la Familia Caricaceae, una familia que incorpora 35 especies que contienen latex, agrupados en cuatro géneros (*Carica*, *Cylicomorpha*, *Jarilla* y *Jacaratia*), de los cuales el más importante es *Carica* (Teixeira da Silva, *et al.*, 2007).

El género *Carica* se constituye de 22 especies y es el único miembro de las Caricaceae que es cultivada como árbol frutal, mientras que los otros géneros se cultivan como ornamentales (Kumar y Srinivasan, 1944). Ecológicamente hablando es una especie de la vegetación primaria y secundaria que puede ser usada con fines de restauración en terrenos degradados y ha sido utilizada para rehabilitar sitios donde hubo explotación minera (Vázquez-Yanes *et al.*, 1999)



Figura 1.2 Plantas de papaya en una plantación comercial del estado de Yucatán (Foto: Miguel Ángel Vallejo-Reyna, 2009).

1.1.3 La papaya como cultivo agrícola

La papaya (*Carica papaya* L.) es una fruta tropical muy apetecible por su suave y agradable sabor, aunado a sus propiedades nutritivas y digestivas tiene un alto impacto socio-económico debido a que es cultivada en zonas marginales. (Flores-Revilla *et al.* 1995; Massieu-Trigo, 2009). Sin embargo, las grandes plantaciones comerciales están enfocadas a productores de gran escala, o productores asociados, pues se necesitan cerca de \$100 mil pesos para el establecimiento de una hectárea. El cultivo es atractivo para los agricultores, ya que cuenta con una rentabilidad alta, pues por cada peso invertido se recuperan \$2.74 pesos (Plan Rector del Sistema Nacional Papaya, 2006), además, mantiene un periodo corto entre siembra y cosecha y se obtiene un alto rendimiento por hectárea (Arango y Roman, 2000).

La papaya puede cultivarse desde el nivel del mar hasta los 1000 m.s.n.m (Vázquez-Yanes, 1999), pero los frutos de mejor calidad y los rendimientos más altos se obtienen en altitudes por debajo de los 800 metros (Plan Rector del Sistema Nacional Papaya,

2006). Se puede cultivar entre latitudes de 32° N y Sur. El rango de temperatura que permite que este frutal se desarrolle es entre 22° y 30°C, pero su óptima es entre 23° y 26°C. Las temperaturas bajas inhiben su crecimiento y alteran el sabor de la fruta mientras que temperaturas muy altas, le provocan abscisión floral y bajas en la producción (Claridades Agropecuarias, 1999). Al ser un cultivo sensible a daño por bajas temperaturas, se limita su almacenamiento a temperaturas menores a 8°C más sin embargo, sin refrigeración, el fruto decolora y deteriora en pocos días como resultado de una rápida maduración e infecciones de microorganismos (Perez-Carrillo y Yahia. 2003).

El agua es fundamental para la papaya, compone alrededor del 85% de su volumen y le afecta tanto su falta como su exceso. La planta requiere agua en la plantación para que enraíce bien y en época de floración para que las flores no sean estériles. Caso contrario, el exceso de agua provoca la muerte de la planta por la aparición de enfermedades de la raíz como la pudrición por hongos (Claridades Agropecuarias, 1999). El agua es el contribuyente principal de la planta; tanto en el proceso de germinación, vivero y primeros meses después de plantada, se necesita para su crecimiento y desarrollo una gran cantidad de agua, por lo cual en esta fase se deben realizar riegos semanales. En la época seca y cuando la lluvia no es adecuada, se debe recurrir al riego para mantener las plantas con un buen desarrollo (Claridades Agropecuarias, 1999).

Las plantaciones de papaya necesitan abundante luz debido a su gran actividad fotosintética. Es imposible desarrollar plantaciones con restricciones de luz, pues las plantas serían alargadas y amarillas, sintomatología esta de desnutrición, lo que trae como consecuencia un inadecuado desarrollo de las plantas. También requieren de suelos con buen drenaje y bien aireados, alto contenido de materia orgánica, un pH que fluctúe entre 5.5 y 6.7, suelos fértiles, profundos y húmedos (Claridades Agropecuarias, 1999).

En países como Cuba, Brasil, Taiwán y otros del África del Sur se ha desarrollado variedades e híbridos que han podido adaptarse a diferentes países de condiciones climáticas similares (Plan Rector del Sistema Nacional Papaya, 2006). En Hawai se desarrolló la variedad "Solo", una variedad de fruto pequeño y dulce que dominó el

mercado hawaiano desde los años 30's hasta finales de los 60's. De esta variedad se originaron otras como el Kapoho (tolerante a sequía) y Waimanalo (más grande, precocidad y tolerancia al quemado de fruta) que dominaron el mercado a partir de los años 70's (Plan Rector del Sistema Nacional Papaya, 2006). En los últimos años se han desarrollado otras variedades hawaianas como "sunset" y sunrise" con pulpa rosada que las hace mas aceptadas en el mercado mundial. Las variedades Rainbow y SunUp fueron desarrolladas exitosamente en Hawaii a mediados de los 90's mediante técnicas biotecnológicas de ADN recombinante para generar individuos tolerantes al Virus de la mancha anular de papaya (PRSV) (Gonsalves, 1998). En Taiwan desarrollaron la variedad Tainung (lugar donde se desarrolló la variedad) que a partir de progenitores de la variedad Cari-flora, le otorgaron características que le permitían resistir al PRSV, al frío.

El agricultor cubano Adolfo Rodríguez Rivera, a través de selecciones realizadas durante 11 años, obtuvo como resultado las variedades Maradol roja y Maradol amarilla, una fruta de sabor dulce y consistencia relativamente dura (Plan Rector del Sistema Nacional, 2006). El nombre Maradol se forma de los nombres María y Adolfo. El señor Adolfo Rodríguez Rivera, compuso el nombre de la variedad juntando el nombre de su esposa y el suyo. Esta variedad de origen cubano, se clasifica por su tamaño en semienana; desarrolla un tronco grueso, exuberante follaje y entrenudos cortos. Se caracteriza por presentar descendencia compuesta por plantas hermafroditas para frutas alargadas y plantas femeninas para frutas redondas. La semilla certificada presenta un 66% de plantas hermafroditas y 33% de plantas hembras. El 1% corresponde a plantas hermafroditas pentandrias, intermedias, estériles y machos. Es una planta de porte bajo con floraciones y fructificaciones tempranas. Con buen manejo, el primer corte se realiza de 130 a 150 días después del trasplante. Esta variedad tiene la desventaja de ser muy susceptible al virus PRSV, pero con manejo de productos de tecnología actualizada, de sistemas de inmunidad y de resistencia sistémica adquirida se puede convivir con el virus y disminuir su efecto total. El color externo es amarillo-naranja-brillante y presenta un intenso color interior rojo-salmón que la hace muy apreciable al consumidor. Su sabor exquisito y la consistencia agradable de su pulpa la distingue del resto con una concentración de 12% brix. En la papaya Maradol certificada prevalecen las frutas alargadas y su peso oscila entre 1.5 a 2.6 kgs (3.3 a 5.7 libras por fruta). Por su

consistencia posee una larga vida de anaquel y resistencia al manejo post-cosecha y transporte (Plan Rector del Sistema Nacional Papaya, 2006).

En México, la falta de semilla mejorada de papayo, y la susceptibilidad a plagas y enfermedades de las variedades e híbridos introducidos motivó a desarrollar el primer híbrido de este frutal, que se ha denominado "Azteca". Este híbrido se adapta a las condiciones ambientales tropicales de México con alta humedad ambiental (60 a 80%), temperaturas promedio de 23 a 26 °C y reúne las características necesarias de calidad comercial que demanda el mercado, como fruto intermedio (de 2 a 3 Kg), aromático, mayor contenido de sólidos solubles, consistencia de la pulpa, mayor resistencia al transporte, y además tolera al PRSV y PapMV (Mirafuentes-Hernández y Azpeitia-Morales, 2008).

Las variedades que han sido cultivadas en México, en forma experimental, son: Maradol Roja, Maradol Amarilla, y algunos cultivares hawaianos como Red Lady y Cariflora entre otros; sin embargo, sólo la Maradol Roja se ha cultivado comercialmente en México desde 1988, y ha ido desplazando paulatinamente a los tipos criollos (Claridades Agropecuarias, 1999; Massieu-Trigo, 2009). Ante la introducción a nuestro país de variedades mejoradas o con ventajas comparativas contra los tipos criollos, es importante la conservación del material genético nativo, ya que las especies silvestres podrían haber desarrollado características de interés para el ser humano y su germoplasma constituye un gran acervo genético para el desarrollo de nuevas variedades que ofrezcan características de interés como mayor productividad y sanidad principalmente y con esto hacer más largo el periodo de vida económica del producto. Esto se podrá lograr en gran medida con trabajos de ingeniería genética, sin embargo la investigación sobre papaya no cuenta con una transferencia de tecnología adecuada, aunque probablemente sea de los cultivos que más la presentan (Claridades Agropecuarias, 1999; Massieu-Trigo, 2009). De hecho la papaya es la primer fruta tropical transgénica en ser comercializada, las variedades Rainbow y SunUp fueron desarrolladas exitosamente en Hawaii para generar individuos tolerantes al PRSV problema que presentaba enormes pérdidas en dicha región; desafortunadamente no son tolerantes a otras cepas del mismo virus por lo que no funcionan contra cepas de otras regiones incluyendo México.

1.1.4 Producción de papaya a nivel mundial

Según datos de la FAO, hasta el año 2008 en el mundo existen 386,379 Ha dedicadas al cultivo de papaya. El 70% de estas tierras se concentran en países en vías de desarrollo en regiones tropicales y subtropicales. Nigeria e India dedican 92,500 y 80,300 Ha respectivamente al cultivo de este producto, le siguen Brasil (36,750 Ha), Bangladesh (19,919 Ha), México (16,084 Ha), República Democrática del Congo (13,500 Ha) y Etiopía (12,500 Ha).

La producción mundial de este cultivo ha aumentado de 7 millones de toneladas en el año 2000 hasta poco más de 9 millones en 2008. En este apartado India y Brasil son los mayores productores (2,685,900 y 1,900,000 T respectivamente) les siguen Nigeria (765,000 T), Indonesia (653,276 T), México (638,237 T), Congo (223,770 T) y Bangladesh (103,609 T).

En las exportaciones, a pesar del área dedicada al cultivo, y la cantidad de toneladas producidas, se presenta México como el principal exportador con 101,306 Toneladas; Belice (33,341 T), Brasil (32,267 T), Malasia (26,938 T) India y China (10,000 T) completan la lista de los principales exportadores. México también cuenta con el mayor valor de exportaciones de papaya (US\$ 55 millones; Figura 1.3) casi el doble de Brasil que ocupa el segundo lugar a pesar de que México solo cuenta con un tercio de la producción total de Brasil. Esto puede deberse a la cercanía que tiene México con Estados Unidos que es el principal importador mundial (138,115 T).

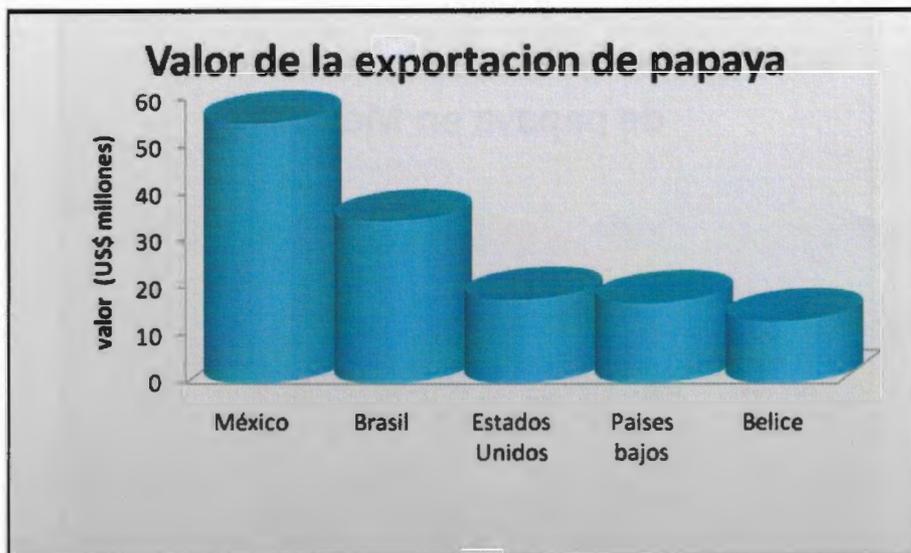


Figura 1.3 Valor de Exportación de Papaya a nivel mundial (Fuente:FAOSTAT 2010)

1.1.5 Producción de papaya en México

En México según datos del Servicio de Información Agroalimentaria y Pesquera (SIAP) de la SAGARPA en 2009 se dedicaron 17,745 Ha al cultivo de papaya de las cuales fueron cosechadas 16,033 Ha en 20 estados de la Republica. Los principales estados productores son Veracruz (179,133 T), Chiapas (119,873 T) y Oaxaca (93,534 T) (Figura 1.4), los cuales generan el 57% de la producción nacional, siendo Chiapas y Oaxaca los que obtienen mayores rendimientos con 69 y 64 Toneladas por hectárea respectivamente.

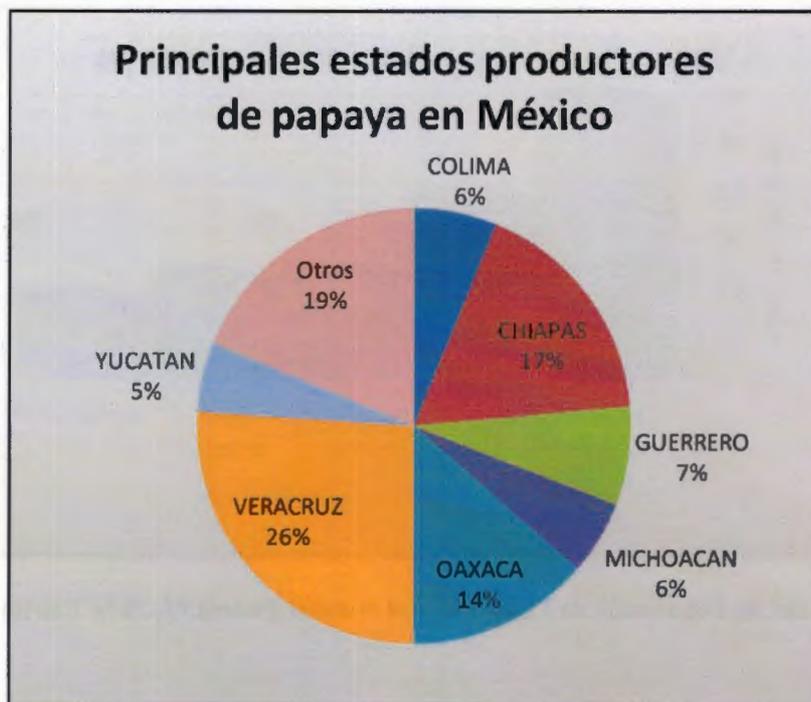


Figura 1.4 Porcentaje de producción de Papaya de los diferentes estados de la República Mexicana (Fuente: SIAP-SAGARPA 2010)

La producción nacional de papaya de 2000 a 2008 ha sido muy variable oscilando entre un mínimo de 638,237 T en 2008 hasta un máximo de 955,694 T en 2003. Sin embargo la cantidad de exportación se ha mantenido estable a lo largo del mismo periodo de tiempo promediando alrededor de 80,000 T que se exportan principalmente a Estados Unidos (Figura 1.5). De acuerdo a datos del servicio de información agroalimentaria y pesquera de la SAGARPA, Yucatán cuenta con 968 has en producción de papaya, colocando al estado como el séptimo estado productor teniendo un alto rendimiento de 52.47 T/Ha



Figura 1.5 Producción y Exportación en miles de toneladas de papaya en México del año 2000 hasta el año 2008 (Fuente. FAOSTAT 2010)

1.1.6 Problemáticas de la producción de papaya

En los últimos años el cultivo de papaya se ha visto afectado por problemas fitosanitarios que limitan su producción y comercialización y que tienen un efecto directo en los costos de producción y en los ingresos, ya que se utilizan aproximadamente \$ 46,000 por Ha en insumos de fertilizantes y plaguicidas lo que equivale a casi el 50% de los insumos necesarios para la producción de una hectárea. En Veracruz el costo del manejo de enfermedades oscilaba entre el 20% del costo del cultivo por hectárea mientras que en Chiapas se incrementa hasta en casi 60% del costo del cultivo (Claridades Agripecuarias, 1999, Plan Rector Sistema Nacional Papaya, 2005).

Dentro de las enfermedades que atacan a este cultivo se encuentran virales como el virus de la mancha anillada de papaya (PRSV por sus siglas en inglés) o bien el virus mosaico de papaya (PMV por sus siglas en inglés); que son de las más diseminadas a nivel mundial y que han sido combatidas con el desarrollo de nuevas variedades como la hawaiana SunUp, la mexicana Azteca o la taiwanesa Tainung. Sin embargo este cultivo es atacado por diferentes organismos como nematodos, principalmente *Rotilenchus*

reniformis que causa daños mecánicos y susceptibilidad a ataques por hongos; enfermedades bacteriales causadas por bacterias del género *Pseudomonas* y *Erwinia* que causan clorosis en hojas y las enfermedades fúngicas originadas por especies como *Alternaria alternate*, *Mycosphaerella sp.* *Cercospora papayae* Hansf., *Fusarium solani* (Mart.) Sacc. *Cladosporium sp.* etc (Poghosyan et al., 2005).

Dentro de las más importantes se encuentra la enfermedad de la antracnosis causada por el hongo hemibiótrofo *Colletotrichum gloeosporioides* (Poghosyan et al. 2005, Tapia-Tussell, 2008) y que es la principal causa de pérdida a nivel post-cosecha. *Colletotrichum gloeosporioides* (Penz.) Sacc., es la mayor causa de pérdidas por antracnosis en diversos frutos tropicales, como aguacate, cítricos, mangos etc. En papaya no se han generado variedades tolerantes a este patógeno pero existen algunas alternativas empleadas en el control de esta enfermedad de post-cosecha que incluyen tratamientos físicos y químicos, éstos últimos implican el uso constante de productos agroquímicos que pueden originar resistencia en el patógeno y algunos incluso pueden causar daños a la salud y/o al medio ambiente, deteriorando en ocasiones la calidad del producto con manchas, olores y sabores desagradables (Gonzalez-Aguilar, et al. 2007).

1.1.7 Biotecnología aplicada en papaya

La papaya se ha convertido en un modelo para otros cultivos frutales con respecto a usos de la biotecnología. Esto se debe en gran parte al excelente desarrollo del cultivo *in vitro* ganado durante tres décadas de investigación. La micropropagación de *Carica papaya* ha sido altamente desarrollada en regiones subtropicales para estabilizar y propagar genotipos seleccionados. Técnicas de rescate de embriones se han desarrollado eficientemente de varias clases de tejidos inmaduros en presencia de auxinas (Manshardt y Drew 1998).

La facilidad que tiene el género *Carica* para la producción de híbridos interespecíficos mediante el aislamiento de embriones y óvulos ha permitido que se realicen estudios que

logran hacer híbridos de papaya con otras seis especies del género *Carica* logrando de esta manera algunos híbridos útiles que logran heredar características de resistencia a enfermedades (Manshardt y Drew, 1998).

También existen análisis filogenéticos basados en polimorfismos de isoenzimas y ADN que ubican a papaya como el miembro más distante del género *Carica*.

Los protocolos para transformación genética de papaya están bien establecidos tanto para transformación mediada por biobalística (Cai *et al.*, 1999) así como mediada por *Agrobacterium* (Zhu *et al.*, 2006). Pero sin duda el mayor impacto biotecnológico realizado en papaya se realizó cuando a fines de los noventa se logró crear una variedad transgénica, que por medio de la expresión de un gen que codifica para la proteína de la cápside del virus PRSV lograron crear dos variedades resistentes al virus de la mancha anular. Las variedades transgénicas de papaya son en la actualidad el único caso de un fruto tropical transgénico que se ha aprobado su comercialización, y también es el único caso de un cultivo transgénico creado al 100% en instituciones públicas (Gonsalves, 1998).

Cuando Ming y colaboradores publicaron en 2008 la secuencia genómica de *Carica papaya* se abrió la posibilidad de analizar su genoma en busca de genes que permitan el desarrollo de nuevas variedades que resuelvan problemas específicos de esta importante especie como puede ser el desarrollo de variedades resistentes a la antracnosis. El genoma de papaya consta de 372 Mega bases (Mb) que contienen un total de 24,746 genes (Ming *et al.*, 2008), de los cuales cerca del 10% (2438 genes) corresponden a genes que codifican para factores de transcripción de 66 familias distintas (Ming *et al.*, 2008). La disponibilidad pública del genoma de una variedad transgénica de papaya (*Carica papaya* L. var. SunUp) brinda un enorme abanico de posibilidades en la búsqueda de genes de interés mediante análisis bioinformático, que puedan ser usados en la misma especie para el mejoramiento genético para el desarrollo de variedades cisgénicas. En los últimos años se han identificado genes que de manera natural han desarrollado las plantas para defenderse de enfermedades o diversas situaciones de estrés, que pueden

servir como referencia para detectar genes de interés biotecnológico en el genoma de papaya.

1.1.8 Mecanismos moleculares de resistencia de las plantas en respuesta al ataque de patógenos.

Las plantas son organismos sésiles, por lo que están en contacto con un sin número de condiciones ambientales para las cuales deben tener adaptaciones que les permitan sobrevivir en condiciones de estrés, incluyendo los ataques por patógenos. En el siglo pasado, el modelo de respuesta “gen por gen” proponía una explicación basada en la existencia de una proteína de resistencia por cada uno de los microorganismos que atacaban a las plantas; sin embargo, el tamaño del genoma requerido para albergar tal cantidad de genes fue la principal causa para buscar explicaciones más robustas que permitieran elucidar un mejor modelo de respuesta de las plantas al ser infectadas por organismos patógenos. Fue así como se empezó a desarrollar el modelo del “sistema inmune” en plantas (Bent y Mackey, 2007; Boller y Yang He, 2009).

En mamíferos, el sistema inmune se basa en el establecimiento de la denominada *memoria inmunológica*, donde cada patógeno es “recordado” por un antígeno característico y propio de ese patógeno en particular. Si un patógeno infecta a un organismo más de una vez, estas células de memoria desencadenan una respuesta específica para ese patógeno que han reconocido, con el fin de eliminarlo rápidamente. En contraste, el sistema inmune de las plantas se caracteriza por activar una respuesta basal contra patógenos, resultado del reconocimiento de patrones moleculares asociados a microorganismos (MAMPs o PAMPs-patrones moleculares asociados a patógenos) que se encuentran conservados en una extensa variedad de organismos y permite de esta manera generar una respuesta de defensa de amplio espectro. Los MAMP's también llamados elicitores, al ser reconocidos por moléculas receptoras del hospedero, desencadenan una cascada de señalización de proteínas cinasas que conllevan hacia la activación de genes de resistencia que permiten la llamada inmunidad activada por PAMP's (PTI; figura 1.6 a) (Bent y Mackey, 2007; Boller y Yang He, 2009).

Esta respuesta basal que permite la PTI es poco duradera, ya que los microorganismos evolucionan rápidamente y desarrollan moléculas efectoras las cuales bloquean la activación de los genes de defensa y permiten al patógeno entrar a la célula hospedera desarrollando en el hospedero una susceptibilidad activada por efectores (ETS; figura 1.6 b). A su vez y como resultado de un proceso coevolutivo, la planta desarrolla proteínas de resistencia (R) que reconocen directa o indirectamente a las moléculas efectoras de los microorganismos. Dichas proteínas generalmente cuentan con dominios de regiones ricas en leucina y de unión a nucleótidos (NBS-LRR) que permiten la inmunidad activada por efectores (ETI; figura 6c). Finalmente el microorganismo, al ser reconocida su proteína efectora, modifica, elimina o genera una nueva proteína efectora que le permite infectar nuevamente a la célula hospedera (figura 6d) generando así un ciclo coevolutivo en el que la frecuencia de los genes R y de los genes efectoras cambia a través de las poblaciones de hospederos y patógenos permitiendo a cada uno desarrollarse dependiendo de dichas frecuencias alélicas (figura 1.7).

Una vez que las plantas reconocen a los elicitores de los patógenos, este sistema de defensa puede generar mecanismos de resistencia que incluyen respuestas locales y sistémicas. En algunos casos, estas interacciones conllevan a la llamada respuesta hipersensible, la cual se caracteriza por una rápida muerte celular programada de la célula infectada y de las células vecinas, evitando de esta manera que se propague el microorganismo por toda la planta. Esta respuesta se caracteriza por ser de amplio espectro (defenderse contra varios tipos de microorganismos), permitiéndole a la planta defenderse de patógenos fúngicos, bacterianos, virales o incluso nemátodos (Bowles, 1990).

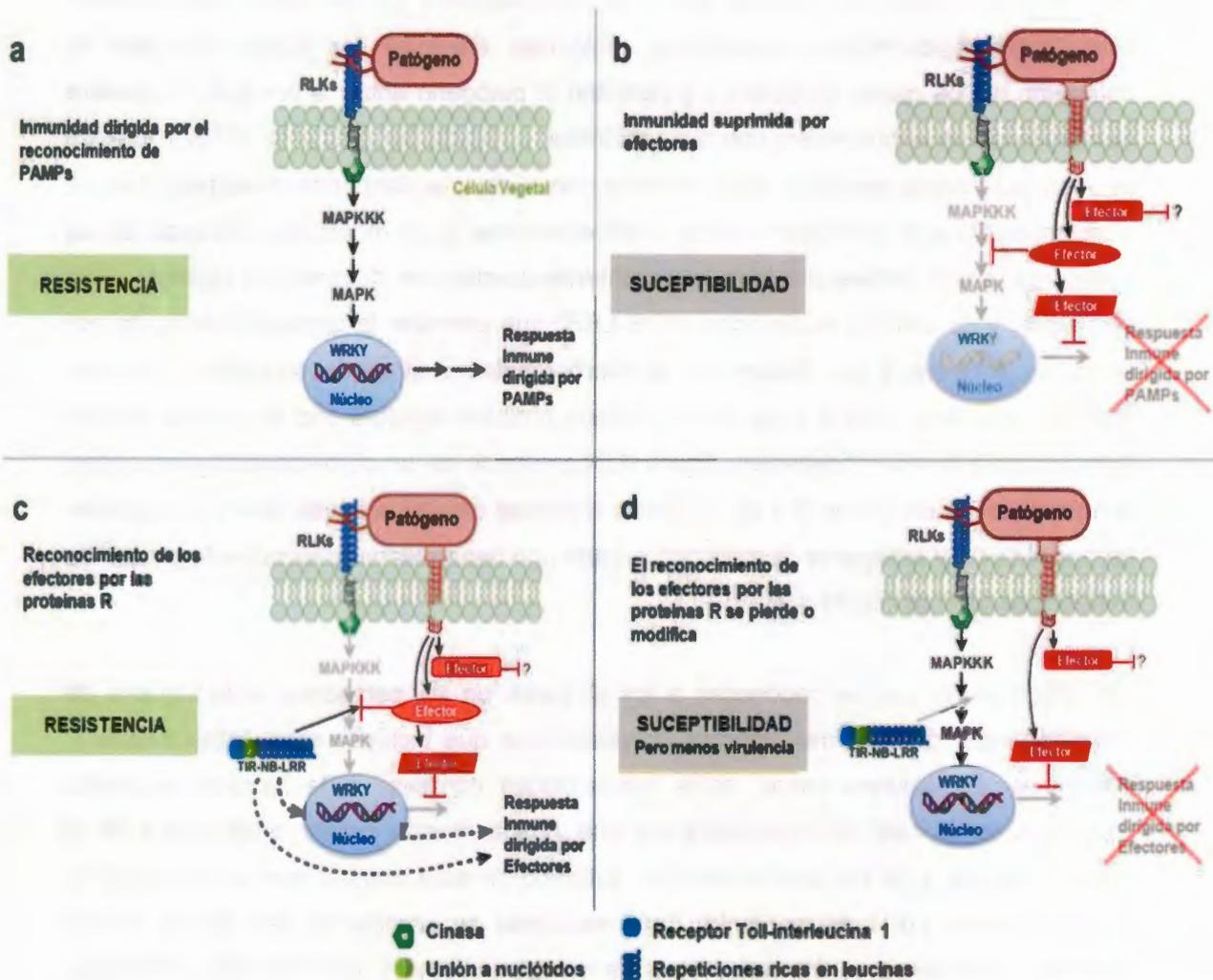


Figura 1.6 Modelo de evolución de la resistencia a microorganismos en plantas. a) reconocimiento de patrones moleculares asociados a patógenos (PAMPs) por receptores tipo cinasas (RLKs) que generan una ruta de señalización de cinasas hasta activar factores de transcripción tipo WRKY que activan la inmunidad dirigida por PAMPs; b) El patógeno desarrolla proteínas efectoras que suprimen la ruta de señalización por cinasas; c) El hospedero desarrolla proteínas R que reconocen los efectores y activan la respuesta inmune; d) Los efectores mutan y evaden el reconocimiento de las proteínas R para bloquear nuevamente la activación del sistema inmune de las plantas (Bent y Mackey, 2007).

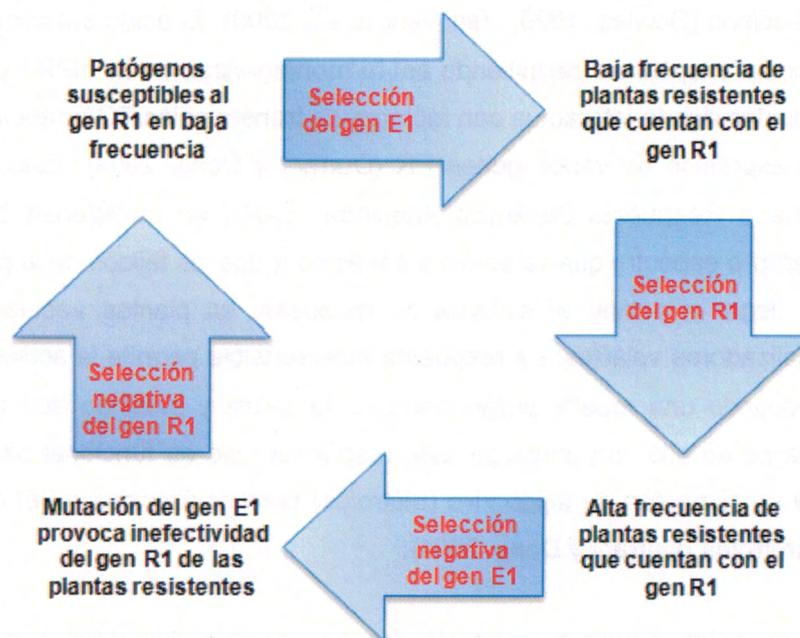


Figura 1.7 Co-evolución entre los genes R de la planta hospedera y los efectores del patógeno (Jones y Dang, 2006).

1.1.9 Rutas de señalización en los mecanismos de defensa de plantas

En los últimos años, los análisis de expresión genética en plantas de *Arabidopsis* han mostrado que la diferencia entre plantas susceptibles y resistentes, se encuentra en la velocidad y magnitud en la cual se activen los genes de defensa. Entre los genes de defensa que se pueden activar en respuesta a las infecciones se encuentran aquellos que codifican para (1) proteínas que participan en la construcción de barreras físicas o que permitan un confinamiento físico del patógeno; (2) enzimas que participen en la biosíntesis de metabolitos secundarios antimicrobiales o bien (3) proteínas relacionadas con patogénesis (PR) (Van Verk *et al.*, 2009).

Las proteínas PR se definen como aquellas que se expresan en las plantas al ser infectadas por un organismo patógeno o por una situación asociada al patógeno como el incremento en los niveles de hormonas señalizadoras como el ácido salicílico, el ácido

jasmónico y el etileno (Bowles, 1990; Van Verk *et al.*, 2009). El ácido salicílico cambia el potencial redox del citoplasma permitiendo así la monomerización de NPR1 y facilitando su entrada al núcleo donde interactúa con factores de transcripción TGA para incrementar los niveles de expresión de varios genes PR (Durrant y Dong, 2004). Este sistema de defensa nombrado Respuesta Sistémica Adquirida (SAR), se caracteriza por ser una respuesta de amplio espectro que se activa a través de todos los tejidos de la planta y que incluso puede llegar a activar el sistema de respuesta en plantas vecinas mediante moléculas señalizadoras volátiles. La respuesta hipersensible permite la activación rápida del SAR, provocando una muerte programada de la célula y evitando que el patógeno pueda alimentarse de ella, sin embargo esta respuesta solo es funcional para aquellos patógenos que se alimentan de tejido vivo (biótrofos) pero no funciona en el caso de los patógenos necrótrofos (Durrant y Dong, 2004).

Para el caso de estos organismos necrótrofos se prenden dos rutas que actúan de manera sinérgica activando genes PR que codifican para proteínas β -1,3-glucanasas (PR-2), quitinasas básicas (PR-3), proteínas tipo Hevein (PR-4), o bien los PDF's (plant defensin) que ayudan a las plantas a defenderse de organismos fúngicos. Dichas rutas, activadas por señales de etileno y ácido jasmónico actúan de manera antagónica con la ruta del ácido salicílico, por lo que existen mecanismos de regulación que permiten la conexión entre ambas rutas (figura 1.8).

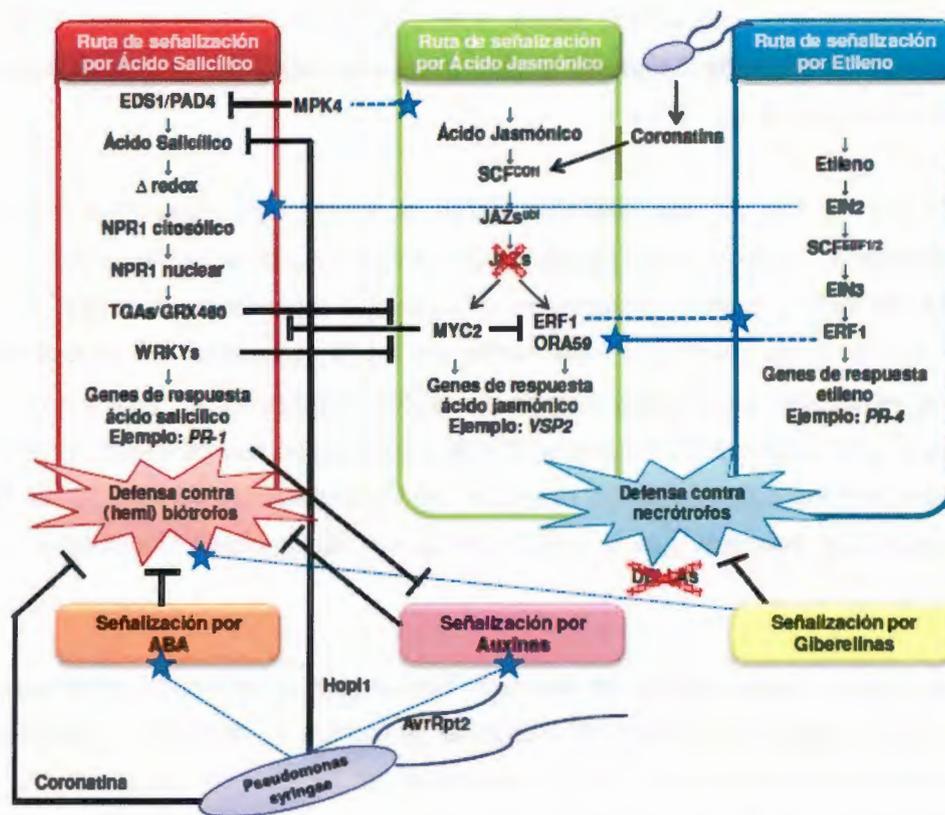


Figura 1.8 Interacción entre las diferentes rutas de señalización reguladas por fitohormonas durante la respuesta de las plantas al ataque de patógenos. Se ha visto que patógenos como *Pseudomonas syringae* producen proteínas efectoras (Coronatina, HopI1 y AvrRpt2) que suprimen la respuesta de defensa de las plantas. Las rutas de señalización guiadas por ácido salicílico, ácido jasmónico y etileno representan los principales mecanismos bioquímicos para activar los genes de defensa. Las barras “|” representan efectos negativos, mientras que las líneas punteadas con estrellas representan efectos positivos (Pieterse *et al.*, 2009).

