



Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

**ESTUDIOS MOLECULARES DE SECUENCIAS
CANDIDATAS A GENES DE RESISTENCIA EN
COCOTERO**

Tesis que presenta

CARLOS ALBERTO PUCH HAU

En opción al título de

DOCTOR EN CIENCIAS

(Ciencias Biológicas: Opción Biotecnología)

Mérida, Yucatán, México

2015



RECONOCIMIENTO

Por medio de la presente, hago constar que el trabajo de tesis titulado: ESTUDIOS MOLECULARES DE SECUENCIAS CANDIDATAS A GENES DE RESISTENCIA EN COCOTERO fue realizado en los laboratorios de la Unidad de Biotecnología del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del Dr. Luis Alfonso Sáenz Carbonell, dentro de la Opción biotecnología, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Atentamente,

Dr. Manuel Martínez Estevez

Director de Docencia

DECLARACIÓN DE PROPIEDAD

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

Firma: _____



Nombre: Carlos Alberto Puch Hau

DEDICATORIA

Esta tesis es dedicada a toda mi familia, en especial a mi esposa **Areli Gamboa** y a mis hijos **Ancel Arturo** y **Alberto Adonahi**, por su apoyo incondicional durante este proceso de mi vida profesional. Les agradezco, cada momento y cada actitud que han alimentado con motivación mi vida académica.

A mis **padres**, por todo su apoyo incondicional para lograr un paso más en mi vida profesional y porque este trabajo es resultado de un proceso de formación lleno de valores y actitudes que fueron transmitidos por ustedes.

AGRADECIMIENTOS

- A mi asesor, el **Dr. Luís Alfonso Sáenz Carbonell**, por darme la valiosa oportunidad y confianza de trabajar en este proyecto. Por su asesoría en el diseño de experimentos, en el análisis de resultados y en la publicación de artículos científicos. Le agradezco incondicionalmente todo el apoyo brindado durante todo mi proceso de formación profesional.
- Al proyecto **Common Found for commodities** (CFC FIG00/22) y al proyecto de CONACYT (CB 129717) por brindarme los recursos necesarios para realizar este proyecto de investigación.
- Al **CONACYT** por la beca otorgada No **240244** para la realización de los estudios de doctorado.
- Al **Dr. Carlos Oropeza Salín**, por todas las facilidades brindadas, por sus valiosas aportaciones en este trabajo de investigación, pero sobre todo por la gran confianza brindada dentro del grupo de cocotero.
- Al **Dr. Daniel Zizumbo Villareal**, por su asesoría brindada sobre el germoplasma de cocotero para la búsqueda de marcadores moleculares.
- Al **Dr. Santy Peraza Echeverría**, por sus valiosas aportaciones y revisiones durante el desarrollo de este proyecto de investigación.
- A comité tutorial conformado por el **Dr. Luis Alfonso Sáenz Carbonell**, **Dr. Carlos Oropeza Salín**, **Dr. Felipe Sánchez Teyer**, **Dr. José María Tun Suárez** y al **Dr. Daniel Zizumbo Villareal** por sus observaciones, comentarios y sugerencias que fueron de gran importancia para dirigir y concluir apropiadamente este proyecto.
- A mi comité revisor conformado por el **Dr. Luis Alfonso Sáenz Carbonell**, **Dr. Carlos Oropeza Salín**, **Dr. Felipe Sánchez Teyer**, **Dr. José María Tun Suárez**, **Dr. Daniel Zizumbo Villareal**, **Dr. Santy Peraza Echeverría** y a la **Dra. Mayra Montero Cortez** por sus valiosas observaciones y aportaciones en este presente documento.
- Al **M. en C. Iván Córdova Lara**, por toda la asesoría técnica brindada para el muestreo de tejido vegetal para la extracción de ácidos nucleicos, la reacción en cadena de la polimerasa en tiempo real y convencional y el análisis de resultados. También te agradezco incondicionalmente todos los conocimientos transmitidos durante esta y cada una de las etapas de mi formación profesional.

- A la **M en C. Adriana Quiroz Moreno**, por su asesoría técnica en la implementación de la técnica de AFLP-RGCs, en el uso de la técnica de electroforesis capilar y en el análisis de resultados.
- A la **M. en C. María Narváez** por el apoyo técnico y logístico en la preparación y envío de muestras para la secuenciación, así como también por todas las facilidades brindadas.
- A la **Biol. Verónica Limones Briones**, por su apoyo técnico en el desarrollo de la técnica de electroforesis en geles de poliacrilamida.
- A los técnicos: **Ing. Miguel Fernández Barrera, Ing. José Luis Chan Rodríguez** y al **Tec. Agric. Nelson Torres Hernández** por su apoyo técnico para la colección de tejido vegetal de cocotero.
- A los trabajadores de la plantación experimental de “Caridad del Cobre”, Santiago Chale Poot y Manuel Chale Pech por su valioso apoyo en la colecta de material vegetal.
- Al Posgrado en Ciencias Biológicas opción Biotecnología del Centro de Investigación Científica de Yucatán por darme las herramientas solidas y por dirigirme de manera congruente para obtener el grado de doctor en Ciencias Biológicas.
- Al **Centro de Investigación Científica de Yucatán**, por todas las instalaciones prestadas para llevar a cabo este trabajo de investigación y por darme la valiosa oportunidad para formarme como Doctor en Ciencias Biológicas.
- A todos los compañeros y excompañeros del laboratorio, **German Nic Matos, Manuel Góngora Paredes, Celso Reyes Martínez, Antonio Andrade Torres, Ana Colli Rodríguez, Yahima Osorno Borst, Graciano Puch Hau, Damaris Ortegón, Rafael Sánchez Borges, Gabriela Sandoval, Gustavo Rivera Solís y pablo palma** por todo su apoyo brindado.

ÍNDICE

ÍNDICE.....	i
ABREVIATURAS.....	ii
LISTADO DE FIGURAS.....	iii
LISTADO DE TABLAS.....	iv
RESUMEN.....	v
ABSTRACT.....	vi
INTRODUCCIÓN.....	1
CAPITULO I.....	3
1.1 ANTECEDENTES.....	3
1.1.1 Importancia del cocotero.....	3
1.1.2 Problemática del cocotero.....	3
1.1.3 Germoplasma resistente al amarillamiento letal.....	6
1.1.4 Interacción planta-patógeno.....	7
1.1.4.1 Sistema inmune innato.....	8
1.1.4.2 Respuesta hipersensible.....	10
1.1.4.3 Resistencia sistémica adquirida.....	11
1.1.5 Proteínas de resistencia.....	13
1.1.5.1 Evolución de genes NBS-LRR en plantas.....	15
1.1.6 Estudios sobre los mecanismos de defensa en cocotero.....	16
1.1.7 Secuencias candidatas a genes de resistencia.....	17
1.1.8 Marcadores moleculares AFLP.....	18
1.1.9 Estudios sobre marcadores moleculares AFLP basados en secuencias NBS-LRR.....	19
1.10. HIPÓTESIS.....	21
1.10.1. HIPÓTESIS 1.....	21

1.10.2. HIPÓTESIS 2.....	21
1.10.3. HIPÓTESIS 3.....	21
1.11. OBJETIVOS.....	21
1.11.1 OBJETIVO GENERAL.....	21
1.11.2 OBJETIVOS ESPECÍFICOS.....	22
1.12. ESTRATEGIA EXPERIMENTAL.....	23
1.13. REFERENCIAS BIBLIOGRAFICAS.....	24
CAPÍTULO II.....	37
MOLECULAR CLONING AND CHARACTERIZATION OF DISEASE- RESISTANCE GENE CANDIDATES OF THE NUCLEOTIDE BINDING SITE (NBS) TYPE FROM <i>Cocos nucifera</i> L.....	37
2.1. ABSTRACT.....	37
2.2. INTRODUCTION.....	37
2.3. MATERIALS AND METHODS.....	39
2.3.1. PLANT MATERIAL AND DNA EXTRACTION.....	39
2.3.2. DEGENERATE PRIMERS AND PCR AMPLIFICATION.....	39
2.3.3. CLONING AND SEQUENCING OF PCR PRODUCTS.....	40
2.3.4. DNA SEQUENCE ANALYSIS AND PHYLOGENETIC TREE CONSTRUCTION.....	40
2.3.5. EXPRESSION ANALYSIS OF RGCs.....	41
2.3.5.1 EXPRESSION ANALYSIS IN PLANTS TREATED WITH SA.	41
2.3.5.1.1. SALICYLIC ACID TREATMENT.....	41
2.3.5.2. EXPRESSION ANALYSIS OF FOUR NBS SEQUENCES OF COCONUT PALMS INFECTED WITH LY-PHYTOPLASMA.....	41
2.3.5.2.1. DIAGNOSTIC OF LY PHYTOPLASMA.....	42
2.3.6. RNA EXTRACTION AND cDNA SYNTHESIS.....	42
2.3.7. TRANSCRIPT EXPRESSION ANALYSIS BY QUANTITATIVE RT- PCR.....	42

2.4. RESULTADOS.....	44
2.4.1. AMPLIFICATION AND IDENTIFICATION OF RGCS FROM COCONUT PALM.....	44
2.4.2. PHYLOGENETIC RELATIONSHIPS OF CnRGCS WITH TIR AND NON-TIR-NBS SEQUENCES.....	50
2.4.3. EXPRESSION PROFILES OF RGCS IN PLANTLET TREATED WITH SALICYLIC ACID.....	55
2.5. DISCUSSION.....	58
2.6. CONCLUSIONS.....	62
2.7. APPENDIX A. SUPPLEMENTARY DATA.....	62
2.8. REFERENCES.....	65
CAPITULO III.....	73
NEW INSIGHTS IN THEIR EVOLUTIONARY HISTORY OF RESISTANCE GENE CANDIDATES (RGCS) IN COCONUT PALM (<i>COCOS NUCIFERA</i> L.).....	73
3.1. ABSTRACT.....	73
3.2. INTRODUCTION.....	74
3.3. MATERIALS AND METHODS.....	75
3.3.1 STUDY OF NBS SEQUENCES.....	75
3.3.1.1 IDENTIFICATION OF NBS SEQUENCES.....	75
3.3.1.2 PHYLOGENETIC RELATIONSHIPS AND CALCULATION OF KA/KS RATIOS.....	76
3.4 RESULTS.....	77
3.4.1. IDENTIFICATION OF NBS SEQUENCES FROM TRANSCRIPTOMIC SEQUENCES.....	77
3.4.2. ANALYSIS FOR THE CONSERVED MOTIF STRUCTURES.....	78
3.4.3 ORGANIZATION AND PHYLOGENY OF NBS SEQUENCES AMONG SPECIES OF ARECACEAE.....	79
3.4.4. COMPARING EVOLUTIONARY RATES AMONG NBS SEQUENCES OF COCONUT PALM.....	82

3.4.5. PHYLOGENETIC RELATIONSHIPS BETWEEN SPECIES OF ARECACEAE, POACEAE AND BRASSICACEAE.....	84
3.4.6. EXPRESSION ANALYSIS OF FOUR NBS SEQUENCES OF COCONUT PALMS INFECTED WITH LY-PHYTOPLASMA.....	83
3.5. DISCUSSION.....	87
3.6. REFERENCES.....	90
CAPITULO IV.....	95
DESARROLLO DE MARCADORES MOLECULARES EN COCOTERO.....	95
4.1. INTRODUCCIÓN.....	95
4.2. MATERIALES Y METODOS.....	98
4.2.1. MATERIAL VEGETAL Y EXTRACCIÓN DEL ADN.....	98
4.2.2. DESARROLLO DE LA TÉCNICA DE AFLP-RGC.....	99
4.2.3. PURIFICACIÓN Y CLONACIÓN DE LOS FRAGMENTOS POLIMÓRFICOS.....	100
4.2.4. DISEÑO DE CEBADORES SCAR Y ANÁLISIS DE PCR.....	101
4.2.5. VALIDACIÓN DEL MARCADOR AFLP-RGC 1-05.....	101
4.3. RESULTADOS.....	103
4.3.1. DESARROLLO DE LA TÉCNICA DE AFLP-RGC.....	103
4.3.2. VALIDACIÓN DE LOS FRAGMENTOS POLIMORFISMOS.....	106
4.3.3. VALIDACIÓN DEL MARCADOR SCAR AFLP-RGC 1-05.....	107
4.3.4. EVALUACIÓN DEL MARCADOR AFLP-RGC 1-05 EN SEGREGANTES Y EN UNA PLANTACIÓN.....	108
4.4. DISCUSIÓN.....	112
4.5. CONCLUSIONES.....	115
4.6. REFERENCIAS BIBLIOGRAFICAS.....	116
CAPITULO V.....	121
5.1. DISCUSIÓN GENERAL.....	121

5.2. CONCLUSIONES GENERALES.....	125
5.3 PERSPECTIVAS.....	127
5.4 REFERENCIAS BIBLIOGRAFICAS.....	129
5.5 ANEXOS.....	131

ABREVIATURAS

AL	Amarillamiento letal	MAMPs	Patrones moleculares asociados a microbios
AFLP	Polimorfismo de la longitud de fragmentos amplificados	MAST	Motif Alignment and Search Tool
ANK	Ankirrina	MEGA	Molecular Evolutionary Genetic Analysis software
AS	Ácido salicílico	MgCl ₂	Cloruro de magnesio
ATP	Adenosin trifosfato	MEME	Multiple Expectation Maximization for Motif Elicitation
CC	Coiled-Coil	MSAP	Methylation Sensitive Amplified Polymorphism
cDNA	Ácido ribonucleico complementario	MXAT	Alto del Atlántico Mexicano
Ct	Ciclo umbral	MXP	Alto del Pacifico Mexicano
Cyt c	Citocromo C	MYD	Enano Malayo Amarillo
DNA	Ácido desoxirribonucleico	NBS	Sitio de unión a nucleótidos
EFR	EF-TU Receptor	NJ	neighbor-joining
EST	Susceptibilidad disparada por efectores	NPR1	No expresivo de genes PR-1
ETI	Inmunidad disparada por efectores	ORF	Open Reading farmer
FLS2	Flagellin Sensitive 2	PAMPs	Patrones moleculares asociados a patógenos
genes <i>Avr</i>	Genes de avirulencia	PCD	Muerte celular programada
Genes <i>R</i>	Genes de resistencia	PCR	Reacción en cadena de la polimerasa
Ka	non-synonymous site	PRRs	Receptores de reconocimiento del patrón
Ks	substitutions per synonymous	PTI	Inmunidad disparada por PAMPs
LRR	Regiones ricas repetidas en Leucina	RAPD	ADN polimórfico amplificado al azar
RGC	Candidatos a genes de resistencia	Proteína R	Proteína de resistencia
RH	Respuesta hipersensible	PR	Proteínas relacionadas con patogénesis
RFLP	Polimorfismo de la longitud de fragmentos de restricción	SAR	Resistencia sistémica adquirida
RLKs	Cinasas tipo receptor	SD	Desviación estandar

RNA	Ácido ribonucleico	SSR	Simple Sequence Repeats or Short Tandem Repeats
rRNA	Ácido ribonucleico ribosomal	TIR	Drosophila Toll and mammalian Interleukin-1 Receptor
RT-PCR	Reacción en cadena de la polimerasa usando la transcriptasa reversa		

LISTADO DE FIGURAS

CAPÍTULO I	3
Figura 1.1 Principales limitantes biológico-ambientales que afectan al cocotero. A) Enfermedad de la decadencia foliar causada por virus. B) Enfermedad de tinangaja y cadang-cadang, causada por viroides. C) Enfermedad de la marchitez letal causada por protozoarios. D) Pudrición del cogollo causada por el oomiceto <i>Phytophthora palmivora</i> . E) Enfermedad del Anillo Rojo causado por el nematodo <i>Radinaphelenchus cocophilus</i> . F) Daños ocasionados por la plaga <i>Rhychonoporus palmarum</i> y F) enfermedad del Amarillamiento Letal ocasionado por fitoplasmas.....	4
Figura 1.2 Esquema representativo de la forma de transmisión del AL. 1) Palma enferma con amarillamiento letal, 2) adquisición del fitoplama del AL por el vector <i>Haplaxius crudus</i> al momento de alimentarse con el contenido del floema, 3) inoculación del fitoplasma a una palma sana e 4) infección sistémica de la planta (Editado: Carlos Puch).....	5
Figura 1.3 Mortalidad total acumulada por ecotipo debido a la enfermedad del Amarillamiento Letal en San Crisanto, Yucatán, México (Zizumbo-Villareal <i>et al.</i> , 2008).....	7
Figura 1.4 Diagrama conceptual del sistema inmune innato de las plantas. 1, 2 y 3, Inmunidad disparada por PAMPs (PTI), 4 y 5, susceptibilidad disparada por los efectores (EST) y 6 y 7, inmunidad disparada por los efectores (ETI) (Pieterse <i>et al.</i> , 2009) (editado: Carlos Puch).....	10
Figura 1.5. Modelo que ilustra la regulación de la actividad de la proteína NPR1 (Pieterse y Van Loon, 2004).....	12
Figura 1.6. Proteínas R, clasificadas de acuerdo al dominio común que comparten (Dangl y Jones, 2001; Chisholm <i>et al.</i> , 2006) (editado: Carlos Puch).....	14
	37

CAPÍTULO II	
Figure 2.1 Conserved motifs within the NBS domain targeted using degenerate RGC primers. The arrows show the position (P-loop, kinase-2, GLPL, RNBS-D and 13R1) and direction of the primers, with tips indicating the 3' primer end.....	44
Fig. 2.2 Phylogenetic tree based on the neighbor-joining method (Saitou and Nei, 1987), of the deduced amino sequences of 143 coconut palm RGCs. Seven different RGCs clades (I-VII) were obtained. The numbers on the branches indicate the percentage of 1000 bootstrap replications supporting the particular nodes.....	46
Fig. 2.3. Aminoacid sequence alignment between P-loop, RNBS-A, Kinase-2, RNBS-B and GLPL of non-TIR-NBS-LRR with the NBS domains of known R genes of TIR and Non-TIR-NBS-LRR family gene N from tobacco (AAA50763.1), L6 from flax (AAA91022), RPP5 (AAF08790), RPS5 (AEE28852), RPM1 (CAA61131) and RPP8 (AAC78631) from <i>Arabidopsis</i> . Conserved domains are boxed. The alignment was constructed utilizing the CLUSTALX program (Thompson et al. 1997) and one representative deduced amino acid sequence from each clade (CnRGC-I: CnRGC121, CnRGC-II: CnRGC36, CnRGC-III: CnRGC118, CnRGC-IV: CnRGC62, CnRGC-V: CnRGC34, CnRGC-VI: CnRGC21, CnRGC-VII: CnRGC17). Similar aminoacids residues are shaded in grey and identical aminoacid residues are shaded in black. The arrow indicate the last aminoacid residue in kinase-2 motif.....	49
Fig. 2.4. Phylogenetic analysis based on alignment of the deduced aminoacid sequences of coconut palm with known <i>R</i> genes. The tree was constructed by the neighbor-joining method using MEGA 5.0 software. Number on the branches indicate the percentages of 1000 bootstrap replications and only those with >50% support are shown. Dots indicate coconut palm sequences isolated in this study and the dot in grey the coconut palm sequences registered in NCBI.....	52

Fig. 2.5. Expression profiles of six CnRGCs. The gene expression was determined by real-time PCR in different tissues from plantlets of 12 months of <i>in vitro</i> culture treated and non-treated with salicylic acid at different times. 18S was used as a reference gene.....	54
Figure 2.6. Relative transcript levels of four CnRGCs obtained from quantitative real-time PCR experiments in different tissues of healthy coconut palm and affected with lethal yellowing disease. Black bars: LY- coconut palms, gray bars: healthy coconut palm and white bars: <i>in vitro</i> coconut palm plantlets. R: Root, S: Stem and L: Leaf. 18S rRNA was used as a reference gene for the expression analysis in coconut palm. The analysis was conducted in three independent palms for each condition, of each sample by triplicate.....	57
Additional file 1. Figure A1. PCR products amplified with three pairs of degenerate primers (PIB/RNBS-D, P-loop/GLPL and 3F2/13R1) using the genomic DNA mix of each ecotype. Lane 1: genomic DNA mix from resistant Mexican Pacific Tall 2 (MXPT2), lane 2: genomic DNA mix from the resistant Malayan Yellow Dwarf (MYD) and lane 3: genomic DNA mix from the susceptible Mexican Atlantic Tall (MXAT). The 1 Kb ladder was used as size markers. The arrow indicates the band cloned.....	63
CAPÍTULO III	73
Figure 3.1. Phylogenetic tree of RGAs obtained of three transcriptome of coconut palm (embryo, endosperm and leaf). . The RGAs from embryo transcriptome are show as solid black circles, the endosperm in white squares and the leaf in solid black triangles. The tree was constructed by the neighbour-joining method with a statistical analysis of bootstrap with 1,000 replicates. The values of bootstrap >50% are shown on the branches. APAF-1 was using as an outgroup.....	78

Figure 3.2. Phylogenetic relationship of RGCs from coconut palm, date palm and oil palm. The 234 NBS sequences were group in 9 main clades (I to IX), 7 of which were conformed by sequences orthologous of the 3 species and 2 clades were species-specific of coconut palm. The black rhombuses represent RGAs of coconut palm. The tree was constructed by the neighbour-joining method with a statistical analysis of bootstrap with 1,000 replicates. The values of bootstrap >50% are shown on the branches. APAF-1 was using as an outgroup.....	81
Figure 3.3 Neighbor-joining tree showing the phylogenetic relationship among RGCs from different species. The complete tree was based on 434 sequences. The clades species-specific, family- specific and complex clades were collapsed into single branches and are show in different colors. Number on the branches indicate the percentages of 1000 bootstrap replications and only those with >50% support are shown. APAF-1 was using as an outgroup.....	85
CAPÍTULO IV	98
Figura 4.1. Perfiles de la técnica de AFLP-RGA entre ecotipos de cocotero resistentes y susceptibles la enfermedad del amarillamiento letal en gel de poliacrilamida al 6% usando las combinaciones de iniciadores <i>EcoR1</i> + AA/P3A y <i>Mse1</i> + CAT/PIA. Carriles del 1 al 5 ecotipos resistentes y carriles del 6 al 10 ecotipos susceptibles. Las flechas señalan las bandas polimórficas entre ambos ecotipos de cocotero.....	103
Figura 4.2. Imagen representativa de la validación del marcador <i>AFLP-RGA 1-05</i> en palmas resistentes, susceptibles y enfermas de amarillamiento letal. A) carriles 1 al 7 palmas resistentes (MYD), carriles del 8 al 14 palmas susceptibles (WAT), B) carriles 1, 3 y 5, palmas resistentes (MYD) y carriles 2, 4 y 6 palmas enfermas de amarillamiento letal. La banda amplificada que corresponde al marcador <i>AFLP-RGA 1-05</i> es señalado con una flecha roja.....	107

Figura 4.3. Validación del marcador <i>AFLP-RGC 1-05</i> en plantas segregantes, resultado de la cruce entre ecotipos resistentes y susceptibles al amarillamiento letal. Carril 1, marcador molecular, carriles del 2 al 9, familia 4 y carriles del 10 al 14, familia 6. La flecha en color rojo indica la banda del tamaño esperado para el marcador <i>AFLP-RGC 1-05</i>	110
Figura 4.4. Plantación ubicada en la localidad de Calkiní Campeche, México, en donde la enfermedad del amarillamiento letal ha devastado cerca del 80 % de la plantación y en donde el marcador <i>AFLP-RGA 1-05</i> fue validado en 12 individuos.....	111
Figura 4.5. Validación del marcador <i>AFLP-RGC 1-05</i> en 12 palmas de la plantación de Calkiní Campeche, México. Las palmas de la 1 a la 10 no presentaban ningún síntoma del AL, el diagnóstico por PCR tiempo real fue negativo y no se observó la presencia del marcador. Palmas 11 y 12 presentaban síntomas de AL, el diagnóstico fue positivo y fueron las únicas en las que se observa la presencia del marcador <i>AFLP-RGC 1-05</i> . La flecha roja indica en tamaño del fragmento esperado.....	111

LISTADO DE TABLAS

CAPÍTULO II	37
Table 1. RGCs amplified using degenerate primers from coconut ecotypes.....	45
Table 2. Sequences that conform each clade of RGCs in coconut.....	47
Table 3. Percentage identity derived from pairwise comparisons between isolated coconut resistance gene candidates and the NBS domain of NBS-LRR resistance genes.....	51
Table 4. Detection of LY-phytoplasma DNA from leaf tissues of asymptomatic and symptomatic of coconut palms.....	55
Additional file 1. Table A1. Degenerate oligonucleotide primers used to amplify resistance gene candidates from coconut.....	62
Additional file 1. Table A2. Primer sequences used to determine the expression profiles of CnRGCs from coconut.....	64
Additional file 1. Table A3. Results of similarity searches between the amplified products by Real-Time PCR and GenBank accessions using <i>blastx</i>	56
CAPÍTULO III	73
Table 1. Number of RGCs identified in coconut palm.....	77
Table 2. NB-LRR-specific amino acid motifs identified in RGCs of coconut palm using the program MEME.....	79
Table 3. Ka/Ks ratios for pairwise comparisons among members of each clade and among RGCs of the different clades of coconut palm.....	83
Table 4. Organization of NBS sequences among species of Arecales, Poales and Brassicales.....	86
CAPÍTULO IV	93
Tabla 4.1. Descripción de genotipos de cocotero empleados para el desarrollo de marcadores moleculares <i>AFLP-RGC</i>	98
Tabla 4.2. Secuencia de los iniciadores empleados para la técnica de AFLP-RGCs.....	100

Tabla 4.3. Bandas polimórficas identificadas usando la técnica de AFLP-RGCs, indicando la combinación de iniciadores, el tamaño del marcador, el ecotipo donde fue amplificado y las accesiones en el GenBank al que presentaron identidad.....	105
Tabla 4.4. Características de los marcadores SCAR validados.....	106
Tabla 4.5. Validación del marcador AFLP-RGA 1-05 en plantas segregantes.....	109

RESUMEN

El cocotero es un cultivo importante en términos económicos y ecológicos en las regiones tropicales del mundo. Desafortunadamente, está sujeto al ataque de diversos agentes, como el fitoplasma '*Candidatus Phytoplasma palmae*', causante de la enfermedad del amarillamiento letal (ALC), por lo tanto, la caracterización de genes de resistencia (*R*) a enfermedades en este cultivo es sumamente importante para su mejoramiento genético. Esta familia de genes, codifican para proteínas con un sitio de unión a nucleótidos (NBS) y un dominio de regiones ricas repetidas en leucina (LRR). En este estudio, nosotros logramos caracterizar más de 200 secuencias con homología a genes de resistencia del tipo NBS-LRR. Los perfiles de expresión de algunas de estas secuencias, mostraron cambios en respuesta a la aplicación exógena del ácido salicílico, así como también en palmas de campo sanas y enfermas de amarillamiento letal. En los estudios de evolución pudimos identificar clados específicos para los RGCs de cocotero, lo cual sugiere que existe una evolución de genes especie-específica en esta palma. Por otra parte, las relaciones filogenéticas encontradas entre RGCs de la familia de Arecales indican una mayor conservación de RGCs en ella, sugiriendo un origen monofilético para los tres géneros. Las relaciones filogenéticas entre los RGCs de arecáceas y brasicáceas indican que estas secuencias posiblemente emergieron antes de dividirse entre monocotiledóneas y dicotiledóneas. Finalmente, empleando la técnica de AFLP-RGC pudimos identificar 17 polimorfismos entre ecotipos de cocotero resistentes y susceptibles al ALC, de los cuales el marcador AFLP-RGC 1-05 resultó ser específico y reproducible para el ecotipo susceptible al amarillamiento letal. Este estudio representa una base importante para el mejoramiento genético de la resistencia a enfermedades en cocotero y en general en especies de arecáceas.

ABSTRACT

Coconut palm is an economically and ecologically important crop in tropical regions of the world crop. Unfortunately, it is subject to attack by various agents, such as phytoplasma '*Candidatus Phytoplasma Palme*', which causes lethal yellowing disease (LY), therefore, characterization of disease resistance genes (*R*), opens the possibility for improving pathogen resistance in this crop. The large family of plant *R* genes encode proteins with a nucleotide binding site and leucine-rich repeat domain. In this study, we characterize more than 200 sequences with homology to NBS-LRR-type sequences. The expression profiles of some of these sequences, showed changes in response to exogenous application of salicylic acid, as well as in healthy coconut palm and those affected with lethal yellowing disease. In evolutionary studies, we could identify specific clades for coconut RGCs, suggesting that there is an evolution of species-specific genes in the palm. Moreover, strikingly similar RGCs between species of Arecales indicates high conservation of specific RGCs in this family, suggesting a monophyletic origin of three genera. The phylogenetic relationship between RGCs of Arecales and Brassicales suggests that these sequences possibly emerged before monocot and dicots divergence. Finally, using the AFLP-RGC technique, we could identify 17 polymorphisms between resistant and susceptible ecotypes to LY coconut, of which the marker AFLP-RGC 1-05 proved to be associated to LY susceptibility. This study provides new insights for future efforts towards improvement of disease resistance in coconut palm and other species of Arecaceae.