1.1.10 Factores de transcripción y su participación en la resistencia a patógenos

Las proteínas que tienen la capacidad de unirse al ADN juegan un papel muy importante en la biología, por ejemplo replicar el genoma, transcribir genes y reparar ADN dañado. Una de las familias más grandes de proteínas que se unen al ADN son los factores de transcripción (FT), los cuales están implicados en los procesos de regulación del

desarrollo, diferenciación y crecimiento celular y al estar implicados en la activación y represión de la transcripción de los genes; son los principales reguladores de la expresión genética (Paez-Redondo *et al.*, 1992).

La biotecnológica de los FT se vislumbra como el futuro para desarrollar nuevas variedades de plantas, estos FT han jugado un papel muy importante en el origen de la domesticación de muchos cultivos (Doebley *et al.* 2006). De esta manera, cambios en los patrones de expresión de diversos FT han redituado en la generación de importantes características fenotípicas en diversos cultivos de interés (Century *et al.*, 2008;) como el gen *shattering4* (*sh4*) que codifica para un factor de transcripción que en estado silvestre permite que las semillas caigan al suelo, mientras que las mutaciones de este gen en las plantas domesticadas permiten que la semilla se quede en la planta (Doebley *et al.*, 2006).

Dentro de las plantas, algunos genes se expresan todo el tiempo (expresión constitutiva), mientras que otros sólo se expresan al responder a estímulos específicos (expresión inducible). Ambos patrones dependen de la interacción de FT con elementos *cis* de las regiones promotoras y/o de la interacción con otros FT para regular la expresión o represión de un gen determinado (Liu, *et al.*, 1999). Por tanto, la alteración en los niveles de expresión de los genes que codifican para FT resulta en un cambio dramático en el fenotipo de la planta y los cambios estructurales representan una gran fuerza evolutiva dentro de estas proteínas (Liu, *et al.*, 1999).

El sistema inmune de plantas consiste de una multitud de reacciones fisiológicas inducidas por el reconocimiento del patógeno. Algunas de estas reacciones conllevan a la muerte celular programada (respuesta hipersensible), modificaciones de la pared celular o bien producción de proteínas antibióticas y metabolitos tóxicos para los microbios. La mayor parte de estas reacciones están asociadas con la regulación de la expresión de los genes (PR). En plantas se han descubierto varios tipos de FT que están implicados en la resistencia a enfermedades (cuadro 1.1). Estos pueden responder a estímulos directos de un patógeno por medio del reconocimiento de elicitores, o bien, por medio de moléculas

mensajeras como ácido salicílico, ácido jasmónico o etileno, regulando de esta manera la expresión de los genes PR (Eulgem, 2005).

Cuadro 1.1 Factores de transcripción asociados con la defensa a patógenos en <i>Arabidopsis thaliana</i> (Eulgem, 2005)				
Tipo de Factor de transcripción	Tamaño de la familia en <i>Arabidopsis</i>	Características Clave	Motivo consenso y sitios de unión	Comentarios
ERF	56	Un Dominio AP2/ERF de Unión a ADN	GCCGCC (caja GCC)	Subfamilia de los FT AP2; activadores y represores
R2R3 Myb	125	Dos repeticiones del dominio Myb	Tipo I: (T/C)AAC(T/G)G Tipo II: G(G/T)T(A/T)G(G/T)T	Subfamilia predominante de los factores Myb de plantas
TGA Bzip	10	Un dominio básico de unión a ADN; un motivo de dimerización de proteína cierre de leucina	TGACGTCA (caja TGA)	Subfamilia de bZIP's; activadores y represores; forman homo- o heterodímeros
NPR1	6	Dominios repetidos Ankyrin; dominio BTB-POZ	Sin sitios de Unión al ADN	Interacción con TGA-bZIP's
Whirly	3	Dominio Whirly	GTCAAAA/T	Forman homotetrámeros; se unen a ADN de cadena simple
WRKY	74	Uno o dos dominios de unión a ADN WRKY	(T)GACC/T (Caja W)	Activadores y represores

Se ha visto que diversos miembros de las familias de factores de transcripción ERF, Myb, TGA, Whirly y WRKY tienen la capacidad de unirse a elementos *cis* de las regiones promotoras de varios genes PR, contribuyendo a la activación o represión de estos genes (Eulgem, 2005).

La acumulación de los FT WRKY parece una característica general en situaciones de defensa. La mayor parte de los miembros de la familia de FT WRKY de *Arabidopsis*

exhibe un incremento en los niveles de transcripción, después de una infección por patógenos o bien por tratamientos con elicitores. (Eulgem, 2005). Se ha visto también que la sobreexpresión o silenciamiento de algunos de ellos, afecta de manera considerable la resistencia o susceptibilidad de la planta (Gutterson y Reuben, 2004; Eulgem, 2005)

Los factores de transcripción han sido estudiados por métodos de cristalografía y espectroscopía de resonancia magnética nuclear (RMN). En plantas, la mayor parte de los análisis se basan en los estudios de las secuencias de aminoácidos. De esta manera se ha observado que en general, la mayor parte de los FT de plantas contienen regiones de unión al ADN, sitios de oligomerización, dominios de regulación de la transcripción y sitios de localización nuclear (NLS). Los factores de transcripción ERF (Ethylene Response Factors) pertenecen a una familia de proteínas exclusiva de plantas, que contienen el dominio AP2/ERF de unión a ADN. Dicho dominio consta de una secuencia altamente conservada de 59 aminoácidos que forman una estructura terciaria única (Ohme-Takagi y Shinshi, 1995). El alineamiento del dominio AP2/ERF de las 122 secuencias de *Arabidopsis* mostró que los aminoácidos Gly-4, Arg-6, Glu-16, Trp-28, Leu-29, Gly-30 y Ala-38 (los números corresponden a los aminoácidos del dominio AP2/ERF partiendo del 1 en el extremo amino hasta el 59 en el extremo carboxilo) son completamente conservados en las 122 secuencias, además, más del 95% de las secuencias contienen los residuos Arg-8, Gly-11, Ile-17, Arg-18, Arg-26, Ala-39, Asp-43, y Asn-57 (Nakano, *et al.* 2006). El arreglo tridimensional del dominio AP2/ERF consta de tres estructuras β -plegada en la región amino terminal y una estructura de α -hélice en el extremo carboxilo terminal (figura 1.9) (Allen, *et al.* 1998). A diferencia de otras familias de factores de transcripción que por lo general se unen al ADN por medio de α -hélices, los FTs de la familia ERF se unen al ADN por medio de residuos de triptófano y arginina (Arg-6, Arg-8, Trp-10, Glu-16, Arg-18, Arg-26 y Trp-28) localizados en las estructuras β -plegadas del dominio AP2/ERF a través de diversos puentes de hidrógeno que se dan entre los aminoácidos con el esqueleto y bases del ADN (Allen, *et al.* 1998; Hao, *et al.*, 1998).

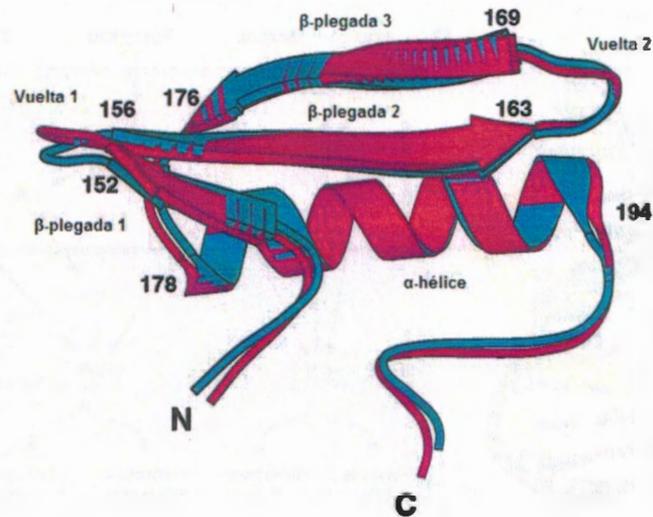


Figura 1.9 Estructura tridimensional del dominio AP2/ERF del FT AtERF1 de *Arabidopsis thaliana* donde se observan las 3 estructuras β -Plegadas y la α -hélice características de este dominio (Allen, *et al.*, 1998).

El contacto entre el del dominio AP2/ERF y el ADN se da en los elementos "GCC box" (figura 1.10) (Allen, *et al.* 1998; Ohme-Takagi y Shinshi, 1995). Dicha caja es un elemento *cis* compuesto por una secuencia de 7 pb (AGCCGCC; Ohme-Takagi y Shinshi, 1995) que se encuentra altamente conservada en las regiones promotoras de genes PR como aquellos que codifican para quitinasas básicas del tipo I o proteínas del tipo β -1-3 glucanasas las cuales se caracterizan por ser enzimas líticas capaces de degradar polisacáridos que se encuentran en las paredes celulares de muchos hongos, haciendo que tengan actividad antifúngica (Mauch *et al.*, 1998; Sela-Buurlage *et al.*, 1993; Melchers *et al.*, 1994; Niderman *et al.*, 1995). En algunos casos miembros de los FTs con el dominio AP2/ERF se unen también a elementos *cis* del tipo CRT/DRE (TACCGACAT; Yamaguchi-Shinozaki y Shinozaki, 1994), los cuales están presentes en las regiones promotoras de genes involucrados con tolerancia a estrés abiótico como el *rd29A*, *COR47*, *COR6.6* y *COR78*, los cuales responden a condiciones de bajas temperaturas, sequias y alta salinidad (Yamaguchi-Shinozaki y Shinozaki, 1994).

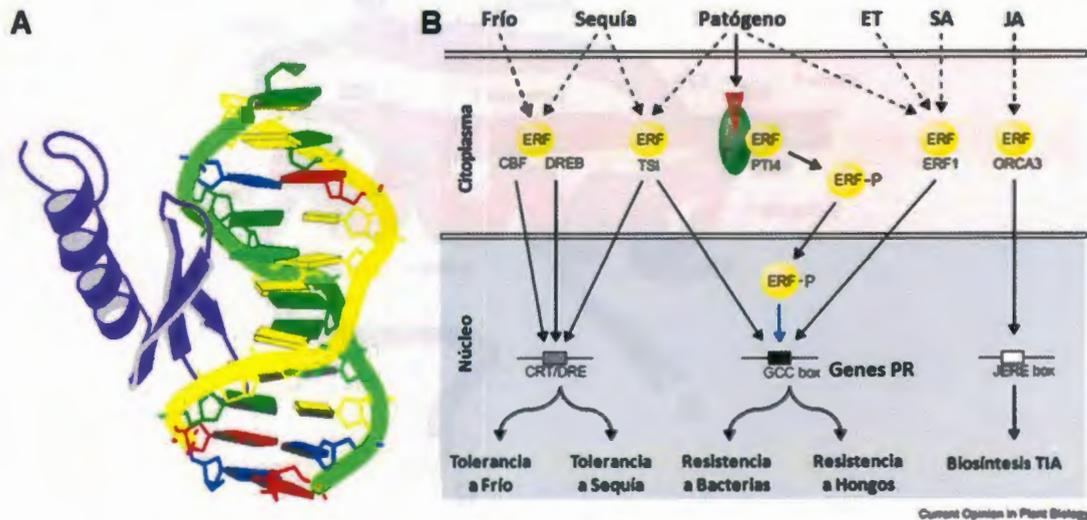


Figura 1.10 A. Representación esquemática de la interacción entre el dominio AP2/ERF y el elemento C/S caja GCC (Protein Data Bank: <http://www.pdb.org/pdb/explore/explore.do?structureId=1GCC>). B. Activación de los genes PR por medio de factores ERF inducidos por factores bióticos, abióticos y fitohormonas

Las proteínas que contienen el dominio AP2/ERF son exclusivas de plantas; Magnani y colaboradores (2004) presentaron la hipótesis de que dicho dominio se originó de proteínas endonucleasas con el dominio HNH presentes en bacterias y virus que fueron transferidas de manera horizontal cuando se dió la endosimbiosis de la cianobacterias para formar los cloroplastos, o bien por infecciones virales o algún otro evento de transferencia horizontal. Dichas endonucleasas HNH se propagaron por el genoma de plantas a través de procesos de transposición y procesos "homing". Algunas de estas endonucleasas divergieron y perdieron el dominio HNH pero retuvieron el dominio AP2, adquiriendo nuevas funciones. La transposición y duplicación de estas proteínas fueron cambiando la actividad endonucleasa hasta eventualmente transformarse en la superfamilia de factores de transcripción AP2/ERF y en las diferentes familias que contienen dicho dominio (Magnani *et al.*, 2004).

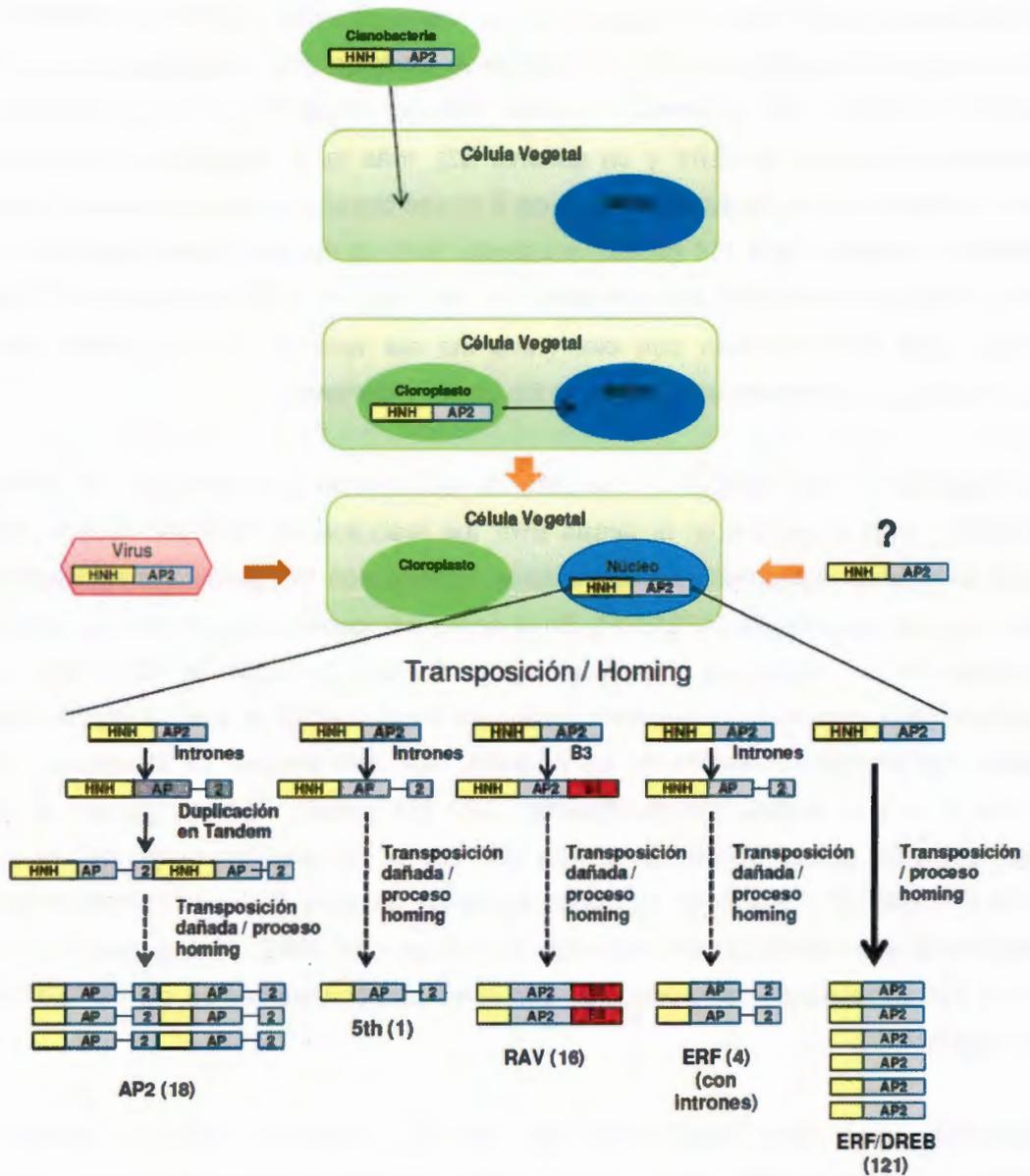


Figura 1.11 . Modelo de evolución del dominio de unión a ADN AP2/ERF en plantas. (Magnani et al., 2004)

La superfamilia de proteínas AP2/ERF fue agrupada por primera vez por Riechman *et al.* (2000), quienes encontraron 144 secuencias en *Arabidopsis* con el dominio AP2/ERF que dividieron en tres subfamilias: AP2 (14 miembros; contiene dos repeticiones del dominio AP2/ERF), EREBP (124 miembros; un solo dominio AP2/ERF) y RAV (6 miembros; contienen un dominio AP2/ERF y un dominio B3), más tarde Sakuma *et al.* (2002) las clasificó dentro de una familia subdividida en 5 clases basadas en la similitud del dominio AP2/ERF: subfamilia AP2 (14 genes), subfamilia RAV (6 genes), subfamilia DREB (55 genes; grupo A; interactúan con elementos *cis* del tipo CRT/DRE), subfamilia ERF (65 genes; grupo B; interactúan con elementos *cis* del tipo GCC box) y otras cuatro secuencias que no entraron en ninguna de las clases anteriores.

Una clasificación más detallada y completa de las proteínas que contienen el dominio AP2/ERF y más específico en la familia ERF fue realizada en 2006 por Nakano *et al.* quienes analizando nuevamente estas proteínas, identificaron 147 genes que se agruparon dentro de una superfamilia (AP2/ERF). En el análisis filogenético que ellos realizaron no solo tomaron en cuenta las diferencias en el dominio de unión al ADN, sino que consideraron la presencia de intrones/exones, así como también el análisis de diferentes motivos estructurales presentes en las proteínas. De esta manera, se organizaron tres familias y un gen aislado sin clasificación: AP2 (18 genes), RAV (6 genes), el gen At4g13040 y los grupos DREB (Grupo A) y ERF (Grupo B) quedaron dentro de una sola familia (Familia ERF) dividida en 10 clados donde los grupos I, II, III y IV corresponden al grupo DREB que interactúa con elementos *cis* del tipo CRT/DRE; y los grupos V, VI, VII, VIII, IX y X corresponden al grupo ERF que interactúa con elementos *cis* del tipo Caja GCC (figura 1.12).

Finalmente, una última clasificación fue realizada utilizando inferencia bayesiana (Goremykin y Moser, 2009). En la cual se confirman aspectos reportados por Nakano *et al.* (2006) y Sakuma *et al.* (2002), sin embargo le dan mayor resolución al árbol ya que a diferencia de los autores anteriores, toman la secuencia codificante completa que incluye en su análisis las sustituciones sinónimas. Goremykin y Moser (2009) confirman la mayoría de los grupos señalados por Nakano *et al.* (2006) y agrega dos grupos más, el XI y XII que no habían sido resueltos en los estudios anteriores.

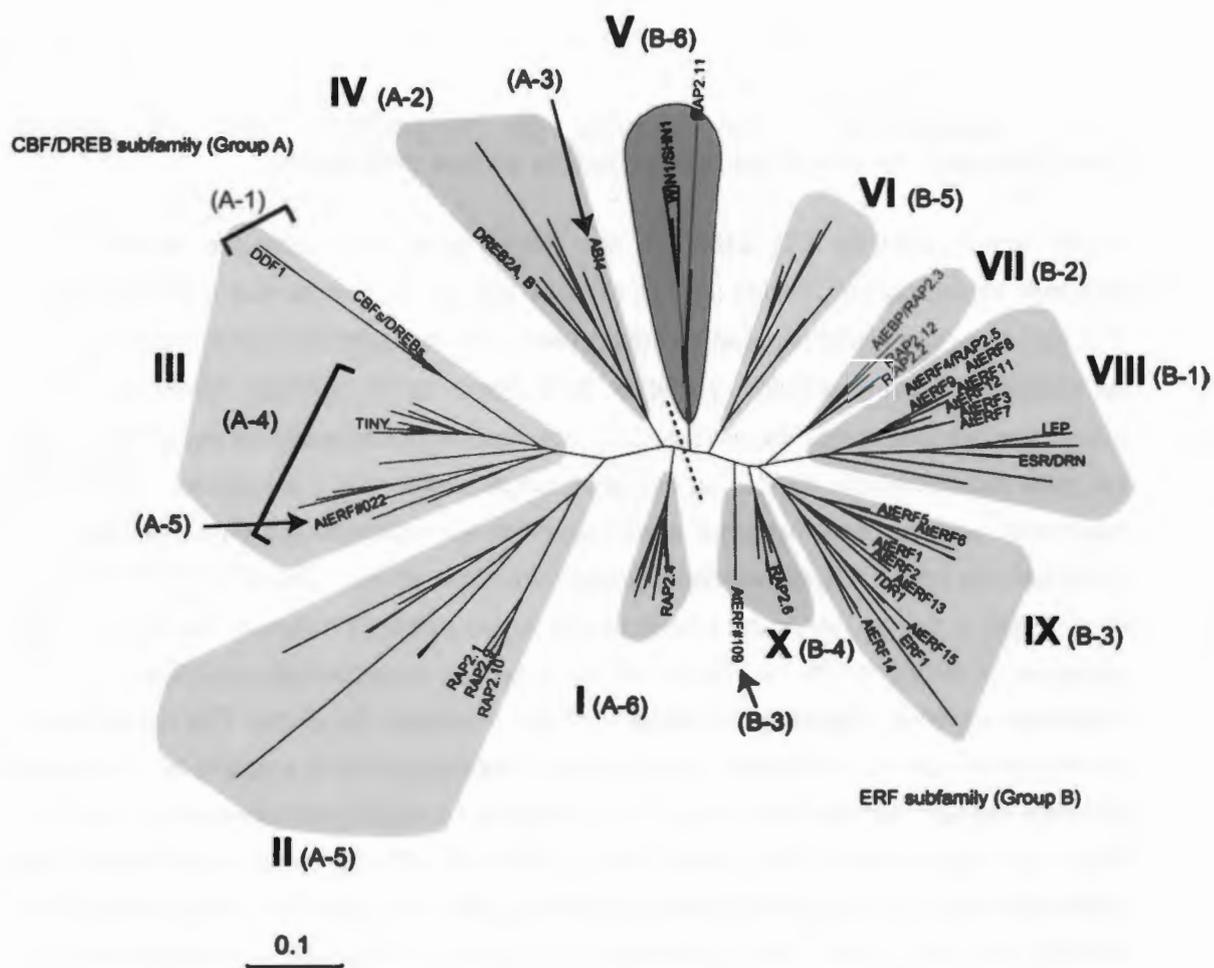


Figura 1.12 Árbol filogenético de los miembros de la familia ERF de *Arabidopsis thaliana* (Nakano et al., 2006)

1.1.11 Sobreexpresión de Factores de transcripción ERF en plantas transgénicas y su efecto en la resistencia contra patógenos.

Como hemos visto los FTs ERF son una familia grande de proteínas dentro de los genomas de las plantas. En los últimos años se han ido estudiando sus implicaciones en diversos procesos biológicos que incluyen cuestiones de desarrollo, así como cuestiones de tolerancia a estrés abiótico y biótico. Este último es de especial interés ya que al interactuar con elementos cis del tipo GCC box regulan la expresión de genes PR lo que los hace atractivos para regular los mecanismos de resistencia a patógenos. También se han hecho estudios que muestran la activación de su expresión por diversas hormonas señalizadoras como el ácido salicílico, ácido jasmónico y etileno (Gu, *et al.*, 2000; Lorenzo *et al.*, 2003) y se han generado evidencias de como estos FT's pueden llegar a generar tolerancia a patógeno. En los últimos 13 años se han caracterizado este tipo de FT en diferentes especies vegetales. El primero en ser estudiado fue el gen *Pti4* de tomate, el cual fue descubierto al investigar proteínas que interactúan con la proteína de resistencia Pto para conferir resistencia al ataque de la bacteria *Pseudomonas syringae* (Zhou *et al.*, 1997), al interactuar con la cinasa Pto, la proteína Pti4 aumenta su afinidad a los elementos GCC para activar la transcripción de genes PR y generar resistencia (Figura 10.10b) (Wu *et al.*, 2002). Pti4 es el factor de transcripción tipo ERF más estudiado y se ha visto involucrado también como parte del represoma que regula negativamente la transcripción del gen PR10 en papa (González-Lamothe *et al.*, 2008) por lo que este gen resulta muy interesante en cuestión de regulación de diversos genes PR ya que parece actuar como un gen pivotante que puede regular positiva y negativamente estos genes. Otro de los primeros genes en ser descubiertos fue el gen *ERF1* de *Arabidopsis thaliana* (Solano *et al.*, 1998) cuya proteína forma parte de la vía de señalización dependiente de etileno y que se encuentra involucrada en la resistencia contra patógenos necrótrofos, pero también funciona como una ruta que modula el antagonismo entre las rutas de señalización dependientes de ácido salicílico y la ruta dependiente de ácido jasmónico (Figura 10.10b) (Van Verk *et al.*, 2009). El gen de soya GmERF3 fue recientemente descrito (Zhang *et al.*, 2009) y forma parte del grupo VII de factores tipo ERF descrito por

Nakano (2006) cuya característica estructural es la presencia de un intrón anterior al dominio AP2/ERF y al que pertenecen otro los genes como *TaERF1* de trigo, *OsEREBP1* de arroz, *CaPF1* y *CaERFLP1* de chile los cuales están involucrados en la resistencia tanto a estrés biótico como abiótico (Cheong, *et al.*, 2003; Lee, *et al.*, 2004; Yi, *et al.*, 2004; Xu, *et al.*, 2007). El gen *GmERF3* se ha visto que aumenta su expresión al ser inducido por las hormonas ácido salicílico, etileno y ácido jasmónico, así como el virus del mosaico de la soya, este gen al hacer contacto con elementos GCC, genera resistencia contra hongos, virus y bacterias al ser sobreexpresada en plantas transgénicas de tabaco (Zhang *et al.*, 2009). También se han identificado el gen *Tsi1* de tomate homólogo al gen *Pti6* el cual interactúa con una proteína denominada Tsip para conferir resistencia a patógenos y salinidad (Park, *et al.*, 2001; Ham *et al.*, 2005); finalmente el caso del gen *JcERF* de *Jatropha curcas* se presenta como el primer gen de la familia ERF clonado de una planta leñosa que también responde a etileno y podría conferir resistencia a patógenos (Tang, *et al.*, 2007). En el cuadro 1.2 se muestran FTs de la familia ERF de diversas especies vegetales que al ser sobreexpresadas en plantas de *Arabidopsis* y tabaco lograron conferir resistencia a patógenos bacterianos, fúngicos o incluso virales (Fischer y Droge-Laser, 2004). La mayor parte de los experimentos de sobreexpresión de estos factores de transcripción se han realizado en plantas modelo bajo el control del promotor constitutivo 35S lo que en ciertos casos a causado un claro efecto pleiotrópico y si bien el resultado confiere resistencia a patógenos también comprometen un desarrollo normal de la planta (Berrocal-Lobo y Molina, 2004; Gu, *et al.*, 2002). De tal forma, es deseable que la sobreexpresión de estos FT este regulada mediante promotores que conduzcan una alta expresión pero que no generen los efectos negativos del promotor 35S o bien mediante el uso de promotores inducibles.

Cuadro 1.2 Factores de transcripción de la subfamilia ERF que han generado tolerancia a patógenos cuando son sobreexpresados en plantas de *Arabidopsis* o tabaco

Factor de transcripción	PLANTA	PATÓGENO	Cita
Pti4, Pti5 y Pti6	<i>Solanum lycopersicum</i>	<i>Erysiphe orontii</i> <i>Pseudomonas syringae</i>	Zhou <i>et al.</i> , 1997; Gu <i>et al.</i> , 2000; Gu, <i>et al.</i> , 2002
Tsi1	<i>Nicotiana tabacum</i>	<i>Pseudomonas syringae</i>	Park, <i>et al.</i> , 2001
OsEREBP1	<i>Oriza sativa</i>	<i>Phytophthora parasitica</i> <i>Pseudomonas syringae</i>	Cheong, <i>et al.</i> , 2003
CaERFLP1	<i>Capsicum annum</i>	<i>Pseudomonas syringae</i>	Lee, <i>et al.</i> , 2004
CaPF 1	<i>Capsicum annum</i>	<i>Xanthomonas axonopodis</i>	Yi, <i>et al.</i> , 2004 Tang, <i>et al.</i> , 2005
ERF1	<i>Arabidopsis thaliana</i>	<i>Botrytis cinérea</i> <i>Plectosphaerella cucumerina</i> . <i>Fusarium oxysporum</i>	Solano <i>et al.</i> , 1998 Berrocal-Lobo <i>et al.</i> , 2002 Berrocal-Lobo y Molina; 2004
TSRF1	<i>Solanum lycopersicum</i>	<i>Ralstonia solanacearum</i>	Zhang <i>et al.</i> , 2004
NtERF5	<i>Nicotiana tabacum</i>	<i>Tobacco mosaic virus</i>	Fischer y Droge-Laser, 2004
OPBP1	<i>Nicotiana tabacum</i>	<i>Pseudomonas syringae</i> <i>Phytophthora parasitica</i>	Guo, <i>et al.</i> , 2004
HvRAF	<i>Hordeum vulgare</i>	<i>Ralstonia solanacearum</i>	Jung, <i>et al.</i> , 2007
JcERF	<i>Jatropha curcas</i>	Etileno	Tang, <i>et al.</i> , 2007

TaERF1	<i>Triticum aestivum</i>	<i>Botrytis cinérea</i> <i>Pseudomonas syringae</i>	Xu, et al., 2007
GbERF2	<i>Gossypium barbadense</i> ;	<i>Alternaria longipes</i>	Zuo, et al., 2007
GmEREBP1	<i>Glycine max</i>	<i>Heterodera glycines</i>	Mazerai et al., 2007
GmERF3	<i>Glycine max</i>	<i>Ralstonia solanacearum</i>	Zhang et al., 2009

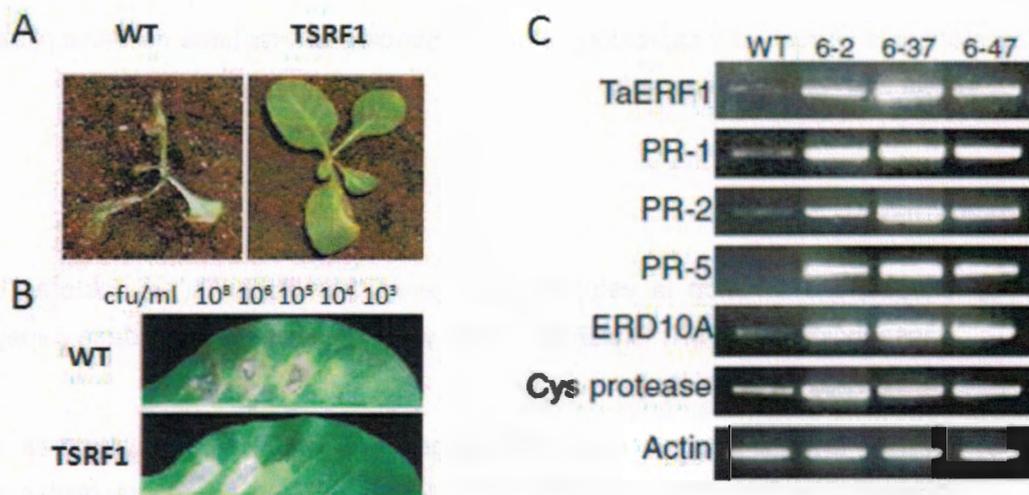


Figura 1.13 La sobreexpresión de factores de transcripción tipo ERF aumentan la tolerancia a patógenos mediante la activación de genes PR. A. Plantas de tabaco sometidas al patógeno *Ralstonia solanacearum*; WT: Línea silvestre, TSRF: Línea transgénica que sobreexpresa el gen TSRF1. B Hojas de tabaco inducidas en puntos específicos a diversas concentraciones de unidades formadoras de colonias de *R. solanacearum*. C. Patrones de expresión de diferentes genes PR en líneas transgénicas de tabaco que sobreexpresan el gen TaERF1 (Zhang et al., 2004; Xu et al., 2007).

1.2 HIPÓTESIS

Homólogos de los genes *Pti4*, *Tsi1*, *GmERF3*, *ERF1* y *JcERF* de la familia ERF probablemente se encuentran presentes en el genoma de *Carica papaya* y se expresan en diferentes tejidos de esta planta.

1.3 OBJETIVOS

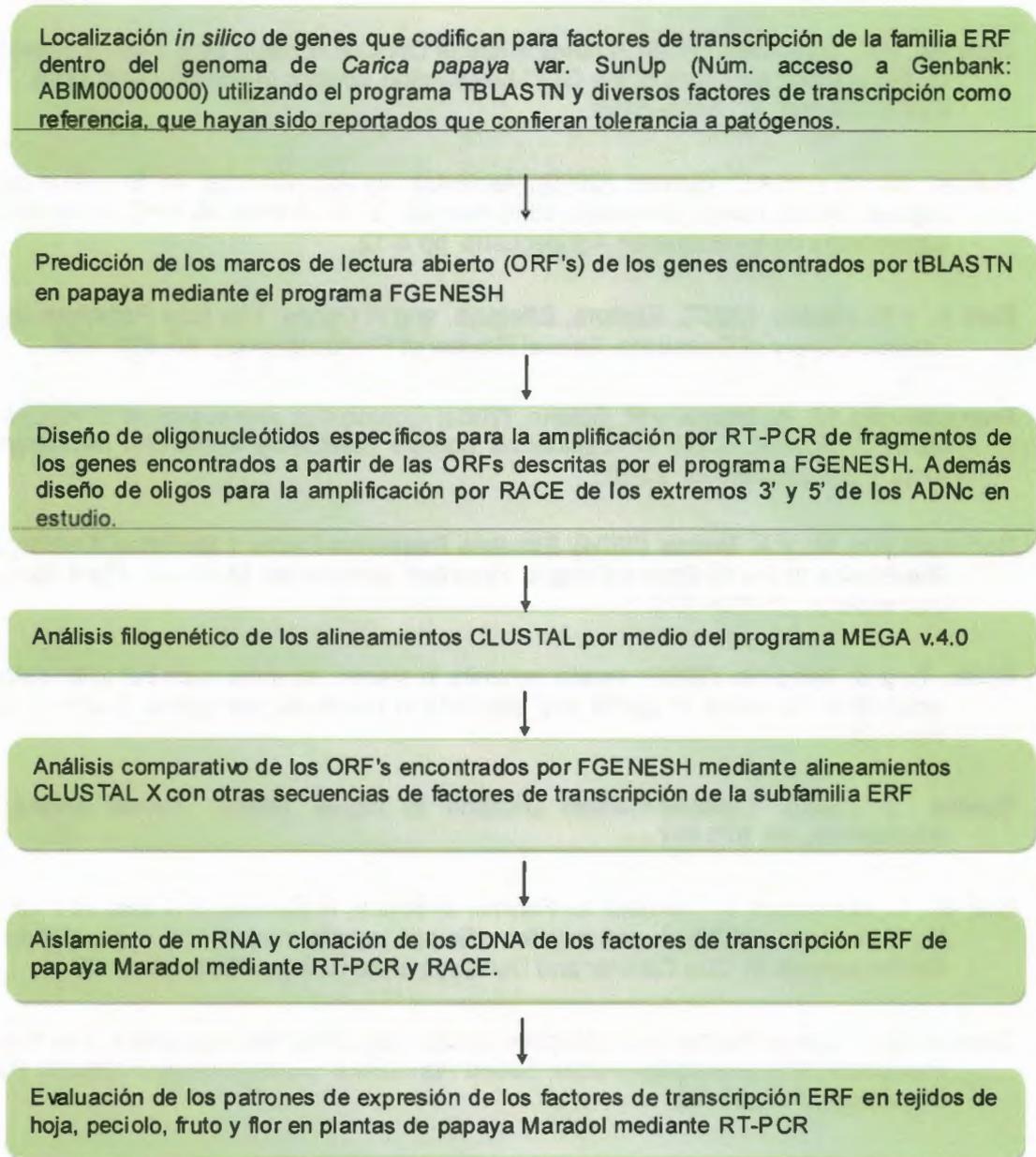
1.3.1 General

Caracterizar la estructura y expresión de genes homólogos a factores de transcripción de la familia ERF en *Carica papaya*.