INTRODUCCIÓN

El cultivo de cocotero está ampliamente distribuido en el mundo y posee un alto valor económico. En México, cubre una superficie de 100 mil ha de las cuales dependen aproximadamente 50 mil familias que viven directamente del cultivo (www.conacoco.com.mx, 2008). De esta palma, se puede aprovechar muchas de sus partes para la obtención de diferentes productos, como es el aceite de cocotero, el cual está siendo utilizado para la producción de biodiesel, como una opción viable para las substitución de combustible fósil, principalmente en las islas de Papua y Nueva Guinea (Lao, 2008). Desafortunadamente, esta palma está sujeta al ataque de diversos agentes que ocasionan enfermedades y que afectan su producción, entre los que se encuentran los virus (Rohde *et al.*, 1990), viroides (Hanold y Randles, 1991), protozoos (Parthasarathy *et al.*, 1978), hongos (Joseph y Rhada, 1975), nematodos (Griffith, 1987) y mollicuotas, como el fitoplasma causante de la enfermedad del amarillamiento letal (ALC). En México, el ALC ha eliminado virtualmente al cocotero Alto del Atlántico en toda la costa de Quintana Roo, Yucatán y Campeche (Domínguez *et al.*, 1999) y se estima que ha devastado aproximadamente unas 13 mil ha de la plantación quedando actualmente una superficie cultivada cercana a 100 mil ha (www.conacoco.com.mx, 2008). A pesar de la gran importancia de este cultivo y a la diversidad de patógenos que afectan su producción, los mecanismos de defensa ante patógenos han sido poco estudiados en comparación con otras plantas las cuales son capaces de reaccionar y defenderse por sí solas. Dentro de estos mecanismos de defensa, las proteínas de resistencia (R) codificadas por genes de resistencia (*R*), proveen una efectiva inmunidad a la planta hospedera, por medio del reconocimiento de los patógenos y la activación de las vías de transducción que dirigen a la resistencia ante la enfermedad (Belkhadir *et al.*, 2004). Por lo tanto, la introducción de genes de resistencia dentro de cultivares susceptibles confieren eficiente inmunidad ante patógenos que expresen el mismo efector. La mayoría de estos genes, codifican proteínas con un dominio de unión a nucleótidos más un dominio de regiones ricas en leucina (NBS-LRR). La técnica de PCR utilizando iniciadores degenerados diseñados en base a los motivos del dominio NBS altamente conservados, ha sido una de las principales estrategias utilizadas para el aislamiento de secuencias candidatas a genes de resistencia (RGCs) en diferentes cultivos (Kanazin *et al.*, 1996; Yu *et al.*, 1996; Graham *et al.*, 2000; Leister *et al.*,

1996; Shen *et al.*, 1998; Leister *et al.*, 1998; Seah *et al.*, 2000; Ayele-Gedil *et al.*, 2001; Rivkin *et al.*, 1999; López *et al.*, 2003; Martínez-Zamora *et al.*, 2004; Calenge *et al.*, 2005; Pei *et al.*, 2007; Peraza-Echeverria *et al.*, 2008; Miller *et al.*, 2008; Wan *et al.*, 2012; Saha *et al.*, 2013). Los RGCs, son secuencias de función desconocida pero que codifican para dominios conservados característicos de las proteínas R. Muchas de estas secuencias han sido mapeadas y se ha demostrado que están ligadas a los genes *R* conocidos, proporcionando evidencia que ciertas regiones genómicas probablemente codifican para resistencia ante patógenos, representando de esta manera una fuente valiosa para el descubrimiento de nuevos genes de resistencia (Pflieger *et al.*, 2001). En cocotero, en un estudio previo se lograron caracterizar 80 RGCs que presentaron homología a genes de resistencia descritas en otras especies de plantas (Puch-Hau, 2012). Esta búsqueda de RGCs se realizó en ecotipos que han mostrado resistencia ante la enfermedad del amarillamiento letal y representan solamente un pequeño número de las 653 reportados en *Oryza sativa* L. spp. *indica* (Shang *et al.*, 2009) y aproximadamente el 50% de las 144 reportadas en la palma *Phoenix dactylifera* L (Al-Mssallem *et al.*, 2013). Por lo tanto, este proyecto contempla: (1) ampliar la caracterización de RGCs en ecotipos de cocotero tanto resistentes como susceptibles a la enfermedad del Amarillamiento letal, (2) estudios de expresión en respuesta al ácido salicílico y en palmas enfermas de ALC, (3) análisis de las relaciones filogenéticas de los RGCs entre especies de Arecales, Poales y Brassicales y (3) la búsqueda de marcadores moleculares para la identificación de germoplasma resistente o susceptible. Los resultados obtenidos en este estudio serán fundamentales para la generación de recursos genómicos útiles en los programas de mejoramiento genético de la inmunidad del cocotero, ante el ALC.

CAPITULO I

1.1 ANTECEDENTES

1.1.1 IMPORTANCIA DEL COCOTERO

El cultivo de cocotero está ampliamente distribuido en el mundo y posee un alto valor económico. La producción mundial del cocotero se estima en 40 billones de nueces, principalmente para la producción de aceite (Harries, 1995).

En América y en el Caribe la producción del cocotero contribuye con cerca del 10% a la producción mundial, siendo México y Brasil los países productores más importantes. En México, el cocotero es uno de los cultivos de mayor importancia económica, cubre una superficie de 100 mil ha de las cuales dependen aproximadamente 50 mil familias que viven directamente del cultivo (www.conacoco.com.mx, 2008). De la palma de cocotero, se puede aprovechar muchas de sus partes para la obtención de diferentes productos, como es el aceite de cocotero, el cual se ha utilizado en múltiples aplicaciones en la industria alimentaria y en la actualidad. Por otra parte, el aceite de cocotero también está siendo utilizado para la producción de biodiesel, como una opción viable para la sustitución de combustible fósil, principalmente en las islas de Papua, Nueva Guinea (Bradley, 2006), y en Filipinas está siendo producido comercialmente donde es usado como aditivo de diesel o gasolina (Lao, 2008).

1.1.2 PROBLEMÁTICA DEL COCOTERO

A nivel mundial, México ocupa el séptimo lugar en la producción de copra y el primero en América, detrás de los grandes productores asiáticos como Indonesia, Filipinas y la India (FAO, 1998). Sin embargo, el cultivo está sujeto a limitantes biológicos ambientales, como el virus; causante de la decadencia foliar del cocotero (Rohde *et al.* 1990), los viroides; causantes de la enfermedad de tinangaja en la isla de Guam y el causante de la enfermedad de cadang-cadang en Filipinas (Hanold y Randles, 1991), los protozoos; causantes de la marchitez letal de las palmas (Parthasarathy *et al.* 1978), oomicetos; como *Phytophthora*

palmivora causante de la pudrición del cogollo ((Joseph y Radha, 1975), el nematodo; *Radinaphelenchus cocophilus* causante de la enfermedad del Anillo Rojo (Griffith, 1987), el insecto conocido como picudo negro *Rhychonoporus palmarum* (Bastidas y Zavala, 1995) y mollicuotas como los fitoplasmas entre los que se encuentra el causante de la enfermedad del amarillamiento letal (ALC), Figura 1.1 (Howard y Barrant, 1989).

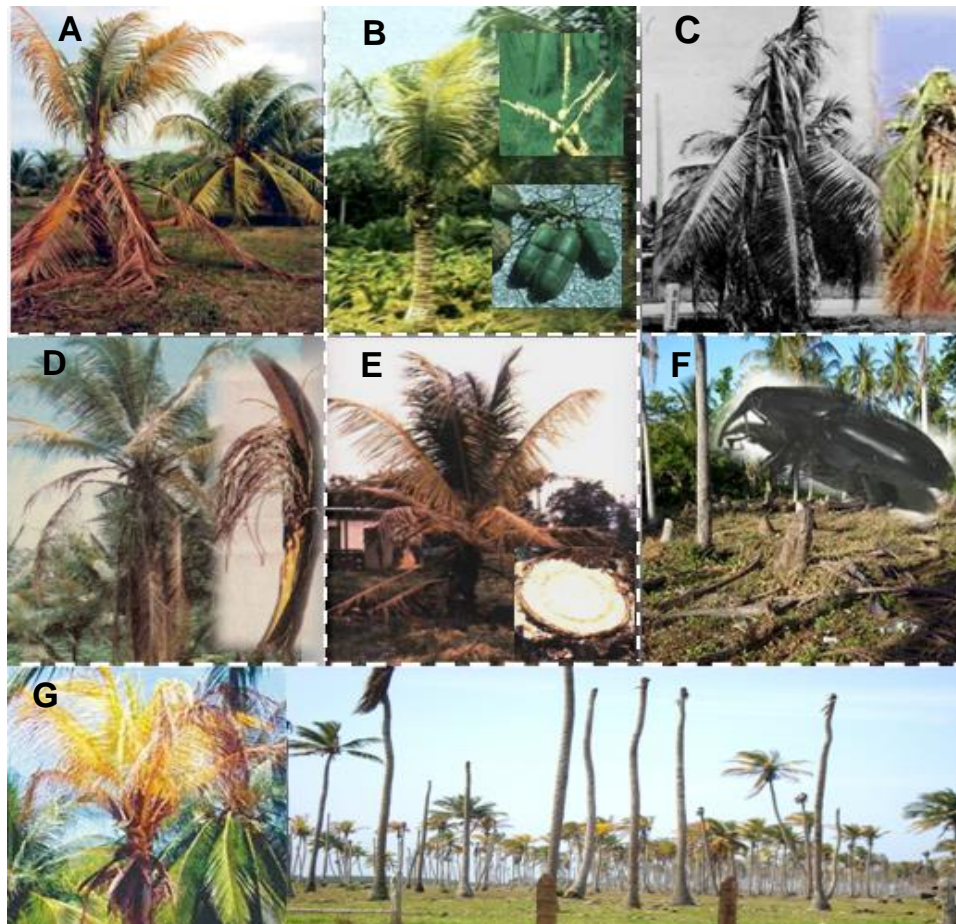


Figura 1.1 Principales limitantes biológico-ambientales que afectan al cocotero. A) Enfermedad de la decadencia foliar causada por virus. B) Enfermedad de tinangaja y cadang-cadang, causada por viroides. C) Enfermedad de la marchitez letal causada por protozoarios. D) Pudrición del cogollo causada por el oomiceto *Phytophthora palmivora*. E) Enfermedad del Anillo Rojo causado por el nematodo *Radinaphelenchus cocophilus*. F) Daños ocasionados por la plaga *Rhychonoporus palmarum* y F) enfermedad del Amarillamiento Letal ocasionado por fitoplasmas.

Entre las enfermedades que afectan al cocotero, el amarillamiento letal es una de las más importantes. En los últimos años, la enfermedad del AL ha destruido millones de palmas en Jamaica, Florida, Belice y Honduras (Harrison *et al.*, 1999). En cuanto a México, ha eliminado virtualmente al cocotero Alto del Atlántico en toda la costa de Quintana Roo, Yucatán y Campeche (Domínguez *et al.*, 1999) y se estima que ha devastado aproximadamente unas 13 mil ha de la plantación (Doyle, 1998) quedando actualmente una superficie cultivada cercana a 100 mil ha (www.conacoco.com.mx, 2008). La principal forma de transmisión de los fitoplasmas entre las plantas, es a través de insectos que se alimentan del contenido del floema, Figura 1.2 (Harrison y Jones, 2004; Broschat *et al.*, 2002). Para la enfermedad del amarillamiento letal, existen evidencias que demuestran que puede ser transmitido por el insecto *Haplaxius crudus* Van Duzee (Homoptera: Cixidae) (Howard *et al.*, 1984).

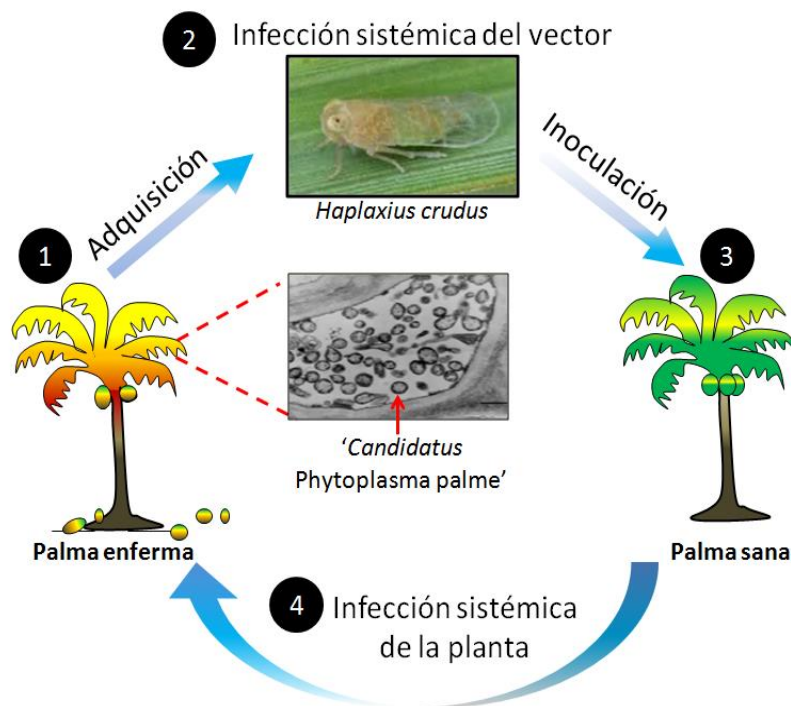


Figura 1.2 Esquema representativo de la forma de transmisión del ALC. 1) Palma enferma con amarillamiento letal, 2) adquisición del fitoplasma del ALC por el vector *Haplaxius crudus* al momento de alimentarse con el contenido del floema, 3) inoculación del fitoplasma a una palma sana e 4) infección sistémica de la planta (Editado: Carlos Puch).

1.1.3 GERMOPLASMA RESISTENTE AL AMARILLAMIENTO LETAL

En México, la mortalidad debido al Amarillamiento Letal se registró en cinco ecotipos de cocotero (Enano Malayo Amarillo, Alto del Atlántico, Alto del Pacífico Mexicano 1, Alto del Pacífico Mexicano 2 y Alto del Pacífico Mexicano 3) representantes de la diversidad de germoplasma cultivada en el país. En base a esto, se realizó un estudio para identificar germoplasma resistente al ALC como una alternativa para el control de dicha enfermedad, lográndose identificar ecotipos tales como el Enano Malayo Amarillo (MYD, por sus siglas en inglés) y tres ecotipos Altos del Pacífico Mexicano (MXP1, MXP2 y MXP3, por sus siglas en inglés) con mediana y alta resistencia. Para el caso del ecotipo Alto del Atlántico (MXAT) las palmas fueron altamente susceptibles (Figura 1.3) (Zizumbo-Villareal *et al.* 2008). Actualmente la replantación con germoplasma resistente al Amarillamiento Letal se convirtió en la única forma más eficaz para enfrentar esta enfermedad. Sin embargo, la pregunta a ¿cuál o cuáles pudieran ser los mecanismos por los cuales diferentes ecotipos responden de manera resistente a la enfermedad del ALC? queda sin responder, por lo tanto, es importante conocer las bases moleculares que pudieran estar implicadas en la resistencia a esta enfermedad, los cuales han sido poco estudiadas.

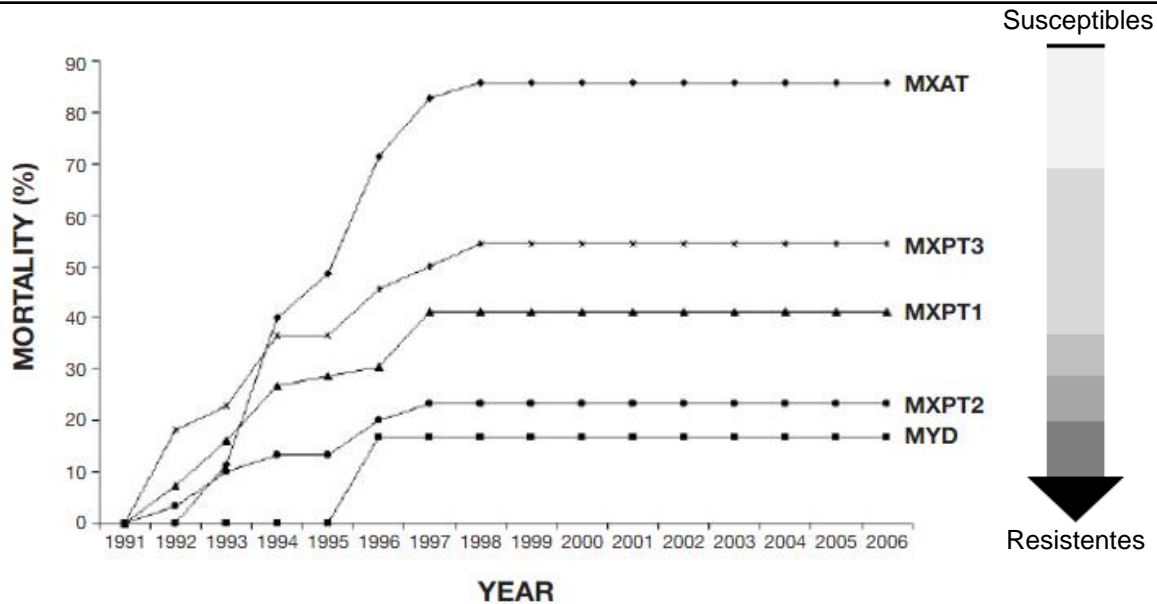


Figura 1.3 Mortalidad total acumulada por ecotipo debido a la enfermedad del Amarillamiento Letal en San Crisanto, Yucatán, México (Zizumbo-Villareal *et al.*, 2008).

1.1.4 INTERACCIÓN PLANTA-PATÓGENO

Las plantas en su ambiente natural se enfrentan al ataque de diferentes organismos, incluyendo bacterias, hongos, oomicetos, virus, nematodos e insectos. Al ser organismos sésiles, dependen de la rápida detección de patógenos y de la adecuada activación de las vías de transducción de señales que le permitan responder y defenderse frente a estos ataques.

Los patógenos de plantas, de acuerdo a su ciclo de vida son generalmente divididos en biotróficos, necrotrofoicos y hemibiotrofoicos (Glazebrook, 2005). Los primeros son patógenos que penetran y establecen contactos cercanos con las células del hospedante para crecer y reproducirse durante su ciclo de vida, son considerados microorganismos obligados, poseen haustorios, causan un daño mínimo a la planta, y en general los síntomas se presentan como resultado de una disminución de nutrientes en el tejido vegetal, pueden secretar enzimas líticas, tienen rango angosto de hospedantes y se considera que son controlados por las vías de defensa dependientes de ácido salicílico (SA) (Oliver y Ipcho, 2004). Los necrotrofoicos, en contraste, son microorganismos no obligados que dependen

del tejido muerto del hospedante para adquirir nutrientes y reproducirse. Éstos a menudo secretan enzimas y toxinas que degradan y matan a las células del hospedante para tener acceso a los nutrimentos, tienen un rango amplio de hospedantes y se considera que son controlados por las vías de defensa dependientes de etileno (ET) y ácido jasmónico (JA) (Oliver y Ipcho, 2004).

Muchos de los patógenos de plantas presentan ambos tipos de ciclo de vida, los cuales son llamados hemibiotróficos. Los hemibiotróficos exhiben una fase biotrófica asintomática, estableciéndose en las células del hospedante y eludiendo la detección por parte del mismo, posteriormente durante el proceso de infección éstos patógenos se dispersan rápidamente, se presenta el desarrollo de los síntomas de la enfermedad y el patógeno mata activamente las células del hospedante (Lee y Rose, 2010; Münch *et al.*, 2008; Qutob *et al.* 2002).

Durante los millones de años de coevolución entre las plantas y sus patógenos, ha habido una serie de adaptaciones que le han permitido lograr adaptarse y sobrevivir. Para lograrlo, las plantas, han desarrollado un sistema inmune innato que es capaz de reconocer a los patógenos y activar las respuestas de defensa apropiadas. En cocotero, se desconoce el papel del sistema inmune innato en respuesta al fitoplasma causante del amarillamiento letal.

1.1.4.1. SISTEMA INMUNE INNATO

Durante la interacción entre una planta y un patógeno, es necesario que la planta cuente con un mecanismo que le permita percibir la presencia del patógeno y poder responder a dicho ataque. El reconocimiento es llevado a cabo, a través de receptores que pueden estar localizados en las superficies celulares o en el citosol y que pueden activar la respuesta inmune.

El primer paso del sistema inmune innato, está basado en la percepción de patrones moleculares asociados a patógenos o microbios (PAMPs o MAMPs, por sus siglas en inglés), a través de receptores de reconocimiento del patrón (PRRs, por sus siglas en inglés) conocidos también como proteínas de resistencia (R), localizadas en la superficie celular. Los PRRs reconocen a los denominados elicitores generales, cuyas estructuras

evolutivamente conservadas se encuentran en una amplia gama de microorganismos, los cuales incluyen a los lipopolisacáridos de bacterias gram-negativas, los péptidos pequeños derivados de la flagelina bacteriana o el factor de elongación Tu, y la quitina y el ergosterol de los hongos (Zipfel y Felix, 2005). Dentro de estos receptores se encuentran el FLS2 (Flagelin Sensitive2, por sus siglas en inglés) y el EFR (EF-TU Receptor), los cuales son cinasas tipo receptor localizados en la membrana plasmática. La activación de los PRRs por parte de los PAMPs, resulta en una cascada de señalización corriente abajo que induce la activación transcripcional de genes relacionados con defensa y la producción de fitoalexinas (Nürnberger *et al.*, 1994). Esta primera respuesta de defensa es conocida comúnmente como “inmunidad disparada por PAMPs o MAMPs” (PTI, por sus siglas en inglés) (Pieterse *et al.* 2009).

A pesar de este tipo de mecanismo de defensa de las plantas, los patógenos exitosamente producen efectores que pueden inhibir PTI (Figura 1.4b), permitiendo de esta manera que la planta entre en un estado de susceptibilidad disparada por tales efectores (EST, por sus siglas en inglés). Esto a la vez, permite que las plantas activen otra clase de receptores ubicados intracelularmente conocidos como proteínas de resistencia (R). Estas proteínas, codificadas por genes *R*, permiten el reconocimiento del patógeno mediante la unión directa o indirecta (hipótesis del guardián) de los productos moleculares específicos de un patógeno particular. Este reconocimiento, conduce a la resistencia mediante la inducción de una serie de respuestas como la Respuesta de Hipersensibilidad (RH) y la expresión de genes relacionados con patogénesis (PRs). Esta segunda barrera de defensa es llamada “inmunidad disparada por efectores” o ETI (Dangl y Jones, 2001; van der Biezen y Jones, 1998) (comúnmente conocida como gen a gen) (Flor, 1971) (Figura 1.4).

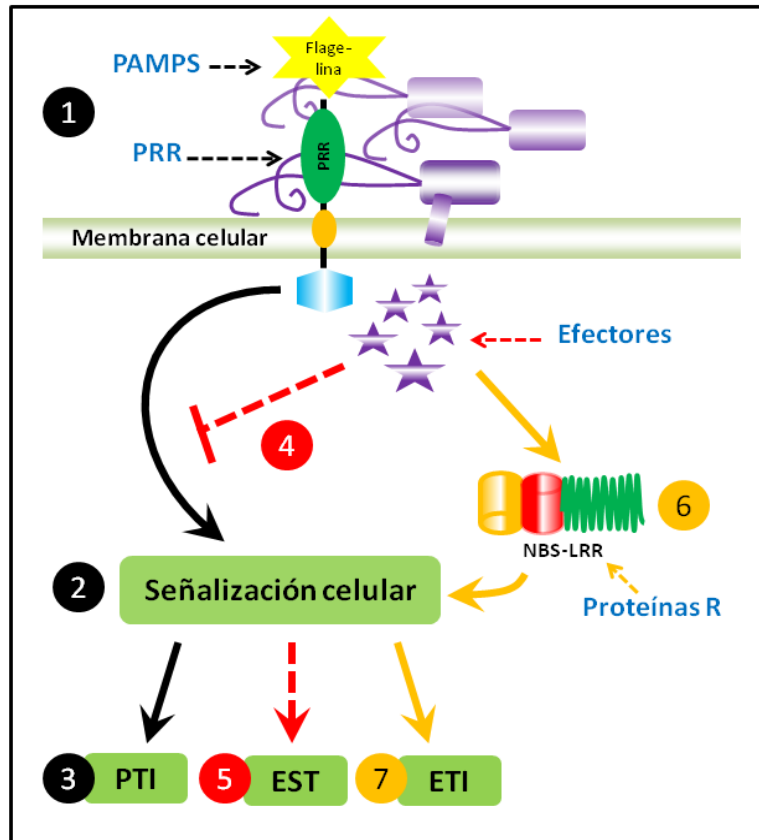


Figura 1.4 Diagrama conceptual del sistema inmune innato de las plantas. 1, 2 y 3, Inmunidad disparada por PAMPs (PTI), 4 y 5, susceptibilidad disparada por los efectores (EST) y 6 y 7, inmunidad disparada por los efectores (ETI) (Pieterse *et al.*, 2009) (editado: Carlos Puch).

1.1.4.2. RESPUESTA HIPERSENSIBLE

La respuesta de hipersensibilidad es el mecanismo de resistencia más eficiente e inmediata que presenta una planta frente al ataque por patógenos. La respuesta HR, es caracterizada por la presencia de manchas cafés formadas por las células muertas en el sitio de infección. Está involucrada en la resistencia ante patógenos por la creación de una barrera física que puede impedir la proliferación y propagación de algunos patógenos, limitando el suministro de nutrientes debido a que el tejido muerto rápidamente llega a deshidratarse (Goodman y Novacky, 1994; Alfano y Collmer, 1996; Heat, 2000). Durante

esta respuesta, se disparan señales locales y sistémicas para la activación de las defensas en las células no infectadas.

La RH está sujeta a un control genético en el que participan diversos factores que actúan como reguladores positivos o negativos. La morfología de las células que presentan RH se ve modificada, las membranas celulares sufren daño y se vuelven disfuncionales (Bennett *et al.* 1996). Así mismo, ocurren cambios intracelulares como la disminución del citoplasma, condensación de la cromatina, vacuolización, disrupción del cloroplasto, entre otros. Para el caso de la mitocondria, sufre una transición en su permeabilidad, lo que influye en la liberación del citocromo C (Cyt c) al citoplasma. El Cyt c induce la activación de complejos proteicos en los que participan las metacaspasas, un grupo de proteasas necesarias para la degradación de proteínas durante la muerte celular programada (PCD, por sus siglas en inglés) (Yao *et al.* 2004).

1.1.4.3 RESISTENCIA SISTÉMICA ADQUIRIDA

La resistencia sistémica adquirida (SAR, por sus siglas en inglés) es un mecanismo de defensa inducido, el cual le confiere a la planta protección de larga duración contra un amplio espectro de microorganismos (Tiryaki y Tunaz, 2004). La SAR se caracteriza por la acumulación tanto de AS (ácido salicílico) como de PRs. La acumulación de SA se presenta localmente, y en menor grado a nivel sistémico. Recientemente se ha propuesto que el éster metilado del SA, el salicilato de metilo, desempeña funciones importantes en el establecimiento de la SAR (Park *et al.*, 2008). La SAR induce un subjuogo de genes relacionados con patogénesis que codifican para proteínas tales como β -glucanasas y quitinasas (Van Loon y Van Kammen, 1970; Ward *et al.*, 1991). En esta respuesta, la proteína denominada NPR1 (No expresivo de genes PR-1, por sus siglas en inglés) emerge como un importante traductor de la señal del ácido salicílico. Antes de ocurrir la inducción de la respuesta sistémica adquirida, la proteína NPR1 se encuentra formando un complejo oligomérico en el citoplasma a través de puentes de disulfuro intermoleculares. Después de la inducción de SAR, la acumulación de AS dispara un cambio de redox en el interior de la célula que lleva a la monomerización de NPR1 (Cao *et al.*, 1997; Aravind y Koonin, 1999). En forma monomérica, NPR1 es traslocada al

núcleo, donde interactúa con factores de transcripción tipo TGA, actuando de esta manera, como un coactivador transcripcional de la expresión de genes de defensa involucrados en la activación de la SAR (Dong, 2004) (Figura 1.5).

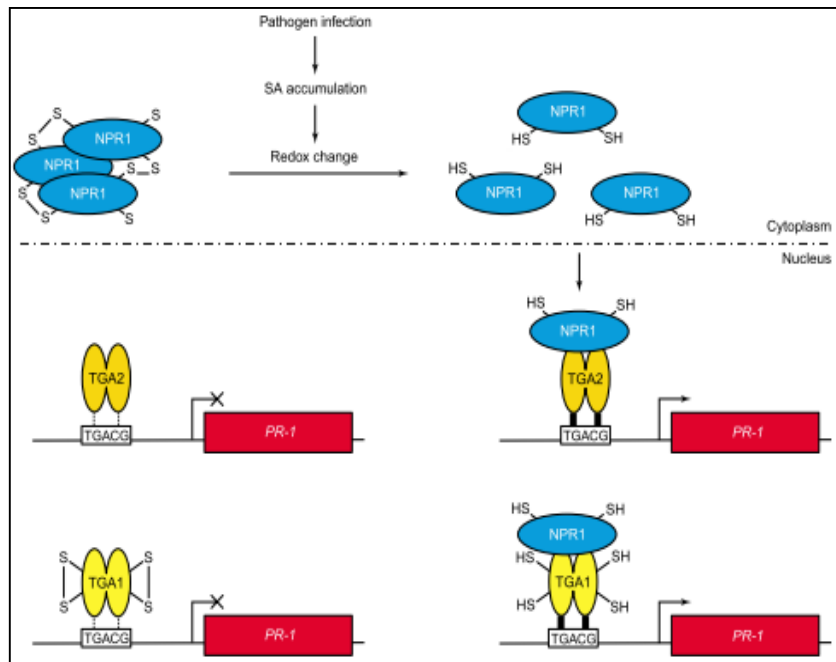


Figura 1.5. Modelo que ilustra la regulación de la actividad de la proteína NPR1 (Pieterse y Van Loon, 2004).

1.1.5. PROTEÍNAS DE RESISTENCIA

Como se ha mencionado anteriormente como requisito para que una planta sea capaz de reconocer de manera eficiente a un patógeno, y por lo tanto de activar las respuestas de defensa apropiadas, se necesita de ciertos receptores denominados proteínas de resistencia (R) (Figura 1.6) (Staskewicz *et al.*, 1995).

Las proteínas R, productos de genes de resistencia (*R*) han sido clasificadas en 6 clases de acuerdo al dominio común que comparten (Ellis y Jones, 1998).

1) Genes *R* que codifican para proteínas con dominios repetidos ricos en leucina (LRR) anclados con un dominio transmembranal (Ellis *et al.*, 2000; Meyers *et al.*, 1999; van Der Biezen *et al.*, 2002).

2) Genes *R* que codifican para proteínas con un dominio con actividad de proteína cinasa de serina/treonina (Ellis *et al.*, 2000).

3) Genes *R* que codifican proteínas con dominio con actividad de proteína cinasa de serina/treonina y un dominio extracelular LRR (Ellis *et al.*, 2000).

4) Genes *R*, que codifican proteínas con un dominio transmembranal anclado a un dominio de hélice enrollada (coiled coil, CC) (Meyers *et al.*, 1999; van Der Biezen *et al.*, 2002).

5) Genes *R* que codifican para proteínas con un dominio de unión a nucleótidos (NBS) y un dominio de regiones ricas en leucina (LRR) (Martin *et al.*, 2003). Dentro de la clase NBS-LRR, la proteína RRS1-R es un nuevo miembro, el cual contiene un sitio de localización nuclear (NLS) y un dominio con homología a los factores de transcripción WRKY (Deslandes *et al.*, 2003).

6) Gen *R* que codifica para una proteína denominada Xa27 en arroz y que no muestra homología con otras proteínas R (Gu *et al.*, 2005).

Las proteínas de resistencia del tipo NBS-LRR son las que en su mayoría han sido secuenciadas y es una de las clases más grandes dentro de los genomas de las plantas. La región NBS es importante para la unión a ATP y la funcionalidad general de los productos de los genes *R* (Saraste *et al.* 1990), mientras que el dominio LRR puede ser el principal

determinante en el reconocimiento de productos de los genes *Avr* (Kobe y Deisenhofer, 1995; Jones y Jones, 1997). Esta clase de proteínas, a su vez, pueden dividirse en dos familias de acuerdo a la característica estructural de la región N-terminal, las cuáles son: TIR-NBS-LRR y los CC-NBS-LRR (no-TIR). El nombre del dominio TIR, es debido a que presentan homología con el dominio de señalización intracelular denominado Toll de *Drosophila* y el receptor de interleucina-1 en mamíferos (Meyers *et al.*, 1999; Young, 2000).

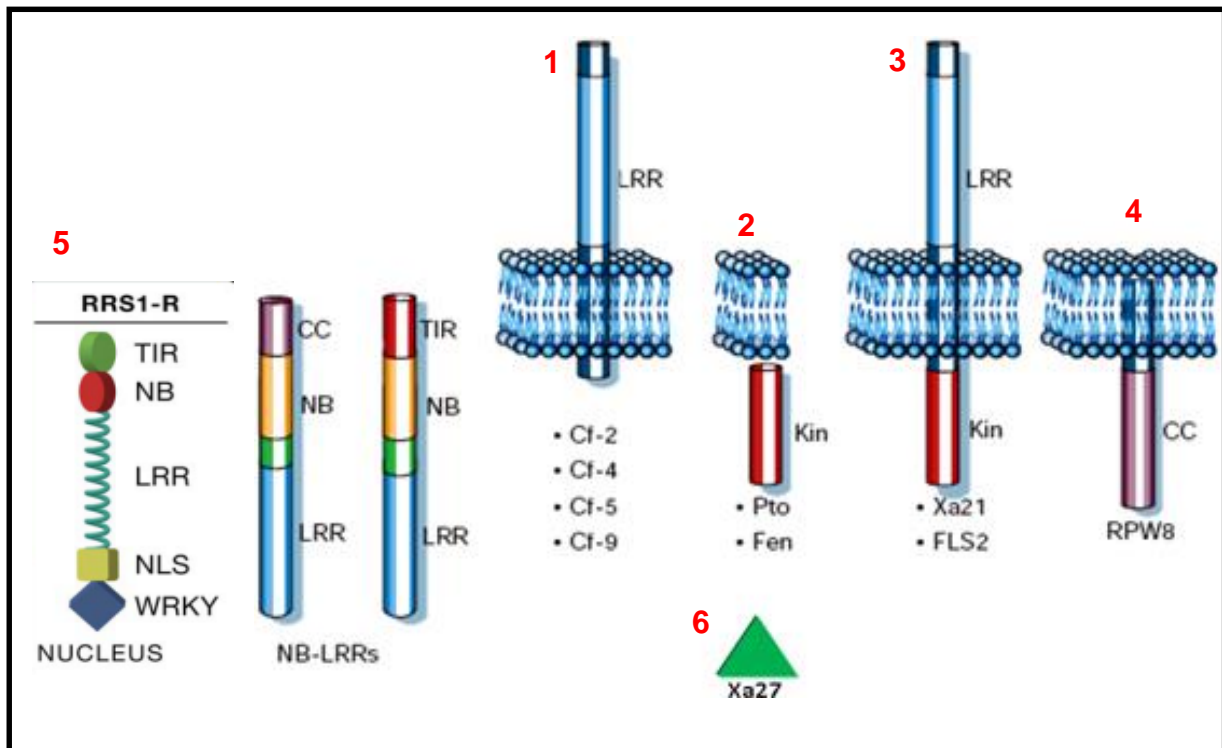


Figura 1.6. Proteínas R, clasificadas de acuerdo al dominio común que comparten (Dangl y Jones, 2001; Chisholm *et al.*, 2006) (editado: Carlos Puch).

1.1.5.1 EVOLUCIÓN DE GENES NBS-LRR EN PLANTAS

El estudio de genes NBS-LRR en los genomas de plantas secuenciadas han contribuido al entendimiento de los principales eventos responsables de su evolución. Se ha determinado que los miembros de esta familia de genes pueden encontrarse como genes aislados o como grupos de genes ligados, lo cual facilita su rápida evolución (Hulbert *et al.* 2001).

A partir de los análisis bioinformáticos de una gran diversidad de genes NBS-LRR se han propuesto diferentes hipótesis sobre la historia evolutiva de estos genes en plantas (Cannon *et al.* 2002), en donde, tanto los procesos de duplicación, como los eventos de pérdida de genes son factores que han cambiado la composición de esta familia de genes dentro de los genomas de las plantas (Plocik *et al.* 2004).

Los genes de resistencia, en comparación con otros genes, siempre muestran un fuerte patrón de selección, evolucionando de manera rápida a través de ciclos repetidos de duplicación, divergencia y pérdida eventual por medio de la formación de pseudogenes o delección, sugiriendo que estos genes se encuentran sometidos a una evolución adaptativa en respuesta a diversos patógenos (Zhang *et al.* 2009; Bakker *et al.* 2006).

Las formas de evolución de estos genes, no siempre muestran los mismos rasgos de selección. Muchos de ellos aparecen por tener un modo innovador de rápida evolución, caracterizado por los rápidos cambios en la secuencia y una fuerte diversificación de selección en regiones que codifican para el dominio LRR, lo que sugiere que los patógenos están impulsando a la planta para evolucionar nuevos genes y de esta manera evitar la infección. Otros genes *R* se encuentran en un modo conservativo, con cambios poco frecuentes en la secuencia y una selección de purificación, sugiriendo que estos genes no pueden ser fácilmente superados por la evolución de los genes de virulencia, por lo tanto, representan una fuente importante

para la generación de la resistencia duradera en plantas susceptibles (Chen *et al.* 2010; McDowell y Simon, 2008).

1.1.6. ESTUDIOS SOBRE LOS MECANISMOS DE DEFENSA EN COCOTERO

En cocotero, los mecanismos de defensa ante patógenos, es un área poco estudiada. Sin embargo, uno de los primeros estudios realizados por Lizama *et al.* (2007), nos da un acercamiento sobre algunos genes que podrían ser expresados en respuesta al ataque de un patógeno. El estudio se basó en un sistema *in vitro*, en donde callos de cocotero fueron tratados con quitosano para simular las interacciones moleculares que ocurren con los patógenos. En este estudio, se pudo observar que la aplicación de quitosano modificó diferencialmente la expresión de algunos genes relacionados con defensa y que presentaron identidad con receptores extracelulares de señalización conocidos como cinasas tipo receptor (RLKs, por sus siglas en inglés), conformados por un dominio extracelular de regiones ricas en Leucina (LRRs) y un dominio intracelular con actividad de cinasa serina/treonina.

La adición de quitosano en callos de cocotero también ocasionó: (1) la expresión de un gen que presentó identidad con el gen de resistencia a *Verticillium dahliae* en tomate (Kawchuk *et al.* 2001); (2) la acumulación de peróxido de hidrogeno (una especie reactiva de oxígeno involucrado en la respuesta hipersensible) y (3) la estimulación en la actividad de una β -1,3-glucanasa, proteína relacionada con patogénesis. Por otra parte, la adición de ácido salicílico modificó los niveles de expresión de estos genes, mientras que el jasmonato de metilo no influyó en la expresión, sugiriendo que ellos pueden estar jugando un papel importante en las respuestas de defensa dependientes de la vía de ácido salicílico.

Por otra parte, Narvaez, (2010) reportó en cocotero, la caracterización de una secuencia con identidad al gen *NPR1* de *Musa* cultivar Dongguan Dajiao. La secuencia deducida de aminoácidos reveló la presencia de un dominio conservado característico de la proteína NPR1 denominado ankirina (ANK) importante para la interacción con otras proteínas. Posteriormente Nic-Matos, (2012) logra aislar y caracterizar el extremo 3' de la proteína tipo NPR1, partiendo de la secuencia obtenida por Narváez, (2010).

Todos estos estudios contribuyen de manera importante en el entendimiento de los mecanismos de defensa en cocotero y su posible papel en respuesta al ataque de patógenos.

1.1.7. SECUENCIAS CANDIDATAS A GENES DE RESISTENCIA

Las búsquedas moleculares en las proteínas R y las señales de transducción rio abajo pueden proporcionar información relevante para el control de las enfermedades (John y Bonnie, 2003).

Empleando iniciadores degenerados diseñados para amplificar el dominio NBS de las proteínas de resistencia tipo NBS-LRR (Saraste *et al.*, 1990; Traut, 1994; Meyers *et al.*, 1999), se han logrado aislar y caracterizar secuencias candidatas a genes de resistencia (RGCs, por sus siglas en inglés) (Martin *et al.*, 2003). Los RGCs, son secuencias con función desconocida pero que codifican para dominios conservados característicos de los genes *R*. Muchas de estas secuencias han sido mapeadas y se ha demostrado que están relacionadas filogenéticamente a los genes *R* conocidos, proporcionando evidencia que ciertas regiones genómicas probablemente codifican para resistencia ante patógenos.

En un estudio previo (Puch-Hau, 2012), logramos de manera exitosa el aislamiento y la caracterización de 80 secuencias candidatas a genes de resistencia (RGCs) a partir de ADN genómico proveniente de ecotipos de cocotero resistentes a la enfermedad del amarillamiento letal (Alto del Pacífico y Enano Malayo Amarillo). A pesar del gran número de secuencias que se caracterizaron en cocotero, estas aparentemente representen un pequeño número de las 653 reportadas en *Oryza sativa* L. spp. *indica* (Shang *et al.*, 2009) y aproximadamente el 50% de las 144 reportadas en la palma *Phoenix dactylifera* L (Al-Mssallem *et al.*, 2013). Por lo tanto, una segunda búsqueda de RGCs en cocotero, probablemente nos permita identificar secuencias no reportadas en el estudio anterior.

Aunque en los últimos años un número creciente de genes *R* han sido clonados, poco se sabe sobre la regulación de su expresión en plantas. Reportes en la literatura, revelan que los genes de resistencia NBS-LRR no son inducibles, pero son expresados de manera constitutiva, así mismo, los análisis de Northern o RT-PCR en diferentes genes *R*, han

revelado la presencia de bajos niveles de transcritos en plantas que no han sido infectados con patógenos (Grant *et al.*, 1995; Salmerón *et al.*, 1996; Parker *et al.*, 1997; Milligan *et al.*, 1998; Mes *et al.*, 2000).

Por otra parte, se ha observado que el ácido salicílico, una fitohormona clave en el establecimiento de la resistencia sistémica adquirida en las plantas, modifica la expresión de estas secuencias (Wan *et al.*, 2012; You-Xiong *et al.*, 2009). Con base a esto, es posible que la aplicación exógena del ácido salicílico en plántulas de cocotero, modifique la expresión de las secuencias candidatas a genes de resistencia. Esto nos daría más evidencias sobre su funcionalidad y su posible papel en los mecanismos de defensa activados por esta molécula en plantas de cocotero.

1.1.8 MARCADORES MOLECULARES AFLP

En recientes años, los marcadores moleculares y especialmente los basados en el ADN, han sido extensivamente usados en diversos estudios, tales como el mapeo y etiquetamiento de genes (Kliebenstein *et al.*, 2001; Karp y Edwards, 1997), el tipo de sexo (Flachowsky *et al.*, 2001; Martinez *et al.*, 1999), análisis de diversidad genética (Erschadi *et al.*, 2000; Palacios *et al.*, 1999) o relaciones genéticas (Mace *et al.*, 1999). El uso de estos marcadores, permite una comparación eficiente debido a que son detectados en todos los estados de desarrollo en los organismos y no son afectadas por condiciones ambientales.

Entre los marcadores moleculares más frecuentemente utilizados, se encuentran los AFLP (polimorfismos de la longitud de los fragmentos amplificados). Los AFLP, son buenos marcadores debido a su capacidad de detectar mayor cantidad de polimorfismos y a su reproducibilidad y robustez, en comparación con otras técnicas tales como los RFLP (Polimorfismos de la longitud de los fragmentos de restricción) y los RAPD (ADN polimórfico amplificado al azar). Estas ventajas son debidas, a que no dependen de la hibridación o la digestión parcial del ADN, las cuales son fuentes débiles de reproducibilidad tal y como ocurre con los RFLP (polimorfismo de la longitud de fragmentos de restricción). El empleo de estrictas temperaturas de PCR en los AFLP hacen a este método más reproducible y robusto que inclusive los RAPD (ADN polimórfico amplificado al azar) (Bleas *et al.*, 1998).

1.1.9 ESTUDIOS SOBRE MARCADORES MOLECULARES AFLP BASADOS EN SECUENCIAS NBS-LRR

El uso de secuencias candidatas o análogas a genes de resistencia del tipo NBS-LRR como marcadores fue propuesto por Chen *et al.* (1998) y fue utilizado con éxito por Zhang *et al.* (2004) para el mapeo de genes de resistencia a enfermedades. Este método es basado en el uso del dominio conservado NBS así como también de la región LRR. En base a esto, Zhang *et al.*, (2007), desarrollaron un sistema de marcadores denominados AFLP-RGA, que se basa en el uso de iniciadores degenerados diseñados a partir de diferentes regiones de los genes de resistencia, incluyendo los dominios NBS y LRR, en combinación con un iniciador selectivo AFLP en una reacción de PCR.

En cocotero se han utilizado diversos marcadores moleculares para estudiar el polimorfismo genético e implementado diferentes estrategias en la búsqueda de marcadores moleculares relacionados con la resistencia al ALC partiendo de las evaluaciones realizadas en campo utilizando RAPDs (Cardaña *et al.* 2003) y Microsatelites (Zizumbo-Villareal *et al.* 2006). Para en el caso de los RAPDs, solamente se pudieron determinar los polimorfismos entre palmas resistentes y susceptibles, sin embargo, no se desarrollaron los marcadores SCAR para ser validados y confirmar su asociación ya sea con la resistencia o susceptibilidad a esta enfermedad. En el caso de los microsatelites, han sido utilizados para establecer relaciones filogenéticas entre diferentes ecotipos de cocotero, sin embargo, su aplicación para la identificación de ecotipos resistentes o susceptibles aún no ha sido evaluado, lo cual sería muy interesante dentro de plantaciones segregantes, resultado de la cruce entre palmas resistentes y susceptibles al ALC.

Por otra parte, el desarrollo de marcadores moleculares en cocotero, basados en secuencias candidatas a genes de resistencia puede ser una estrategia más efectiva para la búsqueda de marcadores moleculares asociados a características de interés como la resistencia a plagas y ha enfermedades, así como también el desarrollo de mapas genéticos (Meyers *et al.*, 1999; Timmerman-Vaughan *et al.*, 2000; Miller *et al.*, 2008).

1.10. HIPÓTESIS

1.10.1. HIPÓTESIS 1

Las secuencias candidatas a genes de resistencia de cocotero modifican su expresión en respuesta a la aplicación exógena del ácido salicílico.

1.10.2. HIPÓTESIS 2

Las secuencias candidatas a genes de resistencia muestran diferentes perfiles de expresión en diferentes tejidos: hoja, tallo y raíz de palmas sanas y enfermas de amarillamiento letal.

1.10.3. HIPOTESIS 3

Los análisis filogenéticos de genes NBS-LRR en diversas especies de plantas demuestran que existen clados específicos de cada especie, así como también clados específicos de una familia de plantas, por lo tanto, es posible que estas relaciones filogeneticas se puedan observar en los RGCs de cocotero.

1.10.4. HIPÓTESIS 4

Empleando la técnica de AFLP-RGC se podrán identificar bandas polimórficas entre ecotipos de cocotero resistentes y susceptibles a la enfermedad del amarillamiento letal.

1.11. OBJETIVOS

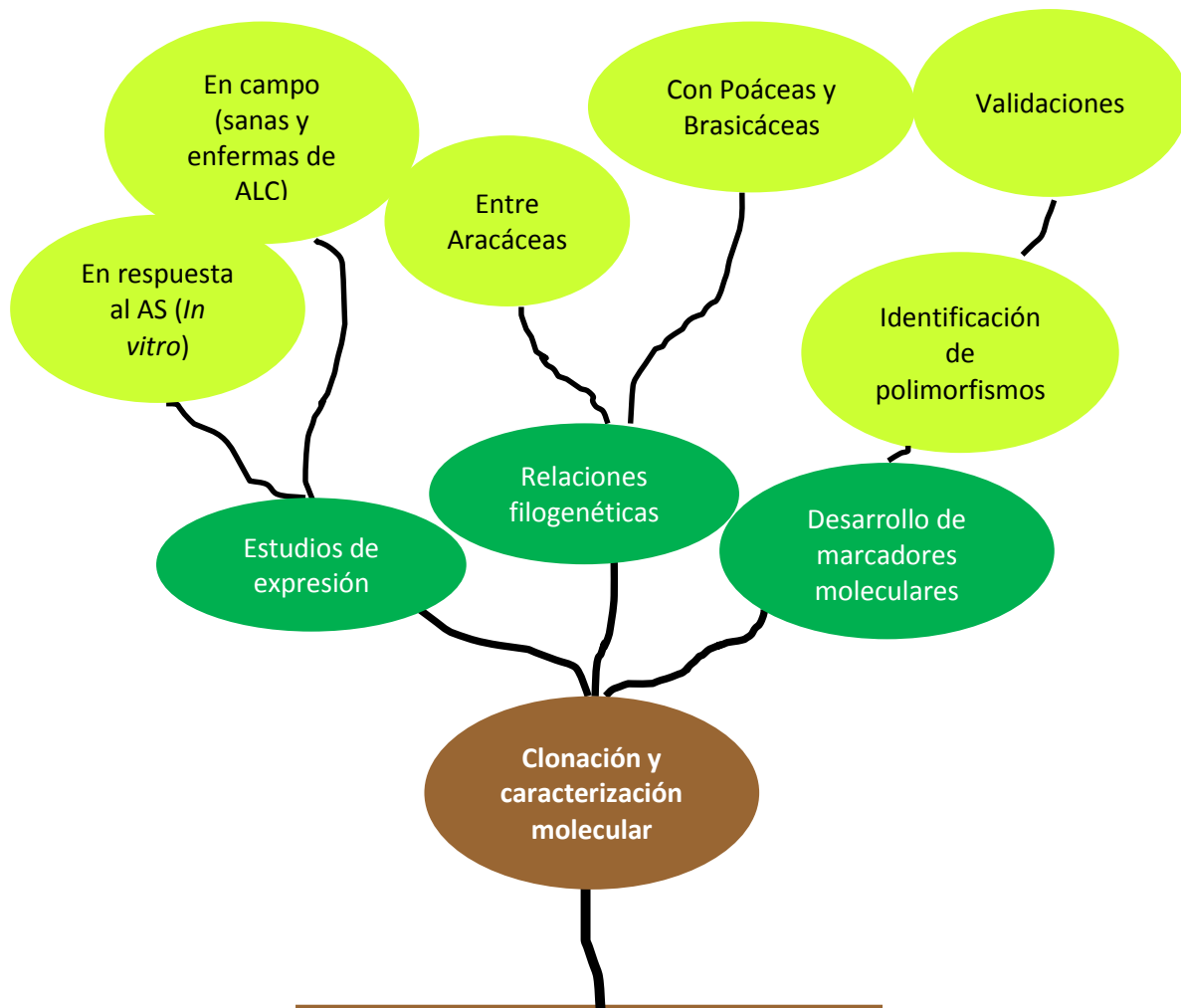
1.11.1. OBJETIVOS GENERALES

- Clonar y analizar los perfiles de expresión y las relaciones filogenéticas de las secuencias candidatas a genes de resistencia en diferentes ecotipos de cocotero.
- Desarrollar marcadores moleculares para la identificación de ecotipos de cocotero resistentes o susceptibles a la enfermedad del amarillamiento letal.

1.11.2. OBJETIVOS ESPECIFICOS

- Clonar y caracterizar secuencias candidatas a genes de resistencia en ecotipos de cocotero resistentes y susceptibles al amarillamiento letal.
- Análizar los perfiles de expresión de RGCs en plántulas de cocotero *In vitro* tratadas y no tratadas con ácido salicílico.
- Análizar los perfiles de expresión de RGCs en palmas de campo sanas y enfermas de amarillamiento letal.
- Análizar las relaciones filogenéticas de los RGCs entre especies de la familia de Arecaceae, Poaceae y Brassicaceae
- Implementar la técnica de AFLP-RGC para la identificación de polimorfismos entre ecotipos de cocotero resistentes y susceptibles a la enfermedad del amarillamiento letal.
- Clonar, secuenciar y caracterizar los fragmentos polimórficos identificados.
- Validar los polimorfismos en palmas resistentes y susceptibles al amarillamiento letal, así como también en una población segregante (F2) resultado de la cruce entre ecotipos resistentes y susceptibles.

1.12. ESTRATEGIA EXPERIMENTAL



Estudios moleculares sobre secuencias candidatas a genes de resistencia en la palma de cocotero

1.13 REFERENCIAS BIBLIOGRAFICAS

Alfano, J.R., and A. Collmer (1996). Bacterial pathogens in plants: life up against the wall. *Plant Cell*, 8, 1683–1698.

Al-Mssallem, I.S., Zhang Hu S., Q. Lin, W. Liu, J. Tan, X. Yu, J. Liu, L. Pan, T. Zhang, Y. Yin, C. Xin, H. Wu, G. Zhang, M.M. Ba Abdullah, D. Huang, Y. Fang, Y.O. Alnakhli, S. Jia, A. Yin, E.M. Alhuzimi, B.A. Alsaihati, S.A. Al-Owayyed, D. Zhao, S. Zhang, N.A. Al-Otaibi, G. Sun, M.A. Majrashi, F. Li, Tala, J. Wang, Q. Yun, N.A. Inassar, L. Wang, M. Yang, R.F. Al-Jelaify, K. Liu, S. Gao, K. Chen, S.R. Alkhaldi, G. Liu, M. Zhang, H. Guo y J. Yu: Genome sequence of the date palm *Phoenix dactylifera* L. *Nature Communications* 2013; 4:2274.

Aravind, L. y E.V. Koonin (1999). Fold prediction and evolutionary analysis of the POZ domain: Structural and evolutionary relationship with the potassium channel tetramerization domain. *J. Mol. Biol*, 285, 1353–1361.

Ayele-Gedil M, M.B. Slabaugh, S. Berry, R. Johnson, R. Michelmore, J. Miller, T. Gulya, S. Knapp (2001). Candidate disease resistance genes in sunflower cloned using conserved nucleotide-binding site motifs: genetic mapping and linkage to the downy mildew resistance gene PI1. *Genome*, 44: 205–212.

Bastidas R. y Y. Zavala (1995). "Principios de Entomología Agrícola", Ediciones Sol de Barro, ISBN 980-245-006-5.

Bakker, E.G., C. Toomajian, M. Kreitman y J. Bergelson (2006). A genomewide survey of R gene polymorphisms in Arabidopsis. *Plant Cell* 18:1803–1818.

Bennett, M., M. Gallagher, J. Fagg, C. Bestwick, T. Paul, M. Beale y J. Mansfield (1996). The hypersensitive reaction, membrane damage and accumulation of autofluorescent phenolics in lettuce cells challenged by *Bremia lactucae*. *The plant Journal*, 9 (6), 851-865.

- Bleas M., S.A. Grandis, Lee y J.T. Trevors (1998). Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial Microbiology & Biotechnology*, 21: 99–114.
- Belkhadir, Y., R. Subramaniam y J.L. Dangl (2004). Plant disease resistance protein signalling: NBS-LRR proteins and their partners. *Curr Opin Plant Biol*, 7:391–399.
- Bradley, W. (2006). Proceedings of the International conf. on Renewable Energy for Developing countries. USA.
- Broschat, T., N. Harrison y H. Donselman (2002). Losses to lethal yellowing cast doubt on coconut cultivar resistance. *Palms* 46:185-189.
- Calenge, F., C.G. Van der Linden, E. Van de Weg, H.J. Schouten, G. Van Arkel, C. Denance y C.E. Durel (2005). Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. *Theor Appl Genet*; 110:660–668.
- Cannon, SB., H.Y Zhu, A.M. Baumgarten, R. Spangler, G. May, D.R. Cook y N.D Young (2002). Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *J Mol Evol*, 54:548–562.
- Cao, H., J. Glazebrook, J.D. Clarke, S. Volko y X. Dong (1997). The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell*, 88, 57–63.
- Cardeña, R., G.R. Ashburner y C. Oropeza (2003). Identification of RAPDs associated with resistance to lethal yellowing of the coconut (*Cocos nucifera* L.) palm. *Scientia Horticulture*, 98; 257-263.
- Chen, X.M., R.F. Line y H. Leung (1998). Genome scanning for resistance-gene analogs in rice, barley, and wheat by highresolution electrophoresis. *Theor. Appl. Genet.* 97:345–355.

- Chen, Q., Z. Han, H. Jiang, D. Tian y S. Yang (2010). Strong positive selection drives rapid diversification of R-genes in Arabidopsis relatives. *J Mol Evol* 70:137–148.
- CONACOCO (2008). Consejo Nacional del cocotero, www.conacoco.com.mx.
- Chisholm, S.T., G. Coaker, B. Day y B.J. Staskawicz (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124:803-814.
- L. Dangl y J. Jones (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411:826–833
- Deslandes, L., J. Olivier, N. Peeters, D.X. Feng, M. Khounlotham, C. Boucher, I. Somssich, S. Genin y Marco (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc. Natl. Acad. Sci. USA*, 100, 8024–8029
- Dominguez, E., J. López, R. Castillo y P. Ruiz (1999). El cocotero (*Cocos nucifera* L.) manual para la producción en México. INIFAP. Libro técnico, 6,133 p.
- Dong, X. (2004). NPR1, all things considered. *Curr. Opin. Plant Biol*, 7, 547–552.
- Doyle, P. (1998). The impact of lethal yellowing type coconut diseases on small farmers. Natural Resources Institute, UK.
- Ellis, J. y D. Jones (1998). Structure and function of proteins controlling strain-specific pathogens resistance in plants. *Current Opinion in Plant Biology*, 1, 288-293.
- Ellis, J., P. Doods y T. Pryor (2000). Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, 3, 278-284.
- Erschadi S., G. Haberer, M. Schoniger y R.A. Torres Ruiz (2000). Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP). *Theoretical and Applied Genetics*, 100:633-640.

- FAO. (1998). Production yearboock. Database Results.
- H. Flachowsky, E. Schumann, W.E. Weber, A. Peil (2001). Aplications of AFLP for the detection of sex-specific markers in hemp. *Plant Breeding*, 120:305-309.
- Flor, H.H. (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* 9, 275–296.
- Glazebrook, J. (2005). Constrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology*, 43, 205-227.
- Goodman, R.N. y A.J. Novacky (1994). *The Hypersensitive Reaction in Plants to Pathogens*. St. Paul, MN: APS Press.
- Graham M, L. Mareck, D. Lohnes, P. Cregan y R. Shoemaker (2000). Expression and genome organization of resistance gene analogues in soybean. *Genome*, 43:86–93.
- Grant M, L. Godard, E. Straube, T. Ashfield, J. Leward, A. Sattler, R. Innes , J. Dangl (1995). Structure of the Arabidopsis RPM1 enabling dual specificity disease resistance. *Science* 269:843–846.
- Griffith, R. (1987). Redring disease of coconut palm. *Plant Dis*, 71: 193-196.
- Gomez-Gomez, L., y T. Boller (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Mol. Cell*, 5, 1003–1011.
- Gu, K., B. Yang, D. Tian, L. Wu, D. Wang, C. Sreekala, F. Yang, Z. Chu, G.L. Wang, F.F White y Z. Yin (2005). *R* gene expression induced by a type-III effector triggers disease resistance in rice. *Nature*, 435, 1122–1125.
- Hanold, D. y J. Randles (1991). Coconut cadang-cadang disease and its viroid agent. *Plant Dis*, 75, 330-335. Lao, D.A. (2008). *Bioenergy Forum*. Bangkok.

- Harries, H. (1995). Coconut. In: J. Smart and D.W. Simmonds. Evolution of crop plants, Second Edition. Logman Scientific & technical. London. 351-357.
- Harrison, N., I. Cordova, p. Richardson y Dibonito (1999). Detection and diagnosis of lethal yellowing. In: C. Oropeza, J.L. Verdeil, G.R. Ashburner, R. Cardeña and J.M. Santamaría (Eds) Current Advance in Coconut Biotechnology. Pp 183-196. Kluwer Academic Publishers. Dordrecht, The Netherland (In press).
- Harrison, N. y P. Jone (2004). Lethal yellowing. In: Compendium of Ornamental Palm Diseases and Disorders. Elliott ML, Broschat TK, Uchida JY, Simone GW (eds). American Phytopathological Society, St. Paul MN pp. 31-4.
- Heat, M.C. (2000). Hypersensitive response-related death. *Plant Molecular Biology*, 3,321–334.
- Howard F. y C. Barrant C. (1989). Question and answers about lethal yellowing disease. *Principies*, 33, 163-171.
- Howard, FW, D.S. Williams y R.C. Norris (1984a). Insect transmission of lethal yellowing to young palms. *International Journal of Entomology* 26:331–338.
- Hulbert, SH, C.A. Webb, S.M. Smith y Q. Sun (2001). Resistance gene complexes: Evolution and utilization. *Annu. Rev. Phytopathol*, 39, 285–312.
- Jones, D. y J. Jones (1997). The roles of leucine rich repeats in plant defences. *Adv. Bot. Res. Adv. Plant Pathol.* 24, 90–167.
- John, M. y J.W. Bonnie (2003). Plant disease resistance genes: recent insights and potential applications. *TRENDS in Biotechnology*, 21, 178-183.
- Joseph, T. y K. Radha (1975). Role of *Phytophthora palmivora* in but rot of coconut. *Plant Dis Repr*, 59, 1014-1017.

- Kanazin, V., L. Frederick Marek y R.C. Shoemaker (1996). Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci USA*, 93:11746–11750.
- Karp A y K.J. Edwards (1997). DNA markers: a global overview. Pages 1-13 *in* Caetano-Anollés G and PM Gresshoff eds. *DNA markers: protocols, applications and overviews*. Wiler-VCH, New York.
- Kawchuk, L.K., J. Hachey, D.R. Lynch, F. Kulcsar, G. van Rooijen y D.R. Waterer (2001). Tomato Ve disease resistance genes encode cell surface-like receptors. *Proc Natl Acad Sci USA*, 98, 6511–5.
- Kliebenstein Dj., J. Gershenzon y T. Mitchell-Olds (2001). Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics*, 159:359-370.
- Kobe, B. y J. Deisenhofer (1995). A structural basis of the interaction between leucine-rich repeats and protein ligands. *Nature*, 374:183-86.
- Lao, D.A. (2008). *Bioenergy Forum*. Bangkok.
- Lee, S. y J.K.C. Rose. (2010). Mediation of the transition from biotrophy to necrotrophy in hemibiotrophic plant pathogens by secreted effector proteins. *Plant Signaling and Behavior* 5, 769-772.
- Leister D, A. Ballvora, F. Salamini y C. Gebhardt (1996). A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat Genet*, 14:421–429.
- Leister D, J. Kurth, D. Laurie, M. Yano, T. Sasaki, K. Devos, A. Graner y P. Schulze-Lefert (1998). Rapid organization of resistance gene homologues in cereal genomes. *Proc Natl Acad Sci U*, 95:370–375.
- Lizama G., I. Estrada, M. Caamal, R. Souza, C. Oropeza, I. Islas y J. Zuñiga (2007). Chitosan activates a MAP-kinase pathway and modifies abundance of defense-related

transcripts in calli of *Cocos nucifera* L. *Physiological and Molecular Plant Pathology*, 70, 130-141.