1.3.2 Particulares

- a) Caracterizar *in silico* la estructura de genes homólogos a los factores de transcripción *Pti4*, *Tsi1*, *GmERF3*, *ERF1* y *JcERF* en el genoma de la papaya transgénica var. SunUp.
- b) Evaluar los patrones de expresión de genes homólogos a los factores de transcripción *Pti4*, *Tsi1*, *GmERF3*, *ERF1* y *JcERF* en diferentes tejidos de papaya var. Maradol
- c) Mapear los extremos 5' y 3' de los ADNc de los genes homólogos tipo ERF de papaya var. Maradol.

1.4 ESTRATEGIA EXPERIMENTAL



1.5 BIBLIOGRAFÍA

- Allen, M., y K. Yamasaki. (1998) A novel mode of DNA recognition by a b-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, 17, 5484–5496.
- Arango, W.L.V. y H.A.C. Roman. (2000). Aspectos agroeconómicos. *In: El cultivo de la papaya en los llanos Orientales de Colombia*. L. V. Arango W (ed). Corporación Colombiana de Investigación Agropecuaria. pp:8-12.
- Bent A., y D. Mackey. (2007). Elicitors, Effectors, and *R* Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology*, 45, 399–436.
- Berrocal-Lobo, M., A. Molina, y R. Solano. (2002) Constitutive expression of ETHYLENE-RESPONSE-FACTOR1 in *Arabidopsis* confers resistance to several necrotrophic fungi. *Plant Journal*, 29, 23-32.
- Berrocal-Lobo, M., y A. Molina. (2004). Ethylene Response Factor 1 Mediates *Arabidopsis* Resistance to the Soilborne Fungus *Fusarium oxysporum*. *Molecular Plant-Microbe Interactions*, 17, 763–770.
- Boller, T., y S. Yang-He. (2009). Innate immunity in plants: An arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science*, 324, 742-743.
- Bowles, J. (1990). Defense-Related proteins in higher plants. *Annual review of biochemical*, 59, 873-907.
- Cai, W., C. Gonsalves, P. Tennant, G. Fermin, M Souza, N Sarindu, F.J. Jan, H.Y Zhu, y D. Gonsalves. (1999) A protocol for efficient transformation and regeneration of *Carica papaya*. *In Vitro Cellular and Developmental Biology-Plant*, 35, 61-69.
- Century, K., T. Lynne-Reuber, y O. Ratcliffe. (2008) Regulating the Regulators: The Future Prospects for Transcription-Factor-Based Agricultural Biotechnology Products. *Plant Physiology*, 147, 20–29.
- Cheong, Y.H., B.C. Moon, J.K. Kim, M.C Kim, I.H. Kim, C.Y Park, J.C Kim, B.O. Park, S. C. Koo, H. W. Yoon, W. S. Chung, C. O. Lim, S. Y. Lee, y M. J. Cho. (2003). BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcriptor factor. *Plant Physiology*, 132, 1961-1972.

- Claridades Agropecuarias. (1999). Papaya un mercado en expansión. Claridades agropecuarias. México. Marzo. pp 3-24.
- Doebley, J.F., B.S. Gaut, B.D. Smith. (2006). The molecular genetics of crop domestication. *Cell*, 127, 1309–1321.
- Durrant, W.E. y X. Dong. (2004). Systemic Acquired resistance. *Annual review of phytopathology*, 42, 185-209.
- Eulgem, T. (2005). Regulation of the *Arabidopsis* defense transcriptome. *Trends in Plant Science*, 10, 71-78.
- FAO, (2008). Estadísticas de FAO. Sistema FAOSTAT
Disponible en: <http://faostat.fao.org/site/567/default.aspx#ancor>
[Acceso 10 Agosto 2010].
- Fischer, U. y W. Dröge-Laser. (2004). Overexpression of *NtERF5*, a New Member of the Tobacco Ethylene Response Transcription Factor Family Enhances Resistance to *Tobacco mosaic virus*. *Molecular Plant-Microbe Interactions*, 17, 1162–1171.
- Flores-Revilla, C., E. Garcia, D. Nieto-Angel, D. Teliz-Ortiz, y J.A. Villanueva-Jimenez. (1995). Integrated management of papaya in Mexico. *Acta Horticulturae*, 370, 151-158.
- Gonsalves, D. (1998). Control of papaya ringspot virus in papaya: a case study. *Annual Review of Phytopathology*, 36, 415–437.
- González-Aguilar G.A., R. Cruz-Valenzuela, J.F. Ayala-Zavala, M. Robles-Sánchez, S. Ruiz-Cruz, W. Tejedor-Espinosa, E. Álvarez-Parrilla. (2007). Efecto de tratamiento con antioxidantes sobre la vida de anaquel de frutos tropicales mínimamente procesados: Cambios fisiológicos y bioquímicos. In: *Procesado mínimo de frutas*. Editores: Lobo M.G., González M. Editorial: Instituto Canario de Investigaciones Agrarias. pp. 37-54.
- Goremykin, V., y C. Moser. (2009) Classification of the *Arabidopsis* *ERF* Gene Family Based on Bayesian Inference. *Molecular Biology*, 43, 729–734.
- Gu, Y.Q., C. Yang, T. Venkatappa, J. Zhou y G. Martin. (2000). *Pti4* Is Induced by Ethylene and Salicylic Acid, and Its Product Is Phosphorylated by the Pto Kinase. *The Plant Cell*, 12, 771–785.

- Gu, Y.Q., M.C. Wildermuth, S. Chakravarthy, Y.T. Loh, C. Yang, X. He, Y. Han y G.B. Martin. (2002). Tomato transcription factors *pti4*, *pti5*, and *pti6* activate defense responses when expressed in *Arabidopsis*. *Plant Cell*, 14, 817–831.
- Guo, Z.J., X. J. Chen, X.L. Wu, L.Q. Ling, y P. Xu. (2004). Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. *Plant Molecular Biology*, 55, 607–618.
- Gutterson, N. y T.L. Reuber. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Current Opinion in Plant Biology*, 7, 465–471.
- Hao, D., M. Ohme-Takagi y A. Sarai. (1998). Unique Mode of GCC Box Recognition by the DNA-binding Domain of Ethylene-responsive Element-binding Factor (ERF Domain) in Plant. *The Journal of Biological Chemistry*, 273, 26857–26861.
- Jones, J. y J. Dangl (2006). The plant immune system. *Nature Reviews*, 444, 323-329.
- Jung, J., S.Y. Won, S.C. Suh, H.R. Kim, R. Wing, Y. Jeong, I. Hwang, y M. Kim. (2007). The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in *Arabidopsis*. *Planta*, 225, 575–588.
- Kumar, L.S.S., V.K Srinivasan. (1944). Chromosome number of *Carica dodecaphylla* Vell. *Fl. Flum. Current Science*, 13, 15.
- Lee, J.H., J.P. Hong, S.K. Oh, S. Lee, D. Choi, y W.T. Kim. (2004). The ethylene-responsive factor like protein 1 (CaERFLP1) of hot pepper (*Capsicum annuum* L.) interacts in vitro with both GCC and DRE/CRT sequences with different binding affinities: Possible biological roles of CaERFLP1 in response to pathogen infection and high salinity conditions in transgenic tobacco plants. *Plant Molecular Biology*, 55, 61–81.
- Lorenzo, O., R. Piqueras, J.J. Sánchez-Serrano, R. Solano, (2003). ETHYLENE RESPONSE FACTOR1 Integrates Signals from Ethylene and Jasmonate Pathways in Plant Defense. *The Plant Cell*, 15, 165–178.
- Liu, L., M. White, y T. MacRae (1999). Transcription factors and their genes in higher plants: Functional domains, evolution and regulation. *European Journal of Biochemistry*, 262, 247-257.
- Magnani, E., K. Sjölander, y S. Hake. (2004). From Endonucleases to Transcription Factors: Evolution of the AP2 DNA Binding Domain in Plants. *The Plant Cell*, 16, 2265–2277.

- Manshardt, R.M. y R.A. Drew. (1998). Biotechnology of papaya. *Acta Horticulturae*, 461, 65-73.
- Massieu-Trigo, Y. (2009). Exportaciones no tradicionales (ENT) y papaya transgénica. La diversidad genética en la globalización. Ponencia para el XXVIII Congreso de Latin American Studies Association Rio de Janeiro, Brasil 11 al 14 de junio de 2009.
- Mauch, F., L. A. Hadwiger y T. Boller. (1988). Antifungal Hydrolases in Pea Tissue : I. Purification and Characterization of Two Chitinases and Two -1,3-Glucanases Differentially Regulated during Development and in Response to Fungal Infection, 87, 325–333.
- Mazarei, M., A. Elling, T. Maier, D. Puthoff, y T. Baum. (2007). GmEREBP1 Is a Transcription Factor Activating Defense Genes in Soybean and *Arabidopsis*. *Molecular Plant-Microbe Interactions*, 20, 107-119.
- Melchers, L. S., M. A.-d. Groot, J. A. v. d. Knapp, A. S. Ponstein, M. B. Sela-Burlage, J.F. Bol, B.J. Cornelissen, P.J. Van den Elzen, y H.J. Linthorst. (1994). A new class of tobacco chitinases homologous to bacterial exo-chitinases displays anti-fungal activity. *Plant Journal*, 5, 469–480.
- Ming, R., S. Hou, Y. Feng, Q. Yu, A. Dionne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis, S.L. Salzberg, L. Feng, M.R. Jones, R.L. Skelton, J.E. Murray, C. Chen, W. Qian, J. Shen, P. Du, M. Eustice, E. Tong, H. Tang, E. Lyons, R.E. Paull, T.P. Michael, K. Wall, D.W. Rice, H. Albert, M.L. Wang, Y.J. Zhu, M. Schatz, N. Nagarajan, R.A. Acob, P. Guan, A. Blas, C.M. Wai, C.M. Ackerman, Y. Ren, C. Liu, J. Wang, J. Wang, J.K. Na, E.V. Shakhov, B. Haas, J. Thimmapuram, D. Nelson, X. Wang, J.E. Bowers, A.R. Gschwend, A.L. Delche, R. Singh, J.Y. Suzuki, S. Tripathi, K. Neupane, H. Wei, B. Irikura, M. Paidi, N. Jiang, W. Zhang, G. Presting, A. Windsor, R. Navajas-Pérez, M.J. Torres, F.A. Feltus, B. Porter, Y. Li, A.M. Burroughs, M.C. Luo, L. Liu, D.A. Christopher, S.M. Mount, P.H. Moore, T. Sugimura, J. Jiang, M.A. Schuler, V. Friedman, T. Mitchell-Olds, D.E. Shippen, C.W. dePamphilis, J.D. Palmer, M. Freeling, A.H. Paterson, D. Gonsalves, L. Wang y M. Alam. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452, 991–996.
- Mirafuentes-Hernández, F. y A. Azpeitia-Morales. (2008). "Azteca", primer híbrido de papaya para el trópico de México. *Revista Fitotecnia Mexicana*, 31, 291-293.
- Nakano, T., K. Suzuki, T. Fujimura, y H. Shinshi. (2006). Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. *Plant Physiology*, 140, 411–432.

- Niderman, T., I. Genetet, T. Bruyere, R. Gees, A. Stintzi, M. Legrand, B. Fritig y E. Mosinger. (1995). Pathogenesis-Related PR-1 Proteins Are Antifungal (Isolation and Characterization of Three 14-Kilodalton Proteins of Tomato and of a Basic PR-1 of Tobacco with Inhibitory Activity against *Phytophthora infestans*). *Plant Physiology*, 108, 17–27.
- Ohme-Takagi, M., y H. Shinshi. (1995). Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element. *The Plant Cell*, 7, 173-182.
- Oviedo. (1851). *Historia general y natural de las indias, islas y tierra firme del mar océano / por el capitán Gonzalo Fernández de Oviedo y Valdés, primer cronista del Nuevo Mundo. Publica la Real Academia de la Historia, cotejada con el código original. Tomo I. Disponible en: http://cdigital.dgb.uanl.mx/la/1080017176_C/1080017176_T1/1080017176_114.pdf.*
- Paez-Redondo, A. Pabo y R. Sauer. (1992). Transcription factors: Structural families and principles of DNA recognition. *Annual Review of Biochemistry*, 61, 1053-95.
- Park, J.M., C.J. Park, S.B. Lee, B.K. Ham, R. Shin, K.H. Paek. (2001) Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell*, 13, 1035-1046.
- Perez-Carrillo, E. y E. Yahia. (2003). Effect of postharvest hot air and fungicide treatments on the quality of 'maradol' papaya (*Carica papaya* L.). *Journal of Food Quality*, 27, 127-139.
- Pieterse, A., A. Leon-Reyes, S. Van der Ent y S. Van Wees (2009). Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, 5, 308-316.
- Plan Rector Sistema Nacional Papaya, (2006). SAGARPA, México.
- Poghosyan, A., G. Zehfuss, y M. Jimenez. (2005). Papaya Diseases: Ecology and Control. *Encyclopedia of Pest Management*, 1, 1–5.
- Riechmann, J.L., J. Heard, G. Martin, L. Reuber, C. Jiang, J. Keddie, L. Adam, O. Pineda, O.J. Ratcliffe, R.R. Samaha, R. Creelman, M. Pilgrim, P. Broun, J.Z. Zhang, D. Ghandehari, B.K. Sherman, G. Yu. (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science*, 290, 2105–2110.

SAGARPA, (2010). SISTEMA PRODUCTO.

Disponible

en:

<http://www.sagarpa.gob.mx/agricultura/Publicaciones/SistemaProducto/Paginas/default.aspx>

Acceso: 30 de Agosto de 2010.

Sakuma, Y, Q. Liu, J.G. Dubouzet, H. Abe, K. Shinozaki y K. Yamaguchi-Shinozaki. (2002). DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold-inducible gene expression. *Biochemical and Biophysical Research Communications*, 290, 998–1009.

Schery, R.W. (1956). *Plantas útiles al hombre*. Salvat. Barcelona. 756 p.

Sela-Buurlage, M.B., A.S. Ponstein, S.A. Bres-Voemans, L.S. Melchers, P.J.M. Van Den Elzen, y B.J.C. Cornelissen. (1993) Only specific tobacco (*Nicotiana tabacum*) chitinases and b-1,3-glucanases exhibit antifungal activity. *Plant Physiology*, 101, 857–863.

Servicio de información Agroalimentaria y Pesquera, Acceso en <http://www.siap.gob.mx/>

Solano, R., A. Stepanova, A. Chao y J. Ecker. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ethylene-insensitive3 and ethylene-response-factor1. *Genes and Development*, 12, 3703–3714.

Tang, W., T. Charles, y R. Newton. (2005). Overexpression of the pepper transcription factor CaPF1 in transgenic Virginia pine (*Pinus virginiana* Mill.) confers multiple stress tolerance and enhances organ growth. *Plant Molecular Biology*, 59, 603–617.

Tang, M., J. Sun, Y. Liu, F. Chen, y S. Shen. (2007). Isolation and functional characterization of the JcERF gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant *Jatropha curcas*. *Plant Molecular Biology*, 63, 419–428.

Tapia-Tussell, R., A. Quijano-Ramayo, A. Cortes-Velazquez, P. Lappe, A. Larque-Saavedra y D. Perez-Brito. (2008). PCR-Based Detection and Characterization of the Fungal Pathogens *Colletotrichum gloeosporioides* and *Colletotrichum capsici* Causing Anthracnose in Papaya (*Carica papaya* L.) in the Yucatan Peninsula. *Molecular Biotechnology*, 40, 293–298.

Teixeira da Silva, J., Z. Rashid, D. Tan Nhut, D. Sivakumar, A. Gera, M. Teixeira Souza y P. Tennant. (2007). Papaya (*Carica papaya* L.) *Biology and Biotechnology*. *Tree and Forestry Science and Biotechnology*, 1, 47–63.

- Van Verk, M., C. Gatz y H. Linthorst. (2009). Transcriptional Regulation of Plant Defense Responses. *Advances in Botanical Research*, 51, 397-438.
- Vázquez-Yanes, C., A.I. Batis Muñoz, M.I. Alcocer Silva, M. Gual-Díaz y C. Sánchez Dirzo. (1999). *Árboles y arbustos potencialmente valiosos para la restauración ecológica y la reforestación*. Reporte técnico del proyecto J084. CONABIO - Instituto de Ecología, UNAM. Acceso en: http://www.conabio.gob.mx/conocimiento/info_especies/arboles/doctos/introd-J084.html.
- Xu, Z.S., L.Q. Xia, M. Chen, R.Y. Zhang, L.C. Li, Y.X. Zhao, Y. Lu, Z.Y. Ni, L. Liu, Z.G. Qiu y Y.Z. Ma. (2007). Isolation and molecular characterization of the *Triticum aestivum* L. Ethylene-responsive factor 1 (*TaERF1*) that increase multiple stress tolerance. *Plant Molecular Biology*, 65, 719-732.
- Yamaguchi-Shinozaki, K., y K. Shinozaki. (1994). A Novel cis-Acting Element in an *Arabidopsis* Gene Is Involved in Responsiveness to Drought, Low temperature, or High-Salt Stress. *The Plant Cell*, 6, 251-264.
- Yi, S.Y., J.H. Kim, Y.H. Joung, S. Lee, W.T. Kim, S.H. Yu y D. Choi. (2004). The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in *Arabidopsis*. *Plant Physiology*, 136, 2862–2874.
- Zavala-León, M.J., J.M.Tun-Suárez, J. Cristóbal-Alejo, E. Ruiz-Sánchez, O. Gutiérrez-Alonso, M. Vázquez-Calderón y R. Méndez-González. (2005). Control postcosecha de la antracnosis en papaya y sensibilidad de *Colletotrichum gloeosporoides* (Penz.) Sacc. a fungicidas organosintéticos. *Revista Chapingo. Serie horticultura*, 11, 251-255.
- Zhang, H., D. Zhang, J. Chen, Y. Yang, Z. Huang, D. Huang, X.C. Wang y R. Huang. (2004). Tomato stress-responsive factor TSRF1 interacts with ethylene responsive element GCC box and regulates pathogen resistance to *Ralstonia solanacearum*. *Plant Molecular Biology*, 55, 825–834.
- Zhang, G., M. Chen, X. Chen, Z. Xu, S. Guan, L. Li, A. Li, J. Guo, L. Mao y Y. Ma. (2008). Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *Journal of Experimental Botany*, 59, 4095–4107.
- Zhang, G., M. Chen, L. Li, Z. Xu, X. Chen, J. Guo, y Y. Ma. (2009). Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *Journal of Experimental Botany*, 60, 3781–3796.

- Zhou, J., X. Tang y G. Martin. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that binds a cis-elements of pathogenesis-related genes. *The EMBO Journal*, 16, 3207-3218.
- Zhu, Y.J., M.M.M. Fitch y P.H. Moore. (2006). "Agrobacterium-mediated transformation of papaya (*Carica papaya* L.)". *Methods in Molecular Biology*. In: *Agrobacterium Protocols: 2/e*, volume 2 (ed. Kan Wang), The Humana Press Inc., Totowa, NJ, 209-217.
- Zuo, K.J., J. Qin, J.Y. Zhao, H. Ling, L.D. Zhang, Y.F. Cao y K.X. Tang. (2007). Over-expression GbERF2 transcription factor in tobacco enhances brown spots disease resistance by activating expression of downstream gene. *Gene*, 391, 80–90.

CAPÍTULO II

CARACTERIZACIÓN *IN SILICO* DE GENES HOMÓLOGOS A FACTORES DE TRANSCRIPCIÓN DE LA FAMILIA ERF EN EL GENOMA SECUENCIADO DE LA PAPAYA TRANSGÉNICA “SUNUP”

2.1 INTRODUCCIÓN

La papaya es uno de los cultivos tropicales de mayor importancia del mundo, se cultiva en regiones tropicales y subtropicales de los cinco continentes produciendo más de 9 millones de toneladas en una superficie de 386,379 Ha (FAO, 2008), de esta cantidad el 50 % se produce en el continente americano donde destacan Brasil con 1.9 millones de toneladas y México con 638,000 toneladas anuales. Este último exporta más de 100,000 toneladas que lo convierten en el principal país exportador de papaya del mundo (FAO, 2008). Sin embargo, como la mayor parte de los cultivos tropicales, la papaya es susceptible a un gran diversidad de enfermedades y plagas en las que destaca la Antracnosis causada por el hongo *Colletotrichum gloeosporioides* (Poghosyan, *et al.*, 2004; Ventura *et al.*, 2004), el cual causa severas pérdidas post-cosecha y un aumento del costo de producción debido al gasto que significa la compra de fungicidas para combatir la enfermedad. En las últimas décadas, el mejoramiento genético basado en la biotecnología, se ha mostrado como una estrategia eficaz en el desarrollo de variedades resistentes a diversos fitopatógenos (Punja, 2001; Campbell *et al.*, 2002; Gurr SJ, y Rushton PJ. 2005a; Gurr SJ, y Rushton PJ. 2005b).

La papaya se ha convertido en un modelo de estudio para otros cultivares frutales con respecto a investigación biotecnológica. Esto se debe en gran parte al desarrollo del cultivo *in vitro* generado durante dos décadas de investigación. (Teixeira da Silva *et al.*, 2007). Además, es la única fruta tropical de la que se han creado variedades transgénicas resistentes a virus y que se han liberado para su comercialización con éxito (Gonsalves, 1998). Ming y colaboradores publicaron en abril de 2008 en la revista Nature la descripción del genoma completo del cultivar transgénico “SunUp” de *Carica papaya*

Linnaeus, lo que convierte a este importante cultivo en la primera fruta tropical en ser secuenciada (Genbank Accession no. ABIM00000000).

La secuenciación de genomas completos como el de *Arabidopsis* ha ayudado mucho en la búsqueda de genes que generen tolerancia a patógenos, dentro de los cuales se destacan los factores de transcripción de la familia ERF, que se han caracterizando en diferentes especies en los últimos años y que por su naturaleza pueden encender genes relacionados con la patogénesis (PR) (Gutterson y Reuber, 2004). Las proteínas ERF son factores de transcripción exclusivos del reino vegetal, se encuentran tanto en plantas monocotiledóneas como dicotiledóneas. Son proteínas que cuentan con un dominio muy conservado llamado AP2/ERF el cuál consta de 59 aminoácidos que adquieren un arreglo tridimensional muy característico formado por tres estructuras β -plegadas en su extremo amino terminal y una α -hélice en su extremo carboxilo. El dominio AP2/ERF es un dominio de unión a ADN que reconoce elementos regulatorios que actúan en cis en las regiones promotoras de genes PR y permiten a la planta defenderse contra el ataque de patógenos biótrofos o necrótrofos (Gutterson y Reuber, 2004; Ohme-Takagi y Shinshi, 1995). Los factores de transcripción ERF responden al ataque de patógenos o a la aplicación exógena de hormonas relacionadas con la defensa de plantas como el ácido salicílico o el etileno (Xu *et al.*, 2008). A fines de los años noventa varios genes del tipo ERF fueron clonados y caracterizados, tal es el caso de los genes *Pti4/5/6* de tomate (Zhou *et al.*, 1997), los cuales interactúan con la proteína de resistencia Pto para activar genes PR que permiten a la planta resistir al ataque la bacteria *Pseudomonas syringae*. Otro gene identificado fue el gen *ERF1* de *Arabidopsis thaliana* (Solano *et al.*, 1998) cuya proteína forma parte de la vía de señalización dependiente de etileno y que se encuentra involucrada en la resistencia contra patógenos necrótrofos, pero también funciona como una ruta que modula el antagonismo entre las rutas de señalización dependientes de ácido salicílico y la ruta dependiente de ácido jasmónico (Van Verk *et al.*, 2009). A partir de entonces, en la última década se han descrito diversos genes que codifican para factores de transcripción tipo ERF que responden a condiciones de estrés. Se han identificado los genes *TaERF1* de trigo, *OsEREBP1* de arroz, *CaPF1* y *CaERFLP1* de Chile y *GmERF3* de soja que pertenecen a un grupo de factores tipo ERF cuya característica estructural es la presencia de un intrón anterior al dominio AP2/ERF y los

cuales están involucrados en la resistencia tanto a estrés biótico como abiótico (Cheong, *et al.*, 2003; Lee, *et al.*, 2004; Yi, *et al.*, 2004; Xu, *et al.*, 2007; Zhang *et al.*, 2009). También se han identificado el gen *Tsi1* de tomate homólogo al gen *Pti6* (Park, *et al.*, 2001); el gen de soya *GmEREBP1* que activa genes de resistencia en respuesta al nematodo *Heterodera glycines* (Mazarei *et al.* 2007); los genes *TSERF1* de tomate y *OPBP1* de tabaco, ambos homólogos al gen *ERF1* (Guo, *et al.*, 2004; Zhang *et al.*, 2004). Los genes *GbERF2* y *HvRAF* en algodón y cebada respectivamente (Jung, *et al.*, 2007; Zuo, *et al.*, 2007). El gen de tabaco *NtERF5* que aumenta la resistencia al virus mosaico del tabaco (Fischer y Droge-Laser, 2004). El caso del gen *JcERF* de *Jatropha curcas* se presenta como el primer gen de la familia ERF clonado de una planta leñosa que también responde a etileno (Tang, *et al.*, 2007). La sobreexpresión de estos genes en plantas modelo como *Arabidopsis* y tabaco ha mostrado que tienen la capacidad de incrementar la resistencia contra diversos fitopatógenos, de ahí el interés de explorar la presencia de homólogos de estos genes en papaya.

La disponibilidad pública del genoma de papaya ha abierto la posibilidad de acelerar la identificación y caracterización de genes tipo ERF por medio de herramientas bioinformáticas y también ha abierto la posibilidad de acelerar al mejoramiento genético de este cultivo mediante la manipulación genética de este tipo de factores de transcripción. Un ejemplo de este mejoramiento sería la generación de nuevas variedades resistentes a patógenos como *C. gloeosporioides*.

Por todo lo anterior, el objetivo del presente capítulo de investigación fue identificar genes de papaya que presenten homología a FT-ERF involucrados en conferir resistencia a fitopatógenos.

2.2 MATERIALES Y MÉTODOS

2.2.1 Identificación de secuencias homólogas a FT-ERF en papaya.

Para identificar probables genes en papaya que codifiquen para FT-ERF que pudieran estar implicados en conferir resistencia a patógenos, se realizó una búsqueda bibliográfica de FT-ERF previamente caracterizados en otras especies; el criterio para elegir las secuencias fue que se haya comprobado por medio de la sobreexpresión de las mismas, que confirieran resistencia a patógenos y/o demostraran una mayor expresión al ser inducidas por un patógeno, ácido salicílico o etileno. De 17 genes involucrados en mecanismos de resistencia a patógenos (Cuadro 1.2), se eligieron cinco secuencias de diferentes especies, con características distintivas entre ellas y que tuvieran mayor información bibliográfica sobre su estructura y función. Las secuencias de aminoácidos de Pti4 de tomate (AAC50047), Tsi1 de tabaco (AAC14323), GmERF3 de soya (ACD47129), ERF1 de *Arabidopsis* (NP_188965) y JcERF de *Jatropha curcas* (AAZ14831) se sometieron como secuencias de referencia (query) en el programa TBLASTN (Gertz *et al.* 2006) del sitio web del National Center for Biotechnology Information (NCBI) con la finalidad de encontrar secuencias homólogas en el genoma de la papaya transgénica "SunUp". Para la búsqueda se utilizaron los parámetros preestablecidos por TBLASTN. Para cada gen de referencia se analizaron las secuencias con mayor puntaje (Score), menor valor E (probabilidad de que el alineamiento BLAST no haya ocurrido al azar) y un porcentaje de identidad (número de nucleótidos iguales) y porcentaje de similitud (número de aminoácidos similares) relativamente altos para cada una de las secuencias de estudio.

2.2.2 Obtención de los marcos de lectura abierta y medición del porcentaje de identidad

A partir de estas secuencias obtenidas mediante el análisis BLAST, se determinaron los posibles marcos de lectura abierta (ORF por sus siglas en inglés) mediante el uso del programa FGENESH (Salamov y Solovyev, 2000). Estas secuencias de ADN fueron

traducidas a secuencia de aminoácidos por medio del programa Bioedit v. 7.0 (Hall, T.A. 1999). Posteriormente se evaluó el porcentaje de identidad entre las secuencias de estudio y las secuencias previamente caracterizadas con el programa MegAlign v. 7.2.

2.2.3 Alineamiento múltiple de secuencias

Se realizó un alineamiento múltiple con las secuencias de aminoácidos utilizando el algoritmo Clustal (Higgins y Sharp, 1988); para esto, cada una de las secuencias tipo ERF de papaya fueron alineadas junto con secuencias caracterizadas previamente. La secuencia de papaya CpERF1 fue alineada con las secuencias Pti4 de tomate (No. Acceso: AAC50047), NtERF1 de tabaco (Q40476) y ERF2 de *Arabidopsis* (BAA32419) que pertenecen al grupo IXa descrito por Nakano (2006). La secuencia CpERF2 fue alineada con las secuencias TSI1 de tabaco (AAC14323), Pti6 de tomate (AAC49741) y CRF4 de *Arabidopsis* (At5g51990), las cuales forman parte del grupo VI. CpERF3, perteneciente al grupo VII, fue alineada con diversas secuencias que incluyen JERF1 de tomate (AAK95687), CaPF1 (AAP72289) y CaERFLP1 (AAS20427) de Chile, NtCEF de tabaco (AAP40022), RAP2.12 de *Arabidopsis* (At1g53910), GmERF3 de soya (ACD47129), OsEREBP1 de arroz (Q6K7E6) y TaERF de trigo (AAX13280). CpERF4 fue alineada con las secuencias OPBP1 de tabaco (U81157), TSRF1 de tomate (AAN32899) y ERF1 de *Arabidopsis* (At3g23240) que se agruparon en el grupo IX dentro del análisis filogenético. Finalmente, la secuencia de papaya CpERF5 fue alineada con las secuencias JcERF de *Jatropha* (AAZ14831), DREB3 de tomate (AAO13360), RAP2.4 (At1g78080) y At1g22190 de *Arabidopsis*. El alineamiento múltiple fue realizado en el software Clustal X (Larkin, *et al.*, 2007) y el sombreado de los aminoácidos conservados se realizó utilizando el programa Boxshade v. 3.21.

2.2.4 Modelaje por homología de la estructura tridimensional de las secuencias tipo ERF de papaya

Para observar si las secuencias de aminoácidos de los factores ERF de papaya podrían presentar una conformación tridimensional similar a la estructura típica del dominio AP2/ERF descrito por Allen *et al.* (1998) se realizó un modelaje por homología de la

estructura tridimensional de las secuencias ERF de papaya utilizando el programa Swiss-Model (Arnold *et al.*, 2006) desarrollado por el instituto suizo de bioinformática. Dicho programa desarrolla una predicción del modelo tridimensional basándose en la homología encontrada por el alineamiento entre la secuencia sometida al programa y la base de datos del "Protein Data Bank", el cual contiene la estructura tridimensional de miles de proteínas que se han modelado por métodos experimentales de espectrofotometría o de resonancia magnética nuclear (Bordoli *et al.*, 2009). Se utilizaron los parámetros predeterminados por el programa.

2.2.5 Análisis filogenético.

Para realizar el análisis filogenético se utilizaron 119 secuencias de aminoácidos de la familia ERF de *Arabidopsis* reportadas por Nakano *et al.* (2006) y recopiladas en el sitio web del "The Arabidopsis Information Resource" (TAIR). Al análisis filogenético se agregaron las 17 secuencias del cuadro 1.2 que corresponden a las secuencias previamente caracterizadas en otras especies y se incluyeron los marcos de lectura abierta de las 5 secuencias de papaya identificadas con el programa TBLASTN. La historia evolutiva fue inferida usando el método de Neighbor-Joining (Saitou y Nei, 1987). El árbol filogenético fue linearizado asumiendo una misma tasa evolutiva para todos los linajes (Takezaki *et al.*, 2004). El árbol fue elaborado a escala, lo que indica que la longitud de las ramas es similar a la distancia evolutiva utilizada para inferir el árbol filogenético. Las distancias evolutivas fueron calculadas utilizando el método de corrección de Poisson (Zuckerkanndl y Pauling, 1965) y fueron calculadas en base a las unidades del número de sustituciones de aminoácidos por sitio. Este análisis filogenético fue realizado utilizando el software MEGA v. 4 (Tamura *et al.*, 2007; Kimura *et al.*, 2008)

2.2.6 Predicción de elementos regulatorios en cis de las regiones promotoras putativas de las secuencias tipo ERF de papaya.

Para localizar elementos regulatorios putativos que actúan en cis de los genes que codifican para los factores de transcripción tipo ERF de papaya, se tomaron aproximadamente 2000 pares de bases anteriores al codón de inicio ATG correspondientes a la región promotora de la transcripción de cada uno de los genes de papaya en estudio. Dicha región fue sometida a la base de datos PLACE (Plant Cis-acting regulatory DNA Elements; Higo *et al.*, 1999). De los datos arrojados por el programa se seleccionaron aquellos relacionados con estrés biótico.

2.3 RESULTADOS

2.3.1 Identificación de secuencias homólogas a FT-ERF en papaya.

El análisis BLAST realizado a partir de las 17 secuencias de referencia presentó resultados semejantes para algunas de ellas, lo que nos permitió identificar 5 contigs (fragmentos secuenciados por el método shotgun) de papaya que presentaban alta similitud con las secuencias de estudio. De esta manera, las secuencias Pti4, Pti5 y GmEREBP1 mostraron similitud con el contig ABIM01004358; las proteínas Pti6 y TSI1 presentaron similitud con el contig ABIM01001349; las secuencias TaERF1, OsEREBP1, CaPF1, CaERFLP1, GmERF3, HvRAF y GBERF2 presentaban mayor similitud con el contig ABIM01025280; La secuencia de ERF1 junto con TSRF1, OPBP1 y NtERF5 exhibieron semejanzas con ABIM01010943 y finalmente solo JcERF mostró semejanza con el contig ABIM01009477. Debido a esta redundancia, de aquí en adelante se presentan los resultados en base a *Pti4*, *TSI1*, *GmERF3*, *ERF1* y *JcERF*, ya que estos genes son los que presentan mayor descripción de su estructura y función en la literatura.

Los cuadros 2.1 a 2.5 muestran los resultados del BLAST para las secuencias Pti4, TSI1, GmERF3, ERF1 y JcERF respectivamente; en ellos se observan los 5 contigs del genoma secuenciado de papaya que presentaron mayor puntaje (Score), menor valor E, mayor porcentaje de identidad y similitud.

Los resultados del análisis en Pti4, ERF1 y JcERF (cuadros 2.1, 2.4 y 2.5) muestran resultados con valores de "Score" mayores de 200 indicando que en dichos contigs se encuentran secuencias similares a las estudiadas. La secuencia de *Jatropha curcas* es la única que genera dos secuencias con valores de score por arriba de 200 y genera la secuencia mas similar en papaya, tal es el caso de "ABIM01009477" con porcentaje de identidad de 68%, de similitud de 76%, un "valor E" de solo $2e^{-80}$ y un valor score de 298 (ver cuadro 2.5). En el caso de los resultados para TSI1 y GmERF3 (Cuadro 2.2 y Cuadro 2.3) no arrojan resultados de score superiores a 200, pero si arrojan resultados entre 80 y 200 que según las características del alineamiento BLAST significan secuencias que

aunque hayan perdido identidad y similitud, mantienen características en común con la secuencia de referencia, en este caso Pti6 y TaERF.

Cuadro 2.1 Resultados de la búsqueda de secuencias homólogas a Pti4 en el genoma de *C. papaya* cv. SunUp, utilizando el programa TBLASTN. Se reportan los cinco primeros resultados.

Número de accesión del GenBank	Longitud del alineamiento TBLASN (Núm. de aminoácidos)	% de identidad	% de similitud	Valor E	Score
ABIM01004358	241	54	66	6e ⁻¹⁰⁴	243
ABIM01016208	131	64	74	6e ⁻³⁴	143
ABIM01017326	98	68	98	3e ⁻³³	141
ABIM01017328	78	69	82	1e ⁻²⁰	115
ABIM01016209	64	75	85	2e ⁻²³	115

Cuadro 2.2 Resultados de la búsqueda de secuencias homólogas a TS11 en el genoma de *C. papaya* cv. SunUp, utilizando el programa TBLASTN. Se Reportan los primeros cinco resultados.