López C, I. Acosta, C. Jara, F. Pedraza, E. Gaitan-Solis E, G. Gallego, S. Beebe, J. Tohme (2003). Identifying resistance gene analogs associated with resistance to different pathogens in common bean. *Phytopathology*, 93:88–95.

Martin, G., A. Bogdanove y G. Sessa (2003). Understanding the functions of plant disease resistance proteins. *Annu Rev Plant Biol*, 54, 23-61.

Martinez EA., C. Destombe, M.C. Quillet y M. Valero. (1999). Identification of random amplified polymorphic DNA (RAPD) markers highly linked to sex determination in the red alga *Gracilaria gracilis*. *Molecular Ecology*, 8:1533-1538.

Martínez-Zamora, M.G. A.P Castagnaro, J.C. Díaz Ricci (2004). Isolation and diversity analysis of resistance genes analogues (RGAs) from cultivated and wild strawberries. *Mol. Genet Genomics*, 272:480–487.

Mace, E.S., C.G. Gebhardt y R.N. Lester (1999). AFLP analysis of genetic relationships in the tribe Daturae (Solanaceae). *Theoretical and Applied Genetics*, 99: 634-641.

McDowell, J.M. y S.A Simon (2008). Molecular diversity at the plant–pathogen interface. *Dev Comp Immunol* 32:736–744

Mes J, A. Van Doorn, J. Wijbrandi, G. Simons, B. Cornelissen y M. Haring (2000). Expression of the *Fusarium* resistance gene I-2 colocalizes with the site of fungal containment. *Plant J* 23:183–193.

Meyers B, A. Kozik, A. Griego, H. Kuang y R.W Michelmore (2003). Genome-wide analysis of NBS-LRR encoding genes in *Arabidopsis*. *Plant Cell* 15:809–834.

Meyers, B., A. Dickerman, R. Michelmore, S. Sivaramakrishnan, B. Sobral y Young N. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J*, 20, 317-332.

- Miller, R., D. Bertoli, F. Baurens, C. Santos, P. Alves, N. Martins, R. Togawa, M. Souza y G. Pappas (2008). Analysis of non-TIR NBS-LRR resistance analogs in *Musa acuminata* Colla: Isolation, RFLP marker development, and physical mapping. *BMC Plant Biology*, 15, 1-15.
- Milligan S, J. Bodeau, J. Yaghoobi, I. Kaloshian, P. Zabel y Williamson (1998). The root knot resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding site, leucine-rich repeat family of plant genes. *Plant Cell* 10:1307–1319.
- Münch, S., U. Lingner, D. S. Floss, N. Ludwig, N. Sauer y H. B. Deising (2008). The hemibiotrophic lifestyle of *Colletotrichum* species. *Journal of Plant Physiology*, 165, 41-51.
- Narváez C.M. (2010). Inducción de SAR en plantas de cocotero infectadas con *Phytophthora capsici*. Tesis de maestría. Instituto Tecnológico de Mérida, Yucatán, México. 60 p.
- Nic-Matos G. (2012). Estudio de la expresión de un gen tipo NPR1 en plantas de cocotero tratadas con ácido salicílico. Centro de Investigación Científica de Yucatán. México.
- Nürnbergger, T., D. Nennstiel, T. Jabs, W.R. Sacks, K. Hahlbrock y D. Scheel (1994). High affinity binding of a fungal oligopeptide elicitor to parsley plasma membranes triggers multiple defense responses. *Cell*, 78 (3), 449-460.
- Oliver, R. P. y S.V.S. Ipcho (2004). *Arabidopsis* pathology breathes new life into the necrotrophs-vs.-biotrophs classification of fungal pathogens. *Molecular Plant Pathology*, 5(4), 347-352.
- Palacios C., S. Kresovich and F. González-Candelas (1999). A population genetic study of the endangered plant species *Limonium dufourii* (Plumbaginaceae) based on amplified fragment length polymorphism (AFLP). *Molecular Ecology*, 8: 645-657.
- Peraza-Echeverría, S., J.L. Dale, R.M. Harding, K. Smith, Mike y C. Collet (2008). Characterization of disease resistance gene candidates of the nucleotide binding site

- (NBS) type from banana and correlation of a transcriptional polymorphism with resistance to *Fusarium oxysporum* f. sp. cubense race 4. *Mol Breeding*, 22: 565–579.
- Park, S. W., E. Kumar, S. Mosher y D. F. Klessig. (2008). Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. *Science*, 318, 113-116.
- Parker J, M. Coleman, V. Szabo, V. Frost, R. Schmidt, E. Van der Biezen, T. Moores, C. Dean, M. Daniels y J. Jones (1997). The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll interleukin-1 receptors with N and L6. *Plant Cell* 9:879–894.
- Parthasarathy, M., W. van Slobbe y C. Soudant (1978). Hartrot or fatal wilt of palms. 1. Coconuts (*Cocos nucifera*). *Principes*, 22, 3-14.
- Pei X, Li S Jiang Y, Wang Z y S. Jia (2007). Isolation, characterization and phylogenetic analysis of the resistance gene analogues (RGAs) in banana (*Musa* spp.). *Plant Sci* 2007; 172:1166–1174.
- Pieterse, M.J. y L.C. Van Loon (2004). NPR1: the spider in the web of induced resistance signaling pathways. *Current Opinion in Plant Biology*, 7; 456-464.
- Pieterse, 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.
- Pflieger, S., V. Lefebvre y M. Causse. (2001). The candidate gene approach in plant genetics: a review. *Molecular Breeding*, 7:275-291.
- Puch-Hau, C. (2012). Aislamiento y caracterización de secuencias candidatas a genes de resistencia en cocotero. Tesis de maestría. Centro de Investigación Científica de Yucatán. México. 94 p.
- Rivkin M., C. Vallejos y P. McClean (1999). Disease-resistance related sequences in common bean. *Genome* 1999; 42:41–47.

- Qutob, D., S. Kamoun y M. Gijzen (2012). Expression of a *Phytophthora sojae* necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. *The plant Journal*, 32, 361-373.
- Rohde, W., J. Randles, Langridge, P. y D. Hanold (1990). Nucleotide sequence of a circular single-stranded DNA associated with coconut foliar decay virus, *Virology*, 176:648-51.
- Saha D, R.S Rana, A.K. Sureja, M. Verma, L. Arya, A.D. Munshi (2013). Cloning and characterization of NBS-LRR encoding resistance gene candidates from Tomato Leaf Curl New Delhi Virus resistant genotype of *Luffa cylindrica* Roem. *Physiological and Molecular Plant Pathology*, 81; 107-117.
- Salmeron J, E. Oldroyd, C. Rommens, S. Scofield, H. Kim, D. Lavelle, D. Dahlbeck, B. Staskawicz (1996). Tomato Prf is a member of a leucine-rich repeat class of plant disease resistance gene and lies embedded within the Pto kinase gene cluster. *Cell* 86:123–133.
- Saraste, M., P. Sibbald y A. Wittinghofer (1990). The P-loop a common motif in ATP- and GTP binding proteins. *Trends Biotechnol*, 15, 430–435.
- Shang J, Y. Tao, X. Chen, Y. Zou, C. Lei, J. Wang, X. Li, X. Zhao, M. Zhang, Z. Lu, et al. (2009). Identification of a new rice blast resistance gene, *Pid3*, by genome wide comparison of paired nucleotide-binding site leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. *Genetics*. 182, 1303–1311.
- Seah S, W. Spielmeyer, J. Jahier, K. Sivasithamparam, S. Lagudah (2000). Resistance gene analogs within and introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat. *Mol Plant Microbe Interact*. 13:334–341.
- Shen K, B. Meyers, M. Islam-Faridi, D. Chin, D. Stelly, R. Michelmore (1998). Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. *Mol Plant Microbe Interact*. 11:815–823.

- Staskawiez, B., F. Ausubel, B. Baker, J. Ellis y J. Jones (1995). Molecular genetics of plant disease resistance. *Science*, 268, 661-667.
- Timmerman-Vaughan, G., T. Frew y N. Weeden (2000). Characterization and linkage mapping of R-gene analogous DNA sequences in pea (*Pisum sativum* L.). *Theoretical and Applied Genetics*, 101, 241-247.
- Tiryaki y H. Tunaz (2004). Systemic acquired resistance: characterization of genes associated with plant defence response. *Journal of Cell and Molecular Biology*, 3, 9-14.
- Traut, T. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide binding-sites. *Eur J Biochem*, 222, 9-19.
- Van Der Biezen E, J.D.G. Jones (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem Sci*, 23:454–456.
- Van Der Biezen, E., C. Freddie, K. Kahn, J. Parker y J. Jones (2002). Arabidopsis RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. *Plant J*, 29, 439-451.
- Van Loon, L.C. y A. Van Kammen (1970). Polyacrylamide disc electrophoresis of the soluble proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN." II. Changes in protein constitution after infection with tobacco mosaic virus. *Virology* 40: 199–211.
- Wan, H., W. Yuan, Q. Ye, R. Wang, M. Ruan, Z. Li, Z. Guozhi, Z. Yao, J. Zhao, S. Liu y Y. Yang (2012). Analysis of TIR- and non-TIR-NBS-LRR disease resistance gene analogous in paper: characterization, genetic variation, functional divergence and expression patterns. *BCM Genomics*. 13: 1-15.
- Ward E.R., S.J. Uknes, S.C. Williams, S.S Dincher, D.L Wiederhold, D.C Alexander, P. Ahl-Goy P., J.P. Metraux y J.A Ryals (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* 3, 1085–1094.

- Yao N., Eisfelder, J. Marvin y J. T. Greenberg (2004). The mitochondrion-an organelle commonly involved in programmed cell death in *Arabidopsis thaliana*. *The Journal*, 40, 596-610.
- You-Xiong Q., X. Li-Ping, L. Jian-Wei y C. Ru-Kai (2009). Isolation and Characterization of NBS-LRR Resistance Gene Analogs from Sugarcane. *Acta Agron Sin*, 35(4): 631-639.
- Yu Y.G., G.R. Buss y M.A Saghai (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proc. Natl. Acad. Sci. USA*, 93; 11751–11756.
- Young, N. (2000). The genetic architecture of resistance. *Curr Opin Plant Biol*, 3, 285-290.
- Zhang, Z., J. Xu, Q. Xu, P. Larkin y Z. Xin. (2004). Development of novel PCR markers linked to the BYDV resistance gene *Bdv2* useful in wheat for marker-assisted selection. *Theor. Appl. Genet.* 109:433–439.
- Zhang, J., Y. Yuan, Niu, Chen, D. Hinchliffe, Lu Yingzhi, Yu Shuxun, R. Percy, M. Ulloa y R. Cantrell (2007). AFLP-RGA markers in comparison with RGA and AFLP in cultivated tetraploid cotton. *Crop science*, 17: 180-187.
- Zhang, Y., J. Wang, X. Zhang, J.Q. Chen, D. Tian, S. Yang (2009). Genetic signature of rice domestication shown by a variety of genes. *J Mol Evol* 68:393–402.
- Zipfel C. y Felix (2005). Plants and animals: different taste for microbes? *Current Opinion in Plant Biology*, 8, 853-360.
- Zizumbo-Villareal D., M. Ruiz-Rodriguez, H. Harries y M. Colunga-Garcia (2006). Population genetics, lethal yellowing disease, and relationships among Mexican and imported Coconut Ecotypes. *Crop Sci*, 46: 2509-2516.
- Zizumbo-Villareal, D., M.P. Colunga-García Marín, M. Fernández-Barrera, N. Torres-Hernández y C. Oropeza (2008). Mortality of Mexican coconut germplasm due to lethal yellowing. *Plant Genetic Resources Newsletter, FAO-Bioversity* 156: 23-33.

CAPITULO II

MOLECULAR CLONING AND CHARACTERIZATION OF DISEASE-RESISTANCE GENE CANDIDATES OF THE NUCLEOTIDE BINDING SITE (NBS) TYPE FROM *Cocos nucifera* L.

2.1. ABSTRACT

In this study, degenerate primers were used to amplify nucleotide-binding site (NBS)-type sequences from coconut ecotypes either resistant or susceptible to lethal yellowing. Genomic DNA fragments of approximately 500-700 bp were obtained and sequenced. Phylogenetic analysis of these fragments revealed that they clustered in seven different clades. We also found that all CnRGCs sequences were grouped within the non-TIR-NBS-LRR subclass of NBS-LRR genes. The expression analysis revealed changes in expression profiles in response to salicylic acid (SA) and a constitutive expression profile in plants untreated with SA. Moreover, different expression profiles of these sequences were observed in healthy and palms affected by lethal yellowing disease. This is the first large scale analysis of NBS-type sequences in coconut palm.

2.2. INTRODUCTION

Coconut is widely distributed throughout the humid tropics where it is cultivated over an estimated area of twelve million ha. It is an important perennial crop contributing significantly to food security, employment and income generation. There is a great ethnic diversity in the ways that various coconut resources are produced and used [1]. However, this palm is highly susceptible to several pathogens that hinder yields including viruses [2], viroides [3], protozoa [4], oomycetes [5], nematodes [6] and mollicutes [7] as the 'Candidatus Phytoplasma palme' responsible for the devastating disease known as lethal yellowing (LY) in the Americas [8]. Due to this disease, in recent years in Mexico, the highly susceptible Atlantic Tall (AT) variety has been almost wiped out from the coast of the Yucatan Peninsula and has diminished in abundance along the coast of the Gulf of Mexico [9]. Despite the

importance of coconut and the diversity of pathogens that affect it, to date there are no reports of R genes capable of conferring resistance to the most devastating diseases of this crop. Therefore, the isolation of these genes could provide opportunities for the generation of disease-resistant varieties through their introgression into susceptible coconut cultivars. The majority of plant R genes encode proteins with cytoplasmic nucleotide-binding site and leucine-rich repeat (NBS-LRR) domains that confer resistance to a wide variety of pathogens and pests including viruses, bacteria, fungi, nematodes and insects [10]. The NBS-LRR class is thought to recognize pathogens and respond by activating signal transduction pathways leading to disease resistance [11]. The C-terminal LRR region allows for a direct or indirect recognition of the pathogen through protein-protein interaction, while the N-terminal region, including the NBS, is thought to be involved in signaling [12]. This class of R genes can be divided into two subclasses based on the presence or absence of a domain at the N-terminus of the protein with homology to the *Drosophila* Toll and mammalian Interleukin-1 Receptor (TIR) [10,13]. The non-TIR-NBSLRR subclass can contain an N-terminal coiled-coil (CC) motif, a subset of which code for a leucine zipper sequence (LZ). These two subclasses can also be distinguished (95% accuracy) by the last residue, D (Aspartate) for TIR-NBS-LRR or W (Tryptophan) for non-TIR-NBS-LRR, of the conserved kinase-2 motif within the NBS domain [14]. The subclass non-TIR-NBS-LRR is widely distributed in both monocotyledonous and dicotyledonous species, while the TIRNBS-LRR subclass appears to be restricted to dicot species [14-16].

The highly conserved motifs in the NBS domain have been targeted to isolate NBS-type sequences by PCR-based strategies. Using this approach, RGCs have been isolated from soybean [17-19], potato [20], lettuce [21], rice and barley [22], wheat [23], sunflower [24], common bean [25,26], strawberry [27], apple [28], banana [29-31], pepper [32] and tomato [33]. Also, genome-wide analyses of NBS-LRR genes have been carried out recently in plants whose genomes have been sequenced [34-37].

Since the genome sequence of coconut palm is still not available, the total number of NBS-LRR sequences in this crop is unknown. So far, there are only four NBS-LRR disease resistance candidates of coconut palm deposited in the GenBank (accession numbers FJ666396, KF002584, KC465244 and GQ856288) without being associated with any published study. Therefore, information about the NBS-LRR class of R genes from this important crop is quite limited. In this study, we used a PCR approach using degenerate

primers for amplification and subsequent characterization of RGCs from the coconut LY-resistant Malayan Yellow Dwarf (MYD), Mexican Pacific Tall 2 (MXPT2) or the LY-susceptible Mexican Atlantic Tall (MXAT) ecotypes [38]. We found seven major classes of coconut RGCs based on their phylogenetic grouping patterns. Furthermore, the expression in coconut plantlets of a representative member of each class showed a low and constitutive expression similar to known *R* genes, and treatment with the defense hormone salicylic acid modified their expression patterns, suggesting that these RGCs may play important roles in disease resistance mechanisms in coconut. Moreover, in healthy and palms affected by lethal yellowing disease were observed differences in the expression profiles of these sequences.

The present study provides the first large scale analysis of RGCs in coconut, offering new opportunities for the development of molecular markers associated with pathogen resistance, *R* gene discovery and molecular breeding in this crop.

2.3 MATERIALS AND METHODS

2.3.1. PLANT MATERIAL AND DNA EXTRACTION

Leaf tissue samples were collected from 30 mature coconut palms belonging to three ecotypes (10 palms per ecotype) that showed a differential response against LY in Mexico [38]. The ecotypes included were MYD, MXPT2 and MXAT. DNA was extracted following the protocol described by Harrison and colleagues [77] from 100 mg samples of leaf tissue. Subsequently, a mixture of DNA of the 10 palms of each ecotype was used as template for the PCR reaction.

2.3.2. DEGENERATE PRIMERS AND PCR AMPLIFICATION

Six pairs of degenerate primers were used [17,48-50,31] (Table S1). PCR reactions were performed in 25 μ L volumes with the following concentrations: 200 ng of template genomic DNA, 2 mM of each degenerate primer, 200 μ M of each dNTPs (Invitrogen), 2 mM MgCl₂, 5x Mango *Taq* Colored Reaction Buffer and 1.5 U of Mango *Taq* DNA polymerase (BIOLINE). Thermal cycling conditions were initially denatured at 96 °C for 1 min, followed

by 40 cycles at 94 °C for 1 min, 45 °C for 1 min and 72 °C for 1 min, and a final elongation step at 72 °C for 8 min. PCR products were separated by electrophoresis on 1% agarose gels (Promega, Madison, Wisconsin, USA) stained with ethidium bromide, visualized by a UV transilluminator and photographed with a GelDoc 2000 (Bio-Rad, Hercules, CA, USA).

2.3.3. CLONING AND SEQUENCING OF PCR PRODUCTS

PCR products of expected size were excised from the agarose gel and purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA, USA). The obtained DNA was cloned into the plasmid pGEM-T Easy vector (Promega, Madison, Wisconsin, USA) and transformed into *Escherichia coli* JM109 competent cells according the manufacturer's instructions (Promega, USA). Recombinant plasmid clones were extracted using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA, USA) and quantified using the NANODROP 2000c Spectrophotometer (Thermo SCIENTIFIC). The recombinant plasmid DNA was sequenced by Davis sequencing Company (CA, USA).

2.3.4 DNA SEQUENCE ANALYSIS AND PHYLOGENETIC TREE CONSTRUCTION

Similarity searches were performed using the BLASTX program [78] (www.ncbi.nlm.nih.gov). The open reading frames (ORFs) of the RGCs were predicted by the fr33 translator (<http://www.fr33.net/translator.php>). Predicted protein sequences were aligned using the ClustalX program version 1.81 with the default settings [39]. Identical amino acids in the alignment were shaded with black and conservative substitutions with gray using the Boxshade program (<http://www.ch.embnet.org>). Phylogenetic trees were constructed by the neighbor-joining (NJ) method [79] with Poisson correction using the NJ algorithm implemented with version 5.0 of the Molecular Evolutionary Genetic Analysis software (MEGA) [80]. Bootstrap analysis was performed with 1000 replicates to evaluate the reliability of the interior nodes for a particular grouping pattern in the tree.

2.3.5 EXPRESSION ANALYSIS OF RGCs

2.3.5.1 EXPRESSION ANALYSIS IN PLANTS TREATED WITH SA

2.3.5.1.1 SALICYLIC ACID TREATMENT

We utilized 12-month-old plantlets cultured in liquid medium Y3 [81] without charcoal and plant growth regulators obtained from in vitro germination of zygotic embryos of MXPT2 ecotype. For the treatment with SA (Sigma, USA), a stock was prepared by dissolving EtOH and filtering through a 0.22 µm membrane (Millipore, USA) which was then added to the sterilized medium to give final concentrations of 2.5 mM. Leaf, stem and root tissue samples were collected at 0, 24, 48, 72, 96 and 120 h after treatment (each treatment consisted of three plantlets). The samples were quickly frozen in liquid nitrogen and stored at -80 °C.

2.3.5.2 EXPRESSION ANALYSIS OF FOUR NBS SEQUENCES OF COCONUT PALMS INFECTED WITH LY-PHYTOPLASMA

2.3.5.2.1 PLANT MATERIAL AND DNA EXTRACTION

Three adult asymptomatic coconut palms and field-driven palms with symptoms of lethal yellowing disease were sampled to determine expression profiles of four NBS sequences in leaf, stem and root tissue. The symptomatic plants were naturally infected coconut palms (*Cocos nucifera* L.) showing stages 1 and 2 of LY symptom development, as defined by McCoy *et al.*³⁰ nut fall (1) and appearance of necrotic inflorescences (2). Additionally we used 12-month-old plantlets cultured in liquid medium Y3³¹ as control of expression. The DNA for the diagnostic of LY phytoplasma was extracted following the protocol described by Harrison *et al.*³³ with minor modifications.

2.3.5.2.1 DIAGNOSTIC OF LY PHYTOPLASMA

For the detection of LY phytoplasma DNA from root tissues of asymptomatic and symptomatic of coconut palms the TaqMan LY16S assay reported by Córdova *et al*⁸³ was utilized. Reactions were performed in 20 µl of TaqMan Universal PCR master mix with AmpErase UNG (Uracil N-Glycosylase) (Applied Biosystems, USA), 1 µl of primer mix containing 900 nM of each primer, probe (250 nM) and 50 ng of DNA. Amplification was performed with a CFX96 real-time PCR system (Bio-Rad, USA). PCR was initiated with two steps: 2 min at 50°C to activate AmpErase UNG, 10 min at 95°C to activate AmpliTaq Gold DNA polymerase followed by 40 cycles at 95°C for 15 sec and 1 min at 61°C. All DNA samples including controls were assessed in duplicate. The threshold cycle (Ct) values of each PCR reaction were manually set to intersect the exponential phase of the amplification curves, but the baseline was automatically set by CFX manager software IQ (Bio-Rad, USA).

2.3.6 RNA EXTRACTION AND CDNA SYNTHESIS

Total RNA was extracted from coconut seedling (leaf, stem and root) following the methodology described by Montero-Córtés and colleagues [82]. The RNA samples were treated with DNase-free (Ambion, Foster city, CA, USA) to remove the genomic DNA. The DNase-treated RNA was quantified using the Nanodrop 1000 (Thermo Scientific, USA). First-strand cDNA was prepared using random hexamers (Invitrogen, Carlsbad, CA, USA) and 500 ng of RNA in conjunction with reverse-transcriptase (Qiagen, Valencia, CA, USA) according to the directions of the manufacturer.

2.3.7 TRANSCRIPT EXPRESSION ANALYSIS BY QUANTITATIVE RT-PCR

Real-time RT-PCR was performed using platinum SYBR Green qPCR superMix-UDG (Invitrogen, USA) and analyzed in an iCycler IQ real-time PCR detection system (Bio-Rad, USA), according to the manufacturer's instructions. Specific primers (Table S2) were designed from a representative sequence of each coconut RGC group to quantify its expression levels, which were normalized to those of the 18S rRNA gene (FW 5' CGGCTACCACATCCAAGGAA-30 and RV 5'-GCTGGAATTACCGCGGCT-3') in each

sample. PCR conditions were 95 °C for 2 min, followed by 35 cycles at 95 °C by 15 s and a final step at 64 °C for 1 min. Real-time PCR results were analyzed using the $2^{-\Delta\Delta C_t}$ method [83], with the appropriate validation experiments performed beforehand (Real-Time PCR Applications Guide BIORAD). Results are expressed in relative expression levels. Data represents the mean \pm standard deviation (SD) of two independent biological samples (each by triplicate) by treatment. Amplified cDNAs were cloned and sequenced to check the specificity of the amplification products (Table S3).

2.4. RESULTS

2.4.1. AMPLIFICATION AND IDENTIFICATION OF RGCS FROM COCONUT PALM

Six pairs of degenerate primers were used to amplify and identify RGCs from coconut palm (Table S1). The positions of the degenerate primers were located in five conserved motifs (P-loop, kinase 2, GLPL, RNBS-D and MHDV) of the NBS domain of plant *R* genes (Fig. 2.1). The six pairs of degenerate primers generated a total of three bands of the predicted size (~700, ~530 and ~650 bp) using the genomic DNA mix of each ecotype (Fig. S1).

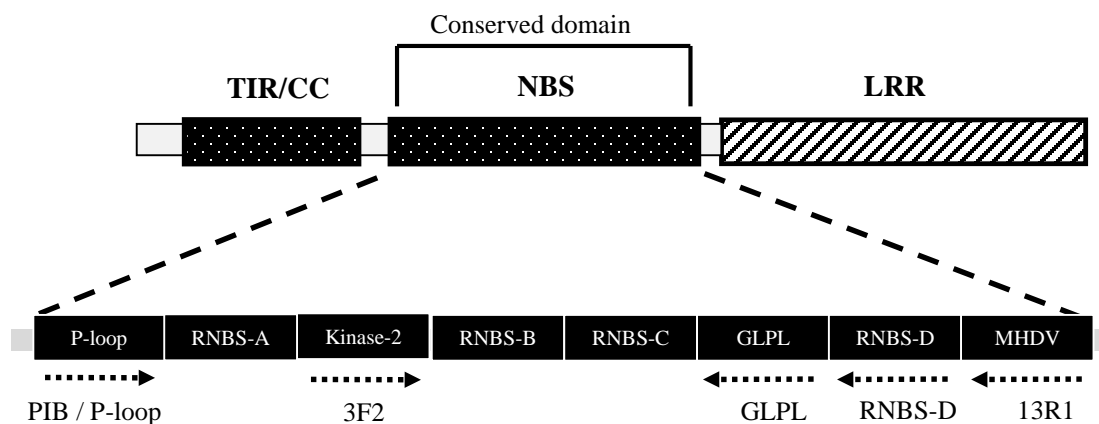


Figure 2.1 Conserved motifs within the NBS domain targeted using degenerate RGC primers. The arrows show the position (P-loop, kinase-2, GLPL, RNBS-D and 13R1) and direction of the primers, with tips indicating the 3' primer end.

The amplicons were cloned and a total of 283 clones were randomly selected for sequencing. The bioinformatics analysis revealed 169 sequences that presented high homology to the NBS-LRR class of *R* genes. The majority of these sequences (143) showed uninterrupted open reading frames (ORFs) and referred to as CnRGCs from here on. The remaining sequences contained premature stop codons so they were excluded from further analysis. The number of clones displaying similarity to RGCs varied between ecotypes and between different primer combinations (Table 1). Some of the sequences showed similarity

to other genes from Genbank, such as retroelements, proteins CIPK (Calcineurin B-Like interacting Protein Kinase), ABC transporter and ATP synthases.

Table 1
RGCs amplified using degenerate primers from coconut ecotypes

Primer combinations	Number of insert-containing plasmids			Number of sequences with homology to R-gene or RGCs by coconut ecotype			Number of sequences with ORF by coconut ecotype		
	MXPT	MYD	MXA	MXPT	MYD	MXA	MXPT	MYD	MXA
PIB-RNBS-D	85	25	13	58	18	0	54	18	0
P-loop-GLPLA	46	79	11	20	8	8	21	6	8
3F2-13RI	80	36	37	38	6	13	27	4	5
PIB-P3A	0	0	0	0	0	0	0	0	0
PIA-P3A	0	0	0	0	0	0	0	0	0
PIA-RNBS-D	na	na	na	Na	na	Na	na	na	Na

MXPT: Mexican Pacific Tall 2

na: no amplification

MYD: Malayan Yellow Dwarf

MXAT: Mexican Atlantic Tall

In order to group these sequences based on their similarities, we constructed a phylogenetic tree based on the sequences with uninterrupted ORFs. This tree grouped 143 RGCs into seven major clades (I-VII) (Fig. 2.2). Every clade varied in the number of RGCs, 8 for clade VII to 52 for clade I (Table 2). The clades I, III and VI included RGCs from the three ecotypes studied, clades II and V included only the resistant ecotypes (MXPT2 and MYD) and clades IV and VII included sequences from one resistant (MXPT2) and one susceptible ecotype (MXAT) to lethal yellowing disease (Table 2). These results indicate high diversity and divergency of RGCs in the coconut palm genome and reflect a difference in the abundance of the members of each clade.