Número de accesión del GenBank	Longitud del alineamiento TBLASN	% de identidad	% de similitud	Valor E	Score
ABIM01001349	264	45	60	3e ⁻⁴⁶	191
ABIM01001452	208	46	59	4e ⁻³⁶	151
ABIM01018971	209	45	58	1e ⁻³⁰	149
ABIM01022576	199	42	58	9e ⁻³³	140
ABIM01012476	78	55	71	2e ⁻²⁰	99

Cuadro 2.3 Resultados de la búsqueda de secuencias homólogas a GmERF3 en el genoma de *C. papaya* cv. SunUp, utilizando el programa TBLASTN. Se Reportan los primeros cinco resultados.

Número de accesión del GenBank	Longitud del alineamiento TBLASN	% de identidad	% de similitud	Valor E	Score
ABIM01025280	159	51	62	4e ⁻³³	142
ABIM01020836	67	83	95	1e ⁻²⁷	123
ABIM01030463	66	77	87	4e ⁻²⁴	112
ABIM01020832	68	73	88	1e ⁻²²	107
ABIM01020835	76	68	81	4e ⁻²²	105

Cuadro 2.4 Resultados de la búsqueda de secuencias homólogas a ERF1 en el genoma de *C. papaya* cv. SunUp, utilizando el programa TBLASTN. Se Reportan los primeros cinco resultados.

Número de accesión del GenBank	Longitud del alineamiento TBLASN	% de identidad	% de similitud	Valor E	Score
ABIM01010943	209	65	77	2e ⁻⁶²	238
ABIM01021369	222	48	63	5e ⁻⁴²	170
ABIM01011742	73	84	93	3e ⁻²³	108
ABIM01004358	93	52	68	2e ⁻²²	105
ABIM01002668	76	63	78	9e ⁻²²	103

Cuadro 2.5 Resultados de la búsqueda de secuencias homólogas a JcERF en el genoma de *C. papaya* cv. SunUp, utilizando el programa TBLASTN. Se Reportan los primeros cinco resultados.

Número de accesión del GenBank	Longitud del alineamiento TBLASN	% de identidad	% de similitud	Valor E	Score
ABIM01009477	230	68	76	2e ⁻⁶⁰	298
ABIM01027225	233	56	70	5e ⁻⁶¹	233
ABIM01009128	64	82	92	4e ⁻²⁷	121
ABIM01015249	127	55	70	5e ⁻²⁷	120
ABIM01014330	123	49	65	1e ⁻¹⁹	96.7

2.3.2 Marcos de lectura abierta y porcentaje de identidad

Las secuencias homólogas a Pti4, TSI1, GmERF3, ERF1 y JcERF fueron denominadas CpERF1, CpERF2, CpERF3, CpERF4 y CpERF5 respectivamente (cuadro 2.6). Todos los ORF's predichos para cada gen de papaya contienen el dominio de unión a ADN AP2/ERF de 59 aminoácidos (figuras 2.1 a 2.5). El ORF con menor longitud fue CpERF4 que cuenta con 627 pb que codifican para una proteína de 208 aminoácidos (aa) (figura 2.4) mientras que la mayor fue CpERF5 con una longitud de 1116 pb que codifican para una proteína de 381 aa (figura 2.5).

El porcentaje de identidad entre los marcos de lectura abierta (ORF) tipo ERF de papaya con respecto a las secuencias Pti4, TSI1, GmERF3, ERF1 y JcERF se muestran en el cuadro 2.6. Para el caso de Pti4, el contig ABIM01004358 contiene un ORF (figura 2.1) que presenta un 58.7% de identidad (cuadro 2.6); El contig ABIM01001349 contiene un ORF (figura 2.2) 50% idéntica a la secuencia de TSI1. El valor más bajo lo presentó la ORF CpERF3 contenida en ABIM01025280 sin embargo esta secuencia no pudo ser predicha en su región 3' y por tanto no presentó el codón de termino (figura 2.3). Por otro lado los valores más altos de identidad fueron reportados en los contigs ABIM01010943 (figura 2.4) y ABIM01009477 (figura 2.5) que contienen ORF's con 68 y 66 % de identidad a ERF1 y JcERF, respectivamente.

Para cada una de las secuencias de estudio, coincidió que el marco de lectura con mayor porcentaje de identidad se encuentra contenido dentro del contig que presentó mayor valor de puntaje en el alineamiento BLAST.

1	atg gcc gta gag tat cac ttt cct cat cat cga agc tca agc ttt	45
1	M A V E Y H F P H H R S S S F	15
46	acc agt ctc atc cct tgt ttg aca gag aac tgg ggt gcc ctg ccg	90
16	T S L I P C L T E N W G A L P	30
91	tta aaa gtt gac gac tct gag gac atg atc ata tac aac tcc ctc	135
31	L K V D D S E D M I I Y N S L	45
136	cgt gac gct ctt cat ttt gga tgg tct ccg tta gat ttg aca gaa	180
46	R D A L H F G W S P L D L T E	60
181	agg aca gtc aaa gcc gag ccc acc gat gag ttg gag atc acg gct	225
61	R T V K A E P T D E L E I T A	75
226	ccg gtc aaa caa gga ata aat cag caa ctt gaa agc gcg tat gtt	270
76	P V K Q G I N Q Q L E S A Y V	90
271	gct cca aag agt gaa aac gga agt gtt tct tca ccg gag aga caa	315
91	A P K S E N G S V S S P E R Q	105
316	gaa aag atc tta cag aag cca aga gga agg cat tat cga ggt gtc	360
106	E K I L Q K P R G R <u>H Y R G V</u>	120
361	agg caa ccg ccg tgg ggg aag ttc gct gct gag att cga gac ccg	405
121	<u>R O R P W G K F A A E I R D P</u>	135
406	gct aaa aac ggc gcg agg gta tgg ctc gga acg tac gaa acg gct	450
136	<u>A K N G A R V W L G T Y E T A</u>	150
451	gag gaa gcg gct ttc gcg tat gac cga gcg gct tat cga atg cga	495
151	<u>E E A A F A Y D R A A Y R M R</u>	165
496	ggc tcg aaa gca ttg ctt aat ttc ccc cac ccg ctt ggc tct aac	540
166	<u>G S K A L L N F P H R L G S N</u>	180
541	gag cct gag ccg gtt aga atc act gcg aaa ccg cga gaa ccg gag	585
181	E P E P V R I T A K R R E P E	195
586	ccc agc acc aga ttg gcc tct gat agt gac ttg gct aag agg ctg	630
196	P S T R L A S D S D L A K R L	210
631	aaa gcg tcg gcg tcg gct aag gaa gct gag ttg gat atg gaa aga	675
211	K A S A S A K E A E L D M E R	225
676	gtg aat aca ctt caa ttt caa tat ccg atg ggt tat gtg cca gtt	720
226	V N T L Q F Q Y R M G Y V P V	240
721	ggc gag caa tta ttg gtt aga tag	744
241	G E Q L L V R *	

Figura 2.1 Marco de lectura abierta de la secuencia CpERF1 situada en el contig ABIM01004358 de papaya var. SunUp. Esta secuencia presentó homología con el gen que codifica para el factor de transcripción Pti4 de tomate. Se subraya el dominio de unión a ADN.

```

1   atg gtg tcc ata cag tcc act ctc aag ctc tct tct cgg gtc ttc 45
1   M  V  S  I  Q  S  T  L  K  L  S  S  R  V  F 15
46  acc acc aac aag ctc ctc cat gat ccc ccg gcg gat gcc aaa gtc 90
46  T  T  N  K  L  L  H  D  P  P  A  D  A  K  V 30
91  cgt cgg aga cgt gtc cag atc ata ctt acc gac gcc gac gcc acc 135
91  R  R  R  R  V  Q  I  I  L  T  D  A  D  A  T 45
136 gat tct tcc agc gag gat gaa gag gaa aga cat ggt ttt cta cga 180
46  D  S  S  S  E  D  E  E  E  R  H  G  F  L  R 60
181 aga gtt aag agg caa gtt caa gtt atc acc ttt gat tca tct tca 225
61  R  V  K  R  Q  V  Q  V  I  T  F  D  S  S  S 75
226 acc cca gac cca cat cca cct cga tct ccg acg cca ctt ccg gct 270
76  T  P  D  P  H  P  P  R  S  P  T  P  L  P  A 90
271 cct act acg tta cct aag caa aag tca cct aga agg aga cca tcg 315
91  P  T  T  L  P  K  Q  K  S  P  R  R  R  P  S 105
316 aga gtg acg aaa tct gac gtc agt cgc ggg caa aaa ttc agg ggt 360
106 R  V  T  K  S  D  V  S  R  G  Q  K  F  R  G 120
361 gtt cgg cag aga cct tgg ggg aag tgg gcc gca gag atc cgg gac 405
121 V  R  O  R  P  W  G  K  W  A  A  E  I  R  D 135
406 cca gcc cag aga aaa cgt atc tgg ctg gga acc ttc aac acc gcc 450
136 P  A  O  R  K  R  I  W  L  G  T  F  N  T  A 150
451 gag gaa gcg gcc gcc gtg tat gac gaa gct gca ttg aag ttt aag 495
151 E  E  A  A  A  A  V  Y  D  E  A  A  L  K  F  K 165
496 ggt tca gac gcc gtt acc aac ttt ccc aac ccg agt gta acg gag 540
166 G  S  D  A  V  T  N  E  P  N  P  S  V  T  E 180
541 ttg acg gct gtg gac agt caa aaa gac gga ttt gac tcg tca agc 585
181 L  T  A  V  D  S  Q  K  D  G  F  D  S  S  S 195
586 aca gcc acg tca tct ccg aca tca gtt ctc cgt tac gag gag ata 630
196 T  A  T  S  S  P  T  S  V  L  R  Y  E  E  I 210
631 aca ccg ttt gat ggc ctc aat tac tgc gac gtt gat acc ttc ggc 675
211 T  P  F  D  G  L  N  Y  C  D  V  D  T  F  G 225
676 ttc gga atc gac gcc gcg tta agc gtg ccg gat ttc atg ttt tcg 720
226 F  G  I  D  A  A  L  S  V  P  D  F  M  F  S 240
721 gga aaa agc ttt gcc gag gaa gag ttc ggg gag ctt gac gtt gac 765
241 G  K  S  F  A  E  E  E  F  G  E  L  D  V  D 255
766 gat ttt ctg gtc gac gcc gtt gcc taa 792
256 D  F  L  V  D  A  V  A  * 263

```

Figura 2.2 Marco de lectura abierta de la secuencia CpERF2 situada en el contig ABIM01001349 de papaya var. SunUp. Esta secuencia presentó homología con el gen que codifica para el factor de transcripción Pti6 de tomate. Se subraya el dominio de unión a ADN.

```

1   atg tgt gga ggt gct ata atc tcc gat ttc ata ccc acc gct act 45
1   M C G G A I I S D F I P T A T 15
46  act cgg tcg tgc aag aat agg aag tct aag aag agc tcc aag agg 90
16  T R S C K N R K S K K S S K R 30
91  tct gag gtt gtc gac ttg gac gat gac ttc gag gct gat ttc cag 135
31  S E V V D L D D D F E A D F Q 45
136 ggc ttc aag gac gac gaa tcc gat att gac gtc gac gaa gat ctc 180
46  G F K D D E S D I D V D E D L 60
181 gag gat att gat gct gtc ttc tct gat att aag ccc ttt gcc ttc 225
61  D D I D A V F S D I K P F A F 75
226 tct gca act cct ctt ccc cgc aaa acc aca gcc tct gct ctc tcc 270
76  S A T P L P R K T T A S A L S 90
271 aat gga tca aag cct gtg aaa gct gtg gaa ttc aat ggg ctg gca 315
91  N G S K P V K A V E F N G L A 105
316 gag aaa tct gca aaa aga aag agg aag aac caa tac agg gga atc 360
106 E K S A K R K R K N Q Y R G I 120
361 agg caa cgc cct tgg ggc aaa tgg gct gct gaa atc cgt gac cca 405
121 R O R P W G K W A A E I R D P 135
406 agg aag ggg gtc agg gtc tgg tta ggt act ttt aac act gct gaa 450
136 R K G V R V W L G T F N T A E 150
451 gaa gct gca aga gct tat gat gct gag gca cgg aga att cgt ggc 495
151 E A A R A Y D A E A R R I R G 165
496 aag aaa gcc aag gtg aac ttc cca gat gaa tct cca cgt gcg tct 540
166 K K A K V N F P D E S P R A S 180
541 cca aag cgt gcc gtg aac tca atg aaa cca gtt gcc aag gca atc 585
181 P K R A V N S M K P V A K A I 195
586 ctg aat tca gca cag cca aat ctg agt cag aat gtt aat tac ttc 630
196 L N S A Q P N L S Q N V N Y F 210
631 aac aac ctt ggt cag gat tac tac aat aca atg gtt ttt gtt gat 675
211 N N L G Q D Y Y N T M V F V D 225
676 gag aaa cca caa atg aat cag ttt gca tca atg aat tca ttt cct 720
226 E K P Q M N Q F A S M N S F P 240
721 gca agt gga aat gct gga gtt aaa ccc ttt gtc cct agt gac aac 765
241 A S G N A G V K P F V P S D N 255
766 acc cat 771
256 T H 257

```

Figura 2.3 Marco de lectura abierta de la secuencia CpERF3 situada en el contig ABIM01025280 de papaya var. sunUp. Esta secuencia presentó homología con el gen que codifica para el factor de transcripción TaERF1 de trigo a pesar de presentarse incompleta en su región 3'. Se subraya el dominio de unión a ADN.

```

1   atg aac tcc tcc tgg aat tca gat ttc tcc tcc gga tca cca gat 45
1   M  N  S  S  W  N  S  D  F  S  S  G  S  P  D  15
46  tct ctc tct ttc gaa aac acc tct ctg cct ttc aac gaa aat gac 90
16  S  L  S  F  E  N  T  S  L  P  F  N  E  N  D  30
91  tcc gag gaa atg ctg ctc tac ggc cta atc gca gaa gcc aat cat 135
31  S  E  E  M  L  L  Y  G  L  I  A  E  A  N  H  45
136 gaa tcg tcg gag tcg acg tat tca aac atg atg atc aag gag gag 180
46  E  S  S  E  S  T  Y  S  N  M  M  I  K  E  E  60
181 gag gtg agt tcc gtg gaa gat ccg aag cag gaa agg tcg tac aga 225
61  E  V  S  S  V  E  D  P  K  Q  E  R  S  Y  B  75
226 ggg gtg agg agg cgg cca tgg ggg aaa ttt gca gca gaa att agg 270
76  G  V  R  R  R  P  W  G  K  F  A  A  E  T  R  90
271 gat tcc acg aga cat gga ata aga gtg tgg tta ggt act ttt gat 315
91  D  S  T  R  H  G  I  R  V  W  L  G  T  F  D  105
316 agt gca gag gca gca gct ctt gcg tat gat caa gcg gcg ttt tct 360
106 S  A  E  A  A  A  L  A  Y  D  O  A  A  F  S  120
361 atg aga gga tca tcg gcg att ctg aat ttt ccg gtg gag aga gtt 405
121 M  R  G  S  S  A  I  L  N  F  P  V  E  R  V  135
406 gag gag tcg ctg agg gag atg aag tgt aac tac gag gaa ggt tgc 450
136 E  E  S  L  R  E  M  K  C  N  Y  E  E  G  C  150
451 tcg cct gtt gtg gct ctg aag aga aaa cac tcg atg agg agg aag 495
151 S  P  V  V  A  L  K  R  K  H  S  M  R  R  K  165
496 atg tgt aac agg aag aag aga gag agg gat gtg agg ata gag aat 540
166 M  C  N  R  K  K  R  E  R  D  V  R  I  E  N  180
541 gta gtg gtt ttc gaa gat ttg gga gca gac tat ttg gag cag ctc 585
181 V  V  V  F  E  D  L  G  A  D  Y  L  E  Q  L  195
586 ttg act tcg tcg tct tct tct gaa atc act atg cct tgg tga 627
196 L  T  S  S  S  S  S  S  E  I  T  M  P  W  *  208

```

Figura 2.4 Marco de lectura abierta de la secuencia CpERF4 situada en el contig ABIM01010943 de papaya var. SunUp. Esta secuencia presentó homología con el gen que codifica para el factor de transcripción ERF1 de *Arabidopsis thaliana*. Se subraya el dominio de unión a ADN.

1	atg gca gct acc atg gat ttt tgc agt agc aga cca gtt caa gca	45
1	M A A T M D F C S S R P V Q A	15
46	ggt ccc ttc ggg gac gag ctt atg gaa gca ctt gag cct ttt att	90
16	G P F G D E L M E A L E P F I	30
91	aaa agt gct tta tct tct tct cct cct cct ctc tcc tca tct	135
31	K S A L S S S P S P P L S S S	45
136	tac cag att cca tct tgt act tcc tcc tct agt tat tct tct gcc	180
46	Y Q I P S C T S S S S Y S S A	60
181	aat gac atc tct ttc tgc cct tca ttt tct acc aac aca aca caa	225
61	N D I S F C P S F S T N T T Q	75
226	ccc agt cta tac tcg ggt gat tgc tgt acc acc tct atg cct caa	270
76	P S L Y S G D C C T T S M P Q	90
271	cct att tca aac gcg cac tcg gtt cat gat ccg tcg gct ctc cag	315
91	P I S N A H S V H D P S A L Q	105
316	caa tcg ggg tcg atc ggg ctg aat cac ctc act cgg tct cag atg	360
106	Q S G S I G L N H L T R S Q M	120
361	aac cag atc cag tcg caa ttt cac ctc cag agt caa caa agc ccg	405
121	N Q I Q S Q F H L Q S Q Q S P	135
406	tct tta ctg tac caa tat tat cct caa cag cag cat gcc ttt cag	450
136	S L L Y Q Y Y P Q Q H A F Q	150
451	ttc ctc agc cca aag cct gtt cca atg aag caa gtt ggt tcc ccc	495
151	F L S P K P V P M K Q V G S P	165
496	cca aag ccc acc aag ctc tac agg ggt gtg agg cag cgg cat tgg	540
166	P K P T K <u>L Y R G V R O R H W</u>	180
541	ggc aag tgg gtc gcc gag atc cgc ctg ccg aag aac cgg acc cgg	585
181	<u>G K W V A E I R L P K N R T R</u>	195
586	ctt tgg ctc ggc acc ttc gat acg gcc gag gaa gca gcc tta gct	630
196	<u>L W L G T F D T A E E A A L A</u>	210
631	tac gac aag gca gct tac aag ctc cga ggt gac ttc gct cgg ctc	675
211	<u>Y D K A A Y K L R G D F A R L</u>	225
676	aac ttc ccc aat ctc cgc cac cgg gga tcc cac atc gac ggc gag	720
226	<u>N F P N L R H R G S H I D G E</u>	240
721	ttc ggc cag tac aag ccc ctt cat tcc tca gtg gac gca aag ctt	765
241	F G Q Y K P L H S S V D A K L	255
766	gat gct atc tgt gaa agt tta gca gag tcg cag aag cag ggg aaa	810
256	D A I C E S L A E S Q K Q G K	270
811	gtg ggg aag caa cat gtt ggc tcc ggc aag aag cga gcc agg cct	855
271	V G K Q H V G S G K K R A R P	285
856	cct cgt atg gag cca gag gtt gag cca ccg cag gct ata caa ggt	900
286	P R M E P E V E P P Q A I Q G	300
901	tca gat tcg aga acg tta gag acg gtt aag ggc gag aag aat tcc	945
301	S D S R T L E T V K G E K N S	315
946	tcc ccg tcg ccg gtg atg acc gag agc gat ggg tct gca gga tct	990
316	S P S P V M T E S D G S A G S	330
991	tca cct tta tcc gag atc acg ttc ggt gaa atg gac aac gag ccg	1035
331	S P L S E I T F G E M D N E P	345
1036	cag tgg agt att gtg ccg gag aac ttc atg ttg cag aag tac cct	1080
346	Q W S I V P E N F M L Q K Y P	360
1081	tca tac gag atc gat tgg gcc tcc ata ttg tca tga	1116
361	S Y E I D W A S I L S *	381

Figura 2.5 Marco de lectura abierta de la secuencia CpERF5 situada en el contig ABIM01009477 de papaya var. SunUp. Esta secuencia presentó homología con el gen que codifica para el factor de transcripción JcERF de *Jatropha curcas*. Se subraya el dominio de unión a ADN.

Cuadro 2.6 Porcentajes de identidad entre secuencias tipo ERF de papaya y cinco secuencias ERF previamente caracterizadas (Pti4, TSI1, GmERF3, ERF1 y JcERF). Los porcentajes de identidad se determinaron con el programa MegAlign (Lasergene ver. 7.2.1). Se resaltan las secuencias de papaya que presentaron mayor porcentaje de identidad con respecto a las secuencias ERF caracterizadas.

Nombre de la secuencia/ No. Acceso al GenBank del contig de papaya	Pti4	TSI1	GmERF3	ERF1	JcERF
CpERF1	58.7	24.5	25.1	32.1	27.6
CpERF2	25.5	50.0	22.6	23.5	26.1
CpERF3	29.2	27.3	45.7	26.9	26.3
CpERF4	33.9	23.3	23.0	68.2	28.5
CpERF5	22.8	21.0	16.3	27.2	66.1

2.3.4 Alineamiento múltiple clustal

Los alineamientos múltiples de los grupos donde se encuentran los factores Pti4, TSI1, GmERF3, ERF1 y JcERF junto con sus homólogos en papaya son mostrados en las figuras 2.7, 2.8, 2.9, 2.10 y 2.11, respectivamente. En dichos alineamientos se muestra la conservación del dominio de unión al DNA AP2/ERF (Resaltando las tres hojas beta y la alfa hélice característicos de este dominio), así como los sitios de localización nuclear (NLS; marcados con triángulos negros)

En el caso del grupo VII se resalta el motivo "MCGGAIISD" caracterizado en los factores de transcripción CaPF1 (Young, *et al.*, 2004) y TaERF1 (Xu, *et al.*, 2007) de función aun desconocida. Este motivo también se encuentra conservado en el homólogo de papaya CpERF3 (ABIMO1025280) (figura 2.8).

En todos los casos se puede observar como los homólogos encontrados en papaya conservan el mismo arreglo entre el dominio AP2/ERF con los aminoácidos que hacen contacto con las cajas GCC (Resaltados con cuadros negros en el alineamiento; Allen *et al.*, 1998), y los NLS.

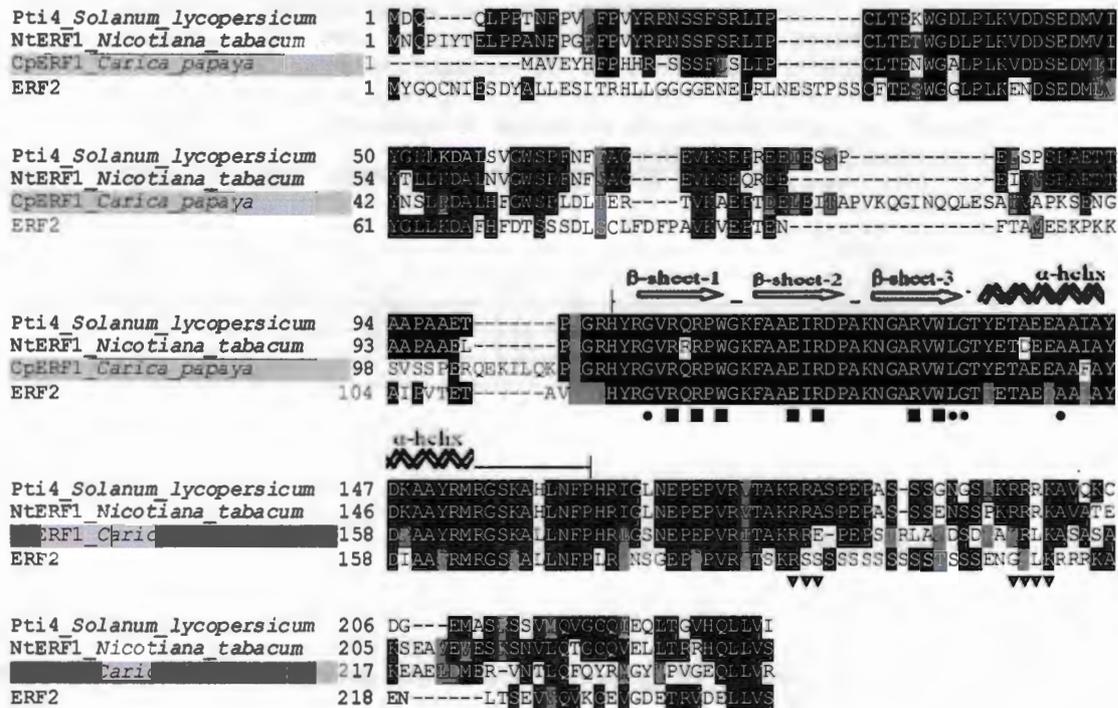


Figura 2.6 Alineamiento clustal X en el que se muestra las similitudes que comparte la proteína CpERF1 con otras proteínas pertenecientes al grupo IXa. Se resalta el dominio de unión AP2/ERF indicando las posiciones de las estructuras β-plegada y α-hélices. Se hacen notar también los residuos de aminoácidos que hacen contacto con el ADN en las cajas GCC (cuadros negros), los aminoácidos conservados en todas las secuencias tipo AP2/ERF (círculos negros) y los sitios de localización nuclear (triángulos negros).

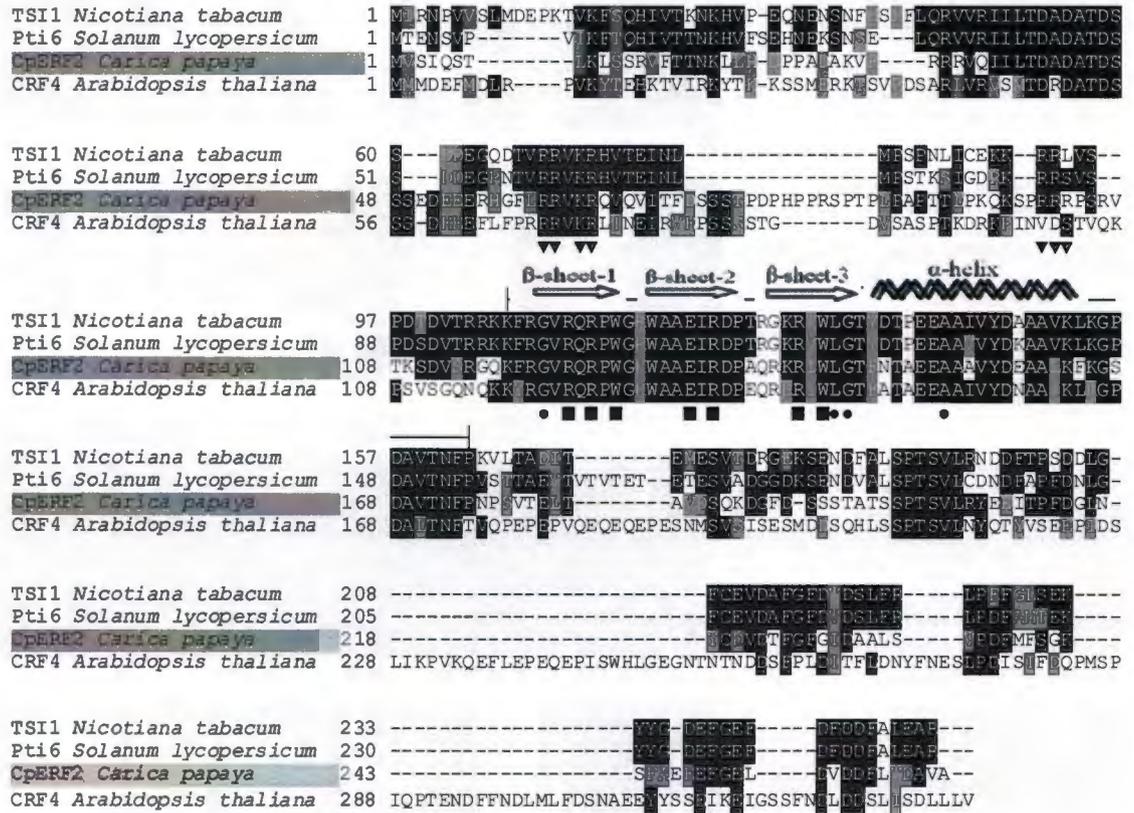


Figura 2.7 Alineamiento clustal X en el que se muestra las similitudes que comparte la proteína CpERF2 con otras proteínas pertenecientes al grupo VI. Se resalta el dominio de unión AP2/ERF indicando las posiciones de las estructuras β-plegada y α-hélices. Se hacen notar también los residuos de aminoácidos que hacen contacto con el ADN en las cajas GCC (cuadros negros), los aminoácidos conservados en todas las secuencias tipo AP2/ERF (círculos negros) y los sitios de localización nuclear (triángulos negros).

Figura 2.8 En la página anterior se observa el alineamiento clustal X en el que se muestra las similitudes que comparte la proteína CpERF3 con otras proteínas pertenecientes al grupo VII. Se resalta el dominio de unión AP2/ERF indicando las posiciones de las estructuras β -plegada y α -hélices. Se hacen notar también los residuos de aminoácidos que hacen contacto con el ADN en las cajas GCC (cuadros negros), los aminoácidos conservados en todas las secuencias tipo AP2/ERF (círculos negros) y los sitios de localización nuclear (triángulos negros).

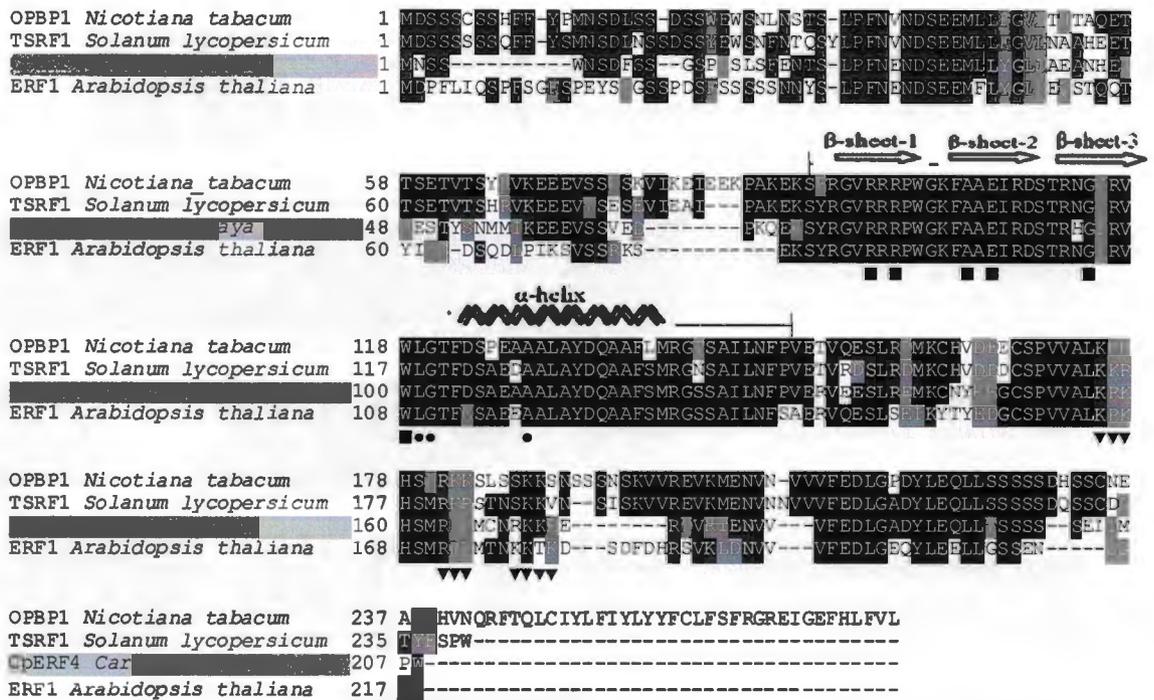


Figura 2.9 Alineamiento clustal X en el que se muestra las similitudes que comparte la proteína CpERF4 con otras proteínas pertenecientes al grupo IXc. Se resalta el dominio de unión AP2/ERF indicando las posiciones de las estructuras β -plegada y α -hélices. Se hacen notar también los residuos de aminoácidos que hacen contacto con el ADN en las cajas GCC (cuadros negros), los aminoácidos conservados en todas las secuencias tipo AP2/ERF (círculos negros) y los sitios de localización nuclear (triángulos negros).

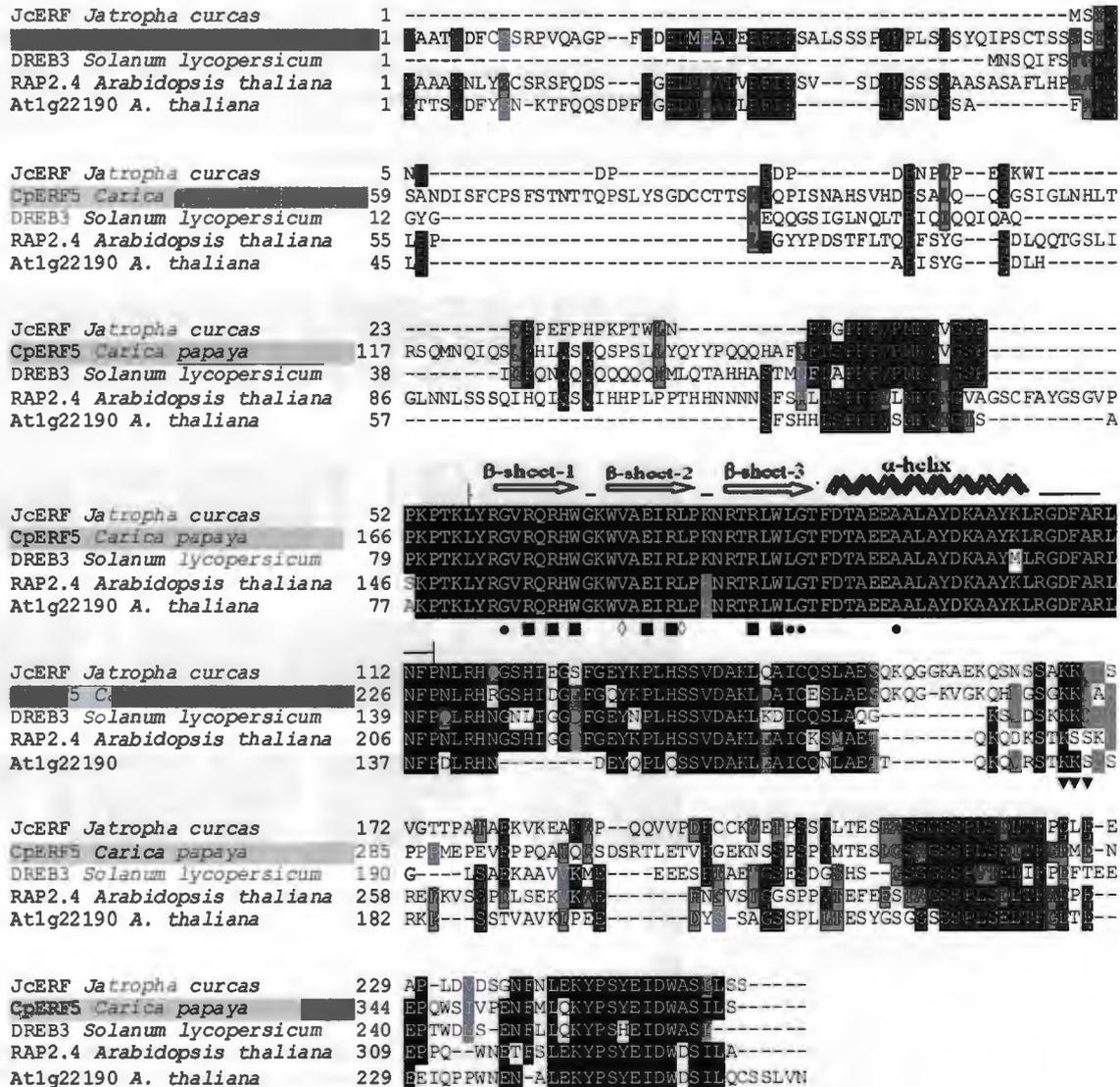


Figura 2.10 Alineamiento clustal X en el que se muestra las similitudes que comparte la proteína CpERF5 con otras proteínas pertenecientes al grupo I. Se resalta el dominio de unión AP2/ERF indicando las posiciones de las estructuras β-plegada y α-hélices. Se hacen notar también los residuos de aminoácidos que hacen contacto con el ADN en las cajas GCC (cuadros negros), los aminoácidos conservados en todas las secuencias tipo AP2/ERF (círculos negros) y los aminoácidos de los FT DREB que hacen contacto los elementos CRT/DRE para activar genes de resistencia a estrés abiótico (rombos blancos).

2.3.5 Estructura tridimensional de las Proteínas CpERF1-CpERF5

La estructura terciaria del dominio AP2/ERF de CpERF1-CpERF5 fue predicha por el programa Swiss-Model. En las cinco proteínas tipo ERF de papaya se observó que las tres estructuras β -Plegadas localizadas en la región amino terminal se arreglan de manera paralela a la estructura de α -hélice localizada en la región carboxilo terminal. Según Allen *et al.* (1998), el arreglo conformacional que adquieren las estructuras β -Plegadas son de gran importancia para la unión de este dominio con los elementos regulatorios en *cis* del tipo "GCC". Los modelos 3D se obtuvieron tomando como molde a la secuencia homóloga AtERF1 (Protein Data Bank ID: 1GCC) de *Arabidopsis thaliana*. La estructura 3D de esta secuencia fue confirmada por estudios de resonancia magnética nuclear heteronuclear y multidimensional (Allen *et al.*, 1998).

2.3.6 Análisis filogenético.

El agrupamiento filogenético por medio de "Neighbor-Joining" muestra como las secuencias de papaya (resaltadas en círculos negros) se agrupan dentro de los mismos clados que las secuencias de otras especies que generan tolerancia a patógenos (rombos grises). De esta manera Pti4 y su homólogo en papaya (CpERF1) se agrupan dentro del Grupo IX con un índice de distancia genética de aproximadamente 0.025. CpERF4 se encuentra agrupado con CpERF1 en el grupo IX junto con las secuencias TSREF1, OPBP1 y ERF1. Por su parte Pti6 y Tsi1 se agrupan junto con la secuencia de papaya CpERF2 en el grupo VI. Mientras que en el grupo VII se aglutinan la mayor cantidad de factores tipo ERF que se ha comprobado su actividad para conferir resistencia a patógenos (OsEREBP1, TaERF1, CaPF1, GmERF3 y CaERFLP1) junto con la secuencia de papaya CpERF3 y finalmente JcERF y CpERF5 se encuentran en el grupo I. Este árbol filogenético concuerda en gran parte con el estudio previo realizado por Xu y colaboradores (2008).

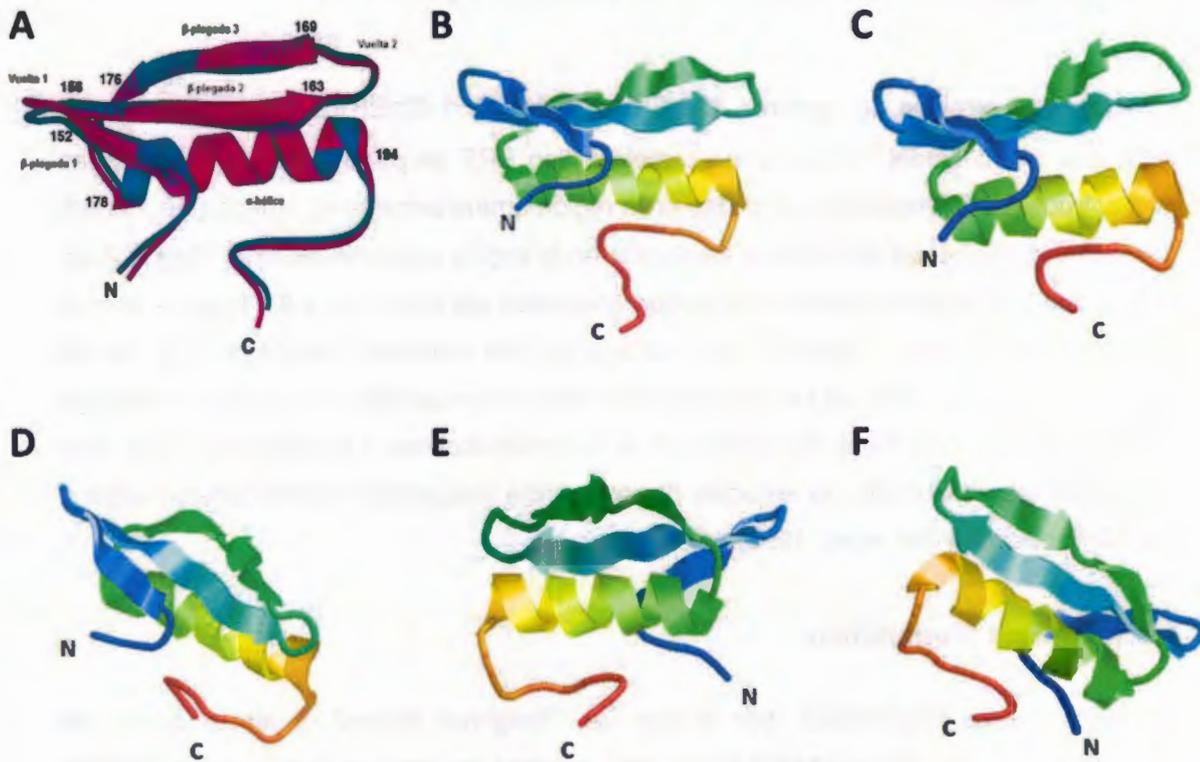


Figura 2.11 En la figura se observa el análisis de la estructura tridimensional de los dominios de unión AP2/ERF de cada secuencia de estudio. **A.** Estructura del dominio AP2/ERF de AtERF1 descrito por Allen *et al.*, 1998. **B.** CpERF1; **C.** CpERF2, **D.** CpERF3, **E.** CpERF4 **F.** CpERF5. En todos los casos se pueden apreciar las 3 estructuras β -plegada y la estructura α -hélice características de este dominio.

2.3.7 Elementos Regulatorios en cis predichos en las regiones promotoras putativas de los genes CpERF1-CpERF5

Se encontraron al menos 8 tipos distintos de elementos regulatorios que se han visto implicados en estrés biótico. Dos de ellos, los elementos Wbox y los elementos GT1 Consensus, fueron encontrados en las regiones promotoras de las cinco secuencias. Las regiones promotoras de los genes CpERF1 y CpERF4 contienen varios elementos en la región próxima al codón de inicio mientras que CpERF2, CpERF3 y CpERF5 los tienen más dispersos en la zona estudiada de 2000 pb anterior al codón de inicio. Solo la región promotora del gen CpERF5 mostró elementos del tipo GCC que pudieran estar implicados en una autorregulación o una regulación guiada por otros factores tipo ERF.

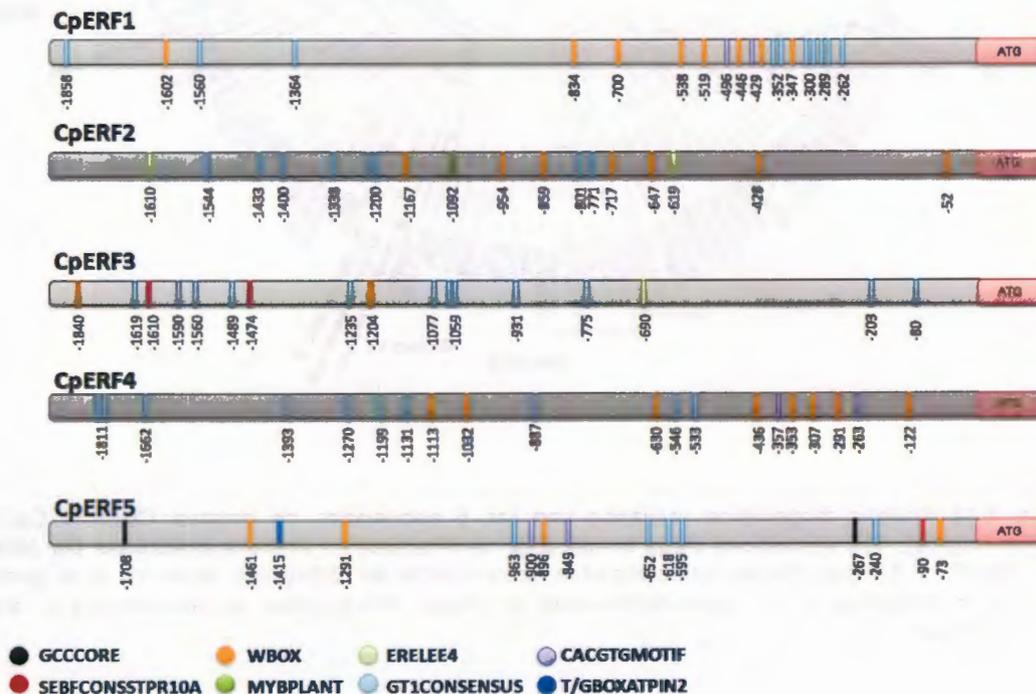


Figura 2.13 Elementos regulatorios que actúan en Cis predichos con el programa PLACE, dentro de las regiones promotoras de los genes *CpERF1-CpERF5* de papaya.

2.4 DISCUSIÓN

Uno de los mayores problemas que presenta el cultivo de la papaya es el relacionado con la pérdida a nivel postcosecha debido al ataque de especies de *Colletotrichum* que causan la enfermedad de la Antracnosis (Poghosyan *et al.*, 2005; Ventura *et al.*, 2004). Cuando Ming y colaboradores publicaron en 2008 la secuencia genómica de *Carica papaya* se abrió la posibilidad de analizar su genoma en busca de genes que permitan el desarrollo de nuevas variedades que resuelvan problemas específicos de esta importante especie como puede ser el desarrollo de variedades resistentes a la antracnosis. El genoma de papaya consta de 372 Mega bases (Mb) que contienen un total de 24,746 genes, de los cuales cerca del 10% (2438 genes) corresponden a genes que codifican para factores de transcripción de 66 familias distintas (Ming *et al.*, 2008). Los factores de transcripción son los reguladores naturales de los procesos de transcripción del genoma, es por esto que el estudio de la función y regulación de estas proteínas se presenta como una opción atractiva para desarrollar cultivos biotecnológicos con nuevas características de interés agronómico, como es el caso de la resistencia a enfermedades (Century *et al.*, 2008). En *Arabidopsis* se ha descubierto que algunos genes de las familias de factores de transcripción ERF, Myb, TGA, Whirly y WRKY participan en la regulación de diferentes rutas que ayudan a las plantas a defenderse contra el ataque de diversos patógenos (Euglem, 2005). Dentro de estas familias se encuentran los factores de transcripción ERF, los cuales se ha visto que se unen a “elementos en cis GCC” que se encuentran en las regiones promotoras de genes PR regulando de esta manera la respuesta de las plantas a estrés biótico (Gutterson y Reuber, 2004). La papaya cuenta con un total de 146 genes que codifican para factores de transcripción de la familia ERF (Ming *et al.*, 2008). En un esfuerzo por identificar secuencias de la familia ERF en papaya que pudieran estar implicadas en el mecanismo de resistencia a patógenos; en este estudio se identificaron cinco secuencias tipo ERF de papaya que presentaron homología con genes cuya función ya ha sido relacionada con la resistencia a patógenos de plantas.

2.4.1 CpERF1: un gen homólogo a Pti4 posiblemente implicado en la interacción con una proteína R del tipo PTO

El gen que codifica para el factor de transcripción Pti4 de tomate fue el primer gen de la familia ERF cuya función fue relacionada con la resistencia de las plantas a estrés biótico, ya que esta proteína es fosforilada por la proteína cinasa Pto que actúa como proteína R (Zhou *et al.*, 1997). La fosforilación de Pti4 aumenta su afinidad de unión a promotores de genes de defensa activados en la ruta del ácido salicílico. Desde entonces este gen se ha convertido en un modelo de estudio para elucidar la función de factores de transcripción tipo ERF en la defensa de las plantas contra patógenos (Xu *et al.*, 2008). El gen de papaya nombrado en este estudio como CpERF1 pertenece junto con el gen Pti4 al grupo IXa de la clasificación de los factores de transcripción ERF propuesta por Nakano *et al.* (2006). Cuenta con el motivo CMIX-3, que es una región de 33 aminoácidos anterior al dominio AP2/ERF (Nakano *et al.*, 2006) que en el caso de CpERF1 comienza en el aminoácido 21. Dicho motivo ha sido identificado como una región ácida en Pti4 (Zhou *et al.*, 1997) que podría funcionar como un posible dominio de transactivación al ser fosforilado por Pto, por lo que este motivo CMIX-3 de CpERF1 podría ser fosforilado por un homólogo de Pto de papaya (Gu *et al.*, 2000). El dominio AP2/ERF de CpERF1 sólo varía en tres aminoácidos con respecto al dominio de unión a ADN de Pti4. Dichas variaciones no se encuentran en los aminoácidos que hacen contacto con elementos Cis del tipo GCC de los genes PR (Allen *et al.*, 1998) por lo que se esperaría que el gen CpERF1 mantuviera esta cualidad. Así mismo, la secuencia CpERF1 cuenta con dos regiones (aa 178-180; aa 196-200) ricas en argininas y lisinas las cuales pueden funcionar como sitios de localización nuclear para esta proteína (Dingwall y Laskey, 1991; Christophe *et al.*, 2000).

En el análisis de identidad realizado con el programa MegAlign se obtiene un 58.7% de identidad, entre la secuencia completa de aminoácidos de la proteína Pti4 y la secuencia completa de papaya CpERF1. Dicho porcentaje se incrementa a un 80% en el motivo CMIX-3 descrito por Nakano y a un 95% en el dominio de unión a ADN AP2/ERF. Este alto porcentaje de identidad entre los motivos de estas secuencias sugiere que la secuencia de papaya CpERF1 mantiene las mismas funciones bioquímicas que la proteína Pti4, ya que en diferentes estudios se ha mostrado que la tasa de evolución

dentro de una secuencia de proteínas ocurre más lento en regiones que le generan funcionalidad a la proteína (Simon *et al.*, 2002) como es el caso del dominio AP2/ERF.

2.4.2 El gen CpERF2 presenta homología con Tsi1 y podría conferir resistencia a estrés biótico y abiótico.

Pti6 y Tsi1 son genes de tomate y tabaco respectivamente que les fue estudiada su capacidad de generar tolerancia a estrés biótico y abiótico (Zhou *et al.*, 1997; Gu *et al.*, 2000; Park, *et al.*, 2001; Gu, *et al.*, 2002; Shin *et al.*, 2002). De acuerdo al alineamiento BLAST utilizando estas proteínas como referencia, se encontró que existe una secuencia dentro del contig ABIM01001349, que presentó 45% de identidad y 60% de similitud con respecto a Pti6. El marco de lectura para este gen de papaya denominado en este estudio CpERF2, cuenta con 50% de identidad con respecto a Pti6 y se agrupa en el análisis filogenético dentro del grupo VI junto con las proteínas Pti6 y Tsi1, en concordancia con el estudio de Nakano y colaboradores (2006). A pesar de estar filogenéticamente cercano a Pti6 y Tsi1, CpERF2 ha sufrido un alto número de mutaciones que hacen que su distancia evolutiva de 0.15 sea relativamente alta dentro del clado. Sin embargo, en el alineamiento múltiple se aprecia que el dominio AP2/ERF se mantiene conservado y mantiene un 77.6% de identidad con respecto a las dos secuencias de referencia; también conserva los 7 aminoácidos descritos por Allen (1998) dentro de las regiones β -Plegadas que le ayudan a unirse a los elementos *cis* del tipo "caja GCC", por lo que se esperaría que mantuviera la función de regular genes PR como en el caso de Pti6 y Tsi1 (Zhou *et al.*, 1997; Gu *et al.*, 2000; Park, *et al.*, 2001). El dominio AP2/ERF de CpERF2 no cuenta con los aminoácidos valina 14 ni ácido glutámico 19 que son importantes para la unión de estas proteínas con elementos *cis* que contengan la secuencia DRE/CRT (TACCGACAT) presentes en las regiones promotoras de genes que responden a sequía o bajas temperaturas (Sakuma *et al.*, 2002), sin embargo se ha comprobado por ensayos de movilidad que el gen Tsi1 se une a este tipo de elementos (Park, *et al.*, 2001) por lo que el gen CpERF2 también podría estar involucrado en conferir tolerancia a estrés abiótico. Nakano describe otros 2 motivos conservados dentro del grupo VI. El motivo CMVI-1 que se mantiene conservado en un 66% en la secuencia de CpERF2 y cuya función es desconocida, y el motivo CMVI-2 rico en argininas y lisinas (RRVKR) que se ha observado

funciona como sitio de localización nuclear en Tsi1 (Park, *et al.*, 2001). Dicho motivo esta conservado en un 100% en CpERF2 con respecto a Pti6 y a Tsi1.

En un principio se había reportado que los últimos 15 aminoácidos de la proteína Tsi1 podrían servir como un sitio de activación transcripcional al unirse con otra proteína (Park, *et al.*, 2001). Sin embargo, al reportarse la proteína Tsp1 que es un dedo de zinc del tipo DnaJ y que sirve como activador transcripcional de Tsi1, se observó que si bien esos 15 aminoácidos no eran necesarios para la unión de Tsi1 con Tsp1, el sitio de unión entre ambas proteínas se encuentra en la región carboxilo posterior al dominio AP2/ERF (Ham *et al.*, 2006). En el 2006 al hacer un estudio más detallado de los motivos conservados de los diferentes grupos de la familia ERF; Nakano *et al.* (2006) reportan el motivo CMVI-3 en la región carboxilo del grupo VI, que en el caso de CpERF2 consta de la secuencia "SPTSVLRY" que se encuentra a partir del aminoácido 200 y que esta conservado en un 87% con respecto a la secuencia Tsi1 y en un 75% con respecto a la secuencia de Pti6. Dicho motivo CMVI-3, al estar tan conservado en este grupo, podría ser la secuencia que interactúa con la proteína Tsp1.

2.4.3 El gen CpERF3 homologo al grupo VII y su posible interacción con proteínas cinasas en la generación de resistencia a estrés biótico y abiótico

El grupo VII es uno de los grupos más interesantes en cuestiones de estrés biótico. La mayor cantidad de genes tipo ERF, que al ser sobreexpresados en plantas de *Arabidopsis* o tabaco generan resistencia a patógenos, se encuentran dentro de este clado. El marco de lectura abierta de CpERF3 presenta el dominio AP2/ERF 100% idéntico al dominio del factor de soya GmERF3. También cuenta con el motivo MCGGAI(I/L), característico de todos los Factores de transcripción ERF del grupo VII (Tournier *et al.*, 2003; Nakano *et al.*, 2006; Xu *et al.*, 2008). El marco de lectura abierta se muestra incompleto debido a que el ORF se encuentra en el extremo del contig de papaya con número de acceso ABIM01025280, por lo que no fue posible por métodos bioinformáticos averiguar la secuencia de la región carboxilo terminal. Sin embargo, la región amino está altamente conservada y no solo mantiene el dominio AP2/ERF y el motivo MCGGAI(I/L) sino que también cuenta con el motivo CMVII-3 el cual contiene residuos de lisina y arginina

(Nakano *et al.*, 2006). El motivo CMVII-3 se encuentra pegado al extremo amino del dominio AP2/ERF y se ha observado que en los genes CaPF1 y TaERF1 funciona como sitio de localización nuclear (Yi *et al.*, 2004; Xu *et al.*, 2007). Según los datos arrojados por el alineamiento BLAST, el gen *CpERF3* cuenta con 2 intrones localizados en la región amino anterior al dominio AP2/ERF. El grupo VII tiene dentro sus características el poseer un intrón en la región amino, anterior al dominio de unión al ADN (Nakano *et al.*, 2006; Xu *et al.*, 2008); dicho intrón también ha sido reportado en el gen de trigo TaERF1 (Xu, *et al.*, 2007). El gen de arroz *OsEBP-89* es el único caso reportado de factores de transcripción del grupo VII que cuenta con dos intrones, sin embargo uno de estos se encuentra del lado amino del dominio AP2/ERF y el otro se encuentra del lado carboxilo; por lo que, de corroborarse en ensayos *in vitro* sería el primer caso de un miembro de este grupo que presenta dos intrones en la región amino terminal. En el aislamiento *in vitro* de la secuencia *CpERF3* se esperaría que la región carboxilo, posterior al dominio de unión a ADN, tuviera el motivo de fosforilación TPD/EI/YS/TS, el cual se encuentra descrito para las proteínas OsEREBP1, CaPF1 y TaERF (Cheong, *et al.*, 2003; Yi *et al.*, 2004; Xu *et al.*, 2007) y que se encuentra en la mayoría de los miembros del grupo VII (Nakano, *et al.*, 2006; Xu, *et al.*, 2008). Dicho motivo de fosforilación se une a una cinasa del tipo BWMK1 que se encuentra en arroz (Cheong, *et al.*, 2003) y que también ha sido descrita en trigo (TaMPK1; Xu *et al.*, 2007), por lo que se esperaría una cinasa homóloga dentro del genoma de papaya. Se ha visto que los miembros del grupo VII se unen tanto a elementos GCC como elementos DRE/CRT, dicha funcionalidad les permite transcribir tanto genes PR como genes COR y de esta manera generar tolerancia tanto a estrés biótico como abiótico (Cheong, *et al.*, 2003; Yi *et al.*, 2004; Xu *et al.*, 2007). Es por esto que, en ensayos de funcionalidad se esperaría que, el gen *CpERF3* pudiera tener esta función dual y generar tolerancia a estrés salino o bajas temperaturas, al igual que conferir resistencia a patógenos.

2.4.4 *CpERF4* el homólogo a *ERF1* de *Arabidopsis* podría estar implicado en la ruta de señalización del etileno y ácido salicílico.

Al igual que *Pti4*, el gen *ERF1* de *Arabidopsis thaliana* fue de los primeros factores de transcripción tipo ERF en comprobarse su actividad de conferir resistencia a patógenos.

Sin embargo a diferencia de Pti4 que es parte de una ruta que va desde el reconocimiento de una proteína efectora por parte de la proteína cinasa Pto que fosforila a Pti4 para aumentar su afinidad por las regiones GCC de algunos genes PR (Zhou *et al.*, 1997), el factor de transcripción ERF1 se ha visto que actúa en la ruta de señalización del etileno y que su activación esta mediada por la dimerización de EIN3 que actúa como factor de transcripción de ERF1 y este a su vez se une a las cajas GCC de algunos genes PR (Solano *et al.*, 1998). Al analizar el genoma de papaya mediante BLAST se pudo localizar la secuencia de CpERF4, cuyo marco de lectura abierta mantiene una identidad con ERF1 de 68%. Ambas secuencias forman parte del grupo IXc descrito por Nakano y colaboradores (2006) y se agrupan dentro del análisis filogenético junto con los factores OPBP1 de tabaco y TSRF1 de tomate, a los que también se ha comprobado su actividad contra patógenos (Guo *et al.*, 2004; Zhang *et al.*, 2004). El dominio de unión AP2/ERF presenta 85% de identidad entre todas las secuencias del alineamiento y mantiene en CpERF4 los siete aminoácidos que hacen contacto con los elementos GCC por lo que se esperaría que el homólogo de papaya también activara genes PR. Los miembros del grupo IXc contienen en su región carboxilo terminal un motivo de 22 aminoácidos llamado CMIX-1 (Nakano *et al.*, 2006) que en el caso de la secuencia de papaya comienza en el aminoácido 180. El motivo CMIX-1 de CpERF4 posee 72% de identidad con ERF1, 90% con OPBP1 y 95% con TSRF1, y a pesar de no haberse estudiado, el alto grado de conservación de este motivo dentro del grupo IXc, supone que posiblemente genera alguna interacción con otra proteína. Así mismo, en el alineamiento se pueden identificar otras regiones conservadas; el motivo CMIX-3 que comienza en el aminoácido 20 de la secuencia de CpERF4, mantiene un 60% de identidad con las secuencias del alineamiento. Este motivo no se ha estudiado tampoco pero se piensa que funciona como un sitio de transactivación (Gutterson y Reuber, 2004); el motivo CMIX-4 descrito por Nakano *et al.*, (2006) contiene residuos de arginina y lisina que tanto en TSRF1 como en OPBP1 confieren localización nuclear. Por todas estas evidencias podemos suponer que al realizarse estudios de funcionalidad de CpERF4, este gen mostrará características funcionales similares a la de los factores ERF1, TSRF1 y OPBP1.

2.4.5 CpERF5 y su homólogo JcERF son miembros de la subfamilia DREB que podrían conferir resistencia a estrés biótico y abiótico.

En el año 2007 Tang y colaboradores describieron al factor de transcripción JcERF de *Jatropha curcas*, en dicho estudio y en base sólo a la secuencia de aminoácidos del dominio AP2/ERF; dicho gen fue catalogado como miembro del grupo B-3 de la subfamilia ERF (Sakuma *et al.*, 2002; Grupo IX en la clasificación de Nakano *et al.*, 2006). En un alineamiento de secuencias, se observó que el factor JcERF compartía un bajo porcentaje de similitud con los miembros del grupo B-3 (entre 12% y 18%; Tang *et al.*, 2007), sin embargo en el presente estudio en base al alineamiento de la secuencia completa de aminoácidos se pudo constatar que la secuencia JcERF pertenece al grupo I de la clasificación de Nakano *et al.* (grupo B-2 de la clasificación de Sakuma *et al.*, 2002). El grupo I pertenece a la subfamilia DREB de los factores ERF que se caracterizan por tener un residuo de valina en la posición 14 y un residuo de ácido glutámico en la posición 19 que les dan la capacidad de unirse a elementos regulatorios del tipo DRE/CRT que contienen la secuencia TACCGACAT (Sakuma *et al.*, 2002). El grupo I, incluyendo la secuencia CpERF5, además de contar con una alta similitud en su dominio AP2/ERF, cuenta con motivos del tipo CMI-1, CMI-2 y CMI-4 localizados en posiciones adyacentes al dominio AP2/ERF (Nakano, *et al.*, 2006). La secuencia del dominio AP2/ERF de CpERF5 mantiene el residuo de valina en la posición 14, sin embargo presenta una leucina en vez de ácido glutámico en la posición 19. El residuo de V14 parece ser más importante para la unión del dominio AP2/ERF a los sitios DRE/CRT (Sakuma *et al.*, 2002) y según el estudio de Li y colaboradores (2005) la mutación de ácido glutámico por leucina en la posición 19, aún permite la unión entre el dominio AP2/ERF y los elementos DRE/CRT, por lo que suponemos que el gen CpERF5 también muestre esta característica. Hasta la fecha el factor JcERF es el único miembro del grupo I que se ha demostrado una función regulatoria en plantas sometidas a estrés. La sobreexpresión de JcERF en plantas de *Arabidopsis* mostró su habilidad para conferir tolerancia a salinidad y bajas temperaturas por lo que se espera que el gen de papaya CpERF5 también cuente con la misma funcionalidad. Por ahora, no se ha comprobado que JcERF o algún otro miembro del grupo I confiera resistencia a patógenos, sin embargo la característica de JcERF de ser inducido por etileno y unirse a elementos GCC (Tang *et al.*, 2007) hace

suponer que tanto JcERF como su homólogo en papaya CpERF5 puedan tener la habilidad dual de conferir resistencia tanto a estés biótico como abiótico.

2.4.6 Posible regulación de la expresión de los factores CpERF1-5

Existen pocos estudios en cuanto a la regulación transcripcional de los factores de transcripción ERF. En el análisis realizado por PLACE se observó que los 5 genes de papaya cuentan en sus regiones promotoras con elementos regulatorios del tipo W BOX (T)(T)TGAC(C/T). Se ha visto que dichos elementos son el sitio de unión de factores de transcripción del tipo WRKY (Ishiguro y Nakamura, 1994; Eugelm *et al.*, 2000; Li *et al.*, 2004). En algunos estudios se ha visto que existen este tipo de elementos W BOX en las regiones promotoras de factores de transcripción de la familia ERF, tal es el caso de AtERF4 que se piensa podría estar regulado por factores de transcripción WRKY (McGrath *et al.*, 2005), también se ha visto un tipo de regulación positiva de la transcripción de la proteína ERF3 llevada a cabo por inducción por heridas o daño mecánico que conllevan a la expresión de factores tipo WRKY que finalmente regulan la transcripción de ERF3 (Nishiuchi *et al.*, 2004). En el mismo estudio se observó que las W boxes más próximas al codón de inicio son más activas que las distantes (Nishiuchi *et al.*, 2004), por lo que se esperaría que CpERF1 y CpERF4 pudieran tener una mayor regulación por estos factores ya que cuentan con varias secuencias W box en las regiones próximas al codón de inicio con respecto a las secuencias CpERF2, CpERF3 y CpERF5 que las tienen en posiciones más distantes.

Otros elementos cis conservados en las regiones promotoras de los cinco genes de papaya son los elementos GT. Los elementos GT generalmente se encuentran en tándem y se encuentran en las regiones promotoras de genes que codifican para proteínas de diversas funciones (Zhou *et al.*, 1999). Como podemos apreciar en la figura 2.13 varios de los elementos GT encontrados en las regiones de los genes CpERF1-5 se encuentran en tándem. Lo que podría explicarse por el hecho de que algunos factores GT se unen al ADN como dímeros o bien porque se pegan en diferentes sitios al mismo tiempo (Zhou *et al.*, 1999). Se ha visto que los elementos GT se encuentran en la región promotora del gen PR-1a de tabaco contribuyendo a la expresión basal de este gen (Buchell *et al.*,

1996). También se ha observado que la unión de los factores GT con sus sitios de unión se reduce en presencia del ácido salicílico, por lo que podría ser que estos elementos estén implicados en la regulación de la transcripción de los genes CpERF1-5 en presencia de ácido salicílico (Buchell *et al.*, 1996; Gu *et al.*, 2000).

Otros elementos que se encontraron en las regiones promotoras y que están relacionados con patogénesis son los elementos "G Box", que contienen un motivo CACGTG y que se ha visto están en regiones promotoras de genes PR (Kim *et al.*, 1992; Menkens *et al.*, 1995). A estas cajas G también se ha visto que se puede unir Pti4 (Chakravarthy *et al.*, 2003) por lo que podrían estar implicados en la autorregulación de estos mismos factores. Los dos elementos GCC que se encontraron en la región promotora de CpERF5 también podrían estar implicados en una autorregulación o una regulación guiada por otro factor tipo AP2/ERF (Chakravarthy *et al.*, 2003).

Otros elementos que se encontraron en menor medida fueron el MYBPlant que se encuentra en algunos genes PR a los que se unen factores MYB y algunos relacionados con la respuesta hipersensible (Daniel *et al.*, 1999; Chakravarthy *et al.*, 2003), el motivo ERELEE4 que está involucrado en la activación de genes inducidos por etileno (Tapia *et al.*, 2005) y el T/GBox que tiene actividad en la ruta del jasmonato (Boter *et al.*, 2004).

2.5 CONCLUSIONES

1. Los análisis bioinformáticos mostraron que las cinco secuencias encontradas dentro del genoma de la variedad SunUp de papaya nombradas en este estudio como CpERF1-CpERF5 presentan una estructura primaria, secundaria y terciaria muy similar a la de factores de transcripción ERF funcionalmente caracterizados. El dominio de unión a ADN (AP2/ERF), algunos motivos que generan interacción con otras proteínas y los motivos de localización nuclear se encuentran altamente conservados en las secuencias de papaya, incluyendo los sitios de interacción del dominio AP2/ERF con elementos regulatorios GCC que inducen la expresión de genes PR.

2. El análisis filogenético confirmó que los genes CpERF1-CpERF5 de papaya son homólogos a los genes descritos en estudios previos, lo que hace suponer que estos genes codifican para factores de transcripción tipo ERF que están regulados por las rutas de señalización de defensa dependientes de ácido salicílico y/o etileno.

3. Las regiones promotoras putativas de los genes CpERF1-CpERF5 tienen varios tipos de elementos regulatorios implicados de alguna manera con mecanismos de defensa en plantas. Lo anterior indica que dichos genes posiblemente estén regulados por rutas bioquímicas inducidas por estrés biótico.

2.6 BIBLIOGRAFÍA

- Allen, M., y K. Yamasaki. (1998). A novel mode of DNA recognition by a b-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, 17, 5484–5496.
- Arnold, K., L. Bordoli, J. Kopp y T. Schwede. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22, 195-201.
- Bordoli, L., F. Kiefer, K. Arnold, P. Benkert, J. Battey y T. Schwede. (2009). Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*, 4, 1-13.
- Boter, M., O. Ruiz-Rivero, A. Abdeen y S. Prat. (2004). Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and *Arabidopsis*. *Genes and Development*, 18, 1577-1591.
- Buchel, A., R. Molenkamp, J. Bol y H. Linthorst. (1996). The PR-la promoter contains a number of elements that bind GT-I-like nuclear factors with different affinity. *Plant Molecular Biology*, 30, 493-504.
- Campbell, M.A, H.A. Fitzgerald y PC. Ronald. (2002). Engineering pathogen resistance in crop plants. *Transgenic Research*, 11, 599–613.
- Century, K., T. Lynne-Reuber y O. Ratcliffe (2008). Regulating the Regulators: The Future Prospects for Transcription-Factor-Based Agricultural Biotechnology Products. *Plant Physiology*, 147, 20–29.
- Chakravarthy, S., R. Touri, M. D'Ascenzo, P. Fobert, C. Després y G. Martin (2003). The Tomato Transcription Factor Pti4 Regulates Defense-Related Gene Expression via GCC Box and Non-GCC Box *cis* Elements. *The Plant Cell*, 15, 3033–3050.
- Cheong, Y.H., B.C. Moon, J.K. Kim, M.C. Kim, I.H. Kim, C.Y. Park, J.C. Kim, B.O. Park, S.C. Koo, H.W. Yoon, W.S. Chung, C.O. Lim, S.Y. Lee y M.J. Cho. (2003). BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcriptor factor. *Plant Physiology*, 132, 1961-1972.
- Christophe, D., C. Christophe-Hobertus, B. Pichon. (2000). Nuclear targeting of proteins: how many different signals?. *Cellular Signalling*, 12, 337–341

- Daniel, X., C. Lacomme, J.B. Morel, y D. Roby. (1999). A novel *myb* oncogene homologue in *Arabidopsis thaliana* related to hypersensitive cell death. *The Plant Journal*, 20, 57–66.
- Dingwall, C., y R. Laskey. (1991). Nuclear targeting sequences - a consensus?. *Trends in Biochemical Sciences*, 16, 478-81.
- Eulgem, T., P.J. Rushton, S. Robatzek, y I.E. Somssich, (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Sciences*, 5, 199–206.
- Eulgem, T. (2005). Regulation of the *Arabidopsis* defense transcriptome. *Trends in Plant Science*, 10, 71-78.
- FAO, 2008. Estadísticas de FAO. Sistema FAOSTAT.
Disponible en: <http://faostat.fao.org/site/567/default.aspx#ancor>
[Acceso 10 Agosto 2010].
- Fischer, U., y Dröge-Laser W. (2004). Overexpression of *NtERF5*, a New Member of the Tobacco Ethylene Response Transcription Factor Family Enhances Resistance to *Tobacco mosaic virus*. *Molecular Plant-Microbe Interaction*, 17, 1162–1171.
- Gertz, M., Y.K. Yu, R. Agarwala, A.A. Schäffery y S.F. Altschul. (2006). Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biology*, 4, 1-14.
- Gonsalves, D. (1998). Control of papaya ringspot virus in papaya: a case study. *Annual Review of Phytopathology*, 36, 415–437.
- Gu, Y.Q., C. Yang, T. Venkatappa, J. Zhou, y G. Martin. (2000). *Pti4* Is Induced by Ethylene and Salicylic Acid, and Its Product Is Phosphorylated by the Pto Kinase. *The Plant Cell*, 12, 771–785.
- Gu, Y.Q., M.C. Wildermuth, S. Chakravarthy, Y.T. Loh, C. Yang, X. He, Y. Han, G.B. Martin. (2002). Tomato transcription factors *pti4*, *pti5*, and *pti6* activate defense responses when expressed in *Arabidopsis*. *The Plant Cell*, 14, 817–831.
- Guo, Z.J., X.J. Chen, X.L. Wu, J.Q. Ling, y P. Xu. (2004). Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. *Plant Molecular Biology*, 55, 607–618.
- Gurr, S.J., y P.J. Rushton. (2005a). Engineering plants with increased disease resistance: what are we going to express?. *TRENDS in Biotechnology*, 23, 275-282.