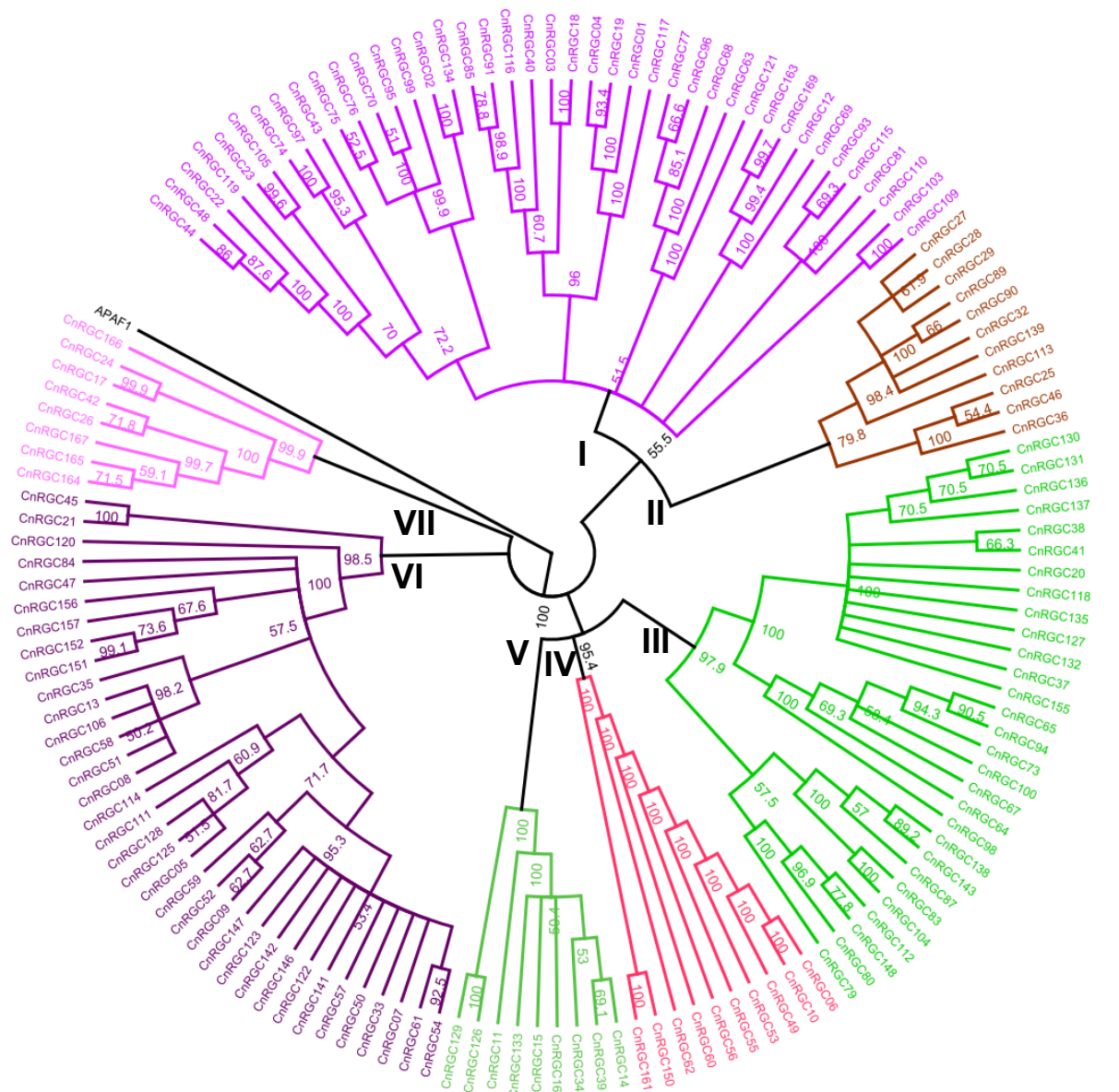
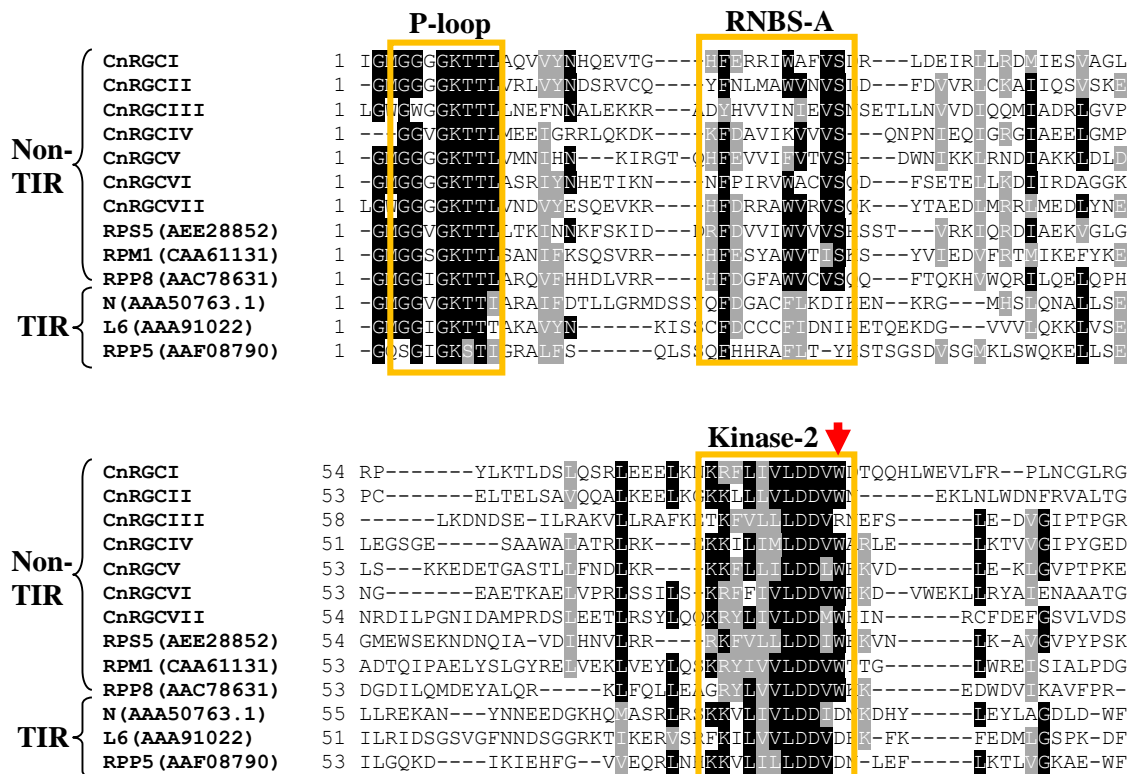


Fig. 2.2 Phylogenetic tree based on the neighbor-joining method (Saitou and Nei, 1987), of the deduced amino sequences of 143 coconut palm RGCs. Seven different RGCs clades (I-VII) were obtained. The numbers on the branches indicate the percentage of 1000 bootstrap replications supporting the particular nodes.

Table 2
Sequences that conform each clade of RGCs in coconut.

Class	Ecotype	Sequence name	Length of encoding nucleotide residues (bp) and amino acids	GenBank accession number		
CnRGC I	MXPT	CnRGC44; CnRGC48; CnRGC22; CnRGC23; 700/233; 700/233; 702/234; 796/232; KM926399; KM926403; KM926379; KM926380				
		CnRGC105; CnRGC74; CnRGC97; CnRGC43; 693/230; 708/236; 702/233; 701/233; KM926450; KM926426; KM926444; KM926398				
		CnRGC99; CnRGC76; CnRGC70; CnRGC95; 687/229; 687/229; 687/229; 699/232; KM926446; KM926428; KM926424; KM926442				
		CnRGC75; CnRGC02; CnRGC85; CnRGC91; 687/228; 691/230; 690/229; 688/229; KM926427; KM926359; KM926435; KM926439				
		CnRGC116; CnRGC40; CnRGC03; CnRGC18; 690/229; 705/234; 705/234; 691/229; KM926459; KM926395; KM926360; KM926375				
		CnRGC04; CnRGC19; CnRGC01; CnRGC77; 701/233; 703/233; 701/233; 706/232; KM926361; KM926376; KM926390; KM926429				
		CnRGC96; CnRGC68; CnRGC63; CnRGC12; 607/202; 696/232; 695/232; 703/234; KM926443; KM926454; KM926418; KM926369				
		CnRGC69; CnRGC93; CnRGC115; CnRGC81; 702/234; 696/231; 716/235; 696/131; KM926423; KM926440; KM926458; KM926432				
		CnRGC110; CnRGC103; CnRGC109 693/230; 681/226; 669/222 KM926453; KM926448; KM926452				
		MYD CnRGC119; CnRGC134; CnRGC117; CnRGC121 693/231; 690/229; 702/233; 707/235 KM926462; KM926476; KM926460; KM926464				
		MXAT CnRGC163; CnRGC169 528/175; 543/180 KM926495; KM926500				
		CnRGC II	MXPT	CnRGC27; CnRGC28; CnRGC29; CnRGC89; 669/221; 669/222; 669/222; 669/222; KM926384; KM926385; KM926386; KM926437		
				CnRGC90; CnRGC32; CnRGC113; CnRGC25; 669/222; 669/222; 723/240; 704/234; KM926438; KM926387; KM926456; KM926382		
				CnRGC46; CnRGC36 696/231; 696/231 KM926401; KM926391		
		MYD	CnRGC139 669/222 KM926481			
CnRGC III	MXPT	CnRGC38; CnRGC41; CnRGC20; CnRGC37; 719/239; 719/239; 721/239; 720/239; KM926393; KM926396; KM926377; KM926392				
		CnRGC65; CnRGC94; CnRGC73; CnRGC100; 720/239; 726/241; 721/239; 723/240; KM926420; KM926441; KM926425; KM926447				
		CnRGC67; CnRGC64; CnRGC98; CnRGC87; 723/241; 724/240; 723/240; 705/234; KM926421; KM926419; KM926445; KM926436				
		CnRGC83; CnRGC104; CnRGC112; CnRGC80; 705/234; 705/234; 708/235; 708/235; KM926433; KM926449; KM926455; KM926431				
		CnRGC79 708/235 KM926430				
		MYD CnRGC130; CnRGC131; CnRGC136; CnRGC137; 723/240; 723/240; 723/240; 723/240; KM926472; KM926473; KM926478; KM926479				
		CnRGC118; CnRGC135; CnRGC127; CnRGC132; 721/240; 719/239; 722/240; 721/239 KM926461; KM926477; KM926469; KM926474				
		CnRGC138; CnRGC143; CnRGC148 705/234; 705/234; 708/235 KM926480; KM926484; KM926487;				
		MXAT CnRGC155 720/239 KM926491				
		CnRGC IV	MXPT	CnRGC06; CnRGC10; CnRGC49; CnRGC53; 499/165; 499/165; 499/165; 499/165; KM926363; KM926367; KM926404; KM926408		
				CnRGC55; CnRGC56; CnRGC60; CnRGC62 499/165; 499/165; 499/165; 499/165 KM926410; KM926411; KM926415; KM926417		
				MYD		
		MXAT CnRGC150; CnRGC161 509/169; 510/169 KM926488; KM926494;				
		CnRGC V	MXPT	CnRGC14; CnRGC39; CnRGC34; CnRGC16 692/230; 690/239; 693/230; 695/231; KM926371; KM926394; KM926389; KM926373		
				CnRGC15; CnRGC11 692/229; 693/230; KM926372; KM926368;		
MYD CnRGC126; CnRGC129; CnRGC133 707/235; 706/234; 693/230 KM926468; KM926471; KM926475						
MXAT						
CnRGC VI	MXPT	CnRGC54; CnRGC61; CnRGC07; CnRGC33; 526/174; 526/174; 511/169; 510/169; KM926409; KM926416; KM926364; KM926388				
		CnRGC50; CnRGC57; CnRGC09; CnRGC52; 510/169; 510/169; 510/169; 510/169; KM926405; KM926412; KM926366; KM926407				
		CnRGC59; CnRGC05; CnRGC111; CnRGC114; 510/169; 699/232; 657/218; 657/218; KM926414; KM926362; KM926486; KM926457				
		CnRGC08; CnRGC51; CnRGC58; CnRGC106; 510/169; 510/169; 510/169; 675/224; KM926365; KM926406; KM926413; KM926451				
		CnRGC13; CnRGC35; CnRGC47; CnRGC84; 701/233; 699/232; 699/232; 672/223 KM926370; KM926422; KM926402; KM926434				
		CnRGC21; CnRGC45 699/232; 700/232 KM926378; KM926400				
		MYD CnRGC141; CnRGC122; CnRGC146; CnRGC142; 510/169; 510/169; 510/169; 510/169; KM926482; KM926465; KM926485; KM926483				
		CnRGC123; CnRGC147; CnRGC125; CnRGC128; 510/169; 510/169; 699/232; 698/232 KM926466; KM926501; KM926467; KM926470				
		CnRGC120 684/228; KM926463				
		MXAT CnRGC151; CnRGC152; CnRGC157; CnRGC156 875/291; 541/179; 327/175; 458/152; KM926489; KM926490; KM926493; KM926492				
		CnRGC VII	MXPT	CnRGC26; CnRGC42; CnRGC17; CnRGC24 731/243; 729/242; 731/243; 732/243 KM926383; KM926397; KM926374; KM926381		
				MYD		
				MXAT CnRGC164; CnRGC165; CnRGC167; CnRGC166 546/181; 528/181; 546/181; 528/175 KM926496; KM926497; KM926499; KM926498		

Representative deduced amino acid sequences, one from each clade CnRGC121 (I), CnRGC36 (II), CnRGC118 (III), CnRGC62 (IV), CnRGC34 (V), CnRGC21 (VI) and CnRGC42 (VII), were aligned using the ClustalX program [39] along with the NBS domain of known resistance proteins such as: RPS5 [40], RPM1 [41], RPP8 [42], RPP5 [43] from Arabidopsis; N from tobacco [44] and L6 from flax [45]. RPS5, RPM1 and RPP8 belong to the subclass non-TIR-NBS-LRR and N, L6 and RPP5 to the subclass TIR-NBS-LRR. The alignment showed that each of the CnRGCs presented the characteristic motifs of NBS domain and the residues associated with the non-TIR-NBS-LRR [14], such as the tryptophan residue (W) that was present in six out of seven sequences at the kinase-2 motif (Fig. 3). In contrast, none of CnRGCs presented an aspartate residue (D) at the end of kinase-2 motif that is characteristic of the sequences that belong to TIRNBS-LRR subclass [14] (Fig. 2.3).



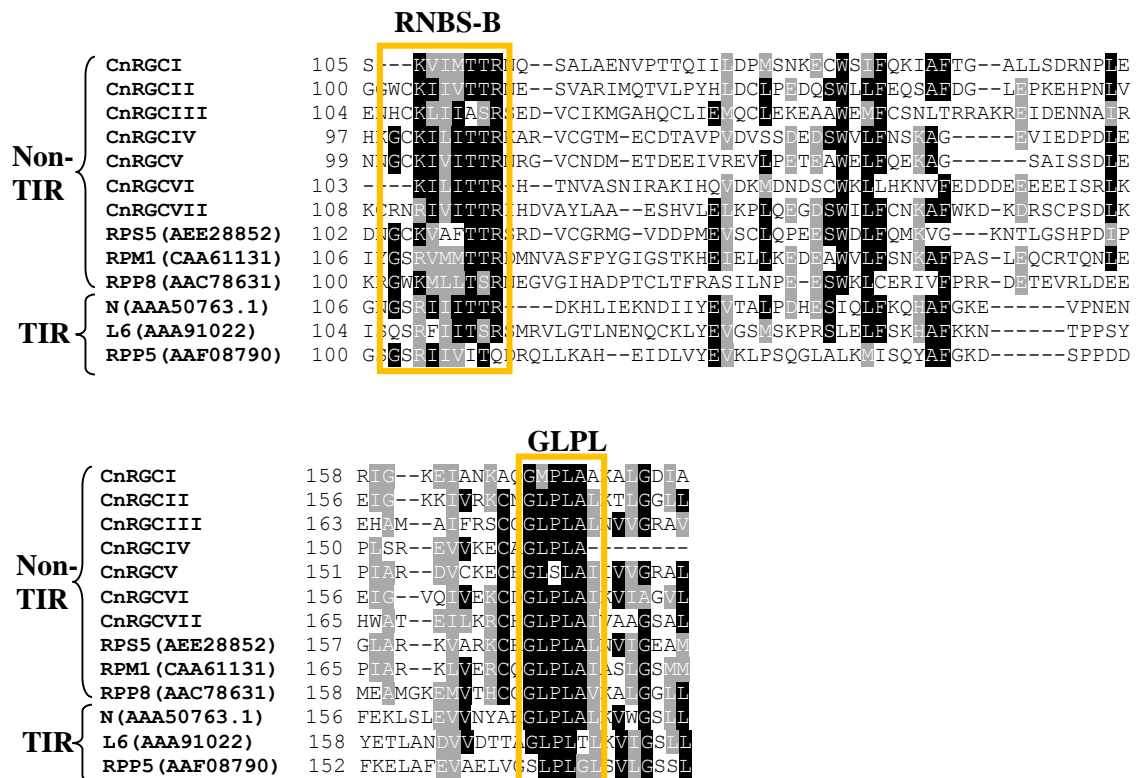


Fig. 2.3. Aminoacid sequence alignment between P-loop, RNBS-A, Kinase-2, RNBS-B and GLPL of non-TIR-NBS-LRR with the NBS domains of known R genes of TIR and Non-TIR-NBS-LRR family gene N from tobacco (AAA50763.1), L6 from flax (AAA91022), RPP5 (AAF08790), RPS5 (AEE28852), RPM1 (CAA61131) and RPP8 (AAC78631) from *Arabidopsis*. Conserved domains are boxed. The alignment was constructed utilizing the CLUSTALX program (Thompson et al. 1997) and one representative deduced amino acid sequence from each clade (CnRGC-I: CnRGC121, CnRGC-II: CnRGC36, CnRGC-III: CnRGC118, CnRGC-IV: CnRGC62, CnRGC-V: CnRGC34, CnRGC-VI: CnRGC21, CnRGC-VII: CnRGC17). Similar aminoacids residues are shaded in grey and identical aminoacid residues are shaded in black. The arrow indicate the last aminoacid residue in kinase-2 motif.

Sequence comparisons of the CnRGCs with themselves were done using the region spanning the P-loop and GLPLA motifs (~170 amino acids) of the NBS domain. When the deduced amino acid sequences of *Cocos nucifera* RGC sequences were compared, the levels of identity ranged from 18.34% (between CnRGC-I and CnRGC-III) to 40.49% (between CnRGC-IV and CnRGC-V) (Table 3), indicating a high degree of divergence between the CnRGCs. CnRGCs were compared with three of the RGCs from coconut palm previously registered in GenBank (KF002584, KC465244, FJ666396) that showed uninterrupted open reading frames. The highest values of identity were observed between CnRGC-II and KF002584 (47.17) and the lowest values were observed between CnRGC-III and FJ666396 (16.46%) (Table 3). These results indicate that KF002584, KC465244, FJ666396 sequences are different from those we are reporting in this study.

We also analyzed sequence identity between the CnRGCs with known R genes from other species (RPP8, RPM1 and RPS5 of *Arabidopsis*), the sequence identity ranged from 19.53% (between CnRGC-III and RPM1) to 43.98% (between CnRGC-V and RPS5) (Table 3). These percent identities were similar to those observed between RGCs or resistance genes from other plant species [26,30,32].

2.4.2. PHYLOGENETIC RELATIONSHIPS OF CnRGCs WITH TIR AND NON-TIR-NBS SEQUENCES

A phylogenetic tree was constructed based on the deduced amino acid sequences of CnRGC-I to VII, including TIR and non-TIR-NBS sequences of known *R* genes reported for other species. The analysis grouped the coconut sequences with R proteins of the non-TIR-NBS-LRR subclass (Fig. 2.4).

The TIR and non-TIR-NBS clades were strongly supported on the basis of the bootstrap analysis. The distribution of the coconut palm RGCs in the different branches of the non-TIR-NBS clade showed a high level of divergence for these sequences. CnRGC-I and II were grouped in a large clade with *R* genes such as *Lr1*, *XA1*, *BPH14*, *Pm3*, *Rdg2* and *RPG1-B*. Other coconut sequences (KC465244, KF002584 and FJ666396) were grouped in the same clade, but presented higher homology with different *R* genes than CnRGCs characterized in this study (Fig. 2.4). The sequences CnRGC-III, IV and V were found in the same clade and have a phylogenetic relationship with the resistance genes *RPS5* and *Rps2* of *Arabidopsis thaliana* involved in the recognition of *Pseudomonas syringae* effectors

(avrPphB and avrRpt2) [46,40,47] (Fig. 2.4). The sequences CnRGC-VI and VII were grouped in two small clades and presented homology with the resistance genes *Pit* and *Pid*.

Table 3
Percentage identity derived from pairwise comparisons between isolated coconut resistance gene candidates and the NBS domain of NBS-LRR resistance genes.

Name	CnRGC I	CnRGC II	CnRGC III	CnRGC IV	CnRGC V	CnRGC VI	CnRGC VII	KF002584	KC465244	FJ666396	RPP8	RPM1	RPS5
CnRGC I	100.00												
CnRGC II		40.00											
CnRGC III			18.34										
CnRGC IV				26.09									
CnRGC V					29.63								
CnRGC VI						39.88							
CnRGC VII							33.33						
KF002584								37.11					
KC465244									39.88				
FJ666396										33.75			
RPP8											24.85		
RPM1												33.14	
RPS5													25.90

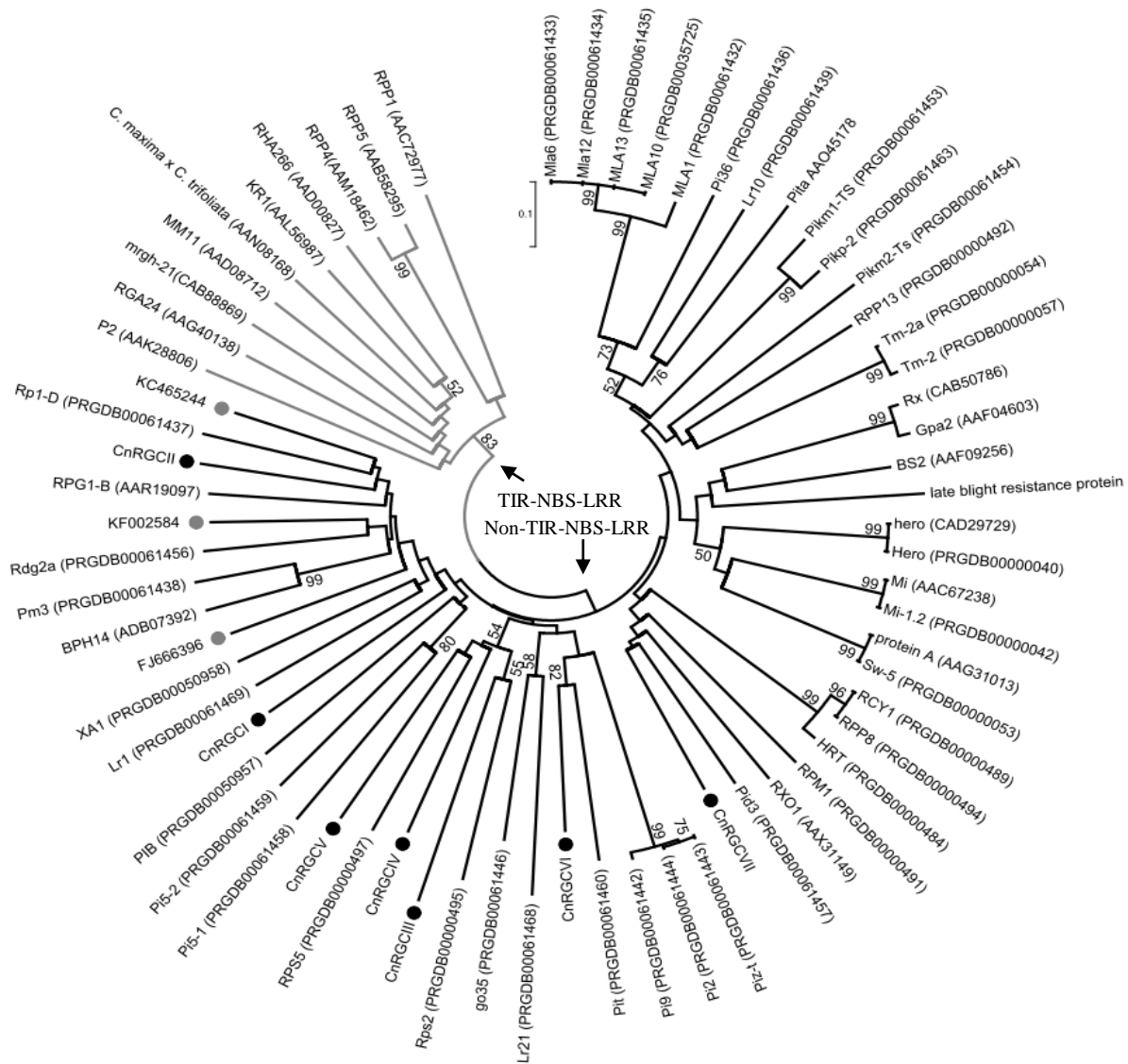


Fig. 2.4. Phylogenetic analysis based on alignment of the deduced amino acid sequences of coconut palm with known *R* genes. The tree was constructed by the neighbor-joining method using MEGA 5.0 software. Number on the branches indicate the percentages of 1000 bootstrap replications and only those with >50% support are shown. Dots indicate coconut palm sequences isolated in this study and the dot in grey the coconut palm sequences registered in NCBI.

2.4.3. EXPRESSION PROFILES OF RGCS IN PLANTLET TREATED WITH SALICYLIC ACID

The gene expression levels of six CnRGCs (I, II, IV, V, VI and VII) were determined in coconut plantlets treated and untreated with SA. Analyses were carried out at 0, 24, 48, 72, 96 and 120 h after treatment in root, stem and leaf tissues. The gene expression profile of CnRGC-III could not be determined because the primers used did not generate a single PCR product.

The gene expression for CnRGC-I, II, IV, V, VI and VII in leaf, stem and root tissues of plantlets not treated with SA, showed for each CnRGC that there was a basal level of expression at time 0 that did not change significantly at any other time of evaluation within the 120 h period of the study (Fig. 2.5A, B, C, D, E, F). In contrast, in the case of plantlets treated with 2.5 mM SA, the level of expression did change for each of the CnRGC studied. In the case of leaf, we studied gene expression for CnRGC-I, II, IV, V, VI and VII. In the cases of stem and root we chose three of these, CnRGC-II, VI and VII. Taken together different expression profiles were observed. One profile was observed in CnRGC-IV in leaf in which there was practically no change with the exception of a small increment at 96 h going back to the basal level (Fig. 2.5C). A second profile (CnRGC-I and II, leaf) was a two-fold increment that peaked at 48 and 72 h, decreasing afterwards (Fig. 2.5A, B). A third profile showed two peaks of activity, one of them being two-fold or greater, with the second peak appearing at the end of the evaluation period at 120 h (CnRGC-VI and VII, leaf, Fig. 2.5E, F; CnRGC-VI and VII, stem, Fig. 2.5H, I; CnRGC-II and VII, root, Fig. 2.5J, L) and two more with the largest peaks of activity observed with 3-9-fold increments (CnRGC-VI and VII, stem, Fig. 2.5H, I). A fourth profile showed two peaks of activity, each one being two-fold or greater, with the second peak decreasing at the end of the evaluation period (CnRGC-V, leaf, Fig. 2.5D; CnRGC-VI, root, Fig. 2.5K). A fifth and last profile showed three peaks of activity with 1.5-2.3-fold increments (CnRGC-II, stem, Fig. 2.5G).

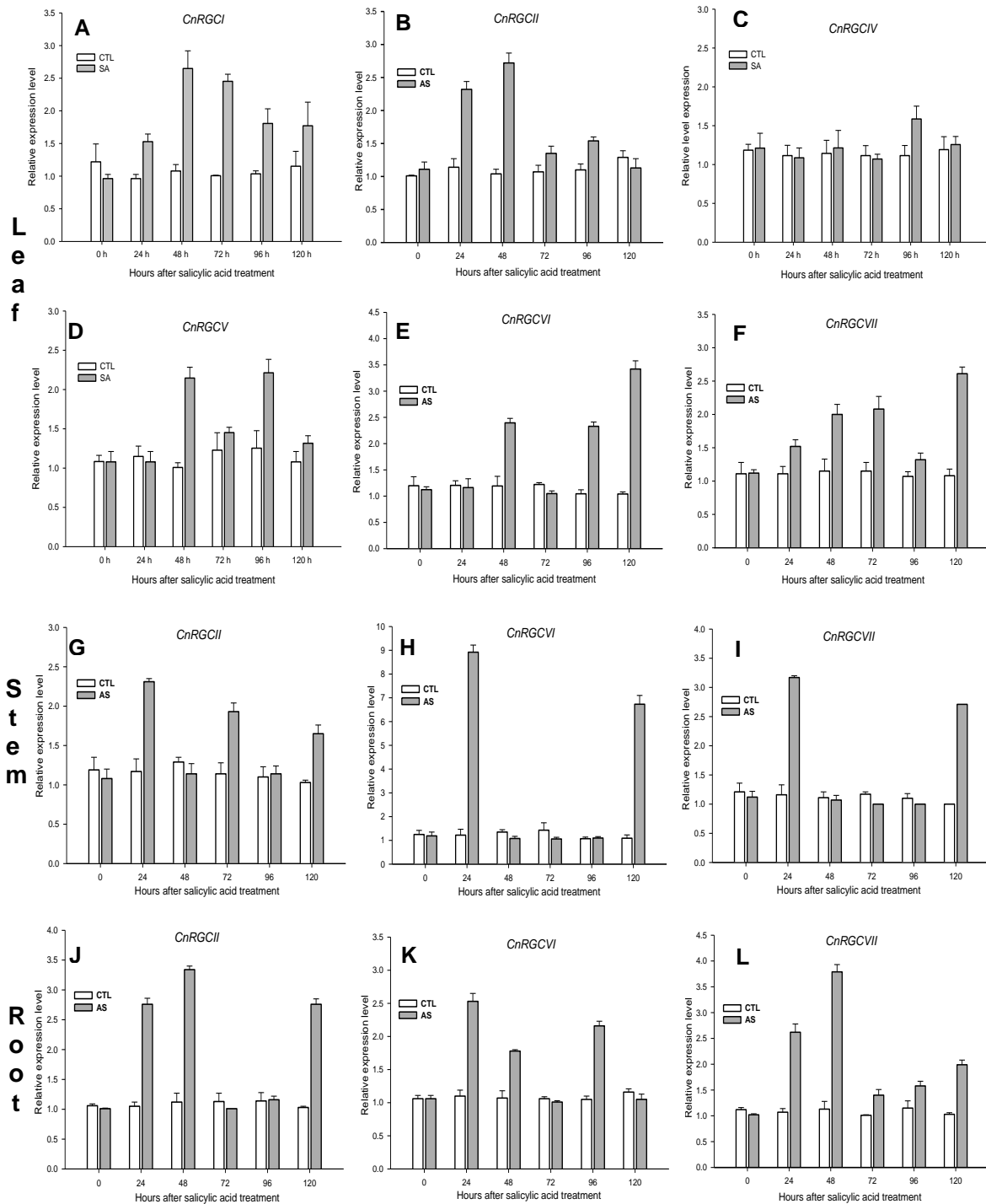


Fig. 2.5. Expression profiles of six CnRGCs. The gene expression was determined by real-time PCR in different tissues from plantlets of 12 months of in vitro culture treated and non-treated with salicylic acid at different times. 18S was used as a reference gene.

3.4.6 EXPRESSION ANALYSIS OF FOUR NBS SEQUENCES OF COCONUT PALMS INFECTED WITH LY-PHYTOPLASMA

With the objective to determine the expression profiles of RGCs during a compatible interaction palm-phytoplasma, four NBS sequences previously reported by our group¹¹ were selected (CnRGCI, CnRGCII, CnRGCVI and CnRGCVII). Symptomatic palms showing disease stages 1 and 2 and asymptomatic palms were collected in the locality of San Crisanto, Yucatan, Mexico. To confirm whether these palms were infected, extracts from trunks were analyzed using LY 16S TaqMan/Real-time PCR assay³³. The three symptomatic palms tested positive for the LY-phytoplasma and the asymptomatics tested negative (Table 4).

Table 4. Detection of LY-phytoplasma DNA from leaf tissues of asymptomatic and symptomatic of coconut palms.