- Gurr, S.J. y P.J. Rushton. (2005b). Engineering plants with increased disease resistance: how are we going to express it?. *TRENDS in Biotechnology*, 23, 283-290.
- Gutterson, N. y T.L. Reuber. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Current Opinion in Plant Biology*, 7, 465-471.
- Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.
- Ham, B., J. Park, S. Lee, M. Kim, I. Lee, K. Kim, C. Kwon y K. Paek. (2006). Tobacco Tsp1, a DnaJ-Type Zn Finger Protein, Is Recruited to and Potentiates Tsi1-Mediated Transcriptional Activation. *The Plant Cell*, 18, 2005-2020.
- Higgins, D.G. y P.M. Sharp. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, 73, 237-244.
- Higo, K., Y. Ugawa, M. Iwamoto y T. Korenaga. (1999). Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research*, 27, 297-300.
- Ishiguro, S. y K. Nakamura. (1994). Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 59 upstream regions of genes coding for sporamin and α -amylase from sweetpotato. *Molecular and General Genetics*, 244, 563-571.
- Jung, J., S.Y. Won, S.C. Suh, H.R. Kim, R. Wing, Y. Jeong, I. Hwang y M. Kim. (2007). The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in *Arabidopsis*. *Planta*, 225, 575-588.
- Kim, S.R., J.L. Choi, M.A. Costa y G. An. (1992). Identification of G-box sequence as an essential element for methyl jasmonate response of potato proteinase inhibitor II promoter. *Plant Physiology*, 99, 627-631.
- Kumar, S, J. Dudley, M. Nei y K. Tamura. (2008). MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*, 9, 299-306.
- Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson y D.G. Higgins. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948.
- Lee, J.H., J.P. Hong, S.K. Oh, S. Lee, D. Choi y W.T. Kim. (2004). The ethylene-responsive factor like protein 1 (CaERFLP1) of hot pepper (*Capsicum annuum* L.)

- interacts in vitro with both GCC and DRE/CRT sequences with different binding affinities: Possible biological roles of CaERFLP1 in response to pathogen infection and high salinity conditions in transgenic tobacco plants. *Plant Molecular Biology*, 55, 61–81.
- Li, J., G. Brader y E. Palva. (2004). The WRKY70 Transcription Factor: A Node of Convergence for Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense. *The Plant Cell*, 16, 319–331.
- Li, X.P, A.G. Tian, G.Z. Luo, Z.Z. Gong, J.S. Zhang y S.Y. Chen. (2005). Soybean DRE-binding transcription factors that are responsive to abiotic stresses. *Theoretical and Applied Genetics*, 110, 1355–1362.
- Maleck, K., A. Levine, T. Eulgem, A. Morgan, J. Schmid, K. Lawton, J.L. Dang y R.A. Dietrich. (2000). The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nature Genetics*, 26, 403–410.
- Mazarei, M., A. Elling, T. Maier, D. Puthoff y T. Baum. (2007). GmEREBP1 Is a Transcription Factor Activating Defense Genes in Soybean and *Arabidopsis*. *Molecular Plant-Microbe Interactions*, 20, 107–119.
- McGrath, K., B. Dombrecht, J. Manners, P. Schenk, C. Edgar, D. Maclean, W. Scheible, M. Udvardi y K. Kazan. (2005). Repressor and Activator Type Ethylene Response Factors Functioning in Jasmonate Signaling and Disease Resistance Identified via a Genome-Wide Screen of *Arabidopsis* Transcription Factor Gene Expression. *Plant Physiology*, 139, 949–959.
- Menkens, A., U. Schindler y A. Cashmore. (1995). The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends in Biochemical Sciences*, 20, 506–510.
- Ming, R., S. Hou, Y. Feng, Q. Yu, A. Dionne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis, S.L. Salzberg, L. Feng, M.R. Jones, R.L. Skelton, J.E. Murray, C. Chen, W. Qian, J. Shen, P. Du, M. Eustice, E. Tong, H. Tang, E. Lyons, R.E. Paull, T.P. Michael, K. Wall, D.W. Rice, H. Albert, M.L. Wang, Y.J. Zhu, M. Schatz, N. Nagarajan, R.A. Acob, P. Guan, A. Blas, C.M. Wai, C.M. Ackerman, Y. Ren, C. Liu, J. Wang, J. Wang, J.K. Na, E.V. Shakhov, B. Haas, J. Thimmapuram, D. Nelson, X. Wang, J.E. Bowers, A.R. Gschwend, A.L. Delche, R. Singh, J.Y. Suzuki, S. Tripathi, K. Neupane, H. Wei, B. Irikura, M. Paidi, N. Jiang, W. Zhang, G. Presting, A. Windsor, R. Navajas-Pérez, M.J. Torres, F.A. Feltus, B. Porter, Y. Li, A.M. Burroughs, M.C. Luo, L. Liu, D.A. Christopher, S.M. Mount, P.H. Moore, T. Sugimura, J. Jiang, M.A. Schuler, V. Friedman, T. Mitchell-Olds, D.E. Shippen, C.W. dePamphilis, J.D. Palmer, M. Freeling, A.H. Paterson, D. Gonsalves, L. Wang y M. Alam. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452, 991–996.

- Nakano, T., K. Suzuki, T. Fujimura, y H. Shinshi. (2006). Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. *Plant Physiology*, 140, 411–432.
- Nishiuchi, T., H. Shinshi y K. Suzuki. (2004). Rapid and Transient Activation of Transcription of the ERF3 Gene by Wounding in Tobacco Leaves. *The Journal of Biological Chemistry*, 279, 55355–55361.
- Ohme-Takagi, M., y H. Shinshi. (1995). Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element. *The Plant Cell*, 7, 173-182.
- Park, J.M., C.J. Park, S.B. Lee, B.K. Ham, R. Shin, K.H. Paek. (2001) Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell*, 13, 1035-1046.
- Poghosyan, A., G. Zehfuss, y M. Jimenez. (2005). Papaya Diseases: Ecology and Control. *Encyclopedia of Pest Management*, 1, 1–5.
- Punja, Z.K. (2001). Genetic engineering of plants to enhance resistance to fungal pathogens—a review of progress and future prospects. *The Canadian Journal of Plant Pathology*, 23, 216–235.
- Saitou, N. y M. Nei. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution*, 4, 406–425.
- Sakuma, Y, Q. Liu, J.G. Dubouzet, H. Abe, K. Shinozaki y K. Yamaguchi-Shinozaki. (2002). DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold-inducible gene expression. *Biochemical and Biophysical Research Communications*, 290, 998–1009.
- Salamov, A. y V. Solovyev. (2000). Ab initio gene finding in *Drosophila* genomic DNA. *Genome Research*, 10, 516-522.
- Shin, R., J. Park, J. An y K. Paek. (2002). Ectopic Expression of Tsi1 in Transgenic Hot Pepper Plants Enhances Host Resistance to Viral, Bacterial, and Oomycete Pathogens. *Molecular Plant-Microbe Interactions*, 15, 983–989.
- Simon, A., E. Stone y A. Sidow. (2002). Inference of functional regions in proteins by quantification of evolutionary constraints. *Proceedings of the National Academy of Sciences*, 99, 2912–2917.

- Solano, R., A. Stepanova, A. Chao y J. Ecker. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ethylene-insensitive3 and ethylene-response-factor1. *Genes and Development*, 12, 3703–3714.
- Takezaki, N., A. Rzhetsky y M. Nei. (2004). Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution*, 12, 823-833.
- Tamura, K., J. Dudley, M. Nei y S. Kumar. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596-1599.
- Tang, M., J. Sun, Y. Liu, F. Chen y S. Shen. (2007). Isolation and functional characterization of the JcERF gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant *Jatropha curcas*. *Plant Molecular Biology*, 63, 419-428.
- Tapia, G., I. Verdugo, M. Yañez, I. Ahumada, C. Theoduloz, C. Cordero, F. Poblete, E. González y S. Ruiz-Lara. (2005). Involvement of Ethylene in Stress-Induced Expression of the TLC1.1 Retrotransposon from *Lycopersicon chilense* Dun. *Plant Physiology*, 138, 2075–2086.
- Teixeira da Silva, J., Z. Rashid, D. Tan Nhut, D. Sivakumar, A. Gera, M. Teixeira Souza y P. Tennant. (2007). Papaya (*Carica papaya* L.) Biology and Biotechnology. *Tree and Forestry Science and Biotechnology*, 1, 47-63.
- Tournier, B., M.T. Sanchez-Ballesta, B. Jones, E. Pesquet, F. Regad, A. Latche, J.C. Pech y M. Bouzayen. (2003). New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. *FEBS Letters*, 550, 149–154.
- Van Loon, L.C., M. Rep y C.M.J. Pieterse. (2006). Significance of Inducible Defense-related Proteins in Infected Plants. *Annual Review of Phytopathology*, 44, 135–62.
- Van Verk, M., C. Gatz y H. Linthorst. (2009). Transcriptional Regulation of Plant Defense Responses. *Advances in Botanical Research*, 51, 397-438.
- Ventura, J.A., H. Costa y J. da Silva-Tatagiba. (2004). Papaya diseases and integrated control. In: Naqvi SAMH (Ed) *Diseases of Fruits and Vegetables* (Vol II), Kluwer Academic Publishers, The Netherlands, 201-268p.
- Xu, Z.S., L.Q. Xia, M. Chen, R.Y. Zhang, L.C. Li, Y.X. Zhao, Y. Lu, Z.Y. Ni, L. Liu, Z.G. Qiu y Y.Z. Ma. (2007). Isolation and molecular characterization of the *Triticum aestivum*

- L. Ethylene-responsive factor 1 (*TaERF1*) that increase multiple stress tolerance. *Plant Molecular Biology*, 65, 719-732.
- Xu, Z.S., M. Chen, L.C. Li y Y.Z. Ma. (2008). Functions of the ERF transcription factor family in plants. *Botany*, 86, 969-977.
- Yi, S.Y., J.H. Kim, Y.H. Joung, S. Lee, W.T. Kim, S.H. Yu y D. Choi. (2004). The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in *Arabidopsis*. *Plant Physiology*, 136, 2862-2874.
- Zhang, H., D. Zhang, J. Chen, Y. Yang, Z. Huang, D. Huang, X.C. Wang y R. Huang. (2004). Tomato stress-responsive factor TSRF1 interacts with ethylene responsive element GCC box and regulates pathogen resistance to *Ralstonia solanacearum*. *Plant Molecular Biology*, 55, 825-834.
- Zhang, G., M. Chen, L. Li, Z. Xu, X. Chen, J. Guo y Y. Ma. (2009). Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *Journal of Experimental Botany*, 60, 3781-3796.
- Zhou, J., X. Tang y G. Martin. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that binds a cis-elements of pathogenesis-related genes. *The EMBO Journal*, 16, 3207-3218.
- Zhou, D.Z. (1999). Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends in plant sciences*, 4, 210-214.
- Zuckerlandl, E. y L. Pauling. (1965). Molecules as Documents of Evolutionary History. *Journal of Theoretical Biology*, 8, 357-366.
- Zuo, K.-J., J. Qin, J.Y. Zhao, H. Ling, L.D. Zhang, Y.F. Cao y K.X. Tang. (2007). Over-expression GbERF2 transcription factor in tobacco enhances brown spots disease resistance by activating expression of downstream gene. *Gene*, 391, 80-90.

CAPÍTULO III

ANÁLISIS DE LA EXPRESIÓN DE LOS GENES *CpERF1-CpERF5* Y MAPEO DE LOS EXTREMOS 5' Y 3' DE SUS CORRESPONDIENTES ADNc EN PAPAYA MARADOL

3.1 INTRODUCCIÓN

A finales de los años 80's se introdujeron a México las primeras semillas de la variedad cubana Maradol, a través de la CONAFRUT, en Xalapa Veracruz y desde entonces este cultivar ha ido desplazando paulatinamente a las variedades criollas hasta ser casi exclusivamente la única variedad que se siembra comercialmente en nuestro país (Claridades Agropecuarias, 1999; Massieu-Trigo, 2009). La papaya Maradol se destaca por su buen sabor y alto valor nutricional, además de su larga vida de anaquel y su resistencia al transporte en largas distancias (Massieu-Trigo, 2009). México es el principal productor y exportador de esta variedad, por lo que en el contexto internacional la papaya Maradol suele ser conocida también como papaya Mexicana (Fundación Produce Chiapas, 2003). La papaya Maradol se ha convertido en uno de los cultivos más importantes del país y se encuentra dentro de la lista de sistemas producto integrados de SAGARPA.

A pesar de que los virus (en especial el virus de la mancha anular) son el principal problema en el cultivo de papaya, es la enfermedad de la antracnosis la principal limitante en la comercialización del fruto, ya que el microorganismo disminuye considerablemente la calidad del producto a nivel post-cosecha y le hace perder su valor comercial (Massieu-Trigo, 2009). De esta manera, la manipulación controlada de genes involucrados en la resistencia a patógenos representa una alternativa atractiva para aumentar la resistencia de papaya a patógenos como *Colletotrichum* que es el agente causal de la antracnosis.

Los factores de transcripción ERF son polipéptidos altamente conservados en el reino vegetal y están presentes tanto en plantas dicotiledóneas como monocotiledóneas. Diversos estudios han comprobado que varios genes de diferentes especies miembros de

esta familia aumentan sus niveles de expresión cuando las plantas son sometidas a estímulos de estrés biótico, ya sea por el ataque de un patógeno (Gutterson y Reuber, 2004) o por la aplicación exógena de hormonas involucradas en la resistencia ácido salicílico, ácido jasmónico o etileno. Este aumento en los niveles de expresión conlleva al aumento de la expresión de genes PR que a su vez permiten a la planta resistir el ataque de un patógeno (Gutterson y Reuber, 2004; Van Verk et al., 2009).

Para el reconocimiento de patógenos, el sistema inmune de plantas mantiene activada una vasta red de proteínas implicadas en el reconocimiento del patógeno o en la activación de proteínas de defensa que en su conjunto permiten a la planta defenderse de la mayor parte de microorganismos con los que está en contacto (interacción incompatible). Sin embargo, cuando los patógenos logran evadir esta respuesta basal y su presencia no es reconocida por los mecanismos moleculares de la planta se produce la enfermedad (interacción compatible). En los últimos años se ha observado que en las interacciones planta-patógeno tanto compatibles como incompatibles se ejerce una serie de regulaciones tanto positivas como negativas de diversos genes implicados en estas interacciones. En muchos casos la diferencia entre la susceptibilidad o resistencia a un patógeno se encuentra asociado con las diferencias en tiempo y magnitud en que ocurren los cambios de expresión de estos genes. Por lo que una respuesta rápida en la regulación de estos genes provoca un aumento en la resistencia al patógeno (Maleck et al., 2000; Van Loon et al., 2006).

En el capítulo anterior se pudieron identificar 5 genes del genoma de la papaya SunUp, que mostraron homología con secuencias tipo ERF que al ser sobre-expresadas en plantas transgénicas generan resistencia a patógenos (cuadro 1.2). Estas secuencias nombradas como CpERF1-CpERF5 conservan el dominio AP2/ERF que hace contacto con elementos GCC de las regiones promotoras de genes PR. También mostraron motivos que podrían funcionar como sitios de localización nuclear, y otros motivos que podrían estar implicados en la interacción con otras proteínas incluyendo proteínas de resistencia (R) del tipo cinasa.

El objetivo del presente capítulo fue investigar la expresión basal de los genes CpERF1-CpERF5 en papaya Maradol. Por medio de RT-PCR se observó la expresión de estos cinco genes en diferentes tejidos de plantas adultas provenientes de una plantación comercial. Posteriormente, se logró la clonación y secuenciación de los ADNc de las cinco secuencias de estudio mediante la técnica de RACE. Los ADNc de CpERF1-CpERF5 fueron comparados con sus correspondientes ORF's predichos y se mapearon los UTRs de cada una de las secuencias de estudio.

3.2 MATERIALES Y MÉTODOS

3.2.1 Material vegetal

El material vegetal de este trabajo consistió de muestras de peciolo, hoja, flor y fruto inmaduro de individuos de papaya Maradol cultivados en la plantación del Rancho Alegre ubicado en el municipio de Tizimín en el estado de Yucatán. Las muestras colectadas fueron congeladas en nitrógeno líquido y almacenadas a -80°C para su posterior uso.

3.2.2 Extracción de ácido ribonucleico

El ARN total de las muestras estudiadas fue aislado mediante el protocolo de Kieffer y colaboradores (2000). Entre 100 y 200 mg de tejido fueron macerados en nitrógeno líquido con un mortero y pistilo. El tejido pulverizado se colocó en un tubo eppendorf de 1.5 ml y se agregaron 600µl de buffer de extracción (CTAB 2%, PVP40 2%, EDTA 25mM, Tris HCl 100mM y NaCl 2M) y 12µl de β-mercaptoetanol. La muestra se incubó a 65°C por 10 minutos. Posteriormente, los ácidos nucleicos se separaron del material proteico agregando 600µl de cloroformo/alcohol isoamílico (24:1), se separaron las fases centrifugando 5 min a 14000 rpm, se tomó el sobrenadante y se precipitó el ácido nucleico con isopropanol frío incubando la muestra en hielo por 5 minutos. La pastilla fue resuspendida en 39 µl de H₂O Ultra pura, seguidamente el ADN fue degradado agregando 5 µl de DNasa (RQ1 RNase-Free DNase 1 µg/µl; Promega), 5 µl de Buffer de la DNasa (RQ1 DNase 10X Reaction Buffer; Promega) y 1 µl de inhibidor de RNasas (RNase OUT

– 40U/μl; Invitrogen). La muestra se incubó con la enzima durante 30 minutos a 37°C. Posteriormente, la DNasa fue retirada mediante un lavado agregando 700 μl del detergente SSTE y 700 μl de Cloroformo-Alcohol isoamílico (24:1), se separaron las fases centrifugando 5 min a 14000 rpm y se transfirió el sobrenadante a un tubo nuevo, seguidamente se precipitó el ARN agregando isopropanol frío y se incubó en hielo la muestra por 5 minutos. La pastilla fue lavada con etanol al 70% 2 veces, se retiró el sobrenadante y la pastilla se dejó secar a temperatura laboratorio (24°C). Finalmente se resuspendió en 60 μl de H₂O ultra pura (Invitrogen).

3.2.3 Cuantificación de ácidos nucleicos mediante espectrofotometría

Para estimar la concentración y el nivel de pureza del ARN extraído se utilizó un espectrofotómetro SmartSpect plus de BioRad. Se tomaron 5μl de muestra más 95μl de UPH₂O (dilución 1:20) para cuantificar la concentración y 5μl de muestra más 95μl de Tris HCl 10mM pH 7.5 para estimar la pureza (radio A260/A280) (Quiagen, 2001).

3.2.4 Síntesis de ADN complementario

La síntesis de ADNc partió de una muestra de 5 μg de ARN total, se agregaron 1 μl (50 μM) de oligo dT primer y la cantidad de agua necesaria para llegar a un volumen de 20 μl, se calentó la muestra a 75°C por 5 minutos para linearizar los ARN mensajeros y se enfrió la muestra en hielo por 3 minutos. A los 20 μl se agregaron 8 μl de buffer de la enzima transcriptasa reversa (5 X first strand buffer), 4 μl 0.01M de DTT (0.1M), 2 μl de dNTP's (10 mM) y 2 μl de inhibidor de RNasas (RNase OUT 40U/μl. Invitrogen). Los 38 μl de reacción se dividieron en dos tubos diferentes con 19 μl cada uno. En el primer tubo se agregó 1 μl de enzima reversa transcriptasa (Super Script III 200 U/ μl. Invitrogen) y en el segundo tubo se agregó 1 μl de agua ultra pura como control negativo para descartar contaminación por ADN genómico. La reacción se incubó por 90 minutos a 50°C, después se inactivó la enzima incubando la muestra a 75°C por 15 min. Seguidamente se agregaron 2 μl de RNasa A (10mg/ml) para degradar el ARN remanente incubando la muestra a 37°C por 30 min. Finalmente, se diluyó la muestra 1:5 ajustando la reacción a un volumen final de 100 μl.

3.2.5 Transcripción Reversa y Reacción en Cadena de la Polimerasa (RT-PCR)

La reacción de RT-PCR se llevó a cabo utilizando 5µl de ADNc de cada tejido como plantilla. El programa para llevar a cabo la reacción utilizó una temperatura inicial de desnaturalización de 95°C por 3 minutos, seguido de 30 ciclos de 95°C por 30 s, 55°C por 30 s y 72°C por 30 s; y una extensión final de 72°C por 5 minutos. Como control positivo para comprobar que no haya habido una carga diferencial en el gel y para cerciorarse que los componentes de la reacción funcionen adecuadamente, se utilizaron oligos para el gen del factor de elongación 1 alfa (EF1 α) de papaya, los cuales flanquean dos exones y un intrón. El control negativo también utilizó estos oligos en la muestra sin retrotranscriptasa, en este caso no se esperaría ver amplificación, de ocurrir sería un indicativo de contaminación de la reacción. Los oligos utilizados para llevar a cabo la RT-PCR se muestran en el cuadro 3.2.

Cuadro 3.1 Oligos específicos para la amplificación de fragmentos de ADNc de los genes de papaya homólogos a factores de transcripción tipo ERF.

Secuencia identificada <i>in silico</i> en papaya	Secuencia homóloga	Secuencia de los oligos diseñados	
CpERF1 (ABIMO1004358)	Pti4	Sentido	5' -ACCGATGAGTTGGA GATCAC-3'
		Antisentido	5' -AATCTCAGCAGCGAACTTCC-3'
CpERF2 (ABIMO1001349)	Pti6	Sentido	5' -TAGAAGGAGACCATCGAGAG-3'
		Antisentido	5' -AACTTCAA TGCA GCTTCGTC-3'
CpERF3 (ABIMO1025280)	TaERF1	Sentido	5' -AAGAGCTTATGA TGCTGAGG-3'
		Antisentido	5' -TG TAGTAA TCCTGACCAAGG-3'
CpERF4 (ABIMO1010943)	ERF1	Sentido	5' -TCAGATTTCTCCTCCGATC-3'
		Antisentido	5'-TGCTTCGGATCTTCCACG-3'
CpERF5 (ABIMO1009477)	JcERF	Sentido	5' -ATCCTCAACAGCAGCATGCC-3'
		Antisentido	5' -TGCCTTGTCGTAAGCTAAGG-3'

3.2.6 Mapeo de los extremos terminales del ADNc mediante la técnica de amplificación de los extremos terminales del ADNc (RACE)

La técnica de "Amplificación rápida de los fragmentos terminales de ADNc" (RACE; por sus siglas en inglés) se basa en una amplificación con un oligo específico para la región codificante y un primer universal para los extremos 5' y 3' seguido de una segunda amplificación mediante una PCR anidada con el fin de amplificar los extremos terminales

5' y 3' del inicio y término de los marcos de lectura abierta, así como las regiones no traducibles (UTRs) (Figura 3.1). En el caso de la amplificación del producto del 3'RACE, los oligos universales "3'RACE" y "3'RACE anidado" se alinean a una secuencia universal del oligo dT utilizado en la síntesis de ADNc, mientras que para el caso del 5'RACE es necesario aplicar una variante del RACE llamada RLM-RACE (RACE mediada por ligación del ARN) en la cual es necesario quitar la 7-metil guanosina que sirve como caperusa del extremo 5' del ARNm y luego ligar un oligo conocido al extremo 5' fosfatado de todos los ARNm, es este oligo en el que se alinean los oligos universales "5'RACE" y "5'RACE anidado". Para la realización del 5'RACE y 3'RACE de los factores de transcripción ERF de papaya Maradol se utilizó ADNc de tejido de hoja y el estuche comercial GenRacer de Invitrogen (Núm. Cat: L1500-01). Para el RACE se utilizaron las mismas condiciones descritas en el RT-PCR. Para la primera amplificación de los extremos 5' RACE se usaron como oligos específicos los oligos antisentido del RT-PCR (cuadro 3.2) y los oligos universales "GenRacer 5' " del Kit. La PCR anidada 5' RACE fue realizada mediante los oligos "5' Antisentido RACE" indicados en el cuadro 3.3 y los oligos universales "GeneRacer 5' Nested" del Kit. La primera amplificación de los extremos 3' RACE se realizó con los oligos "Sentido" de la RT-PCR (cuadro 3.2) y los oligos universales "GenRacer 3' " del kit. Finalmente la PCR anidada para el 3' RACE fue realizada con los oligos "3' Sentido RACE" indicados en el cuadro 3.3 y los oligos universales "GeneRacer 3' Nested" del kit.

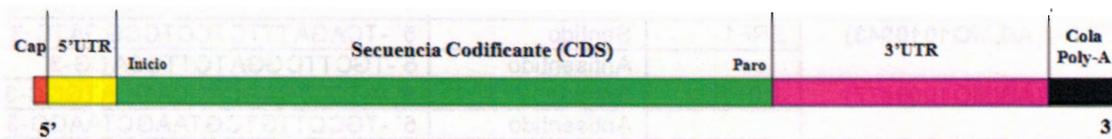


Figura 3.1 Estructura típica de un ARN mensajero incluyendo la región codificante y las regiones no traducibles (UTR)

Cuadro 3.2 Oligos específicos para la PCR anidada 5' RACE (Antisentido 5' RACE) y 3' RACE (Sentido 3' RACE) de CpERF1-CpERF5			
Secuencia identificada <i>in silico</i> en papaya	Secuencia homóloga	Secuencia de los Oligos diseñados	
CpERF1 (ABIMO1004358)	Pti4	Oligo sentido 3' RACE	5' -TATGGCTCGGAACGTACG-3'
		Oligo antisentido 5' RACE	5' -TCATGTCCTCAGAGTCGTC-3'
CpERF2 ABIMO1001349	Pti6	Oligo sentido 3' RACE	5' -ATGACGAAGCTGCATTGAAG-3'
		Oligo antisentido 5' RACE	5' -TCGATGGTCTCCTTCTAGG-3'
CpERF3 (ABIMO1025280)	TaERF1	Oligo sentido 3' RACE	5' -ACTTCAACAACCTTGGTCAG-3'
		Oligo antisentido 5' RACE	5' -ATCGTCCAAGTCGACAAC-3'
CpERF4 (ABIMO1010943)	ERF1	Oligo sentido 3' RACE	5' -TGATCCGGAGGAGAAAATCTG-3'
		Oligo antisentido 5' RACE	5'-TCGAAGATTTGGGAGCAGAC-3'
CpERF5 (ABIMO1009477)	JcERF	Oligo sentido 3' RACE	5' -ATCCGAGATCACGTTCGGTG-3'
		Oligo antisentido 5' RACE	5'-ACTAGAGGAGGAAGTACAAG-3'

3.2.7 Clonación de los extremos 5' y 3' terminales del ADNc en el vector pGEMT-easy

Para la clonación y aislamiento de los productos de RACE se utilizó el vector de clonación pGEM-T Easy (Figura 3.2). La reacción de ligación se llevó a cabo en un volumen final de 10 µl usando 0.5 µl de la enzima T4 ADN ligasa (1U/µl) (NEB), 1µl del vector pGEM-T Easy (50ng/µl), 7.5 µl de producto de PCR y 1µl del amortiguador de la ligasa 10X. La mezcla de ligación se incubó por 16 hrs a 16°C. La ligasa fue inactivada a 65°C por 10 min.

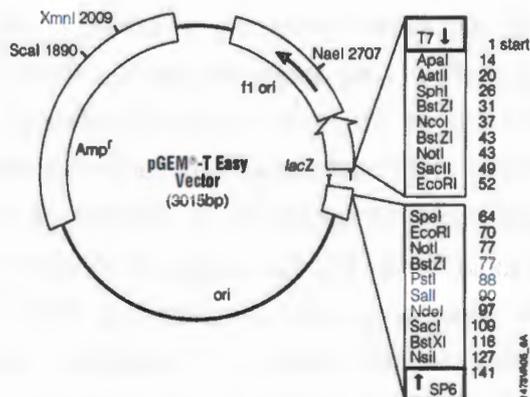


Figura 3.2 Mapa con los principales puntos de referencia del vector de clonación pGEM[®]-T Easy

Se tomarón 3 μ l de la reacción de ligación para transformaron células competentes de *Escherichia coli* mediante choque térmico a 42°C por 50 s. Las bacterias se crecieron en medio LB sólido (Luria Agar) conteniendo ampicilina (100mg/l) como antibiótico de selección, IPTG (0.5 mM) como inductor de la expresión del gen *lacZ* y XGAL (80 mg/l) como substrato de la β -galactosidasa para indicar la expresión de la misma enzima mediante el monitoreo de colonias azules y blancas.

Las colonias blancas, que indican la falta de expresión de la enzima β -galactosidasa debido a la interrupción del gen *lacZ* por la inserción del fragmento de interés (ver figura 3.2) fueron seleccionadas y cultivadas en 5 ml de medio LB líquido con ampicilina (100mg/l) a 37°C y 220 rpm durante toda la noche.

Una vez multiplicado el plásmido recombinante, se procedió a tomar 4 ml del cultivo para extraer el plásmido con el protocolo de lisis alcalina (Sambrook y Rusell, 2001). El cultivo bacteriano fue sedimentado por centrifugación (14000 rpm) para remover todo el medio LB y lavado con 1 ml de amortiguador STE frío (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0 y 100mM NaCl). Una vez retirado el amortiguador STE (14000 rpm por 1min) se resuspendió la pastilla en 100 μ l de "Solución I" (50mM glucosa, 25 mM Tris-HCl pH 8.0, 10 mM EDTA) y se agregaron 200 μ l de "Solución II" (0.2N NaOH, 1% SDS) para lograr la lisis de las células incubando 3 min a temperatura ambiente. Se añadieron 150 μ l de "Solución III" (3M KOAc, 5M AcOH) incubando en hielo por 5 min. Se centrifugó a 14000 rpm por 5 min y el sobrenadante fue mezclado con 900 μ l de etanol 100% con 0.1 vol de NaAc 3M para precipitar los ácidos nucleicos. Al protocolo se le agregó un paso de purificación final agregando 600 μ l del detergente SSTE e incubando 5 minutos a 70°C, después se agregaron 700 μ l de Cloroformo-Alcohol isoamílico (24:1), se separaron las fases centrifugando 5 min a 14000 rpm y el sobrenadante fue transferido a un tubo nuevo para precipitar los ácidos nucleicos agregando un volumen de etanol al 100% frío y 0.1 volumen de acetato de sodio 3M pH 7.0. La muestra se incubó en hielo por 5 minutos, se centrifugó y la pastilla obtenida se lavó con etanol al 70%. Finalmente se retiró el sobrenadante y la pastilla se dejó secar a temperatura laboratorio. La pastilla se resuspendió en 50 μ l de Tris-HCl pH 7.5 y se agregó 2 μ l de RNasa A (10mg/ml).

Una vez aislado y purificado el vector pGEM-T Easy con el fragmento de interés, se tomaron 3 μ l de éste y se digirió con 0.3 μ l de la endonucleasa *EcoRI* (50U/ μ l) (Invitrogen), 2 μ l de amortiguador ReACT3 (invitrogen) y agua ultra pura para un volumen final de 20 μ l. La reacción de digestión se llevó a cabo a 37°C por 30 min. El plásmido digerido se fraccionó por electroforesis en un gel de agarosa 1.2%.

3.2.8 Secuenciación

Antes de realizar la reacción de secuenciación, las muestras fueron purificadas utilizando filtros de sílice (High Pure PCR Product Purification Kit, Roche Cat. No. 11 732 668 001). Para cada fragmento 5' y 3' de las cinco secuencias de estudio se realizaron 2 reacciones de secuenciación con oligos específicos para el vector pGEM-T Easy, una reacción con oligo sentido (5 μ M) y otra reacción con el oligo antisentido (5 μ M) siguiendo el método de Sanger utilizando dideoxinucleótidos (abreviados como ddNTP's, son nucleótidos que carecen de un grupo 3'-hidroxilo (-OH) en la desoxirribose). Las reacciones se llevaron a cabo utilizando de 2-5 μ l de los plásmidos recombinantes, 1 μ l de primer (Sentido o antisentido), 1 μ l de la mezcla de secuenciación (BigDye® Terminator v3.1 Cycle sequencing Kit; Applied Biosystems), 3.5 μ l de amortiguador de la mezcla de secuenciación (5X) y agua ultra pura para un volumen final de 20 μ l. El programa de amplificación fue de 3 minutos a 95°C, seguido de 35 ciclos de 30 segundos a 95°C; 30 segundos a 50°C y 4 minutos de extensión a 60°C. con una extensión final de 10 minutos a 60°C. A continuación se purificó la amplificación de secuenciación agregando 75 μ l de una solución de etanol al 70% conteniendo 0.2mM de MgSO₄, se incubó por 15 minutos a temperatura ambiente y se centrifugó por 20 minutos a 14000 rpm. Se decantaron suavemente las muestras y se dejaron secar toda la noche. Una vez purificadas las muestras se enviaron al servicio de separación capilar del "Australian Genome Research Facility" (AGRF).

3.2.9 Análisis Bioinformático

Los cromatogramas obtenidos tanto sentido como antisentido de cada secuencia de estudio fueron ensamblados con el programa SeqMan (DNASTAR Lasergene v. 7.2.1) con la finalidad de obtener la lectura final de la secuencia sentido-antisentido. Solo los cromatogramas que presentaran una lectura clara de la separación capilar fueron tomados en cuenta para el estudio. Una vez identificada la secuencia de los oligos utilizados en la amplificación anidada del RACE, se realizó un alineamiento entre la secuencia de ADNc de papaya Maradol obtenida *in vitro* (RACE) y la secuencia de la predicción *in silico* del marco de lectura abierta obtenido por FGENESH para los genes tipo ERF de estudio de la papaya transgénica SunUp. Este alineamiento se realizó con la finalidad de verificar el inicio y término de los ORFs de estudio.

3.3 RESULTADOS

3.3.1 Aislamiento de ARN total.

Los resultados del aislamiento de ARN mostraron las bandas ribosomales 25S y 18S integras, lo cual indica que los ARN mensajeros no se encuentran degradados y por lo tanto pudieron ser utilizados para sintetizar ADNc. En la figura 3.3 podemos observar la integridad del ARN aislado de cada uno de los diferentes tejidos. Se cargó 1 µg de ARN de acuerdo a la cuantificación realizada por el espectrofotómetro.

En el cuadro 3.3 se puede notar el rendimiento en µg de ARN total obtenido por gramo de tejido, podemos observar que obtuvimos más rendimiento en hoja, en la cual se aislaron más de 400 µg de ARN total por gramo de hoja, mientras que el rendimiento en peciolo fué de 50 µg de ARN total. En flor y fruto también se obtuvieron buenos rendimientos sin embargo no fueron tan altos como los obtenidos en hoja.

Por otro lado la pureza medida como la razón de la absorbancia a 260nm entre la absorbancia a 280nm siempre se mantuvo en los parámetros requeridos entre 1.8 y 2 (cuadro 3.3).

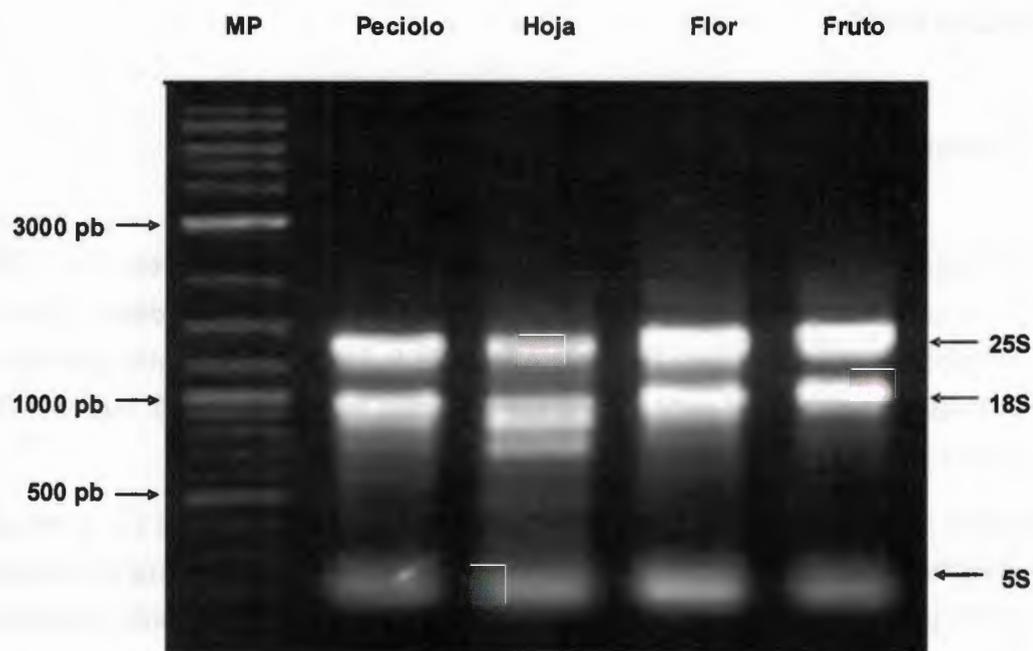


Figura 3.3 Electroforesis de ARN total (1µg) de diferentes tejidos de papaya Maradol. La extracción de ARN total se realizó mediante el protocolo de Kiefer *et al.* (2000) con modificaciones.

Cuadro 3.3 Rendimiento y grado de pureza de los extractos de ARN total de diferentes tejidos de papaya

Tejido	Rendimiento µg ARN/g tejido	Pureza A260/A280
Peciolo	50.5	1.88
Hoja	435	1.8
Flor	248	1.97
Fruto	101.5	1.99

La concentración y grado de pureza se determinaron por espectrofotometría utilizando un equipo SmartSpect Plus BioRad mediante una dilución 1:20. La concentración se determinó diluyendo en H₂O y el grado de pureza diluyendo en Tris-HCl 10 mM pH 7.5 de acuerdo a las recomendaciones de Quiagen (Quiagen, 2001).

3.3.2 Expresión basal de los genes CpERF1-CpERF5 en diferentes tejidos de plantas de papaya Maradol.

Los datos arrojados en el presente capítulo de investigación muestran una expresión basal de los genes estudiados muy similar en todos los tejidos muestreados con excepción del gen *CpERF4*, el cual no mostró una expresión basal visible bajo las mismas condiciones estándar de PCR en las que se llevó a cabo el experimento. El gen *CpERF2* presenta aparentemente una menor expresión que el control en peciolo y hoja, pero incrementa su expresión en tejido de flor. Por su parte, *CpERF3* y *CpERF5* muestran una amplificación similar a la expresión del gen control EF1 α , mientras que el gen *CpERF1* mostró una amplificación más intensa, particularmente en flor y fruto (figura 3.4).

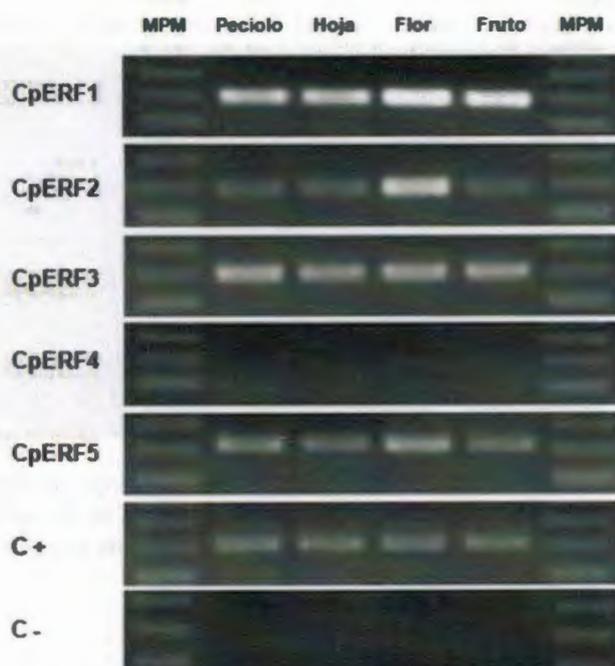


Figura 3.4 Amplificación por medio de RT-PCR de los genes CpERF1-CpERF5 en cuatro tejidos de papaya Maradol. Los amplicones de 200pb se fraccionaron por electroforesis en gel de agarosa al 1.2% teñido con bromuro de etidio. Como controles positivo (C+) y negativo (C-) se utilizaron oligos para el gene EF α de papaya. El control negativo consistió en omitir a la retrotranscriptasa. Las bandas del marcador de peso molecular (MPM.- 2-Log DNA Ladder (0.1–10.0 kb), New England Biolabs) corresponden a 100, 200 y 300 pb. Se fraccionó un volumen de 10 μ l de volumen total de 50 μ l de la RT-PCR para cada uno de los carriles del gel.

3.3.3 Amplificación de los extremos 5' y 3' de los ADNc de CpERF1-CpERF5

La amplificación de los extremos 5' y 3' de los ADNc de CpERF1-CpERF5 se muestran en las figuras 3.5 y 3.6 respectivamente. El tamaño de los amplicones 5' RACE varió de 150 pb en CpERF4 hasta 750 pb en CpERF2 (figura 3.5). En el caso de los genes CpERF2 y CpERF5 se utilizó la enzima "Expand" de Roche en lugar de Taq polimerasa para poder generar un amplicón de mejor calidad. El tamaño de los amplicones 3'RACE varió de 150 pb en el fragmento de CpERF4 hasta los 750 pb en CpERF3 (Figura 3.6). En la figura 3.6 se muestran dos bandas nítidas en las regiones del extremo 3' de los genes CpERF1 y CpERF2.



Figura 3.5 Productos del 5' RACE de los factores de transcripción ERF de papaya. Los amplicones se fraccionaron por electroforesis en gel de agarosa al 2% teñido con bromuro de etidio. Como control positivo se utilizaron oligos para el gen EF₁α de papaya.

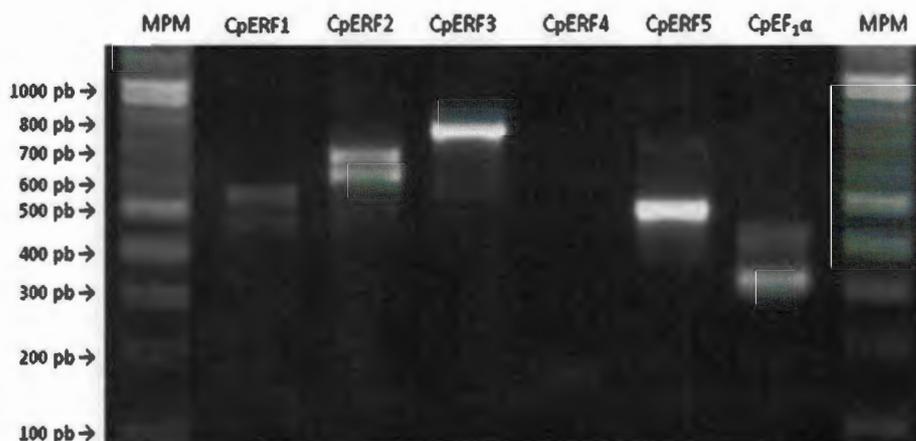


Figura 3.6 Productos del 3' RACE de los factores de transcripción ERF de papaya. Los amplicones se fraccionaron por electroforesis en gel de agarosa al 2% teñido con bromuro de etidio. Como control positivo se utilizaron oligos para el gen EF₁α de papaya.

3.3.4 Clonación de los productos del RACE de CPERF1-CPERF5

Los productos del RACE fueron exitosamente insertados y clonados en el vector pGEMT-Easy. La digestión del vector por la endonucleasa EcoRI fue analizada por fraccionamiento en gel de agarosa al 1.2%. Se pudieron observar los fragmentos del tamaño esperado para los productos del RACE (Figuras 3.7 y 3.8). Solo el caso de CpERF5 mostró un doble bandeo indicando la presencia de un sitio de reconocimiento de EcoRI dentro del extremo 3' (Figura 3.8). Para cada fragmento fueron analizadas al menos 3 clonas para verificar que se insertara el fragmento que correspondiera al tamaño esperado

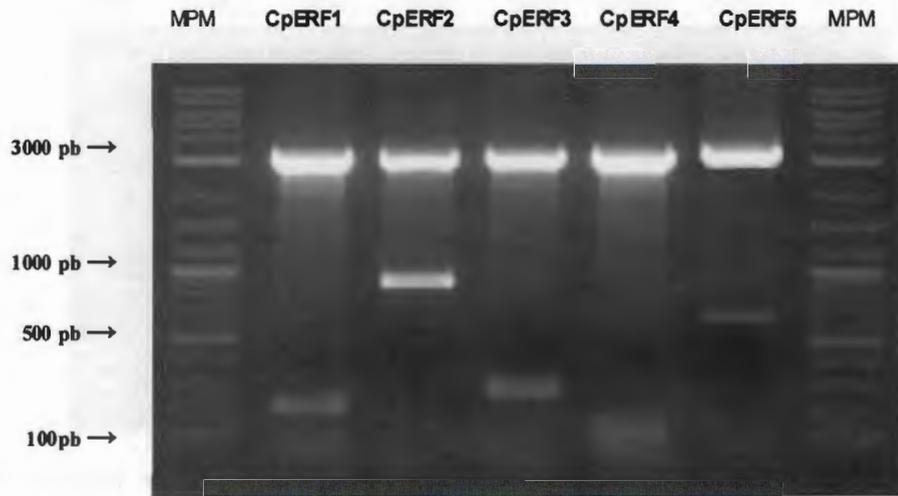


Figura 3.7 Digestión del vector pGEM-T Easy con la enzima de restricción EcoRI para liberar los insertos correspondientes a los productos del 5'RACE de los genes de papaya CpERF1-CpERF5. La banda de 3 kb corresponde al vector pGEM-T Easy.

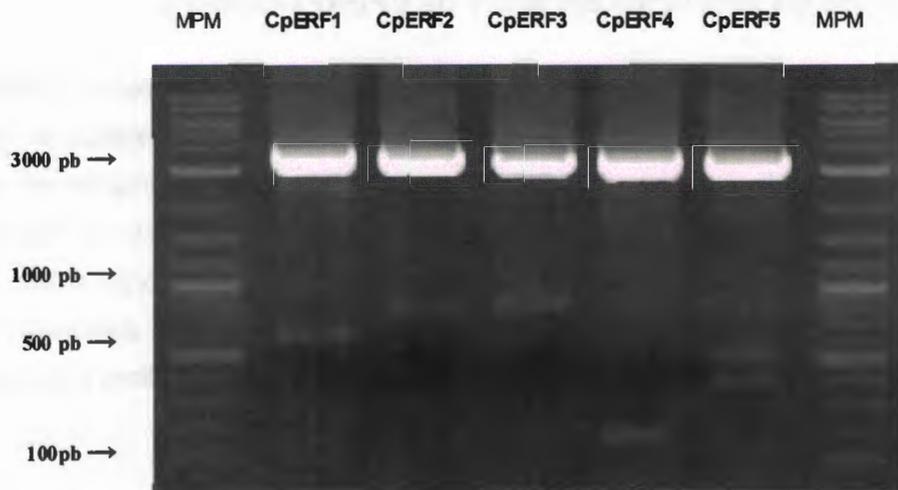


Figura 3.8 Digestión del vector pGEM-T Easy con la enzima de restricción EcoRI para liberar los insertos correspondientes a los productos del 3'RACE de los genes de papaya CpERF1-CpERF5. La banda de 3 kb corresponde al vector pGEM-T Easy.

3.3.5 Mapeo de los extremos 5' y 3' de los ADNc de CpERF1-CpERF5

Las secuencias de los productos de RACE tanto para la región 5' como 3' del ADNc de los cinco genes de papaya se muestran alineadas con la secuencia de la predicción del ORF para cada gen en las figuras 3.9-3.13, en dichas figuras las zonas sombreadas muestran las bases conservadas entre la predicción *in silico* de los ORF's de cada gen y los productos secuenciados. En la región 5' del ORF de CpERF1 de Maradol existe una guanina en vez de una alanina que se presenta en SunUP, mientras que en la región 3' existen 6 nucleótidos en la región del ORF de Maradol que no se encuentran en el gen de SunUP (Figura 3.9). Los genes CpERF2 y CpERF4 no presentaron variación alguna entre ambas variedades de papaya (Figuras 3.10 y 3.12). El aislamiento *in vitro* permitió conocer el ORF completo del gen CpERF3 que había sido imposible detectar por la predicción *in silico* (Figura 3.11). Se observó que la región 5' del marco de lectura abierta para CpERF3 predicho por FGENESH estaba incorrecto lo que hacía parecer que tenía un segundo intrón, sin embargo la secuencia del producto 5' RACE de papaya Maradol coincide 100% con la secuencia del contig ABIM01025280 de papaya SunUp (dato no mostrado). El producto 3' RACE del gen CpERF3 permitió conocer el extremo 3' incluyendo el término del marco de lectura abierta y la región 3' UTR, parte de este fragmento coincidió con una secuencia encontrada en el contig ABIM01025279 de papaya SunUP (dato no mostrado). Finalmente el ADNc de CpERF5 pudo ser secuenciado por completo (Figura 3.13). En este gen la secuencia de Maradol difiere en cinco nucleótidos con respecto a la secuencia de SunUp. Sin embargo se observó que la predicción de FGENESH era incorrecta en el inicio del marco de lectura abierta, ya que en el aislamiento *in vitro* de este gen se comprobó que el codón de inicio ATG se encontraba 174 pb después del codón de inicio predicho por FGENESH. El 5RACE de CpERF3 y el 3RACE de CpERF1 y CpERF2 muestran dobles amplificaciones sugiriendo sitios alternativos de las regiones no traducibles que fueron confirmadas al secuenciar los dos fragmentos amplificados de cada gen.

5 RACE	1	CACCTCCGGAAGCTTCAGTTTATATAAAATGAATGAA	37
5 RACE	38	[REDACTED]	97
ORF <i>in silico</i>	1	[REDACTED]	60
5 RACE	98	[REDACTED]	155
ORF <i>in silico</i>	61	[REDACTED]	118
ORF <i>in silico</i>	119	TCATATACAACCTCCCTCCGTGACGCTCTTCATTTTGGATGGTCTCCGTTAGATTTGACAG	178
ORF <i>in silico</i>	179	AAAGGACAGTCAAAGCGAGCCACC GATGAGTTGGAGATCACGGCTCCGGTCAAACAAG	238
ORF <i>in silico</i>	239	GAAATAAATCAGCAACTTGAAAAGCGGTATGTTGCTCAAAGAGTGAAAACGGAAGTGT	288
ORF <i>in silico</i>	299	CTTACCGGAGAGACAAGAAAAGATCTTACAGAAGCCAAGAGGAAGGCATTATCGAGGTG	358
ORF <i>in silico</i>	359	TCAGGCAACGGCCGTGGGGAAGTTCCTGCTGAGATTCGAGACCCGGCTAAAAACGGCG	418
ORF <i>in silico</i>	419	CGAGGG 424	
ORF <i>in silico</i>	425	[REDACTED]	484
3 RACE	1	[REDACTED]	60
ORF <i>in silico</i>	485	[REDACTED]	544
3 RACE	61	[REDACTED]	120
ORF <i>in silico</i>	545	[REDACTED]	604
3 RACE	121	[REDACTED]	180
ORF <i>in silico</i>	605	[REDACTED]-----[REDACTED]	658
3 RACE	181	[REDACTED]GTCGGC[REDACTED]	240
ORF <i>in silico</i>	659	[REDACTED]	718
3 RACE	241	[REDACTED]	300
ORF <i>in silico</i>	719	[REDACTED] 744	
3 RACE	301	[REDACTED] 326	
3 RACE	327	GTAGTTTCCTAAGTTGTGGATTGTATAACAAGGTAAGAATGAAGTGATGGTTTGATTATT	386
3 RACE	387	GTATGGAGTGAATAGGAACTTGTGGGTTGAAATGCAGACTACAGATGTTGGGTTGAATG	446
3 RACE	447	CATACTCCATATAATTCCGTGAAAAGCAAAAGCCAGAATTGGTTTGTAAATTTCTGA	506
3 RACE	507	AAATGATGGGGTTCGGCTA <u>ATAAA</u> TTCGGTTTATGTATC <u>T</u> AAAAAAAAAAAAAAAA	566

Figura 3.9 Alineamiento del ORF de CpERF1 predicho *in silico* de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol. El producto 5'RACE muestra la región 5'UTR (1-37 pb) y sombreado en negro muestra el alineamiento de la región 5' de los ORFs de Maradol y el predicho de SunUp. El producto 3'RACE muestra la región 3'UTR (327-549 pb) y sombreado en negro el alineamiento de la región 3' de los ORFs de Maradol y el predicho de Sun Up. En la región 3'UTR se remarca con rectángulos los posibles sitios de poliadenilación (Near Upstream Elements) y el sitio de anclaje para la polimerización de la cola poli-A

5 RACE	1	CATTTCTCCCTCTTTCTTCCCACTCTTCTCTCTCACCTACTTTCTCTGTFTT	56
5 RACE	57	CCTATGTATTGGGTTTCTACTAGTTAGGACTGACTCTTTTCCGCCAAAGGTGTGAAAGA	116
5 RACE	117	CCCAATACTCTTTGCACCAGACACACTTCCATATGTAGAAAGGATAGCCAGTGTGCA	176
5 RACE	177	GAGTGTGATTGTCGGTTCITTCACTAAACTTCAAACCTTTGTTCCTCTTTTGT	236
5 RACE	237	CCGCCTGTCTGATCTCTCGTTCCTTCCCAGTTCCTTTCTTCAACATCTTCTT	296
5 RACE	297	TCTTTTCTATAAATACACGC GGATTCTTTACTCTCCTCGAGATTATGTATGTTG	356
5 RACE	357	AAACATCACAGGATCTTCATCC TAGCTTCTTAATCTTGAAAAGCTGTTCTTGCTCCA	416
5 RACE	417	TGCAGTAGCAGTTTGTGTTGCTTTTACGTGCTTTTCTACTCATGTCTGACCAAAGCTCGT	476
5 RACE	477	CTCCTACCCTTGGATCTGCTCCCTCCCAACATCAACCC 514	
5 RACE	515	ATGGTGTCCATACAGTCCACTCTCAAGCTCTCTTCTCGGGTCTTCAACCACCAAGCTC	574
ORF <i>in silico</i>	1	ATGGTGTCCATACAGTCCACTCTCAAGCTCTCTTCTCGGGTCTTCAACCACCAAGCTC	60
5 RACE	575	CTC CAT GA TCC CC CGGCGGATGC CAAAGTCCGT CGGAGACGTGTCCAGATCATACTTACC	634
ORF <i>in silico</i>	61	CTC CAT GA TCC CC CGGCGGATGC CAAAGTCCGT CGGAGACGTGTCCAGATCATACTTACC	120
5 RACE	635	GACGCCGACGCACCGATCTTCCAGCGAGGATGAA GAGGAAA GACATGGTTTCTTACGA	694
ORF <i>in silico</i>	121	GACGCCGACGCACCGATCTTCCAGCGAGGATGAA GAGGAAA GACATGGTTTCTTACGA	180
5 RACE	695	AGAGTTAAGAGCAAGTTCAAGTTATCACTTTGATTCATCTTCAACCCAGACCACAT	754
ORF <i>in silico</i>	181	AGAGTTAAGAGCAAGTTCAAGTTATCACTTTGATTCATCTTCAACCCAGACCACAT	240
5 RACE	755	CCACCTCGATCTCCGACGCCACTTCC 780	
ORF <i>in silico</i>	241	CCACCTCGATCTCCGACGCCACTTCC 266	
ORF <i>in silico</i>	268	GGCTCCTACTACGTTACC TAAGCAAAAGTCAACC TAGAAGGAGACCAATCGAGTGTGACGAA	326
ORF <i>in silico</i>	327	ATCTGACGTCACTCGCGGGCAAAATTCAGGGGTGTTCGGCAGAGACCTTGGGGGAAGTG	386
ORF <i>in silico</i>	387	GGCCGACAGATCCGGGACCCAGCCAGAGAAAACGTATCTGGCTGGGAACTTCAACAC	446
ORF <i>in silico</i>	267	CGCCGAGGAAGCGGCCCGTGT 469	
ORF <i>in silico</i>	470	ATGACGAAGCTGCATTGAAGTTAAGGGTTCAGACGCCGTTACCAACTTTCCCACCCGA	529
3 RACE	1	ATGACGAAGCTGCATTGAAGTTAAGGGTTCAGACGCCGTTACCAACTTTCCCACCCGA	60
ORF <i>in silico</i>	530	GTGTAA CGGAGTTGACGGCTGTGACAGTCAAAAAGACGGATTGACTCGTCAAGCACAG	589
3 RACE	61	GTGTAA CGGAGTTGACGGCTGTGACAGTCAAAAAGACGGATTGACTCGTCAAGCACAG	120
ORF <i>in silico</i>	590	CCACGT CATCTCCGACATCAGTTCTCCGTTACGAGGAGATAACACCGTTTGTGGCCTCA	649
3 RACE	121	CCACGT CATCTCCGACATCAGTTCTCCGTTACGAGGAGATAACACCGTTTGTGGCCTCA	180
ORF <i>in silico</i>	650	ATTACTGC GACGTGTATACCTTCGGCTTCGGAA TCGACGCCCGTTAAGCGTGGCCGGATT	709
3 RACE	181	ATTACTGC GACGTGTATACCTTCGGCTTCGGAA TCGACGCCCGTTAAGCGTGGCCGGATT	240
ORF <i>in silico</i>	710	TCA TGT TTTTCGGGAAAAAGCTTTGCCGAGGAAGAGTTCGGGGAGCTTGACGTTGACGATT	769
3 RACE	241	TCA TGT TTTTCGGGAAAAAGCTTTGCCGAGGAAGAGTTCGGGGAGCTTGACGTTGACGATT	300
ORF <i>in silico</i>	770	TTCTGGTGCACGCCGTTGCCTAA 792	
3 RACE	301	TTCTGGTGCACGCCGTTGCCTAA 323	
3 RACE	324	CGACGAACGTTATTTAGTACCA GTAATCCGAAGATATTTA TGTGGGTGCTGACGGCAAT	383
3 RACE	384	CTCCGT TAGCAAC CCGGGCAGTGAGGYGGTTACGTGCATATACTATTTGTTATAACTA	443
3 RACE	444	CAGTTTTTGAAGGGGGGCTTC CGGTGACTCCGATTA TGCAATCCAAAAATACCCCTCA	503
3 RACE	504	CATATTTTATTGTGTGGTCAAAGTGAAGTGAAGGCCCCCTTAAATTTACGGGAGTGAATTA	563
3 RACE	564	TCTATGAAAGAGTTAAATGTAGACGATTTTGTCTGTGCGAAGAGAGATTACGTTGT	623
3 RACE	624	GGAGTAGTAAATTATTATAAATTTGGTGGTTTGTAAATTAATTATT AATAATGTATGG	683
3 RACE	684	TTGAAAAGTGA AAAAAAAAAAAAAAAAAAAAA 713	

Figura 3.10 Alineamiento del ORF de CpERF2 predicho *in silico* de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol. El producto 5'RACE muestra la región 5'UTR (1-514 pb) y sombreado en negro muestra el alineamiento de la región 5' de los ORFs de Maradol y el predicho de SunUp. El producto 3'RACE muestra la región 3'UTR (324-684 pb) y sombreado en negro el alineamiento de la región 3' de los ORFs de Maradol y el predicho de SunUp. En la región 3' se remarca con rectángulos los posibles sitios de poliadenilación (Near Upstream Elements) y el sitio de anclaje para la polimerización de la cola poli-A

5 RACE	1	AACTCAAACTTTCGTTTCCCTTACACCCTCGTCTTCTACCTTCTGCCATAGTAAGCC	58
5 RACE	59	ATGTGTGGAGGTGCTAT AATCTCCGATTT CAT ACCACCCTACTACTCGGTGGTCAAG	118
ORF in silico	1	ATGTGTGGAGGTGCTAT AATCTCCGATTT CAT ACCACCCTACTACTCGGTGGTCAAG	60
5 RACE	119	TTGACCGCCGACTACTTATGGCCC GATCTGAACAG	153
5 RACE	154	GAATAGGAAATC TAAGAAGAGCTC CAA GAGGTCTGAGGTGTGCGACTTGACCGAT	208
ORF in silico	60	GAATAGGAAATC TAAGAAGAGCTC CAA GAGGTCTGAGGTGTGCGACTTGACCGAT	114
ORF in silico	115	GACTTCGAGGCTGATTTCCAGGGCTTCAAGGACGACGAATCCGATATTGACGTCGACGAA	174
ORF in silico	175	GATCTCGACGATATGATGCTGCTTCTCTGATATTAAGCCCTTGCCTTCTCTGCAACT	234
ORF in silico	235	CCTCTCCCGCAAAACCAAGCCCTCTGCTCTCTCCAA GATCAAAGCCTGTGAAAGCT	294
ORF in silico	295	GTGGAATCAATGGGCTGGCAGAGAAATCTGC AAAAAGAGAGAAACCAATACAGG	354
ORF in silico	355	GGAAATCAGGCAACGCCCTTGGGGCAATGGGCTGCTGAAATCCGTGACC CAAGGAAGGGG	414
ORF in silico	415	GTCAAGGCTCGTGGTATAGTACTTTAACACTGCTGAAGAAAGCTGCAGAGCTTATGATGT	474
ORF in silico	475	GAGGCACGGAGAATTCGTGCAAGAAAGC CAA GGTGAAC TTCAGATGAATCTCCAGCT	534
ORF in silico	535	GCCTCTCAAAGCGTGC CGTGAAC TCAAT GAAAC CAGTTGCCAAGGCAATCCTGAATTC	594
ORF in silico	595	GCACAGCCAAATCTGAGTCAGAATGTTAATT	625
ORF in silico	626	ACTTCAA CAACCTTGGTCAGGATTACTACAATACAATGGTTTGTGTTGATGAGAAACCA	685
3 RACE	1	ACTTCAA CAACCTTGGTCAGGATTACTACAATACAATGGTTTGTGTTGATGAGAAACCA	60
ORF in silico	686	AAATGAATCAGTTTGCATCAATGAATTCATTTCTGCAAGTGGAATGCTGGAGTTAAAC	745
3 RACE	61	AAATGAATCAGTTTGCATCAATGAATTCATTTCTGCAAGTGGAATGCTGGAGTTAAAC	120
ORF in silico	746	CCTTTGTCCCTAGTGACAACACCCAT	771
3 RACE	121	CCTTTGTCCCTAGTGACAACACCCAT	146
3 RACE	147	ATGTAATTCAGTCTGACCAAGGGAGTAACTC GTTTGGGTGTTCTGAATTTGGCTGGGA	206
3 RACE	207	GATCAGGCCACAAAAGACTCCGAAATCTCTCTGTTCTTTGGATCAACCTCAGTTTGTG	266
3 RACE	267	GAGGATTGTAAACCAGAGAAGAAGTTGAAAGTGTAGTTCTGAAAC TATGGTCCCTGTTCAA	326
3 RACE	327	GGCAATGCCAACAGTCTCTGTCTGAGGAGCTGCTAGCTTTGCAATCAAATGAAGTAC	386
3 RACE	387	TTACAAGTGCCTCATCTGACTCGAAC TGGGATT CATCTCTGATGCTTTCCTTAACGGC	446
3 RACE	447	GATGCGCTCAGGATGCTGGGAAC TCAATGGACCTTTGGGCCTTGTATGACCTTCTTCT	506
3 RACE	507	CTGGTTGGGGAGTCTTTGCAAGTCA TCCCTATTTACCAGTCTTGTAAATAAGGC	566
3 RACE	567	TAATGTTCTGATGTTCCAGGAATGCTGTTTCTTCTTTAATGAAGATGGAGTTGAA	626
3 RACE	627	CAGTGGAAAGTTGCTTAAGCCATGTGTTGTGTCGT CAGGAGCTTTAAATAGAAGCGAGTA	686
3 RACE	687	TATAGGTACCATTAAGTCTTTTATTATTGGGTTTGTAGACATGATGAATTGAAC	746
3 RACE	747	TTGGATTAATGACTGATGGATGACTTAAATGTTGACTTGTGATGGAATACTAAAAC	806
3 RACE	807	GATGATGCAATGTTTAAATTTGCTGCTGTTCTCAAAAAAAAAAAAAAAAAAAAA	860

Figura 3.11 Alineamiento del ORF de CpERF3 predicho *in silico* de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol. El producto 5'RACE muestra la región 5'UTR (1-58 pb) y sombreado en negro muestra el alineamiento de la región 5' de los ORFs de Maradol y el predicho de SunUp. El producto 3'RACE muestra la región 3'UTR (528-842 pb) y sombreado en negro el alineamiento de una parte de la región 3' de los ORFs de Maradol y el predicho de SunUp. En la región 3' se remarca con doble subrayado las bases que codifican para un posible sitio de interacción con proteína cinasa descrito para genes miembros del grupo VII (Nakano *et al.*, 2006). Sombreado en gris se indica el sitio donde se encuentra el intrón. Remarcado con línea gruesa se resalta en la región 3' el codón de paro TGA que indica el término del marco de lectura abierto. Los rectángulos indican los posibles sitios de poliadenilación (Near Upstream Elements) y el sitio de anclaje para la polimerización de la cola poli-A

5 RACE	1	ATCATGCTCCCCTCACTGAGAAAACAGAGGACCCCAAAAATCCCTAATCAAATTTAC	57
5 RACE	58	ATGAACTCCTCCTGGAATTCAGATTTCTCCTCCGGATCA	96
ORF <i>in silico</i>	1	ATGAACTCCTCCTGGAATTCAGATTTCTCCTCCGGATCA	39
ORF <i>in silico</i>	40	CCAGATTCCTCTCTTTTCGAAAAACACCTCTCTGCCTTTCAACGAAAATGACTCCGAGGAA	99
ORF <i>in silico</i>	100	ATGCTGCTCTACGGCCTAATCGCAGAAGCCAATCATGAATCGTCGGAGTCGACGTATTCA	159
ORF <i>in silico</i>	160	AACATGATGATCAAGGAGGAGGAGGTGAGTTCCGTGGAAGATCCGAAGCAGGAAAGGTCCG	219
ORF <i>in silico</i>	220	TACAGAGGGGTGAGGAGCGGCCATGGGGGAAATTTGCAGCAGAAATAGGGATCCACG	279
ORF <i>in silico</i>	280	AGACATGGAATAAGAGTGTGGTTAGTACTTTTATAGTGCAGAGGCAGCAGCTCTGCG	339
ORF <i>in silico</i>	340	TATGATCAAGCGGCGTTTCTATGAGAGGATCATCGGCGATTCGAATTTCCGGTGGAG	399
ORF <i>in silico</i>	400	AGAGTTGAGGAGTCGCTGAGGAGATGAGTGTAACTACGAGGAAGGTGTCTCGCTGTT	459
ORF <i>in silico</i>	460	GTGGCTCTGAAGAAAACTCTGATGAGGAGGAAGATGTGTAACAGGAAGAGAGAG	519
ORF <i>in silico</i>	520	AGGGATGTGAGGATAGAGAATGTAGTGGTTT	550
ORF <i>in silico</i>	551	TCGAAGATTGGGAGCAGACTATTGGAGCAGCTCTTGACTTCGTCGTTCTTCTGAAA	610
3 RACE	1	TCGAAGATTGGGAGCAGACTATTGGAGCAGCTCTTGACTTCGTCGTTCTTCTGAAA	60
ORF <i>in silico</i>	611	TCACTATGCCTTGGTGA	627
3 RACE	61	TCACTATGCCTTGGTGA	77
3 RACE	78	TTTTAATGTGCTTACTTCAGTAATTTTCCACTGTCTAAAAA	132

Figura 3.12 Alineamiento del ORF de CpERF4 predicho *in silico* de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol. El producto 5'RACE muestra la región 5'UTR (1-57pb) y sombreado en negro muestra el alineamiento de la región 5' de los ORFs de Maradol y el predicho de SunUp. El producto 3'RACE muestra la región 3'UTR (78-115 pb) y sombreado en negro el alineamiento de una parte de la región 3' de los ORFs de Maradol y el predicho de SunUp. En la región 3' se remarca con rectángulos los posibles sitios de poliadenilación (Near Upstream Elements) y el sitio de anclaje para la polimerización de la cola poli-A

Figura 3.13 En la siguiente página se muestra el alineamiento del ORF de CpERF5 predicho *in silico* de papaya SunUp y los productos 5'RACE y 3'RACE de papaya Maradol. El producto 5'RACE muestra que la región 5'UTR de papaya Maradol comienza en la base 87 del ORF predicho *in silico* y termina en la 262. Sombreado en negro muestra el alineamiento de la región 5' y 3' de los ORFs de Maradol y el predicho de SunUp. En la base 174 del producto 5 RACE se resalta el codón de inicio ATG del ORF de CpERF5. La región 3'UTR comienza a partir de la base 693 del producto 3 RACE. En la región 3' UTR se remarca con rectángulos los posibles sitios de poliadenilación (Near Upstream Elements) y el sitio de anclaje para la polimerización de la cola poli-A

3.4 DISCUSIÓN

3.4.1 Expresión basal y espacial de los genes CpERF1-CpERF5

El análisis por medio de RT-PCR de la expresión de los genes tipo ERF en diferentes tejidos del cultivar Maradol de papaya mostró una expresión basal de los genes *CpERF1*, *CpERF2*, *CpERF3* y *CpERF5* en los 4 tejidos en estudio. Si bien, el análisis no es por sí mismo de naturaleza cuantitativa, si genera un panorama de que genes y en que tejidos se observa una mayor expresión relativa, ya que el análisis fue realizado con las mismas muestras de ADNc para todos los genes, y las reacciones de amplificación fueron llevadas a cabo en las mismas condiciones tanto de cantidad de ADNc, como número de ciclos de amplificación. De esta manera se puede inferir que hay mayor nivel de expresión del gen *CpERF1* con respecto los demás genes incluyendo al gen control *EF₁α*, de la misma manera en que el gen *CpERF2* muestra una banda más intensa en el tejido floral que representa una mayor expresión en comparación con los demás tejidos.

Son pocos los estudios que muestran la expresión basal de los factores tipo ERF en diversos tejidos de la planta. En un estudio realizado en tabaco, se muestra un patrón de expresión diferencial del factor de transcripción *Tsi1*, dicho gen tiene una expresión gradual de mayor a menor en tejidos de hoja, tallo, raíz y flor respectivamente (Park *et al.*, 2001). El gen *Tsi1* es homólogo al gen *CpERF2* y los datos del estudio de Park y sus colaboradores contrastan con los resultados de este estudio ya que en el caso de *CpERF2*, el tejido de flor es el que muestra una mayor expresión. Sin embargo hay que recordar que los patrones de expresión realizados en el presente estudio están hechos en base a muestras de plantas que están sometidas a las condiciones ambientales de una plantación comercial y solo pretenden demostrar que los genes de estudio tienen una expresión basal en estas condiciones.

Con excepción del gen *CpERF4*, los otros cuatro genes presentaron una expresión basal constitutiva. Por estudios previos sabemos que los factores de transcripción tipo ERF están involucrados en aspectos de desarrollo, estrés biótico y abiótico (Xu *et al.*, 2008). En condiciones de una plantación comercial sometida a muchos factores ambientales

sería muy precipitado saber que factor o factores están involucrados en los niveles de expresión de los genes de papaya de estudio, pero podríamos suponer que al estar expresados constitutivamente y estar disponibles para la célula, podría haber una respuesta bioquímica más rápida hacia cualquier factor inductivo ya sea biótico o abiótico.

El gen *CpERF4* no mostró una expresión basal bajo las mismas condiciones de amplificación que el resto de los genes, sin embargo fue posible determinar que cuenta con una muy baja expresión ya que con los análisis llevados a cabo por RACE, que cuentan con una doble amplificación debido al PCR anidado dentro de la técnica, fue posible amplificar los extremos 5' y 3' del ADNc indicándonos con esto, que el gen si se encuentra dentro del transcriptoma de la planta, aunque en una condición tan baja que no fue posible ver su expresión en condiciones estándar de PCR. Estos datos concuerdan con un estudio realizado en tomate en el cual no se pudo comprobar la expresión basal del factor de transcripción *TSRF1* en tejidos de hoja, tallo, flor y raíz de tomate (Zhang *et al.*, 2004). Como vimos en el capítulo anterior, el factor de transcripción *TSRF1* presenta homología con *CpERF4* y es posible que ambos genes se comporten de la misma manera en términos de expresión. Zhang y colaboradores comprobaron en su estudio que el gen *TSRF1* es inducible y su nivel de expresión aumenta considerablemente al ser sometida la planta a tratamientos con ácido salicílico, etileno y patógenos, por lo que podríamos esperar que el gen *CpERF4* esté regulado por las mismas hormonas.