Asymptomatic	LY 16S TaqMan/Real-time PCR assay	Symptomatic	LY 16S TaqMan/Real-time PCR assay
Palm1	ND	Palm1	20.82 (positive)
Palm2	ND	Palm2	22.32 (positive)
Palm3	ND	Palm3	20.60 (positive)

ND: Not detected

The expression profiles of four genes were evaluated in the three different tissues (leaf, stem and root) of symptomatic and asymptomatic coconut palm. The CnRGCVI was highly expressed in the stem tissue of diseased coconut palm while in healthy palm no changes were observed in the three tissues analyzed. The same pattern was observed in the three palms evaluated. CnRGCVII presented the highest level of expression in root tissue of LY-affected coconut palm, while in healthy palms only a minor change in the expression was observed in stem tissue. CnRGCII increased its expression level in leaf tissue of LY-affected coconut palm with respect to healthy palms. However in healthy palms CnRGCII showed a similar pattern all the three tissues analyzed. Interestingly CnRGCI was highly expressed in root of healthy palms, the same pattern was observed in the three coconut palms analyzed.

However, in LY-affected palms CnRGCI modifies its expression profile and was highly expressed in leaf and stem, but not in root. This profile was observed in two coconut palms for leaf and three for stem tissue. No changes were observed for *In vitro* plantlets in the expression for these four sequences (Figure 2.6). These results suggest that during palm-phytoplasma interaction occur changes in the expression profiles of these sequences in the different tissues analyzed.

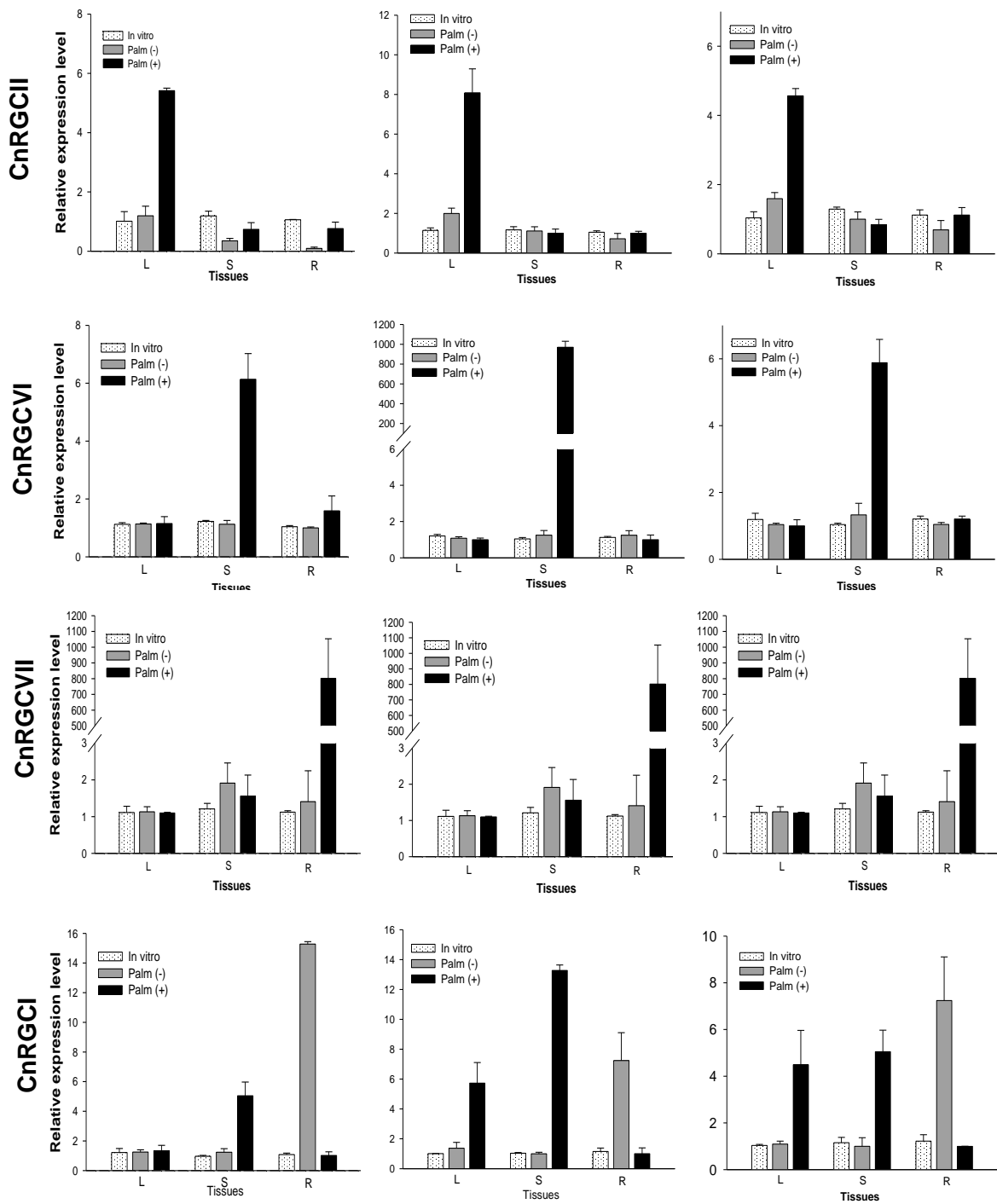


Figure 2.6. Relative transcript levels of four CnRGCs obtained from quantitative real-time PCR experiments in different tissues of healthy coconut palm and affected with lethal yellowing disease. Black bars: LY- coconut palms, gray bars: healthy coconut palm and white bars: *in vitro* coconut palm plantlets. R: Root, S: Stem and L: Leaf. 18S rRNA was used as a reference gene for the expression analysis in coconut palm. The analysis was conducted in three independent palms for each condition, of each sample by triplicate.

2.5 DISCUSSION

Coconut palms in the Americas have been extensively affected by the phytoplasma-associated disease LY. In order to deal with it, resistance ecotypes have been successfully identified and are currently being used for replanting programmes in Mexico [38]. However, LY resistance screening in coconut is a process that takes several years, is very expensive and is labor intensive. So, for this purpose and breeding programmes in general, it is very important to develop alternative approaches, therefore, in this study we report the first large-scale analysis of RGCs of the NBS type from three coconut ecotypes.

Using degenerated primers designed from conserved regions of the NBS domain of *R* genes previously described for other species [17, 48-50, 31], we could amplify RGCs from the genomic DNA from three coconut ecotypes, two LY-resistant and one LY-susceptible. A total of 169 sequences showed homology with *R* genes and RGCs from the NCBI database and 143 presented open reading frames. These could be potentially expressed as protein and might represent *R* genes functional against various pathogens and therefore were designated as CnRGCs. The rest with premature stop codons are probably derived from pseudogenes, probably arising by point mutations, insertions, deletions or substitutions of nucleotides [51, 52]. Pseudogenes with NBS-LRR motifs have been reported in *A. thaliana* from 149 NBS-LRR genes identified, 10 were pseudogenes [34] and in *Oryza sativa* L., variety *indica* from 653 genes NBS-LRR, 184 were pseudogenes [35].

Multiple sequence alignment of the deduced amino acid of CnRGC-I to CnRGC-VII with known TIR and non-TIR *R* genes revealed the presence in these coconut sequences of the NBS characteristic motifs: the P-loop, RNBS-A, kinase-2, RNBS-B and GLPL hydrophobic motifs. The presence of these suggests that these sequences are related to the NBS-LRR class of *R* genes [53-56]. Moreover, this analysis indicated that CnRGCs are related to the subclass of non-TIR-NBS-LRR of *R* genes, presenting a residue of tryptophan (W) in the final position of kinase-2 in 6 of the 7 residue of CnRGCs instead of an aspartic acid (D), which is characteristic of the TIR-NBS-LRR [14]. Phylogenetic analyses also classified CnRGCs within the subclass non-TIR-NBS-LRR and presented homology with *R* genes belonging to the same subclass. So far, the TIR domain has not been found in the structure of NBS-LRR genes of monocot plants, for example, the TIR domain is absent in the rice genome sequence [15, 57, 35] and has not been found in other monocot plants [30, 31, 58].

Our results are in agreement with the hypothesis that the TIR subclass is restricted to dicotyledonous species. This domain is believed to have been lost in NBS-LRR genes in monocots during the divergence between monocots and dicots [59]. Despite the relatively large amount of RGCs that were isolated from coconut, we still do not know the total number of RGCs in its genome. Comparative genomic analyses indicate that there are large differences in the number of *R* genes in plant species and that the number of these neither increase nor decrease in proportion to genome size [32]. For instance, a high number of NBS-LRR genes has been found both in monocot and dicot species, for example in *Oryza sativa* L. spp. Indica there are 653 NBS-LRR genes [35] and in *Vitis vinifera* there are 459 NBS-LRR genes [60]. Whereas there are other species with a lower number of NBS-LRR genes, for instance in *Carica papaya* and *Cucumis sativus* there are 54 and 57 NBS-LRR genes respectively [36, 32]. In the case of palm species, there is a report for date palm (*Phoenix dactylifera*) with 144 NBS-LRR genes found in its genome sequence [37]. This number is very close to the one found for coconut in the present study.

Although, the number of *R* genes that have been cloned in plants over the past two decades has increased, little is known about how their expression is regulated. Apparently, the expression of *R* genes of the NBS-LRR type occurs in a constitutive manner, as observed for various *R* genes such as *RPM1* [41], *Prf* [61], *RPP5* [43], *Mi* [62], *I2* [63] and other studies that have revealed the presence of low levels of transcripts in plants that have not been infected with a pathogen or treated with other stimuli. Similarly, in coconut plantlets that were not treated with any stimulus, a constitutive or basal expression was observed for six CnRGCs at any time within the 120 h period of the study. These findings are consistent with the postulated role of *R* proteins that act as preformed receptors that recognize pathogen effectors in a gene by gene interaction that leads to disease resistance [64].

SA is a signaling factor acting as a second messenger in systemic acquired resistance [65, 66]. After treatment of seedlings with SA, we observed that five out of six CnRGCs (CnRGC-I, II, V, VI and VII), showed five different patterns of expression, one with no change and the other four showing one or more peaks. Direct comparison between different tissues, leaf, stem and root tissues, could be done only for CnRGC-VI and VII, in these cases patterns two patterns were observed, one for each CnRGC. Overall, these differences in expression patterns of the different CnRGCs occurring in the different tissues evaluated, might be indicative of a differential transcriptional regulation due to SA treatment. An observation that

should be further studied. Increasing RGC expression in response to SA has also been reported in *Capsicum annuum* [32] and *Saccharum officinarum* [67]. These findings also support a role of SA as critical signal that regulates the upstream genes during plant-pathogen interaction besides its role in the downstream gene regulation [68, 69, 67, 32].

The NBS-LRR resistance genes are involved in conferring resistance to a wide variety of pathogens and pests including viruses, bacteria, fungi, nematodes and insects.⁵⁸ However, to our knowledge there are no reports of *R* genes conferring resistance to phytoplasma even when phytoplasma effectors in plants have been characterized.⁵⁹ To evaluate the possible role of these genes in response to LY-phytoplasma, we selected four sequences and the gene profile expression was analyzed in healthy palms and those palms infected with LY-phytoplasma. The results showed that in diseased palms all four genes modified their gene expression at tissue level (CnRGCI in leaf, CnRGCVI in stem, CnRGCVII in root, and CnRGCVIII in leaf and stem) with respect to healthy palms. Moreover the tissue specific expression profile could be associated with pathogen distribution as has been reported by Córdova *et al.*³³ Interestingly CnRGCI was found to be differentially expressed in the root tissue of coconut healthy palms with respect to the diseased palms. These results indicate that changes occur in the expression profile of these genes during interaction palm-phytoplasma that probably is not induced at the indicated time to recognize the pathogen and activate defense mechanisms. Moreover, the expression of CnRGCI at the root of healthy palms comparing to diseased ones, may be associated with an early response to infection phytoplasma AL, because the root is one of the first tissue that is affected during the disease. However, this expression could also result from the presence of another pathogen in the root.

RGCs found in different plant species have been associated with resistance to plant pathogens, for instance, R2 and R6 in *Capsicum annuum* were present in varieties with resistance against *Phytophthora capsici* [70], RGC2 from *Musa acuminata ssp. malaccensis* was associated with resistance against *Fusarium oxysporum* [30]; Glyma01g04000, Glyma05g09440 and Glyma06g40980 genes whose expression was associated with resistance to different pathogens such as *Phakopsora pachyrhizi*, *Sclerotinia sclerotiorum* and *Bradyrhizobium japonicum* [71].

Coconut palms can be affected by pathogens such as viruses [2], viroides [3], protozoa [4], oomycetes [5], nematodes [6] and mollicutes as the 'Candidatus Phytoplasma palme' responsible for the devastating disease known as lethal yellowing in America [8], therefore it is possible that the CnRGCs isolated in this study, could be associated with resistance to this pathogens by recognizing pathogen effectors. Recent studies on phytoplasmas revealed the presence of effectors in sequenced genomes of Aster Yellows phytoplasma, wherein the effector protein SAP11 (WB secreted protein AY-11) modulates the defense response in plants providing benefits to the insect vector (*Macrosteleles quadrilineatus*), increasing their offspring [72, 73]. It is possible that resistant genes can recognize these effectors and play an important role in the resistance to these pathogens, or perhaps in the vectors that transmit them. NBS-LRR *R* genes against insects have already been reported [74-76].

2.6 CONCLUSION

This study is the first large-scale analysis of NBS-LRR sequences in coconut, reporting the finding of 143 CnRGCs sequences grouped in seven clades belonging to non-TIR-NBS-LRR subfamily based on phylogenetic tree and sequence analyses. Knowledge of the existence of these contributes to the understanding of disease defense mechanisms in this palm species (and probably others) and also it might be useful for future molecular breeding programs aimed to develop disease resistance to different pathogens, including LY, in current coconut palm cultivars of socio-economic importance.

2.7. APPENDIX A. SUPPLEMENTARY DATA

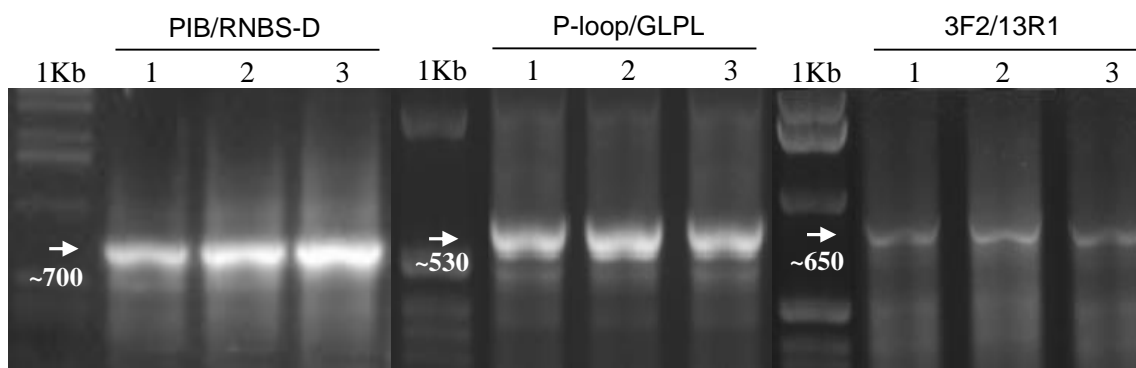
Additional file 1. Table A1. Degenerate oligonucleotide primers used to amplify resistance gene candidates from coconut.

Primer	Sequence	Combination	Reference
PIA (forward)	GGIATGCCIGGIIIIGGIAARACIAC		Bertioli <i>et al.</i> , 2003
PIB (forward)	GGIATGGGIGGIIIIGGIAARACIAC	PIA/P3A	Bertioli <i>et al.</i> , 2003
P-loop(forward)	GGIGGIGTIGGIAAIACIAC	PIA/RNBS-D	Kanazin <i>et al.</i> , 1996
3F2 (forward)	GAGGTACTTCCTGGTGCTGGAYGAYRTBTGG	PIB-P3A	Miller <i>et al.</i> 2008.
GLPLA(reverse)	A(A/G)IGCTA(A/G)IGGIA(A/G)ICC	PIB/RNBS-D	Aarts <i>et al.</i> , 1998.
P3A (reverse)	AIITYIRIIRYIAGIGGYAAICC	P-loop/GLPL	Bertioli <i>et al.</i> , 2003
RNBS-D(reverse)	GGRAAIARISHRCARTAIVIRAARC	3F2/13R1	Penuela <i>et al.</i> , 2002
13R1 (reverse)	CGGCAAGTCGTGCAYVAKRTCRTGCA		Miller <i>et al.</i> 2008

Codes for degenerate positions: **I** = inosine; **R** = A/G; **Y** = C/T; **M** = A/C; **K** = G/T; **W** = A/T; **S**=C/G; **B** = C/G/T; **D** = A/G/T; **H** = A/C/T; **V** = A/C/G; **N** = A/C/G/T.

Additional file 1. Table A2. Primer sequences used to determine the expression profiles of CnRGCs from coconut.

Primer	Sequence	Product (bp)
CnRGC-I	F: 5'-ACTCCATCCTCTTCCCGCAGTTGT-3' R: 5'-AGAGAATGTCCCCACTACCCAGA-3'	210
CnRGC-II	F: 5'-5'-TCCCACCTCTCCTCAGCTGTTTC-3' R: 5'-5'-CCTTGACTGTTTACCAGAAGACCA-3'	192
CnRGC-III	F: 5'-GATGGATATGCCTTCATTGCTTTC-3' R: 5'-CGTAGAGCTAAGAGAGAGATCGACGAAA-3'	176
CnRGC-IV	F: 5'-GGATCCCAACAACAGTTTTCA-3' R: 5'-CAGAATCCAAACATAGAGCAGA-3'	181
CnRGC-V	F: 5'-GCATCCGTAGTGCATTTTCC-3' R: 5'-AGGCAGGTAGTGCATTTCTT-3'	174
CnRGC-VI	F: 5'-TTGCTTCTCTCCTTCGACCTTAAAACC-3' R: 5'-TATTAGAGCGAAGATCCACCAAGTC-3'	210
CnRGC-VII	F: 5'-CTTTCGGACATTGGCATTTTAGTC-3' R: 5'-CTAGCAGTTGAAAGCCGTGTTCTGG-3'	278



Additional file 1. Figure A1. PCR products amplified with three pairs of degenerate primers (PIB/RNBS-D, P-loop/GLPL and 3F2/13R1) using the genomic DNA mix of each ecotype. Lane 1: genomic DNA mix from resistant Mexican Pacific Tall 2 (MXPT2), lane 2: genomic DNA mix from the resistant Malayan Yellow Dwarf (MYD) and lane 3: genomic DNA mix from the susceptible Mexican Atlantic Tall (MXAT). The 1 Kb ladder was used as size markers. The arrow indicates the band cloned.

Additional file 1. Table A3. Results of similarity searches between the amplified products by Real-Time PCR and GenBank accessions using *blastx*.

Name	Size (bp)	% Identity with the original sequence	Highest similarity with GenBank accessions	Query cover (%)	E-value	Identity (%)
CnRGCI	290	99	Putative CC-NBS-LRR disease resistance protein (<i>Rubus sp.</i>).	82	4e-11	40
			NBS-LRR resistance protein (<i>Populus trichocarpa</i>).	82	5e-11	41
CnRGCI	152	98	NBS coding resistance gene protein (<i>Nicotiana tabacum</i>).	86	7e-10	57
			I2, NBS domain region (<i>Nicotiana tabacum</i>).	86	1e-09	57
CnRGCI	182	99	Putative NBS-LRR protein, partial (<i>Pleomele marginata</i>)	93	3e-15	63
			Putative NBS-LRR protein, partial (<i>Sansevieria trifasciata</i>)	93	7e-13	59
CnRGCV	109	98	Disease resistance protein At4g27190-like, partial (<i>Cicer arietinum</i>).	71	3e-04	69
			Disease resistance protein At4g27190-like (<i>Vitis vinifera</i>).	90	7e-04	61
CnRGCVI	205	98	NBS-LRR disease resistance protein (<i>Musa acuminata</i> subsp. <i>burmannica</i>).	92	1e-20	70
			NBS-LRR resistance protein (<i>Musa</i> AAB Group).	95	2e-17	60
CnRGCVII	295	100	Disease resistance protein RPM1-like (<i>Vitis vinifera</i>).	100	9e-22	46
			Disease resistance protein RPM1 (<i>Populus trichocarpa</i>).	100	1e-21	47

2.8 REFERENCES

- [1]. Foale M: An introduction to the coconut palm. In: *Coconut Genetic Resources*. Edited by Batugal P, Ramanatha V, Rao GP, Oliver J. International Plant Genetic Resources Institute – Regional Office for Asia, the Pacific and Oceania (IPGRI-APO), Serdang, Selangor DE, Malaysia; 2005:1-8.
- [2]. Rohde W, Randles JW, Langridge P, Hanold D. Nucleotide sequence of a circular single stranded DNA associated with coconut foliar decay virus. *Virology* 1990; 176:648–51.
- [3]. Hanold D, Randles JW. Coconut cadang-cadang disease and its viroid agent. *Plant Dis* 1991; 75:330–5.
- [4]. Parthasarathy MV, van Slobbe WG, Soudant C. Hartrot or fatal wilt of palms. 1. Coconuts (*Cocos nucifera*). *Principes* 1978; 22:3–14.
- [5]. Joseph T, Radha K. Role of *Phytophthora palmivora* in bud rot of coconut. *Plant Dis Repr* 1975; 59:1014–7.
- [6]. Griffith R. Red ring disease of coconut palm. *Plant Dis* 1987; 71:193–6.
- [7]. Howard FW, Barrant CI. Questions and answers about lethal yellowing disease. *Principes* 1989; 33:163–71.
- [8]. Harrison NA, Oropeza C. Coconut lethal yellowing. In: *Characterization, Diagnosis and Management of Phytoplasmas*. Edited by Harrison NA, Rao GP, Marcone C. Houston TX, USA: Studium Press LLC; 2008: 219-248.
- [9]. Oropeza C, Zizumbo-Villarreal D. The history of lethal yellowing in Mexico. In: *Proceedings of the international workshop on lethal yellowing-like diseases of coconut*. Edited by Eden-Green SJ, Ofori F. Ghana: Elmina 1997:69–76.
- [10]. Dangl L, Jones J. Plant pathogens and integrated defence responses to infection. *Nature* 2001; 411:826–833.
- [11]. Belkhadir Y, Subramaniam R, Dangl JL. Plant disease resistance protein signalling: NBS-LRR proteins and their partners. *Curr Opin Plant Biol* 2004; 7:391–399.
- [12]. Tan S, Wu S. Genome wide analysis of nucleotide-binding site disease resistance genes in *Brachypodium distachyon*. *Comp. Funct. Genomics* 2012;doi:10.1155/2012/418208.

- [13]. Yue JX, Meyers BC, Chen JQ, Tian DC, Yang SH. Tracing the origin and evolutionary history of plant nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes. *New Phytol* 2012; 193:1049–1063.
- [14]. Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J* 1999; 20:317-332.
- [15]. Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Leach JE, Hulbert SH. Diversity in nucleotide binding site-leucine rich repeat genes in cereals. *Genome Res* 2002; 12:1871–1884.
- [16]. Meyers BC, Morgante M, Michelmore RW. TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. *Plant J* 2002; 32:77-92.
- [17]. Kanazin V, Mareck L, Shoemaker R. Resistance gene analogs are conserved and clustered in soybean. *Proc Natl Acad Sci USA* 1996; 93:11746–11750.
- [18]. Yu Y, Buss G, Maroof M. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proc Natl Acad Sci USA* 1996; 93:11751–11756.
- [19]. Graham M, Mareck L, Lohnes D, Cregan P, Shoemaker R. Expression and genome organization of resistance gene analogues in soybean. *Genome* 2000; 43:86–93.
- [20]. Leister D, Ballvora A, Salamini F, Gebhardt C. A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat Genet* 1996; 14:421–429.
- [21]. Shen K, Meyers B, Islam-Faridi M, Chin D, Stelly D, Michelmore R. Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. *Mol Plant Microbe Interact* 1998; 11:815–823.
- [22]. Leister D, Kurth J, Laurie D, Yano M, Sasaki T, Devos K, Graner A, Schulze-Lefert P. Rapid organization of resistance gene homologues in cereal genomes. *Proc Natl Acad Sci USA* 1998; 95:370–375.
- [23]. Seah S, Spielmeyer W, Jahier J, Sivasithamparam K, Lagudah S. Resistance gene analogs within and introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat. *Mol Plant Microbe Interact* 2000; 13:334–341.

-
- [24]. Ayele-Gedil M, Slabaugh MB, Berry S, Johnson R, Michelmore R, Miller J, Gulya T, Knapp S. Candidate disease resistance genes in sunflower cloned using conserved nucleotide-binding site motifs: genetic mapping and linkage to the downy mildew resistance gene PI1. *Genome* 2001;44: 205–212.
- [25]. Rivkin M, Vallejos C, McClean P. Disease-resistance related sequences in common bean. *Genome* 1999; 42:41–47.
- [26]. López C, Acosta I, Jara C, Pedraza F, Gaitan-Solis E, Gallego G, Beebe S, Tohme J. Identifying resistance gene analogs associated with resistance to different pathogens in common bean. *Phytopathology* 2003; 93:88–95.
- [27]. Martínez-Zamora MG, Castagnaro AP, Díaz-Ricci JC. Isolation and diversity analysis of resistance gene analogues (RGAs) from cultivated and wild strawberries. *Mol Genet Genomics* 2004; 272:480–487.
- [28]. Calenge F, Van der Linden CG, Van de Weg E, Schouten HJ, Van Arkel G, Denance C, Durel CE. Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. *Theor Appl Genet* 2005; 110:660–668.
- [29]. Pei X, Li S, Jiang Y, Wang Z, Jia S. Isolation, characterization and phylogenetic analysis of the resistance gene analogues (RGAs) in banana (*Musa* spp.). *Plant Sci* 2007; 172:1166–1174.
- [30]. Peraza-Echeverria S, James-Kay A, Canto-Canché B, Castillo-Castro E. Structural and phylogenetic analysis of Pto-type disease resistance gene candidates in banana. *Mol Genet Genomics* 2008; 278:443–453.
- [31]. Miller R, Bertoli D, Baurens F, Santos C, Alves P, Martins N, Togawa R, Souza C, Pappas G. Analysis of non-TIR NBS-LRR resistance gene analogs in *Musa acuminata* Colla: Isolation, RFLP marker development, and physical mapping. *BMC Plant Biology* 2008; 8: 1-15.
- [32]. Wan H, Yuan W, Ye Q, Wang R, Ruan M, Li Z, Zhou G, Yao Z, Zhao J, Liu S, Yang Y. Analysis of TIR- and non-TIR-NBS-LRR disease resistance gene analogs in pepper: characterization, genetic variation, functional divergence and expression patterns. *BMC Genomics* 2012; 13: 1-15.
- [33]. Saha D, Rana RS, Sureja AK, Verma M, Arya L, Munshi AD. Cloning and characterization of NBS-LRR encoding resistance gene candidates from Tomato Leaf

-
- Curl New Delhi Virus resistant genotype of *Luffa cylindrica* Roem. *Physiological and Molecular Plant Pathology* 2013; 81; 107-117.
- [34]. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* 2003; 15:809–834.
- [35]. Shang J, Tao Y, Chen X, Zou Y, Lei C, Wang J, Li X, Zhao X, Zhang M, Lu Z, et al. Identification of a new rice blast resistance gene, *Pid3*, by genome wide comparison of paired nucleotide-binding site leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. *Genetics* 2009; 182, 1303–1311.
- [36]. Porter BW, Paidi M, Ming R, Alam M, Nishijima WT, Zhu YJ. Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. *Mol Genet Genomic* 2009; 281:609–626.
- [37]. Al-Mssallem IS, Hu S, Zhang X, Lin Q, Liu W, Tan J, Yu X, Liu J, Pan L, Zhang T, Yin Y, Xin C, Wu H, Zhang G, Ba Abdullah MM, Huang D, Fang Y, Alnakhli1 YO, Jia S, Yin A, Alhuzimi EM, Alsaihati BA, Al-Owayyed SA, Zhao D, Zhang S, Al-Otaibi NA, Sun G, Majrashi MA, Li F, Tala, Wang J, Yun Q, Alnassar NA, Wang L, Yang M, Al-Jelaify RF, Liu K, Gao S, Chen K, Alkhaldi SR, Liu G, Zhang M, Guo H and Yu J: Genome sequence of the date palm *Phoenix dactylifera* L. *Nature Communications* 2013;4:2274.
- [38]. Zizumbo-Villareal D, Colunga-García M, Fernández-Barrera M, Torres Hernández N, Oropeza-Salín C. Mortality of Mexican coconut germplasm due to lethal yellowing. *Plant Genetic Resources Newsletter*, 2008; 156: 23-33.
- [39]. Thompson J, Gibson T, Plewniak F, Jeanmougin F, Higgins D. The clustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids Res* 1997; 24:4876-4882.
- [40]. Warren RF, Henk A, Mowery P, Holub E, Innes RW. A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. *Plant Cell* 1998; 10:1439-1452.
- [41]. Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes R, Dangl JL. Structure of the *Arabidopsis* *RPM1* gene enabling dual specificity disease resistance. *Science* 1995; 269:843–846.
- [42]. McDowell JM, Dhandaydham M, Long TA, Aarts MGM, Goff S, Holub EB, Dangl JL. Intragenic recombination and diversifying selection contribute to the evolution of

- downy mildew resistance at the RPP8 locus of *Arabidopsis*. *Plant Cell* 1998; 10:1861–1874.
- [43]. Parker JE, Coleman MJ, Szabo V, Frost LN, Schmidt R, van Der Biezen EA, Moores T, Dean C, Daniels MJ, Jones JD. The *Arabidopsis* downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *Plant Cell* 1997; 9:879-894.
- [44]. Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B. The product of the tobacco mosaic virus resistance gene N: similarity to Toll and the Interleukin-1 receptor. *Cell* 1994; 78:1101–1115.
- [45]. Lawrence GJ, Finnegan EJ, Ayliffe MA, Ellis JG. The L6 gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene RPS2 and the tobacco viral resistance gene N. *Plant Cell* 1995; 7:1195–1206.
- [46]. Kunkel BN, Bent AF, Dahlbeck D, Innes RW, Staskawicz BJ. RPS2, an *Arabidopsis* disease resistance locus specifying recognition of *Pseudomonas syringae* strains expressing the avirulence gene *avrRpt2*. *Plant Cell* 1993; 5:865–875.
- [47]. Shao F, Golstein C, Ade J, Stoutemyer M, Dixon JE, Innes RW. Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. *Science* 2003; 301: 1230-1233.
- [48]. Aarts M, te Lintel H, Holub E, Beynon J, Stiekema W, and Pereira A. Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Mol Plant Microbe Interact* 1998; 11:251-258.
- [49]. Peñuela S, Danesh D, Young ND. Targeted isolation, sequence analysis, and physical mapping of non-TIR NBS-LRR genes in soybean. *Theor Appl Genet* 2002; 104:261-272.
- [50]. Bertioli D, Leal-Bertioli S, Lion M, Santos V, Pappas G, Cannon S, Guimareas P. A large scale analysis of resistance gene homologues in *Arachis*. *Mol Genet Genomics* 2003; 270: 34-45.
- [51]. Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res* 1998; 8:1113-1130.
- [52]. Chandrasekaran, Betrán E. Origins of new genes and pseudogenes. *Nat. Educ.* 2008; 1, 1.
- [53]. Hulbert SH, Webb CA, Smith SM, Sun Q. Resistance gene complexes: evolution and utilization. *Annu Rev Phytopathol* 2001; 39:285-312.