3.4.2 Mapeo de los extremos terminales del ADNc de *CpERF1-CpERF5*

El análisis de RACE permitió identificar el inicio y término de la transcripción para cada gen y confirmar el marco de lectura abierta para cada secuencia estudiada. Las regiones no traducibles (UTR) 5' de los genes *CpERF1-CpERF5* variaron en un rango de longitud de 37 pb en *CpERF1* hasta 514 en *CpERF2*, mientras que las regiones no traducibles 3' variaron de 38 pb en *CpERF4* hasta 371 pb en *CpERF2*.

Para el caso del gen *CpERF1*, el análisis RACE de sus extremos terminales del ADNc mostró variantes con respecto al marco de lectura abierta predicho en el genoma de la variedad SunUp de papaya. En la secuencia de la región 5' de *CpERF1* del cultivar

Maradol existe un cambio de una guanina por una alanina que cambia el tercer aminoácido de la proteína de una valina en SunUp por una isoleucina en Maradol, mientras que en la región 3' existe una inserción de 6 nucleótidos (GTCGGC) que generan que la proteína del cultivar Maradol tenga dos residuos de serina (aa 217) y alanina (aa 218) que no tiene la papaya SunUp. Sin embargo estas variaciones no se encuentran dentro de los motivos conservados que Nakano *et al.*, (2006) reporta para este grupo por lo que no se esperaría que tuvieran un efecto importante en la funcionalidad de la proteína. La corta longitud de la región 5' UTR del gen *CpERF1* de sólo 37 pb concuerda con otras regiones 5' UTR cortas reportadas en los genes *Pti4* que cuenta con 46 pb y *Pti5* que cuenta con 23 pb (Zhou et al., 1997). Por otro lado la longitud de la región 3' UTR de *CpERF1* (224 pb) se asemeja a la longitud de la región 3' UTR de su homólogo *Pti4* que cuenta con 214pb (Zhou et al., 1997).

El ADNc de *CpERF2* no mostró diferencias con la predicción *in silico* realizada en el genoma del cultivar SunUp. La longitud de la región 5' UTR de *CpERF2* (514pb) fue mayor que la longitud de la región 5'UTR de *Pti6* (311pb; Zhou et al., 1997) pero ambas están por arriba del promedio de 103 pb de las regiones 5' UTR reportadas en plantas (Pesolea *et al.*, 2001). De manera general se ha observado que las características estructurales de las regiones 5' UTR tienen un papel importante en el control de la traducción de los ARNm al generarles mayor estabilidad. Se ha reportado que algunos genes que participan en procesos de desarrollo usualmente cuentan con regiones 5' UTR más largas que el promedio (Kozak, 1987; Mignone *et al.*, 2002). Esto puede estar correlacionado con el hecho de que en análisis de expresión basal, el gen *CpERF2* parecía tener una mayor expresión en tejido de flor y probablemente esté involucrado también en procesos de desarrollo como ha sido observado en otros genes tipo ERF (Nakano et al., 2006), y particularmente en desarrollo floral (Elliott et al., 1996).

El mapeo por RACE permitió elucidar correctamente la estructura del gen *CpERF3*. El análisis *in silico* indicó la presencia de 2 intrones en la región amino anterior al dominio AP2/ERF del ORF predicho. El 5' RACE arrojó una secuencia de 35 pb que no había sido identificada por el análisis *in silico* y confirmó que la secuencia *CpERF3* sólo cuenta con un intrón anterior al dominio de unión a ADN como ha sido reportado para varios

miembros del grupo VII (Cheong, *et al.*, 2003; Nakano *et al.*, 2006; Xu, *et al.*, 2007; Zhang *et al.*, 2009). El fragmento 5' RACE tiene 100% de identidad con una región del contig ABIM01025280 de papaya SunUp, lo que indica que no hay variación entre los dos cultivares en esta región. El análisis 3' RACE permitió conocer el marco de lectura abierta completo que consta de una secuencia de 395 aa y cuenta con 60% de identidad con el gen GmERF3 de soya (Zhang *et al.*, 2009) y 41% de identidad con el gen TaERF1 de trigo (Xu *et al.*, 2007). Se observó también el codón de paro y la presencia del motivo de fosforilación TPD/EI/Y/S/TS en el gen CpERF3, el cual se encuentra descrito para las proteínas OsEREBP1, CaPF1 y TaERF (Cheong, *et al.*, 2003; Yi *et al.*, 2004; Xu *et al.*, 2007). La región terminal del marco de lectura abierta y el motivo de fosforilación no habían sido localizados *in silico* debido a que el marco de lectura del gen CpERF3 se encuentra incompleto en el extremo del contig ABIM01025280 de papaya.

El gen *CpERF4* no mostró ninguna diferencia en sus regiones terminales con respecto a la predicción *in silico* del ORF, cuenta con una región 5' UTR de 57 pb similar a la región 5' UTR de *ERF1* que cuenta con 61 pb, sin embargo la región 3' UTR es considerablemente menor que la región 5' UTR ya que sólo cuenta con 38 pb. Esto contrasta con el patrón general que muestra las regiones 3' UTR mayores a las regiones 5' UTR (Mignone *et al.*, 2002). Las regiones no traducibles 3' del ARNm tienen funciones que le generan estabilidad y localización (Mignone *et al.*, 2003). Estas funciones son mediadas por sus interacciones con proteínas regulatorias y con ARN's incluyendo micro ARNs (miRNAs). En mamíferos cerca de la mitad de los genes utilizan sitios alternos de poliadenilación que generan múltiples isoformas del ARNm que difieren sobre todo en sus regiones 3' UTR (Sandberg *et al.*, 2008). Para algunos ARNm tener sitios de interacción con miRNAs pueden ser tan perjudicial que se ejerce sobre ellos una presión selectiva para tener solo las isoformas con UTRs cortos (Farh *et al.*, 2005). De esta manera se ha visto que aumentar la expresión relativa de las isoformas con regiones 3' UTR cortas ayuda a la regulación positiva de la expresión de estos genes (Sandberg *et al.*, 2008). Los análisis de expresión del gen *CpERF4* no mostraron expresión en ningún tejido de la planta, y por tanto podríamos suponer que el tener una región 3' UTR corta, pudiera relacionarse con su falta de expresión al haber posibles miRNAs que pudieran estar

implicados en el silenciamiento de este gen, sin embargo hacen falta más estudios para comprobar esta hipótesis.

El gen CpERF5 al igual que CpERF3 tuvo una predicción incorrecta del inicio del ORF por parte del programa FGENESH. El análisis del fragmento 5' RACE comprobó que la proteína CpERF5 tiene una secuencia de 285 aminoácidos, que representan 87 aminoácidos menos que la predicción *in silico* hecha con el genoma del cultivar SunUp de papaya. Se ha visto en *Arabidopsis* que los programas de predicción de marcos de lectura abierta predicen incorrectamente uno de cada cinco exones en alguno de sus bordes, lo que se observa claramente en este caso (Brendel y Zhu, 2002; Rhee *et al.*, 2006). El marco de lectura predicho *in silico* contenía al menos 3 metioninas que pudieran haber sido el inicio del ORF, sin embargo ninguna de ellas se encuentra en el ADNc de las distintas clonas que fueron secuenciadas. La proteína del gen CpERF5 mantiene los motivos estructurales típicos del grupo I (Nakano *et al.*, 2006) y elimina la región amino terminal de la predicción *in silico* que había sido muy ambigua en el alineamiento del grupo I del capítulo anterior. Las tres variaciones encontradas dentro del marco de lectura con respecto a la predicción no tienen influencia en modificar ningún aminoácido por lo que se esperaba que la proteína CpERF5 tenga las características que fueron discutidas en el capítulo anterior.

Solo el gen *CpERF1* presentó el típico sitio de poliadenilación AAUAAA. Los genes *CpERF2-CpERF5* cuentan con variantes de este sitio lo cual es común que se presente en plantas ya que se ha observado que solo el 10% de los genes en *Arabidopsis* cuentan con el sitio AAUAAA, el cual es mucho más común en animales, mientras que en plantas este sitio llamado NUE (Near Upstream Element) tiene una gran plasticidad a las mutaciones y forma parte de una estructura tripartita que incluye los FUE (Far Upstream Element) y el sitio de anclaje (CS; Cleavage site). Los sitios FUE no fueron identificados debido a que no existe un patrón para estas secuencias y son extremadamente variables mientras que los sitios de anclaje coincidieron con las bases Y/A que han sido reportadas en otras especies (Rhotnie, 1996; Graber *et al.*, 1999; Loke *et al.*, 2005).

3.5 CONCLUSIÓN

1. El presente capítulo de investigación demostró que los genes *CpERF1-CpERF5* están presentes dentro del genoma de papaya Maradol, estos genes se encuentran expresados en todos los tejidos de la parte aérea de una planta sometida a las condiciones ambientales de una plantación comercial, con excepción del gen *CpERF4* cuya expresión fue baja pudiéndose detectar de manera clara únicamente en el PCR anidado del RACE. La expresión de *CpERF4* podría estar regulada de manera positiva por algún inductor como por ejemplo, ácido salicílico o etileno, que son hormonas implicadas en estrés biótico o también por el ataque de algún patógeno.

2. En general se presentaron pocas variaciones de los genes aislados en la variedad Maradol con respecto a las predicciones hechas en el genoma de la variedad SunUp, y estas variaciones no generan cambios estructurales importantes en los motivos o dominios de cada uno de los genes, sin embargo se observaron dos marcos de lectura abierta (*CpERF3* y *CpERF5*) que habían sido mal predichos por el programa FGENESH, por lo que se resalta la importancia de hacer el aislamiento *in vitro* para comprobar las predicciones *in silico* y obtener los ORFs correctos.

3 La disponibilidad de los ADNc completos de *CpERF1-CpERF5* permitirá realizar trabajos posteriores para evaluar tanto su función en relación al ataque de patógenos, así como también evaluar el potencial biotecnológico de estas secuencias para la generación de variedades cisgénicas de papaya Maradol resistentes a la antracnosis y otras enfermedades.

3.6 BIBLIOGRAFIA

- Brendel, V. y W. Zhu. (2002). Computational modeling of gene structure in *Arabidopsis thaliana*. *Plant Molecular Biology*, 48, 49–58.
- Cheong, Y.H., B.C. Moon, J.K. Kim, M.C. Kim, I.H. Kim, C.Y. Park, J.C. Kim, B.O. Park y S.C. Koo. (2003). BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcription factor. *Plant Physiology*, 132, 1961-1972.
- Claridades Agropecuarias. (1999). Papaya un mercado en expansión. Claridades agropecuarias. México. Marzo. pp 3-24.
- Gutterson, N. y T.L. Reuber. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Current Opinion in Plant Biology*, 7, 465–471.
- Graber, J., S. Cantor, A. Mohr y T. Smith. (1999). In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species. *Proceedings of the National Academy of Sciences*, 96, 14055–14060.
- Elliott, R.C., A.S. Betzner, E. Huttner, M.P. Oakes, W.Q. Tucker, D. Gerentes, P. Perez y D.R. Smyth. (1996). AINTEGUMENTA, an APETALA2-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *Plant Cell*, 8, 155–168.
- Farh K.K.H., A. Grimson, C. Jan, B.P. Lewis, W.K. Johnston, L.P. Lim, C.B. Burge y D.P. Bartel. (2005). The Widespread Impact of Mammalian MicroRNAs on mRNA Repression and Evolution. *Science*, 310, 1817-1821.
- Kieffer, E. W. Heller y D. Ernst. (2000). A simple and efficient protocol for isolation of functional RNA from plant tissues rich in secondary metabolites. *Plant Molecular Biology Reporter*, 18, 33-39.
- Kozak, M. (1987). An analysis of 5' –noncodign sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*, 15, 8125-8148.
- Loke, J., E. Stahlberg, D. Strenski, N. Haas, P. Wood y Q. Li. (2005). Compilation of mRNA Polyadenylation Signals in *Arabidopsis* Revealed a New Signal Element and Potential Secondary Structures1. *Plant Physiology*, 138, 1457–1468.

- Massieu-Trigo, Y. (2009). Exportaciones no tradicionales (ENT) y papaya transgénica. La diversidad genética en la globalización. Ponencia para el XXVIII Congreso de Latin American Studies Association Rio de Janeiro, Brasil 11 al 14 de junio de 2009.
- Mignone, F., C. Gissi, S. Liuni y G. Pesole. (2002). Untranslated regions of mRNAs. *Genome Biology*, 3, reviews0004.1-0004.10
- Ming, R., S. Hou, Y. Feng, Q. Yu, A. Dionne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis, S.L. Salzberg, L. Feng, M.R. Jones, R.L. Skelton, J.E. Murray, C. Chen, W. Qian, J. Shen, P. Du, M. Eustice, E. Tong, H. Tang, E. Lyons, R.E. Paull, T.P. Michael, K. Wall, D.W. Rice, H. Albert, M.L Wang, Y.J. Zhu, M. Schatz, N. Nagarajan, R.A. Acob, P. Guan, A. Blas, C.M. Wai, C.M. Ackerman, Y. Ren, C. Liu, J. Wang, J. Wang, J.K. Na, E.V. Shakirov, B. Haas, J. Thimmapuram, D. Nelson, X. Wang, J.E. Bowers, A.R. Gschwend, A.L. Delche, R. Singh, J.Y. Suzuki, S. Tripathi, K. Neupane, H. Wei, B. Irikura, M. Paidi, N. Jiang, W. Zhang, G. Presting, A. Windsor, R. Navajas-Pérez, M.J. Torres, F.A. Feltus, B. Porter, Y. Li, A.M. Burroughs, M.C. Luo, L. Liu, D.A. Christopher, S.M. Mount, P.H. Moore, T. Sugimura, J. Jiang, M.A. Schuler, V. Friedman, T. Mitchell-Olds, D.E. Shippen, C.W. dePamphilis, J.D. Palmer, M. Freeling, A.H. Paterson, D. Gonsalves, L. Wang y M. Alam. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452, 991–996.
- Nakano, T., K. Suzuki, T. Fujimura, y H. Shinshi. (2006). Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. *Plant Physiology*, 140, 411–432.
- Park, J.M., C.J. Park, S.B. Lee, B.K. Ham, R. Shin, K.H. Paek. (2001) Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell*, 13, 1035-1046.
- Pesolea, G., F. Mignonea, C. Gissia, G. Grillob, F. Licciullib y S. Liuni. (2001). Structural and functional features of eukaryotic mRNA untranslated regions. *Gene*, 276, 73–81.
- Quiagen, (2001). The Bench Guide. Protocols, hints, and tips for molecular biology labs. Disponible en: <http://www.qiagen.com/literature/benchguide/default.aspx>
Acceso 30 de Agosto de 2010.
- Rhee, S., J. Dickerson y D. Xu. (2006). Bioinformatics and its applications in plant biology. *Annual Review of Plant Biology*, 57, 335-360.
- Rothnie, H.M. (1996). Plant mRNA 3'-end formation. *Plant Molecular Biology*, 32, 43–61.
- SAGARPA, (2010). SISTEMA PRODUCTO.

Disponible en:

<http://www.sagarpa.gob.mx/agricultura/Publicaciones/SistemaProducto/Paginas/default.aspx> Acceso: 30 de Agosto de 2010.

- Sambrook, J. y D. Rusell. (2001). Preparation of Plasmid DNA by Alkaline Lysis with SDS: Minipreparation. *In* Molecular Cloning. A laboratory manual. 3ra edición. Cold Spring Harbor Laboratory Press. pp 1.32-1.35.
- Sandberg, R., J. Neilson, A. Sarma, P. Sharpn y C. Burge. (2008). Proliferating Cells Express mRNAs with Shortened 3' Untranslated Regions and Fewer MicroRNA Target Sites. *Science*, 320, 1643-1646.
- Van Verk, M., C. Gatz y H. Linthorst. (2009). Transcriptional Regulation of Plant Defense Responses. *Advances in Botanical Research*, 51, 397-438.
- Xu, Z.S., L.Q. Xia, M. Chen, R.Y. Zhang, L.C. Li, Y.X. Zhao, Y. Lu, Z.Y. Ni, L. Liu, Z.G. Qiu y Y.Z. Ma. (2007). Isolation and molecular characterization of the *Triticum aestivum* L. Ethylene-responsive factor 1 (*TaERF1*) that increase multiple stress tolerance. *Plant Molecular Biology*, 65, 719-732.
- Xu, Z.S., M. Chen, L.C. Li y Y.Z. Ma. (2008). Functions of the ERF transcription factor family in plants. *Botany*, 86, 969-977.
- Yi, S.Y., J.H. Kim, Y.H. Joung, S. Lee, W.T. Kim, S.H. Yu y D. Choi. (2004). The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in *Arabidopsis*. *Plant Physiology*, 136, 2862-2874.
- Zhang, H., D. Zhang, J. Chen, Y. Yang, Z. Huang, D. Huang, X.C. Wang y R. Huang. (2004). Tomato stress-responsive factor TSRF1 interacts with ethylene responsive element GCC box and regulates pathogen resistance to *Ralstonia solanacearum*. *Plant Molecular Biology*, 55, 825-834.
- Zhang, G., M. Chen, L. Li, Z. Xu, X. Chen, J. Guo y Y. Ma. (2009). Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *Journal of Experimental Botany*, 60, 3781-3796.
- Zhou, J., X. Tang y G. Martin. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that binds a cis-elements of pathogenesis-related genes. *The EMBO Journal*, 16, 3207-3218.

CAPITULO IV

4.1 DISCUSIÓN GENERAL

El estudio de factores de transcripción implicados en la regulación de rutas metabólicas que generan tolerancia de plantas a condiciones estresantes, se presenta como el futuro en las aplicaciones biotecnológicas para el desarrollo de variedades resistentes a diferentes condiciones ambientales que afectan la productividad y la comercialización de los cultivos agrícolas (Century *et al.*, 2008).

A finales de los años noventa se observó que los factores de transcripción Pti4 y ERF1 formaban parte de las rutas de señalización mediada por ácido salicílico y etileno respectivamente para la activación de genes PR en la respuesta inmune de las plantas (Zhou *et al.*, 1997; Solano *et al.*, 1998). A partir de entonces, se han caracterizado factores de transcripción del tipo ERF en plantas de diferentes taxa y se ha comprobado su función en conferir resistencia a estrés biótico. La familia de factores de transcripción tipo ERF ha sido caracterizada en genomas de *Arabidopsis* (*Arabidopsis thaliana*), arroz (*Oriza sativa*), soya (*Glycine max*) y álamo (*Populus trichocarpa*) y se ha comprobado que miembros de esta familia tienen funciones en el desarrollo de la planta así como en condiciones de estrés biótico y abiótico. Estos factores de transcripción únicos de plantas han sido propuestos como candidatos ideales para mejorar la resistencia de los cultivos a condiciones de estrés biótico (Gutterson y Reuber, 2004; Xu *et al.*, 2008). Con la disponibilidad pública del genoma de papaya (Ming *et al.*, 2008) se observaron 146 secuencias que contienen el dominio AP2/ERF en este cultivo tropical.

En este trabajo se lograron identificar y caracterizar cinco genes tipo ERF en el genoma de la papaya transgénica SunUp que son homólogos a genes de otras especies y de las que se ha observado su funcionalidad en conferir resistencia a patógenos. Dichos genes, denominados en este estudio como *CpERF1-CpERF5* son homólogos a los genes *Pti4*, *Tsi1*, *GmERF3*, *ERF1* y *JcERF* respectivamente; cuentan dentro de las estructuras β -plegadas de su dominio AP2/ERF con residuos de triptófano y arginina (Arg-6, Arg-8, Trp-

10, Glu-16, Arg-18, Arg-26 y Trp-28) que permitirían a estas proteínas unirse a “cajas GCC” en regiones promotoras de genes PR (Allen, *et al.* 1998; Ohme-Takagi y Shinshi, 1995). CpERF1 y CpERF3 cuentan con motivos dentro de su secuencia de nucleótidos que codifican para regiones que interactúan con proteínas del tipo cinasa que aumentan su afinidad a las cajas “GCC” (Xu *et al.*, 2008). También se observaron en todas las secuencias de estudio, motivos que codifican para regiones ricas en argininas y lisinas, las cuales pueden funcionar como sitios de localización nuclear (Dingwall y Laskey, 1991; Christophe *et al.*, 2000), característica muy importante para funcionar como factores de transcripción.

Las proteínas generalmente se clasifican en base a su secuencia, estructura y función. Una familia de proteínas puede estar representada en base a un árbol filogenético que muestre las relaciones evolutivas entre las proteínas. Un análisis filogenético puede ser usado de esta manera para inferir transferencia horizontal de genes, realizar genómica comparativa y/o predecir funciones de proteínas entre otras cosas (Rhee *et al.*, 2006). Los alineamientos de la secuencia de proteínas de CpERF1-CpERF5 y secuencias previamente caracterizadas de factores tipo ERF, permitió observar que las secuencias de estudio se encuentran filogenéticamente relacionadas con factores tipo ERF involucrados en mecanismos de defensa de la plantas contra patógenos, por tanto es muy probable que las secuencias de estudio sean factores de transcripción de la familia ERF que activen genes PR de papaya contribuyendo al sistema de defensa de esta planta. El análisis realizado coincide con estudios filogenéticos realizados por otros autores que proponen a este tipo de proteínas como excelentes candidatos para el mejoramiento genético de plantas (Gutterson y Reuber, 2004; Nakano *et al.*, 2006; Xu *et al.*, 2008).

Con la excepción del gen CpERF4, el cual muestra una expresión tan baja que sólo pudo ser observada mediante la PCR anidada del RACE, los otros cuatro genes de papaya descritos en este estudio mostraron expresión en plantas adultas sometidas a las condiciones ambientales de una plantación comercial. Esto supone que las proteínas se encuentran disponibles para las plantas cuando existe un estímulo que active las rutas bioquímicas donde están involucradas estas proteínas, incluyendo las rutas bioquímicas de defensa que activan la expresión de genes PR (Van Verk *et al.*, 2009). Es decir, es posible que su expresión basal esté implicada en el reconocimiento de patógenos

virulentos (interacción incompatible) para conferir resistencia a los posibles patógenos que se encuentren en la plantación comercial. Sin embargo hay que tener en cuenta que aun cuando estén activados los mecanismos de defensa basales, los patógenos podrían tener mecanismos moleculares para evadir esta respuesta y provocar la enfermedad (interacción compatible). En diversos estudios se ha observado que en las interacciones planta-patógeno tanto compatibles como incompatibles se ejerce una serie de regulaciones tanto positivas como negativas de diversos genes implicados en estas interacciones. A pesar de que estén activados los genes, en muchos casos la diferencia entre la susceptibilidad o resistencia a un patógeno se encuentra asociado con las diferencias en tiempo y magnitud en que ocurren los cambios de expresión de estos genes. Por lo que una respuesta rápida en la regulación de estos genes provoca un aumento en la resistencia al patógeno (Maleck *et al.*, 2000; Van Loon *et al.*, 2006).

Hasta ahora son pocos los estudios que muestran la expresión basal de los factores tipo ERF en diversos tejidos de la planta, y no se ha reportado en ningún caso el comportamiento de estos en condiciones de campo por lo que sería interesante evaluar la expresión de los genes de papaya en condiciones de laboratorio en los cuatro tejidos de estudio. Un caso interesante fue el del gen CpERF4 que no presentó una expresión notable en las condiciones estándar de PCR llevadas a cabo en este estudio aunque su expresión fue detectada en la PCR anidada del RACE. El gen *TSRF1* homólogo a CpERF4 no mostró una expresión basal en tejidos de tallo hoja flor y raíz en estudios por Northern Blot, pero se observó que este gen era inducido al estimular la planta con ácido salicílico, etileno y patógenos y de esta manera logró generar resistencia a patógenos al ser sobre-expresado en plantas de tabaco (Zhang *et al.*, 2004). Por lo que CpERF4, a pesar de mostrar una expresión basal baja podría ser igualmente inducible por alguna fitohormona relacionada con estrés biótico lo que convierte a este gen en el de mayor interés para futuros estudios biotecnológicos ya que esta característica de ser un gen posiblemente inducible podría ser de enorme utilidad para la generación de una variedad de papaya resistente a patógenos mediante la expresión controlada de este gen. La disponibilidad de estas secuencias permitirá llevar a cabo análisis de funcionalidad que comprueban que los genes CpERF1-CpERF5 son factores de transcripción que generan tolerancia a estrés biótico mediante la activación de genes PR.

Encontrar genes de utilidad para el desarrollo biotecnológico de nuevas variedades de plantas se hace cada vez más importante en un mundo en constante crecimiento poblacional que contrasta con una menor producción de alimentos. Los Organismos Genéticamente Modificados (OGM) se presentan como una alternativa para hacer frente a las limitantes ambientales en la producción de alimentos. Por otra parte se ha visto que su comercialización ha tenido fuerte impacto económico en los últimos años generando un mercado de cerca de 7000 millones de dólares en 2006 (Brooks y Barfoot, 2008). También han tenido fuerte impacto ambiental al evitar el uso de cerca de 300 millones de kg de ingredientes activos de distintos pesticidas así como la disminución de 1.2 millones de toneladas de gases que se dejaron de emitir por la aplicación de pesticidas (Brooks y Barfoot, 2008).

Ante el debate nacional e internacional sobre el uso de OGMs en cultivos que puedan transferir los genes introducidos a poblaciones silvestres, encontrar genes propios de la misma especie, que tengan utilidad comercial, es muy importante para la generación de organismos cisgénicos, los cuales tienen la ventaja de no contar con material genético externo a la especie que pueda ser transferido a las poblaciones silvestres. Se ha propuesto que los organismos cisgénicos a pesar de ser igualmente manipulados genéticamente que una variedad transgénica, son similares a los organismos obtenidos por cruces tradicionales, pero con la ventaja de que en los cisgénicos sólo se transfiere el gen de interés (Shouten *et al.*, 2006).

Se espera que los resultados del presente trabajo puedan contribuir al desarrollo biotecnológico de variedades cisgénicas de papaya que tengan las características comerciales de las que goza la variedad Maradol y puedan además ser más tolerantes a las diversas plagas y enfermedades que afectan a este cultivo, incluyendo la antracnosis que es la principal limitante de la comercialización de este producto al causar grandes pérdidas a nivel post-cosecha (Massieu-Trigo, 2009).

4.2 CONCLUSIONES

1. El genoma de la papaya transgénica SunUp cuenta con cinco secuencias altamente similares a los factores de transcripción ERF, Pti4 y Tsi1 de tomate, GmERF3 de soya, ERF1 de *Arabidopsis* y JcERF de *Jatropha*. Estas cinco secuencias de papaya nombradas como CpERF1-CpERF5 contienen el dominio AP2/ERF, así como las estructuras secundaria y terciaria necesarias para unirse a elementos tipo cajas GCC que se encuentran en las regiones promotoras de genes de defensa PR.

2. Los análisis filogenéticos demostraron que las secuencias CpERF1-CpERF5 de papaya son homólogas a factores de transcripción de la familia ERF que han sido previamente caracterizados en otras especies, lo que hace suponer que las secuencias de papaya responden a hormonas involucradas en estrés biótico como ácido salicílico, ácido jasmónico y etileno, así como por el ataque de patógenos.

3. Los ARNm de los genes CpERF1-CpERF5 se encuentran presentes en el transcriptoma de plantas adultas de papaya Maradol provenientes de una plantación comercial. CpERF1, CpERF2, CpERF3 y CpERF5 se expresan de manera constitutiva en tejidos de la parte aérea de la planta, mientras el perfil de expresión constitutivo de estos genes es similar a la de otros genes involucrados en la resistencia a patógenos.

4. La expresión de CpERF4 es muy baja en todos los tejidos estudiados siendo casi nula y solo se detectó mediante el RACE. Debido a que la técnica implica una doble amplificación del ADNc mediante un PCR anidado que ayuda que la baja expresión pueda ser detectada.

4. Se clonaron las regiones 5' y 3' de los ADNc de las secuencias de estudio que incluyeron las regiones 5'UTR y 3'UTR, y se corroboraron las predicciones del inicio y término de los ORFs. Los genes CpERF3 y CpERF5 tuvieron errores en la predicción *in silico* por parte del programa FGENESH, por lo que se comprobó la importancia de corroborar *in vitro* la secuencia de los extremos de los ORFs.

4.3 PERSPECTIVAS

Los genes CpERF1-CpERF5 reúnen todas las características estructurales para actuar como factores de transcripción dentro de las rutas de ácido salicílico y/o etileno que conducen a activar la expresión de genes PR. Sin embargo, es indispensable hacer análisis funcionales que comprueben este supuesto. Se espera que al someter plantas de papaya a tratamientos con ácido salicílico, etileno o algún tipo de patógeno, se aumente la expresión de CpERF1-CpERF5 principalmente de CpERF4 que no mostró expresión basal, lo cual podría ser evaluado mediante RT-PCR de tiempo real. Se puede inferir que el aumento de los niveles de expresión de estos genes aumente proporcionalmente los niveles de expresión de uno o varios genes PR propios de papaya como se ha observado en otras especies de plantas.

Análisis de interacción Proteína-ADN como los ensayos de cambio de movilidad por electroforesis (EMSA) deben comprobar que los dominios AP2/ERF de las proteínas CpERF1-CpERF5 se unen a elementos que actúan en Cis del tipo “GCC” que se encuentran en las regiones promotoras de diversos genes PR y de esta manera generar evidencia de la funcionalidad de estas proteínas como factores de transcripción. Los casos de CpERF3 y CpERF5 se espera que también puedan tener una actividad de unión a elementos del tipo DRE/CRT que se encuentran en las regiones promotoras de genes que responden a estrés abiótico como salinidad y baja temperatura, ya que sus genes homólogos TaERF1 y JcERF han mostrado tener esta habilidad de respuesta dual ante condiciones de estrés biótico y abiótico.

De igual manera se puede esperar que ensayos de interacción proteína-proteína del tipo dos híbridos, demuestren que al menos las proteínas CpERF1 y CpERF3 (homólogos a Pti4 y TaERF1 respectivamente) interactúen con proteínas del tipo cinasas, ya que se ha visto que tanto Pti4 como TaERF1 interactúan con este tipo de proteínas para aumentar su afinidad a las cajas GCC (Van Verck *et al.*, 2009). En general, CpERF1-CpERF5 cuentan con motivos que podrían estar involucrados en la interacción con otras proteínas, por lo que los ensayos de dos híbridos permitirán comprobar con qué tipo de proteínas

interactúan las proteínas tipo ERF de papaya y de esta manera generar más evidencias acerca de su función.

Finalmente, se espera que estudios de sobreexpresión y/o silenciamiento de estos genes, dentro de papaya o en plantas modelo como *Arabidopsis* o Tabaco ayude a tener una mayor comprensión de su funcionamiento y de sus probables implicaciones en conferir resistencia a patógenos. El gen CpERF4 que no mostró una expresión basal considerable se presenta como el más atractivo para uso biotecnológico ya que su expresión aparentemente está fuertemente regulada probablemente por inductores de estrés biótico. De ser así, estos genes y en especial CpERF4 serían candidatos atractivos para ser manipulados por tecnología de ADN recombinante con la finalidad de generar nuevas variedades de papaya que sean resistentes a diversos tipos de patógenos. Se espera que estas variedades cisgénicas de papaya permitan minimizar el uso de pesticidas tóxicos para el medio ambiente y salud humana, así como reducir los costos de producción de este cultivo tropical.

4.4 BIBLIOGRAFIA

- Allen, M., y K. Yamasaki. (1998) A novel mode of DNA recognition by a b-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, 17, 5484–5496.
- Century, K., T. Lynne-Reuber, y O. Ratcliffe. (2008) Regulating the Regulators: The Future Prospects for Transcription-Factor-Based Agricultural Biotechnology Products. *Plant Physiology*, 147, 20–29.
- Christophe D., C. Christophe-Hobertus y B. Pichon. 2000. Nuclear targeting of proteins: how many different signals?. *Cell Signal*, 12, 337.341.
- Dingwall, C., y R. Laskey. (1991). Nuclear targeting sequences - a consensus?. *Trends in Biochemical Sciences*, 16, 478-81.
- Gutterson, N. y T.L. Reuber. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Current Opinion in Plant Biology*. 7. 465–471.
- Maleck, K., A. Levine, T. Eulgem, A. Morgan, J. Schmid, K. Lawton, J.L. Dang y RA. Dietrich. (2000). The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nature Genetics*, 26, 403-410.
- Massieu-Trigo, Y. (2009). Exportaciones no tradicionales (ENT) y papaya transgénica. La diversidad genética en la globalización. Ponencia para el XXVIII Congreso de Latin American Studies Association Rio de Janeiro, Brasil 11 al 14 de junio de 2009.
- Ming, R., S. Hou, Y. Feng, Q. Yu, A. Dionne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis, S.L. Salzberg, L. Feng, M.R. Jones, R.L. Skelton, J.E. Murray, C. Chen, W. Qian, J. Shen, P. Du, M. Eustice, E. Tong, H. Tang, E. Lyons, R.E. Paull, T.P. Michael, K. Wall, D.W. Rice, H. Albert, M.L Wang, Y.J. Zhu, M. Schatz, N. Nagarajan, R.A. Acob, P. Guan, A. Blas, C.M. Wai, C.M. Ackerman, Y. Ren, C. Liu, J. Wang, J. Wang, J.K. Na, E.V. Shakirov, B. Haas, J. Thimmapuram, D. Nelson, X. Wang, J.E. Bowers, A.R. Gschwend, A.L. Delche, R. Singh, J.Y. Suzuki, S. Tripathi, K. Neupane, H. Wei, B. Irikura, M. Paidi, N. Jiang, W. Zhang, G. Presting, A. Windsor, R. Navajas-Pérez, M.J. Torres, F.A. Feltus, B. Porter, Y. Li, A.M. Burroughs, M.C. Luo, L. Liu, D.A. Christopher, S.M. Mount, P.H. Moore, T. Sugimura, J. Jiang, M.A. Schuler, V. Friedman, T. Mitchell-Olds, D.E. Shippen, C.W. dePamphilis, J.D. Palmer, M. Freeling, A.H. Paterson, D. Gonsalves, L. Wang y M. Alam. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452, 991–996.

- Nakano, T., K. Suzuki, T. Fujimura, y H. Shinshi. (2006). Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. *Plant Physiology*, 140, 411–432.
- Ohme-Takagi, M., y H. Shinshi. (1995). Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element. *The Plant Cell*, 7, 173-182.
- Rhee, S., J. Dickerson y D. Xu. (2006). Bioinformatics and its applications in plant biology. *Annual Review of Plant Biology*, 57, 335-360.
- Schouten, H.J., F.A. Krens y E. Jacobsen. (2006) Cisgenic plants are similar to traditionally bred plants. *EMBO Reports*, 7, 750–753.
- Solano, R., A. Stepanova, A. Chao y J. Ecker. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ethylene-insensitive3 and ethylene-response-factor1. *Genes and Development*, 12, 3703–3714.
- Van Loon, L.C., M. Rep y C.M.J. Pieterse. (2006). Significance of Inducible Defense-related Proteins in Infected Plants. *Annual Review of Phytopathology*, 44, 135–62.
- Van Verk, M., C. Gatz y H. Linthorst. (2009). Transcriptional Regulation of Plant Defense Responses. *Advances in Botanical Research*, 51, 397-438.
- Xu, Z.S., M. Chen, L.C. Li y Y.Z. Ma. (2008). Functions of the ERF transcription factor family in plants. *Botany*, 86, 969-977.
- Zhang, H., D. Zhang, J. Chen, Y. Yang, Z. Huang, D. Huang, X.C. Wang y R. Huang. (2004). Tomato stress-responsive factor TSRF1 interacts with ethylene responsive element GCC box and regulates pathogen resistance to *Ralstonia solanacearum*. *Plant Molecular Biology*, 55, 825–834.
- Zhou, J., X. Tang y G. Martin. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that binds a cis-elements of pathogenesis-related genes. *The EMBO Journal*, 16, 3207-3218.