-
- [54]. McDowell JM, Woffenden BJ. Plant disease resistance genes: recent insights and potential applications. *Trends Biotechnol* 2003; 21:178-83.
- [55]. McHale L, Xiaoping T, Koehl P, Michelmore R. Plant NBS-LRR proteins: adaptable guards. *Genome Biology* 2006; 7: 212.
- [56]. Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW. Plant disease resistance genes: current status and future directions. *Physiol Mol Plant Pathol* 2012; 78:51-65.
- [57]. Zhou T, Wang T, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D. Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol Genet Genomics* 2004; 271:402–415.
- [58]. Tarr DE, Alexander HM. TIR-NBS-LRR genes are rare in monocots: evidence from diverse monocot orders. *BMC Res* 2009; Notes, 2, 197.
- [59]. Pan Q, Wendel J, Fluhr R. Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J Mol Evol* 2000; 50:203-213.
- [60]. Yang S, Zhang X, Yue JX, Tian D, Chen JQ. Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mo. Genet Genomics* 2008; 280, 187–198.
- [61]. Salmeron J, Oldroyd E, Rommens C, Scofield S, Kim H, Lavelle D, Dahlbeck D, Staskawicz B. Tomato Prf is a member of a leucine-rich repeat class of plant disease resistance gene and lies embedded within the Pto kinase gene cluster. *Cell* 1996; 86: 123–133.
- [62]. Milligan S, Bodeau J, Yaghoobi J, Kaloshian I, Zabel P, Williamson V. The root knot resistance gene Mi from tomato is a member of the leucine zipper, nucleotide binding site, leucine-rich repeat family of plant genes. *Plant Cell* 1998; 10:1307–1319
- [63]. Mes J, Van Doorn A, Wijbrandi J, Simons G, Cornelissen B, Haring M. Expression of the Fusarium resistance gene I-2 colocalizes with the site of fungal containment. *Plant J* 2000; 23:183–193
- [64]. Van Der Biezen E, Jones JDG. Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem Sci* 1998; 23:454–456.
- [65]. Ajith A, Srinivasa RU, Choong-Min R, Stacy NA, Li K, Yuhong T, Kirankumar S M, Kumar D, Klessig DF. Salicylic acid and systemic acquired resistance play a role in attenuating crown gall disease caused by *Agrobacterium tumefaciens*. *Plant Physiol* 2008; 146:703–715.

- [66]. Thomma BP, Penninckx IA, Broekaer WF, Cammue BP. The complexity of disease signaling in Arabidopsis. *Curr Opin Immunol* 2011; 13:63–68.
- [67]. Que Y-X, Xu L-P, Lin J-W, Chen R-K. Isolation and Characterization of NBS-LRR Resistance Gene Analogs from Sugarcane. *Acta Agron Sin* 2009; 35, 631–639.
- [68]. Shirano Y, Kachroo P, Shah J, Klessig DF. A gain-of-function mutation in an arabidopsis Toll interleukin1 receptor-nucleotide binding site-leucine-rich repeat type R gene triggers defense responses and results in enhanced disease resistance. *Plant Cell* 2002; 14:3149–3162.
- [69]. Tian AG, Luo GZ, Wang YJ, Zhang JS, Gai JY, Chen SY. Isolation and characterization of a Pti1 homologue from soybean. *J Exp Bot* 2004; 396:535–537.
- [70]. Egea-Gilabert C, Dickinson MJ, Bilotti G, Candela ME. Isolation of resistance gene analogs in pepper using modified AFLPs. *Biol Plant* 2003; 47:27–32
- [71]. Kang YJ, Kim KH, Shim S, Yoon MY, Sun S, Kim MY, Van K, Lee SH. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biology* 2012; 12:139
- [72]. Bai X, Correa VR, Toruno TY, Ammar el D, Kamoun S, Hogenhout SA. AY-WB phytoplasma secretes a protein that targets plant cell nuclei. *Mol Plant-Microbe Interact* 2009; 22:18–30.
- [73]. Sugio A, MacLean AM, Kingdom HN, Grieve VM, Manimekalai R, Hogenhout SA. Diverse Targets of Phytoplasma Effectors: From Plant Development to Defense Against Insects. *Annu Rev Phytopathol* 2011; 49:175–95.
- [74]. Magdalena R, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM. The nematode resistance gene Mi of tomato confers resistance against the potato aphid. *Proc Natl Acad Sci USA* 1998; 95: 9750–9754.
- [75]. Nombela G, Williamson VM, Muñoz M. The Root-Knot Nematode Resistance Gene Mi-1.2 of Tomato Is Responsible for Resistance Against the *Whitefly Bemisia tabaci*. *The American Phytopathological Society* 2003; 7: 645–649.
- [76]. Bo D, Weilin Z, Bingfang L, Jing H, Zhe W, Zhenyig S, Ruifeng H, Lili Z, Rongzhi C, Bin H, Guangcun He. Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proc Natl Acad Sci USA* 2009; 52: 22163-22168.

- [77]. Harrison NA, Richardson PA, Kramer JB, Tsai JH. Detection of the mycoplasma-like organism associated with lethal yellowing disease of palms in Florida by polymerase chain reaction. *Plant Pathology* 1994; 43: 998-1008.
- [78]. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; 25:3389–3402.
- [79]. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4:406–425.
- [80]. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; 28:2731-2739.
- [81]. Rillo EP, Paloma MBF. In vitro culture of Makapuno coconut embryos. *Coconuts Today* 1992; 9: 90–101.
- [82]. Montero-Córtes M, Sáenz L, Córdova I, Quiroz A, Verdeil J-L, Oropeza C. GA₃ stimulate the formation and germination of somatic embryos and the expression of a KNOTTED-like homeobox gene of *Cocos nucifera* (L.). *Plant Cell Rep* 2010; 29(9): 1049-1059. DOI 10.1007/s00299-010-0890-0. ISSN: 0721-7714.
- [83]. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 2001; 25: 402-408.

CAPITULO III

NEW INSIGHTS IN THEIR EVOLUTIONARY HISTORY OF RESISTANCE GENE CANDIDATES (RGCS) IN COCONUT PALM.

3.1 ABSTRACT

BACKGROUND: The family of resistance gene candidates (RGCs) with a nucleotide-binding site (NBS) domain accounts for the largest number of disease resistance genes (*R* genes) and is one of the largest gene families found in plants. These sequences are the most comprehensively studied in terms of sequence evolution and genome distribution. However, in coconut palm and in general in the family of Arecaceae, our understanding of the evolution of NBS-LRR genes is rather limited.

RESULTS: In this study specific phylogenetic clades were found for RGCs of coconut palm, indicating species-specific evolution of resistance genes. However, strikingly similar RGCs between species of Arecales indicates high conservation of specific RGCs of this family, suggesting a monophyletic origin of three genera. The phylogenetic relationship between RGCs of Arecales and Brassicales suggests that these sequences possibly emerged before being divided between monocots and dicots.

CONCLUSION: This study provides new insights for future efforts towards improvement of disease resistance in coconut palm and other species of Arecaceae.

Keywords: resistance gene candidates; coconut palm; lethal yellowing; evolution

3.2 INTRODUCTION

The economic importance of the palm family of plants (Arecaceae or Palmae) is second only to the grass family (Poaceae) among monocotyledons and the third among all plant families after legume (Leguminosae). Among the palm crops, the top three out of the five major domesticated palm species in the world are African oil palm (*Elaeis guineensis*), coconut palm (*Cocos nucifera*) and date palm (*Phoenix dactylifera*) the other two are peach palm (*Bactris gasipaes*) and betel palm (*Areca catechu*) (FAO, <http://www.fao.org>).

Coconut palm is widely distributed throughout the humid tropics where it is cultivated over an estimated area of twelve million ha. The coconut tree is highly versatile, with extensive applications in the fields of agriculture and industry: food, fiber, oil, soil fertilizers, spa ingredients, furniture, fashion accessories, garments, construction and building materials, oleochemicals and biofuels¹. Unfortunately this palm is susceptible to several pathogens that hinder yields including viruses², viroides³, protozoa⁴, oomycetes⁵, nematodes⁶, insect pests⁷ and mollicutes⁸ such as the '*Candidatus Phytoplasma palme*' responsible for the devastating disease known as lethal yellowing (LY) in this Americas⁹. Currently genetic studies focused on the search for genes with biotechnological applications, including those involved in the defense mechanism increasingly begin to gain more relevance to this crop^{1,10, 11,12}. Among the defense related functions in the plant genomes sequenced, there is a group of genetic codes for NBS-LRR (nucleotide-binding sites and leucine-rich repeat) proteins that have been associated with effector-triggered immunity (ETI)¹³. The NBS domain is involved in signaling and includes several highly conserved and strictly ordered motifs such as P-loop, kinase-2 and GLPL motifs¹⁴, which has been demonstrated by the binding and hydrolysis of ATP and GTP. However, the LRR motif is typically involved in protein-protein interactions and ligand binding with pathogen-derived molecules, suggesting that this domain may play a pivotal role in defining pathogen recognition specificity¹⁵. In plants, the NBS-LRR genes have been subdivided into two main groups based on the presence or absence of the N-terminal Toll/interleukin-1 receptor (TIR) homology region^{16,17}. Most of those genes, especially in the monocots which lack the TIR, have a coiled coil (CC) motif in the N-terminal region. Studies have focused on this family of proteins because its only known function is disease resistance¹⁸. The past few years have seen remarkable progress in understanding the mechanisms of *R* gene evolution in Arabidopsis, but notable findings have

also emerged from studies in other plant species such as wheat, sorghum, maize, brachypodium, rice and cotton, among others^{19,20,21,22}.

Comparative genomic analyses has indicated that plant genomes can encode several hundred NBS-LRR genes, and that there is a great diversity in the number and distribution of the subclasses of these genes. Some genomes, such as the genome papaya and *Cucumis sativus* contain about 50 NBS-LRR genes^{23,24}, in contrast to 653 identified in *Oryza sativa*²⁵. The great variation in the total number of NBS-LRR genes might indicate the complex evolutionary pattern in different plant genomes and that they are organized either as isolated genes or as linked clusters of varying sizes that are thought to facilitate rapid *R*-gene evolution²⁶.

In coconut palm and Arecales in general NBS sequences have only focused on the cloning and characterization^{11,27}, while their evolutionary process remains largely unknown. In fact, studies of NBS sequence evolution in coconut palm may shed some light on breeding strategies for disease control in coconut palm crops. Annotation and identification of NBS sequences in *Cocos nucifera*^{11,27}, *Phoenix dactylifera*²⁸ and *Elaeis guineensis*²⁹ will permit comparison of resistance gene analogues between species of the family Arecales and provide insight into their evolutionary relationships. The comparative analysis of NBS genes among Arecales and Poales as monocot plants and Brassicales as dicot plants, can allow us to find evolutionary relationships, common features or specific characteristics of *R* genes that have not been reported in this family.

In this study, several RGCs were retrieved from database NCBI and embryo transcriptome, endosperm and coconut leaf as well as a previous study that we developed¹¹. Characterization of these sequences included identification of protein domains, phylogenetic analyses to investigate their evolutionary relationships and the number of substitutions per synonymous (Ks) and non-synonymous site (Ka) to estimate their evolutionary rates. These investigations might serve as a blueprint for further efforts to clone functional NBS-LRR genes from coconut palm and other species of Arecales, because so far there is no reported resistance gene in this family of plants.

3.3 MATERIAL AND METHODS

3.3.1 STUDY OF NBS SEQUENCES

3.3.1.1 IDENTIFICATION OF NBS SEQUENCES

The NBS sequences used in this study were retrieved from the transcriptome of the embryo (accession number GBGL00000000), endosperm (accession number GBGK00000000) and coconut leaves (accession number GBGM00000000) reported by Huang *et al*,¹⁰ as well as a previous study that we developed.¹¹ The motif sequences were identified using the Multiple Expectation Maximization for Motif Elicitation (MEME) software and the Motif Alignment and Search Tool (MAST) (<http://meme.sdsc.edu/meme/website/intro.html>).

3.3.1.2 PHYLOGENETIC RELATIONSHIPS AND CALCULATION OF KA/KS RATIOS

The NBS sequences were downloaded from the database of NCBI. The NBS domain was used for Phylogenetic analysis. Sequences with less than 50 % of the full-length NBS or that did not present the NBS domain were excluded. Multiple alignments of amino acid sequences were performed using a ClustalW with default options³⁴. Phylogenetic trees were constructed by the neighbor-joining (NJ) method³⁵ with Poisson correction using the NJ algorithm implemented with version 6.0 of the Molecular Evolutionary Genetic Analysis software (MEGA)³⁶. Bootstrap analysis was performed with 1000 replicates to evaluate the reliability of the interior nodes for a particular grouping pattern in the tree.

The ratio of non-synonymous substitution per non-synonymous site and synonymous substitution per synonymous site are here represented by K_a and K_s , respectively. Analysis were calculated using the Nei-Gojobori model³⁷ using version 6.0 of MEGA³⁶. The ratio of non-synonymous to synonymous nucleotide substitution (K/K_s) was detected in the NBS region between sequences of same clade and in sequences of deferent clades. It is widely accepted that a K_a/K_s ratio >1 indicates positive selection, the ratio 1 indicates neutral evolution and <1 purifying selection.

3.4 RESULTS

3.4.1 IDENTIFICATION OF NBS SEQUENCES FROM TRANSCRIPTOMIC SEQUENCES

In a previous study we reported 143 NBS sequences with ORF (open reading frame) from coconut palm and 26 NBS with codon stop¹¹. In this study we identified 62 NBS sequences from embryo, endosperm and leaf tissue transcriptome reported by Huang *et al.*¹⁰ Of the 62 NBS sequences 53 present ORF and 9 codon stop. Of the three tissues, the embryo transcriptome had the maximum total number of NBS transcripts (26 sequences), whereas the leaf transcriptome had the minimum total number of NBS transcripts (17 sequences). However, in the leaf transcriptome was found the largest number of NBS transcripts with ORF (17 sequences) compared with endosperm (15 sequences) (Table 1). The NBS sequences with ORF identified in the transcriptome were used to generate a phylogenetic tree. In this tree we observed that all sequences expressed in the three tissues were different, indicating that gene expression can be tissue specific and could be associated with the site of infection of different pathogens (Figure 3.1).

Table 1. Number of RGCs identified in coconut palm.

Ecotype	Source	Total transcripts	Total of NBS-LRR sequences	# sequences with ORF	# sequences with codons stop	Reference
MYD	Genomic DNA	----	32	28	4	Puch-Hau et al. 2015
MXPT	Genomic DNA	----	116	102	14	
MXAT	Genomic DNA	----	21	13	8	
	Total		169	143	26	
GD	Embryo transcriptome	86, 254	26	21	5	In this study
	Endosperm transcriptome	229, 886	19	14	4	
	Leaf transcriptome	159, 509	17	17	0	
	Total		62	53	9	

MYD: Malayan Yellow Dwarf

MXPT: Mexican Pacific Tall

MXAT: Mexican Atlantic Tall

GD: Green Dwarf

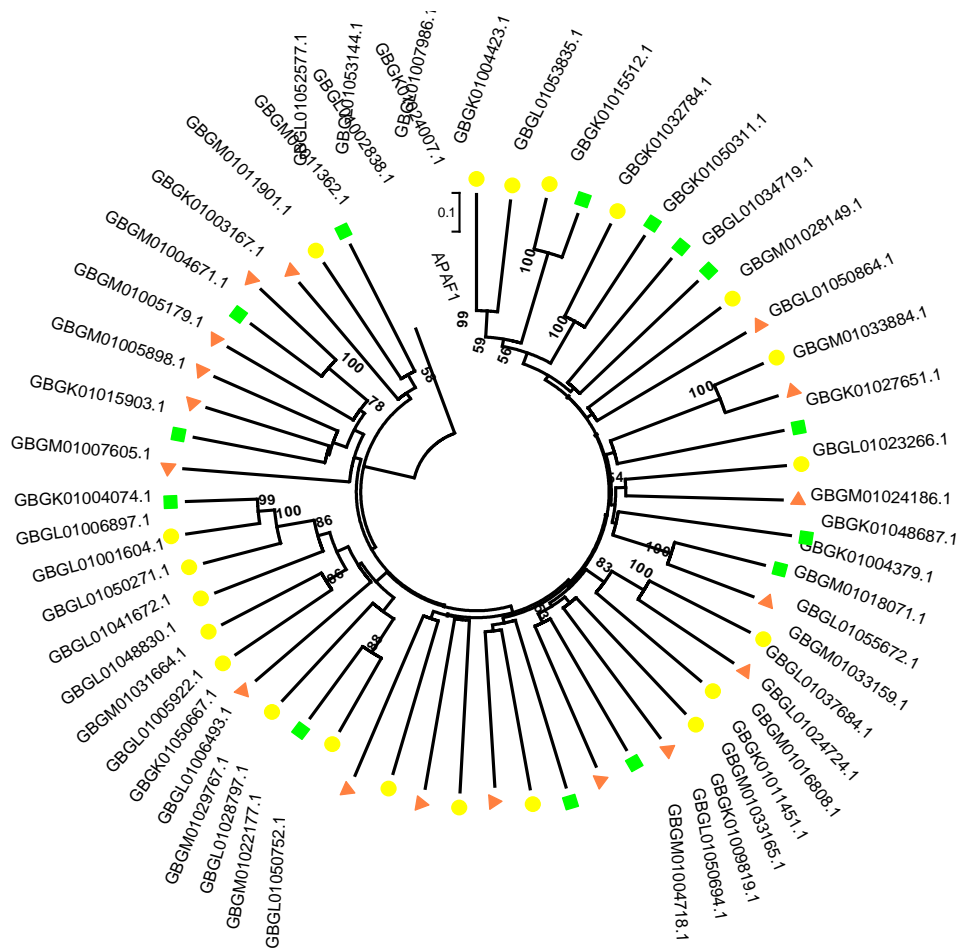


Figure 3.1. Phylogenetic tree of RGAs obtained of three transcriptome of coconut palm (embryo, endosperm and leaf). The RGAs from embryo transcriptome are show as solid yellow circles, the endosperm in green squares and the leaf in solid red triangles. The tree was constructed by the neighbour-joining method with a statistical analysis of bootstrap with 1,000 replicates. The values of bootstrap >50% are shown on the branches. APAF-1 was using as an outgroup.

3.4.2 ANALYSIS FOR THE CONSERVED MOTIF STRUCTURES

To analyze the structural diversity of the conserved motifs in the NBS sequences, we predicted the conserved motif using the MEME motif detection software. All NBS sequences previously characterized¹¹ and those identified in the three transcriptome¹⁰ were used to identify the conserved motifs. We found 20 putative conserved motifs among them. The detailed motif sequences were shown in Table 2. Previous work identified eight major motifs in the NBS region, and most of them have different patterns depending on whether they are

present in the TNL (TIR-NBS-LRR) or CNL (non-TIR-NBS-LRR) groups¹⁶. In this study, the MEME results identified the motifs that matched the eight major motifs (P-loop, RNBS-A-non-TIR, Kinase-2, RNBS-B, GLPL, RNBS-C, RNBS-D-non TIR) identified previously confirming that the NBS domain in the most conserved region among the domains encoded by *R* genes. On the other hand, we identified in some sequences the LDL motif located in the C-terminal region known as LRR domain. In addition, some of the sequences presented unknown motifs (5, 6, 10, 11, 12, 13, 14, 15, 17, 18, 20 and 29 motifs) in the CC and NBS domains.

Table 2. NB-LRR-specific amino acid motifs identified in RGCs of coconut palm using the program MEME

Domain	Motif ^a	Sequence ^b	Domain	Similar to	Length	E-Value
CC	Motif 14	IIESVTGEQCQLTNMDAMQHELKEQLK GK	CC	-	29	2.8e-283
NB-ARC	Motif 4	GGGGKTTLAQEIYNH	NB-ARC	P-loop	15	1.3e-1136
	Motif 7	RRIDYHFHVRIWCVSQDFDV	NB-ARC	RNBS-A-non-TIR	21	2.5e-964
	Motif 17	DLMRRLIKDLYNENRDILPGNIDAMP CDSLAEVLHG YLQQK	NB-ARC	-	41	1.0e-172
	Motif 12	IQQMIADRLGVPWKDNDSEILRAKMLLRA	NB-ARC	-	29	2.1e-419
	Motif 10	INLLKDIYDGTGDLAGDQSKSSLEPKVE	NB-ARC	-	29	9.2e-612
	Motif 3	RYFLVLDDVWNEFVW	NB-ARC	Kinase-2	15	3.7e-1152
	Motif 15	CDLLCNTLKSC	NB-ARC	-	11	8.2e-255
	Motif 2	ANGCKIIITRNEQVCRQMGA	NB-ARC	RNBS-B	21	4.8e-1446
	Motif 18	VGIPRPNK	NB-ARC	-	8	5.9e-171
	Motif 6	LSEEDCW S L FCKKAF	NB-ARC	-	15	9.0e-834
	Motif 1	HQHLEDIGMEIVKKCHGLPLA	NB-ARC	GLPL	21	1.2e-1652
	Motif 11	AKVVG SAMRHETDPREWRHAL	NB-ARC	-	21	3.0e-657
	Motif 20	KAMKAYPSRIP	NB-ARC	-	11	1.6e-155
	Motif 9	MPNDMFPALYYSYDH	NB-ARC	RNBS-C	15	1.2e-615
	Motif 8	LPHHLQQCFPYCCFF	NB-ARC	RNBS-D-non TIR	15	5.1e-687
	Motif 5	PSHLKQCFAYCSMF PKDYHFDKDYLVQCWMAQGF IQPQGSN	NB-ARC	-	41	4.0e-985
Motif 13	MEDIGDEYFNDLIMRNFFQPS	NB-ARC	-	21	1.6e-295	
Motif 16	YQMHDMMHDLA	NB-ARC	MHDV	11	4.7e-232	
LRR	Motif 19	PDSLGNLIHLRYLNMSYTKISVMPESIGNLRNLQFLMMSYC	LRR1	Motif1 LDL	41	3.9e-165

^aMotifs are listed according to their rank derived from the program MEME analysis. ^bConsensus amino acid sequence derived from MEME program analysis.

3.4.3 ORGANIZATION AND PHYLOGENY OF NBS SEQUENCES AMONG SPECIES OF ARECACEAE

One hundred forty one, fifty nine and thirty four NBS sequences from coconut palm, date palm and oil palm respectively, were used to clarify the phylogenetic relationships between Arecaceae species (Figure 3.2). The tree generated showed nine major groups which are quite distinct from each other. Of the 9 groups, 7 were composed of sequences of the three species of palms and showed clear orthologous relationships between each clade, suggesting that these genes are conserved between Arecaceae species. The NBS

sequences from different species always clustered together in the phylogenetic tree. Interestingly, we have revealed two groups (VIII and IX) conformed by sequences only of coconut palm, these homologs were classified into a species-specific gene family.

The III group containing the majority of NBS sequences (50 sequences) of coconut palm, while I group the minority of sequences (1 sequence). For the date palm, the majority of sequences were identified in the VII group (15 sequences) and the minority in the group I (1 sequence). In oil palm, the majority of sequences were identified in the VI group (13 sequences) while there was a minority in the IV and VII (2 sequences). The great variation in the number of NBS sequences in each clade might indicate the complex evolutionary pattern of NBS-LRR genes in the Arecaceae family. This rapid copy number evolution for *R*-genes implicitly or explicitly assumes a fast turnover of gene copies through repeated cycles of duplication, divergence and eventual loss by pseudogene formation or deletion, suggesting that these genes underwent adaptive evolution in response to the diverse pathogens.¹⁹

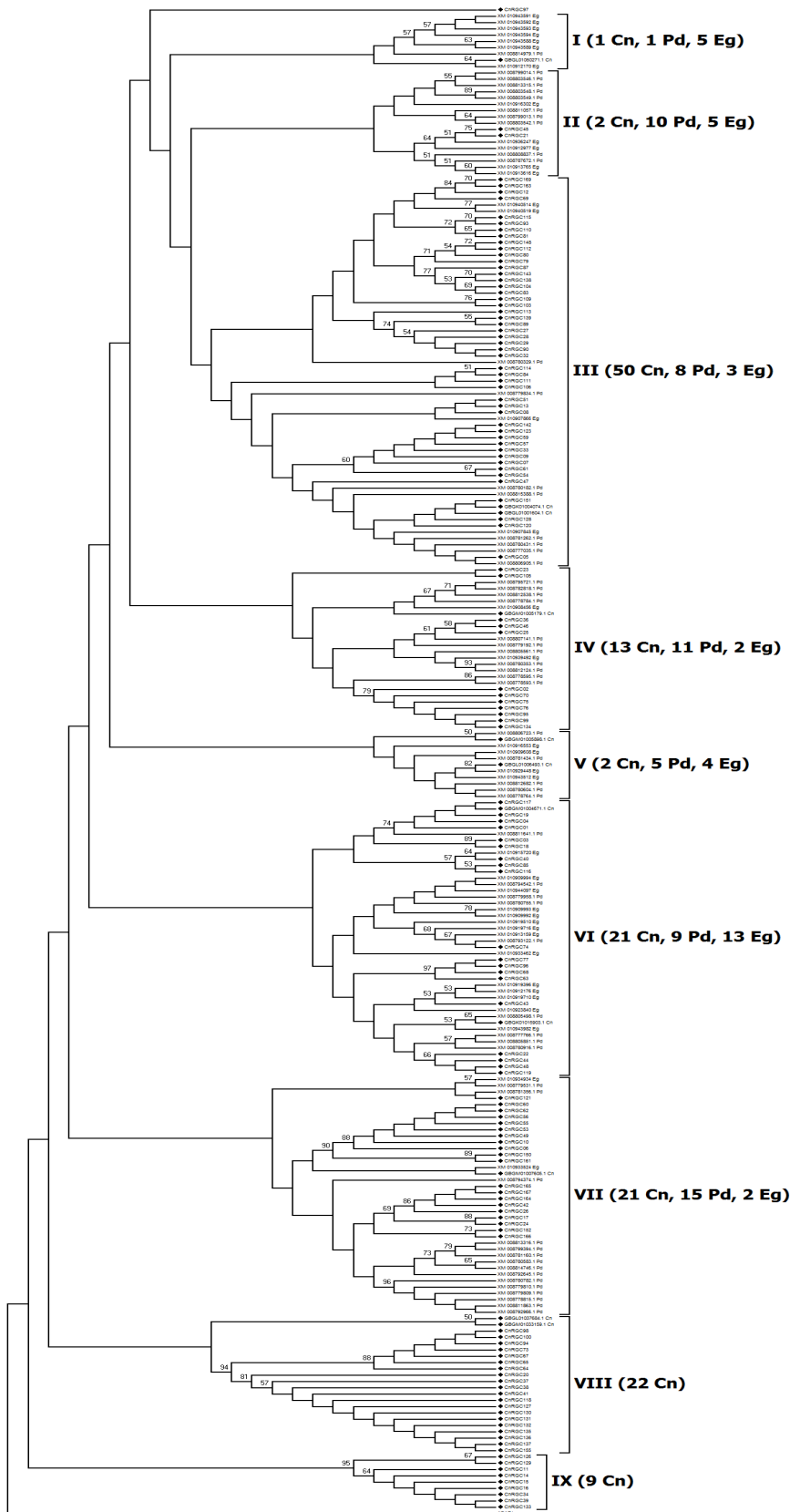


Figure 3.2 Phylogenetic relationship of RGCs from coconut palm, date palm and oil palm. The 234 NBS sequences were group in 9 main clades (I to IX), 7 of which were conformed by sequences orthologous of the 3 species and 2 clades were species-specific of coconut palm. The black rhombuses represent RGMs of coconut palm. The tree was constructed by the neighbour-joining method with a statistical analysis of bootstrap with 1,000 replicates. The values of bootstrap >50% are shown on the branches. APAF-1 was using as an outgroup.

3.4.4 COMPARING EVOLUTIONARY RATES AMONG NBS SEQUENCES OF COCONUT PALM

In genetics, the Ka/Ks ratio is used as an indicator of selective pressure acting on a protein-coding gene.⁴⁰ The groups with less than three sequences were excluded (group I and II) for this analysis. The average Ka/Ks ratio found in the NBS sequences of coconut palm was <1 for the groups III, IV, VI and IX, indicating that these sequences were under strong purifying selection or negative selection implying the elimination of most of the non-synonymous substitutions. However, groups V, VII and VIII were under positive selection or diversifying selection, implying the accumulation of advantageous mutations during the course of evolution. Moreover, we calculated the Ka/Ks ratio between NBS sequences of different clades. The results showed that most Ka/Ks values were <1 in pairwise comparisons, with the exceptions of clade I/clade VII, clade III/clade VII, clade IV/clade VII and clade VII/clade IX with values >1, suggesting that these NBS sequence pairs may be under positive selection. We found that many NBS sequence pairs exhibited Ka/Ks ratios of more than 0.8 (clade I/ clade II, clade I/clade III, clade I/clade V, clade II/clade VII, clade II/clade IX, clade III/clade IX, clade IV/clade VI, clade V/clade VI, clade V/clade VII, clade V/clade VIII and clade VI/clade VII) suggesting potential positive selection (Table 3). This finding indicates that purifying selection and positive selection play a critical role in the evolutionary processes of these NBS sequences in coconut palm.

Table 3. Ka/Ks ratios for pairwise comparisons among members of each clade and among RGCs of the different clades of coconut palm.

	Clade name	Ka	Ks	Ka/Ks	
Between members of each clade	Clade III	0.729	0.980	0.744	
	Clade IV	0.273	1.456	0.188	
	Clade V	0.731	0.677	1.080	
	Clade VI	0.493	1.568	0.314	
	Clade VII	0.596	0.498	1.197	
	Clade VIII	0.378	0.339	1.115	
	Clade IX	0.037	0.086	0.430	
	Between clades	Clade I/Clade II	0.463	0.520	0.890
		Clade I/Clade III	0.390	0.394	0.990
Clade I/Clade IV		0.585	0.964	0.607	
Clade I/Clade V		1.198	1.356	0.883	
Clade I/Clade VI		0.667	1.054	0.633	
Clade I/Clade VII		1.107	1.062	1.042	
Clade I/Clade VIII		0.775	1.166	0.665	
Clade I/Clade IX		0.756	1.197	0.648	
Clade II/Clade III		0.409	0.707	0.579	
Clade II/Clade IV		0.713	1.248	0.575	
Clade II/Clade V		0.967	1.638	0.590	
Clade II/Clade VI		0.739	1.008	0.733	
Clade II/Clade VII		1.181	1.190	0.992	
Clade II/Clade VIII		0.825	1.403	0.588	
Clade II/Clade IX		0.670	0.724	0.925	
Clade III/Clade IV		0.682	0.921	0.740	
Clade III/Clade V		0.900	1.377	0.654	
Clade III/Clade VI		0.623	1.253	0.497	
Clade III/Clade VII		1.323	1.064	1.243	
Clade III/Clade VIII		0.841	1.219	0.690	
Clade III/Clade IX		0.700	0.781	0.896	
Clade IV/Clade V		1.220	1.530	0.797	
Clade IV/Clade VI		0.667	0.676	0.987	
Clade IV/Clade VII		1.218	0.863	1.411	
Clade IV/Clade VIII		0.871	1.474	0.591	
Clade IV/Clade IX		0.818	1.234	0.663	
Clade V/Clade VI		1.027	1.160	0.885	
Clade V/Clade VII		1.606	1.608	0.999	
Clade V/Clade VIII		1.292	1.365	0.947	
Clade V/Clade IX		0.909	1.414	0.643	
Clade VI/Clade VII		1.084	1.344	0.807	
Clade VI/Clade VIII		0.806	1.012	0.796	
Clade VI/Clade IX	0.823	1.099	0.749		
Clade VII/Clade VIII	1.027	1.492	0.688		
Clade VII/Clade IX	1.286	1.000	1.286		
Clade VIII/Clade IX	0.619	2.894	0.214		

3.4.5 PHYLOGENETIC RELATIONSHIPS BETWEEN SPECIES OF ARECACEAE, POACEAE AND BRASSICACEAE

To study the evolutionary relationships among NBS sequences of species of Arecaceae, Poaceae and Brassicaceae, we built a phylogenetic tree using the conserved NBS domain, including the P-loop, kinase-2, kinase-3 and GLPL motifs. The species from the Arecaceae family were *Cocos nucifera*, *Phoenix dactylifera* and *Elaeis guineensis*. For the Poaceae family they were *Oryza sativa* and *Sorghum bicolor*, while those of the Brassicaceae family were *Arabidopsis thaliana*, *Brassica rapa* and *Brassica oleracea*. The sequences that didn't show correct alignment were excluded from this analysis (Figure 3.3).

To clarify the phylogenetic relationships, three types of clades were identified among the eight species: species-specific clade, family-specific clade and complex gene clade (Table 4). If the homolog genes in a clade were only present in the genome of a species but absent in the other, the clade was defined as species-specific. The second type of clade was classified as family-specific, which is composed of homolog genes of species from the same family. The remaining clades were categorized as complex gene clades which include homolog genes of different species. We identified a total of 17 species-specific clades, 6 in the family of Arecaceae (5 in *Cocos nucifera* and 1 in *Phoenix dactylifera*), 10 in Poaceae (4 in *Oryza sativa* and 6 in *Sorghum bicolor*) and only 1 in Brassicaceae (conformed by 1 sequence of *Brassica rapa*). Among the NBS sequences of the three families of plants, 24 family-specific clades were identified, 8 in Arecaceae (3 between *Phoenix dactylifera*-*Elaeis guineensis* and 5 between *Cocos nucifera*-*Phoenix dactylifera*-*Elaeis guineensis*), 13 in Poaceae (between *Oryza sativa*-*Sorghum bicolor*) and 3 in Brassicaceae (1 between *Brassica rapa*-*Brassica oleracea* and 2 between *Arabidopsis thaliana*-*Brassica rapa*-*Brassica oleracea*). For the complex clades, 14 were identified, of which 9 were composed of homolog genes of Arecales and Poales, 3 between Arecales and Brassicales and 2 between Poales and Brassicales, but none composed of NBS sequences of the three families. These results indicate that the sequences of Arecaceae have more phylogenetic relationships with the Poaceae family (9 complex clades) than with the Brassicaceae (3 complex clades). This result is as expected because both families (Arecaceae and Poaceae) belong to the class of monocots, however homolog sequences were also identified between monocots and dicots (Arecaceae-Brassicaceae and Poaceae and Brassicaceae).

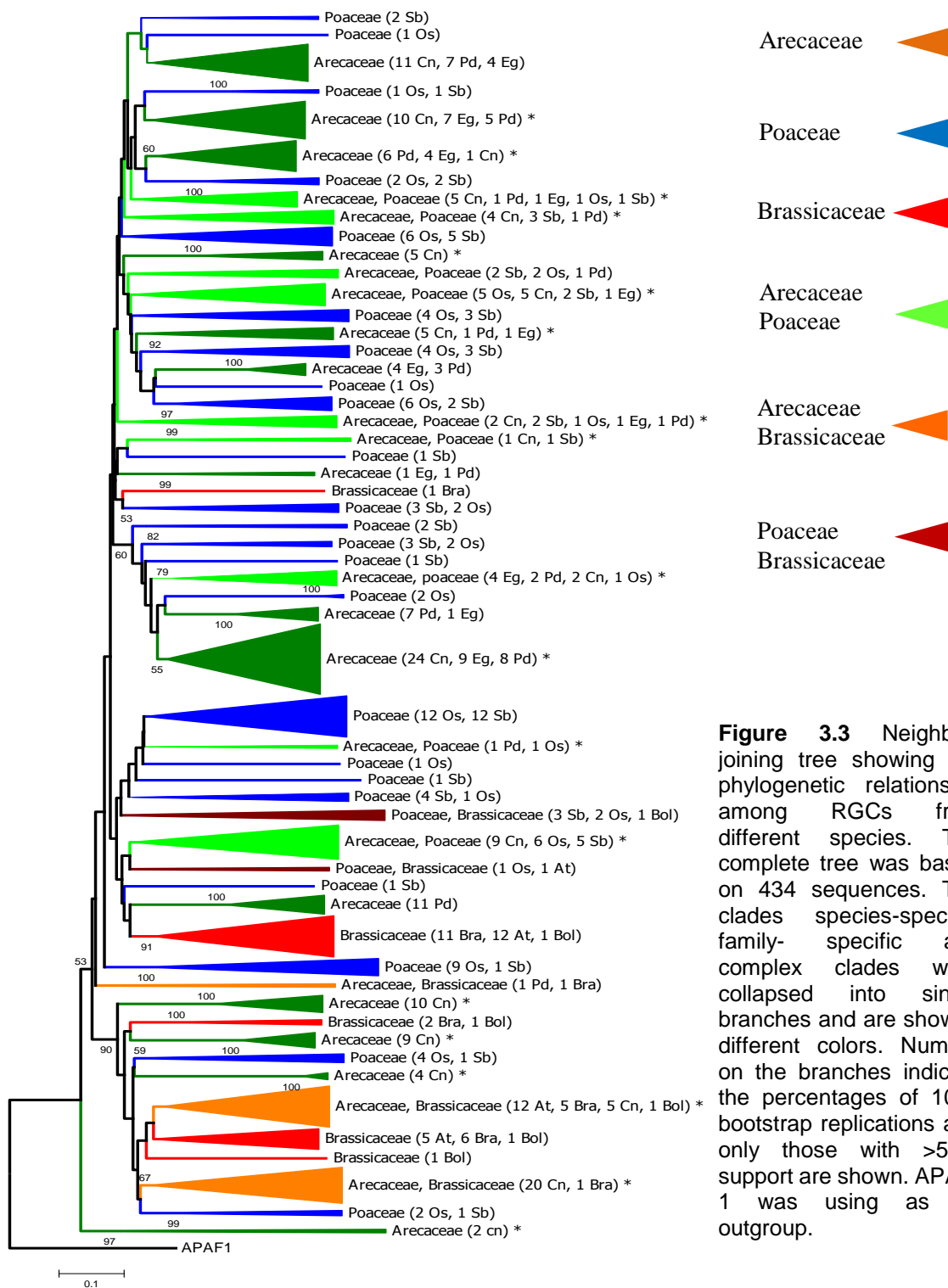


Figure 3.3 Neighbor-joining tree showing the phylogenetic relationship among RGCs from different species. The complete tree was based on 434 sequences. The clades species-specific, family-specific and complex clades were collapsed into single branches and are shown in different colors. Number on the branches indicate the percentages of 1000 bootstrap replications and only those with >50% support are shown. APAF-1 was used as an outgroup.

Table 4. Organization of NBS sequences among species of Arecales, Poales and Brassicales

Type of clade	Areaceae			Poaceae			Brassicaceae			Total	
	<i>Cn</i>	<i>Pd</i>	<i>Eg</i>	<i>Os</i>	<i>Sb</i>	<i>At</i>	<i>Br</i>	<i>Bo</i>			
Species-specific											
No. of clades	5	1	0	4	6	0	1	0	0	17	
No. of genes	30	11	0	5	7	0	1	0	0	54	
Type of clade											
	Areaceae			Poaceae			Brassicaceae			Total	
	<i>Cn-Pd</i>	<i>Cn-Eg</i>	<i>Pd-Eg</i>	<i>Cn-Pd-EG</i>	<i>Os-Sb</i>	<i>At-Br</i>	<i>Br-Bo</i>	<i>At-Br-Bo</i>			
Family-specific											
No. of clades	0	0	3	5	13	0	1	2	2	24	
No. of genes	0	0	11+6	51+27+25	55+41	0	2+1	17+17+2	2	255	
Type of clades											
	Areaceae-Poaceae			Areaceae-Brassicaceae			Poaceae-Brassicaceae			Areaceae-Poaceae-Brassicaceae	
Complex clades											
No. of clade	9			3			2			0	
No. of genes	41+34			26+20			6+2			0	
<i>Cn: Cocos nucifera</i>				<i>Sb: Sorghum bicolor</i>							
<i>Pd: Phoenix dactylifera</i>				<i>At: Arabidopsis thaliana</i>							
<i>Eg: Elaeis guineensis</i>				<i>Bn: Brassica napus</i>							
<i>Os: Oriza sativa</i>				<i>Bo: Brassica oleracea</i>							

3.5 DISCUSSION

The NBS-LRR genes play a vital role in protecting plants from the attacks of various pathogens. Many studies of NBS-LRR genes have been reported in several plant species,^{41,42,23,19,43,44,45,46} however, in coconut palm only two studies of NBS sequences have been reported.^{27,11} These studies only focused on cloning and characterization, but studies of diversity, distribution and phylogenetic relationships of NBS sequences of coconut palm with species of Arecales, poales and brassicales have not been reported.

The number of these genes always differs in different plants, even plants within the same family or same genus. For example, in the grass species, the numbers are quite different, from about 100 in maize to 500 in rice,^{19,47} while that in brassicales the numbers are among 92 in *Brassica rapa* to 149 in *Arabidopsis thaliana*. In palms, of the two sequenced genomes, only 144 NBS genes are reported in *Phoenix dactylifera* and there is no precise data for these genes in the genome of *Elaeis guineensis*. In coconut palm, we identified a total of 231 NBS sequences of which 169 were reported in a previous study (143 with ORF and 23 with stop codons)¹¹ and 62 in this study (53 with ORF and 9 with stop codons), however the total number of NBS genes in its genome remains unknown because it has not been sequenced. Comparative genomic analyses indicate that the number of *R* genes neither increases nor decreases in proportion to genome size⁴⁸, therefore, the likely number of NBS genes is higher in the genome of coconut palm compared with date palm, but this could be more due to the difference in the sizes of genomes (~2.5 Gb for coconut palm and ~671 Mb for date palm),^{49,28} than to the nature of selection pressure imposed by their pathogens and to their very different life histories.

Transcriptome sequencing of coconut palm provide abundant sequence information for identifying the decisive role of these genes in individual tissue for conferring the resistance against pathogens. We identified NBS transcripts from the transcriptome reported by Huang *et al*,¹⁰ in the embryo, endosperm and leaf tissue of coconut healthy palm (Table 1). All the transcripts were different in each tissue (Figure 1). These results indicate a tissue-specific expression and the basal expression observed in three tissues are consistent with the need to activate a rapid response to pathogen attack. Interestingly, the expression observed in the embryo, may indicate that the defense mechanisms are preformed long before germination occurs in the coconut palm.

Comparison of NBS sequences of coconut palm with those of other Arecales indicates the existence of specific clades for *Cocos nucifera*. Genetically defined clusters of *R* usually result from tandem duplications of paralogous genes and have been observed in NBS-LRR encoded in the Arabidopsis and rice genomes.^{40,50,17,41,42,51} In addition, several clades include a mixture of NBS sequences from *Elaeis guineensis* and *Phoenix dactylifera*, indicating that similar RGCs are still shared in different genera of the Arecales.

The variation of RGC family size between species of Arecales could be attributed to gene duplication, deletion, pseudogenization, and functional diversification.^{52,53} The nature of selection pressure imposed by their pathogens is expected to be diverse, therefore the selection is responsible for the drastic variations in numbers of NBS genes. The existence of specific clades is understandable because species-specific pathogens are by definition different. In this scenario, the rapid expansion and/or contraction is a fundamentally important strategy for a species to adapt to the rapidly changing spectrum of pathogens.

Subsequently, when we calculated the Ka/Ks values among sequences of the same clade and sequences of different clades, we could identify >1 values, meaning that positive selection is the force that directs the evolution of new *R* genes. However <1 values were obtained also, suggesting that they were under negative selection or diversifying selection. Our results indicated that NBS sequences of coconut palm are under selective constraints that provide further evidence consistent with a birth-and-death model of *R* gene evolution.

To infer on the evolutionary history among NBS sequences of the Arecales family with respect to Poales and Brassicales, we built a phylogenetic tree. Of the NBS clusters in these three families, the majority of species-specific clades (5 clades) were identified in coconut palm, these clusters frequently consist of tandem duplications of the same species.^{54,55} Family-specific clades were also identified, five in Arecales, suggesting the monophyletic origin of the three species and strong conservation of some NBS sequences in these plants, as has been reported in others species as Poaceae and Rosaceae^{54,22}. Family-specific NBS sequences may also exhibit restricted taxonomic functionality, as shown by the cloned *R* genes *Bs2* from pepper and *RPS2* from *Arabidopsis*, which function in related species within the same family but not in species outside of the family⁵⁶.

Heterogeneous clusters (complex clades), in which sequences belong to different phylogenetic lineages, are also present. The major phylogenetic relationships were identified among species of monocot (9 clades between Arecales and Poales), however, phylogenetic relationship among NBS sequences of monocot and dicot species were also

found (3 clades between Aracaceae and Brassicaceae and 2 clades between Poaceae and Brassicaceae). This result agrees with the proposal by Cannon *et al*⁶⁷ that there are several ancient clades of NBS sequences and that some of these predate the split between monocots and dicots.

In conclusion this paper analyses the NBS sequences of coconut palm and other Areaceae species to reveal the evolutive history of these genes. Phylogenetic reconstruction in the Areaceae family revealed specific clades of NBS sequences for coconut palm indicating genus-specific evolution of resistance genes. However, strikingly similar NBS sequences were shared in different species of coconut palm, date palm and oil palm highlighting a monophyletic origin of these three genera and the high conservation of some RGC in these palms. The phylogenetic relationship with Brassicaceae family suggests that the RGC genes originated before monocotyledonous and dicotyledonous classes split. Importantly, RGCs identified in the present study may be used as candidate genes for engineering pathogen resistance in coconut as well as other related species.

3.6 REFERENCES

1. Fan H, Xiao Y, Yang Y, Xia W, Mason AS *et al.*, RNA-Seq Analysis of *Cocos nucifera*: Transcriptome Sequencing and Subsequent Functional Genomics Approaches. *PLoS ONE* 8(3): e59997 (2013).
2. Rohde W, Randles JW, Langridge P and Hanold D, Nucleotide sequence of a circular single stranded DNA associated with coconut foliar decay virus. *Virology* 176:648-51 (1990).
3. Hanold D and Randles JW, Coconut cadang-cadang disease and its viroid agent. *Plant Dis* 75:330-5 (1991).
4. Parthasarathy MV, van Slobbe WG and Soudant C, Hartrot or fatal wilt of palms. Coconuts (*Cocos nucifera*). *Principes* 22:3-14 (1978).
5. Joseph T and Radha K, Role of *Phytophthora palmivora* in bud rot of coconut. *Plant Dis Repr* 59:1014-7 (1975).
6. Griffith R, Red ring disease of coconut palm. *Plant Dis* 71:193-6 (1978).
7. Sánchez P and Cerda H, The complex *Rhynchosporium palmarum*/*Bursaphelenchus cocophilus* in palms. *Boletín de Entomología Venezolana* 8, 1–18 (1993) (in Spanish).
8. Howard FW and Barrant CI, Questions and answers about lethal yellowing disease. *Principes* 33:163-71 (1989).
9. Harrison NA and Oropeza C, Coconut lethal yellowing. In: Harrison NA, Rao GP, Marcione C, editors. Characterization, diagnosis and management of phytoplasmas. *Houston TX, USA: Studium Press LLC* p. 219-48 (2008).
10. Huang Y, Lee CP, Fu FL, Chang B and Matzke A, De Novo Transcriptome Sequence Assembly from Coconut Leaves and Seeds with a Focus on Factors Involved in RNA-Directed DNA Methylation. *G3: Genes, Genome, Genetics* 4: 2147-2157 (2014).
11. Puch-Hau C, Oropeza C, Peraza-Echeverría S, Gongora-Paredes M, Córdova I, Narváez M, Zizumbo-Villareal D and Sáenz L, Molecular cloning and characterization of disease-resistance gene candidates of the nucleotide binding site (NBS) type from *Cocos nucifera* L. *Physiological and Molecular Plant Pathology* 89: 87-96 (2015).
12. Nejat N, Cahill DM, Vadamalai G, Ziemann M, Rookes J and Naderali N, Transcriptomics-based analysis using RNA-Seq of the coconut (*Cocos nucifera*) leaf in response to yellow decline phytoplasma infection. *Mol Genet Genomics* (2015).

-
13. Jones DA and Jones JDG, The role of leucine-rich repeat proteins in plant defences. *Advances in Botanical Research* 24:89–167 (1997).
 14. Belkhadir Y, Nimchuk Z, Hubert DA, Mackey D and Dangl JL, “Arabidopsis RIN4 negatively regulates disease resistance mediated by RPS2 and RPM1 downstream or independent of the NDR1 signal modulator and is not required for the virulence functions of bacterial type III effectors AvrRpt2 or AvrRpm1”. *Plant Cell* 16(10): 2822–2835 (2004).
 15. Ellis JG, Lawrence GJ, Luck JE, and Dodds PN, “Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity”. *Plant Cell* 11(3): 495–506 (1999).
 16. Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW and Young ND, “Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily”. *Plant Journal* 20 (3): 317–332 (1999).
 17. Richly E, Kurth J, and Leister D, “Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution”. *Molecular Biology and Evolution* 19 (1): 76–84 (2002).
 18. Meyers BC, Kaushik S and Nandety RS, Evolving disease resistance genes. *Curr Opin Plant Biol* 8:129–134 (2005).
 19. Li J, Ding J, Zhang W, Zhang Y, Tang P *et al.*, Unique evolutionary pattern of numbers of gramineous NBS–LRR genes. *Mol Genet Genomics* 283: 427–438 (2010).
 20. Bouktila D, Khalfallah Y, Habachi-Houimli Y, Mezghani-Khemakhem M and Makni M, Makni H, Full-genome identification and characterization of NBS-encoding disease resistance genes in wheat. *Mol Genet Genomics* 290(1):257-71 (2015).
 21. Khan AM, Khan AA, Azhar MT, Amrao L and Cheema HM, Comparative analysis of resistance gene analogues encoding NBS-LRR domains in cotton. *J Sci Food Agric* (2015).
 22. Gu L, Si W, Zhao L, Yang S and Zhang X, Dynamic evolution of NBS–LRR genes in bread wheat and its progenitors. *Mol Genet Genomics* 290:727–738 (2015).
 23. Porter BW, Paidi M, Ming R, Alam M, Nishijima WT and Zhu YJ, Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. *Mol. Genet. Genomics* 281: 609–626 (2009).

-
-
24. Wan H, Yuan W, Bo K, Shen J, Pang X and Chen J, Genome-wide analysis of NBS-encoding disease resistance in *Cucumis sativus* and phylogenetic study of NBS-encoding genes in Cucurbitaceae crops. *BMC Genomics* 14: 109 (2013).
 25. Shang J, Tao Y, Chen X, Zou Y, Lei C, Wang J, Li X, Zhao X, Zhang M, Lu Z *et al.*, Identification of a new rice blast resistance gene, Pid3, by genome wide comparison of paired nucleotide-binding site leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. *Genetics* 182: 1303–1311 (2009).
 26. Hulbert SH, Webb CA, Smith SM and Sun Q, Resistance gene complexes: Evolution and utilization. *Annu. Rev. Phytopathol* 39: 285–312 (2001).
 27. Rajesh MK, Rachana KE, Naganeeswaran SA, Shafeeq R, Thomas RJ, Shareefa M, Merin B and Anitha K, Identification of expressed resistance gene analog sequences in coconut leaf transcriptome and their evolutionary analysis. *Turkish Journal of Agriculture and Forestry* 39: 1-14 (2015).
 28. Al-Mssallem IS, Hu S, Zhang X, Lin Q, Liu W, Tan J *et al.*, Genome sequence of the date palm *Phoenix dactylifera* L. *Nat Commun* 4:2274 (2013).
 29. Foan CC, Lee YW, Tan JS and Alwee SS, Amplification and sequencing of partial-length disease resistance gene homologues coding for NBS-LRR proteins in oil palm (*Elaeis guineensis*). *AsPac J. Mol. Biol. Biotechnol* 20(1): 25-31 (2012).
 30. McCoy RE, Howard FW, Tsai JH, Donselman HM, Thomas DL, Basham HG, Atilano RA, Eskafi FM, Britt L and Collins ME, Lethal yellowing of palms. University of Florida Agricultural Experiment Stations Bulletin 834, Gainesville, USA (1983).
 31. Eeuwens C, Mineral requirements for growth and callus initiation of tissue explants excised from mature coconut palms (*Cocos nucifera*) and date (*Phoenix dactylifera*) palm cultured in vitro. *Physiol Plant* 42:173–178 (1976).
 32. Harrison NA, Richardson PA, Kramer JB, Tsai and JH, Detection of the mycoplasma like organism associated with lethal yellowing disease of palms in Florida by polymerase chain reaction. *Plant Pathol* 43:998-1008 (1994).
 33. Córdova I, Oropeza C, Puch-Hau C, Harrison N, Collí-Rodríguez A, Narvaez M, Nic-Matos G, Reyes C and Sáenz L, A Real-Time PCR assay for detection of coconut lethal yellowing phytoplasmas of group 16SrIV subgroups A, D and E found in the Americas. *Journal of Plant Pathology* 96 (2): 343-352 (2014).

-
34. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez A, Thompson JD, Gibson TJ and Higgins DG, Clustal W and Clustal X version 2.0. *BIOINFORMATICS* 23 (21); 2947–2948 (2007).
 35. Saitou N and Nei M, The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425 (1987).
 36. Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S, MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30: 2725-2729 (2013).
 37. Nei M. and Gojobori T, Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3:418-426 (1986).
 38. Michelmore R and Meyers B, Clusters of resistance genes in plants evolve by divergent selection and birth-and-death process. *Genome Res* 8:1113–1130 (1998).
 39. Meyers BC, Kozik A, Griego A, Kuang H and Michelmore RW, Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* 15:809-834 (2003).
 40. Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J and Tian D, Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol Genet Genomics* 271:402-415 (2004).
 41. Kang YJ, Kim KH, Shim S, Yoon MY, Sun S *et al.*, Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biol* 12 (2012).
 42. Perazzolli M, Malacarne G, Baldo A, Righetti L, Bailey A, Fontana P *et al.*, Characterization of Resistance Gene Analogues (RGAs) in Apple (*Malus x domestica* Borkh.) and Their Evolutionary History of the Rosaceae Family. *PLoS ONE* 9(2): e83844 (2014).
 43. Arya P, Kumar G, Acharya V and Singh AK, Genome-Wide Identification and Expression Analysis of NBS-Encoding Genes in *Malus x domestica* and Expansion of NBS Genes Family in Rosaceae. *PLoS ONE* 9(9): e107987 (2014).
 44. Lozano R, Hamblin M, Prochnik S and Jannink J, Identification and distribution of the NBS-LRR gene family in the Cassava genome. *BMC Genomics*, 16:360 (2015).
 45. Luo S, Zhang Y, Hu Q, Chen J, Li K, Lu C, Liu H, Wang W and Kuang H, Dynamic nucleotide-binding site and leucine-rich repeat-encoding genes in the grass family. *Plant Physiol* 159:197–210 (2012).

-
-
46. Wan H, Yuan W, Ye Q, Wang R, Ruan M, Li Z *et al.*, Analysis of TIR- and non-TIR-NBS LRR disease resistance gene analogous in pepper: characterization, genetic variation, functional divergence and expression patterns. *BMC Genomics* 2012; 13:1-15
 47. Alsaihati B, Lin Q and Al-Mssallem I, Coconut Genome de novo Sequencing. Plant & animal genome XXII. *Conference San Diego, CA* (2014).
 48. Zhu H, Cannon SB, Young ND and Cook DR, Phylogeny and genomic organization of the TIR and non-TIR NBS-LRR resistance gene family in *Medicago truncatula*. *Mol Plant Microbe Interact* 15:529-539 (2002).
 49. Baumgarten A, Cannon S, Spangler R and May G, Genome-level evolution of resistance genes in *Arabidopsis thaliana*. *Genetics* 165:309-319 (2003).
 50. Demuth JP and Hahn MW, The life and death of gene families. *Bioessays* 31: 29–39 (2009).
 51. Seoighe C and Gehring C, Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends Genet* 20: 461–464 (2004).
 52. Yang S, Zhang X, Yue JX, Tian D and Chen JQ, Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol Genet Genomics* 280: 187–198 (2008).
 53. Leister D, Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends Genet* 20: 116–122 (2004).
 54. Tai TH, Dahlbeck D, Clark ET, Gajiwala P, Pasion R, Whalen MC, Stall RE and Staskawicz BJ, Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci. U.S.A* 96:14153-14158 (1999).
 55. Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, and Young ND, Diversity, distribution, and ancient taxonomic relationships within the TIR and Non-TIR NBS–LRR resistance gene subfamilies. *J. Mol. Evol* 54: 548–562 (2002).
 56. Dangl L, Jones J. Plant pathogens and integrated defence responses to infection. *Nature* 411:826-33 (2001).
 57. Sugio A, MacLean AM, Kingdom HN, Grieve VM, Manimekalai R and Hogenhout SA, Diverse targets of phytoplasma effectors: from plant development to defense against insects. *Annu Rev Phytopathol* 49:175-95 (2011).

CAPITULO IV

DESARROLLO DE MARCADORES MOLECULARES EN COCOTERO

4.1 INTRODUCCIÓN

El cocotero (*Cocos nucifera* L.) es un cultivo perenne que contribuye significativamente con la seguridad alimentaria y la generación de empleos e ingresos. Es cultivado en los trópicos en un área estimada de 12 millones de ha (Foale, 2005). Desafortunadamente, esta palma está sujeta al ataque de diversos agentes que producen enfermedades, incluyendo virus (Rohde *et al.*, 1990), viroides (Hanold y Randles, 1991), protozoos (Parthasarathy *et al.* 1978), hongos (Joseph y Radha, 1975), nematodos (Griffith, 1987) y mollicutes como “*Candidatus Phytoplasma palme*”, responsable de la enfermedad devastadora conocida como el amarillamiento letal (Harrison y Oropeza, 2008). Varios de estos patógenos ya han sido controlados y no representan una amenaza para la producción del cocotero, sin embargo, el amarillamiento letal, sigue siendo uno de las principales enfermedades que afectan drásticamente a este cultivo en México y en otras partes del mundo. Actualmente la única forma más eficaz para el control de la enfermedad del amarillamiento letal, es replantando con germoplasma resistente (Been, 1995) las cuales ya han sido identificadas en el país (Zizumbo-Villareal *et al.*, 2008).

En cuanto a la acción y la herencia de la resistencia a las enfermedades ocasionadas por fitoplasmas poco es conocido en cualquier especie de planta y mucho menos ante la enfermedad del amarillamiento letal en cocotero. Con el paso de los años se han propuesto varios modelos sobre la herencia de la resistencia al amarillamiento letal. Primero fue propuesto que la resistencia esta regida por la presencia de un solo locus de un gen dominante (Whitehead, 1986). Sin embargo, la hipótesis fue basada en resultados preliminares y posteriormente fue asignada como incorrecta. Posteriormente fue propuesto, que este caracter es gobernado por un principal gene co-dominante con la influencia de loci de genes como modificadores menores (Harries, 1995).

Ahsburner y Been (1997), realizan un reanálisis de resultados previos e hipótesis con respecto a la herencia de la resistencia al ALC y determinan, que en general, la heredabilidad de la resistencia al amarillamiento letal es muy alta. Por lo tanto puede ser

asumido que la principal influencia en la expresión de la resistencia al LY es a través de afectos genéticos aditivos. Por lo tanto, es probable que la resistencia al amarillamiento letal sea controlada por más de un solo gen, es decir, es poligénica. Actualmente, los estudios moleculares durante la interacción planta-fitoplasma empiezan a florecer y podran generar más evidencias sobre como en realidad esta regido este carácter en las palmas de cocotero. En este estudio empleamos la técnica de AFLP–RGC (Polimorfismos en la longitud de fragmentos amplificados acopladas a secuencias candidatas a genes de resistencia) (Hayes y Saghai-Marrof, 2000; Egea–Gilabert *et al.*, 2003) para la identificación de marcadores moleculares posiblemente asociados con la resistencia o a la susceptibilidad al ALC en cocotero. Esta técnica, es una proceso modificado de AFLP primero propuesto por Hayes y Saghai-Marrof (2000) en soya (*Glycina max*) y se basa en el uso de un iniciador específico de la técnica de AFLP más un iniciador degenerado diseñado a partir del dominio NBS (sitio de unión a nucleótidos) de los genes de resistencia. Esta técnica ha sido usado exitosamente para mapear genes de resistencia en especies de plantas, tales como chile (Egea-Gilabert, 2003) y manzana (Calenge *et al.* 2005). En diferentes programas de mejoramiento genético asistido por marcadores moleculares ha sido posible la infiltración de genes *R* dentro de variedades susceptibles, lográndose de esta manera, generar la resistencia a diferentes patógenos en cultivos como arroz (Singh *et al.* 2012), trigo (Miedaner y Korzun, 2012; Olson *et al.* 2013) y cebada (Miedaner y Korzun, 2012). En cocotero, empleando esta técnica logramos identificar diferentes bandas polimórficas entre ecotipos de cocotero resistentes y susceptibles al ALC, los cuales fueron clonados y secuenciados. Estos polimorfismos fueron validados y solamente uno de ellos (*AFLP-RGC 1-05*) fue específico y reproducible para el ecotipo susceptible (Alto del Oeste Africano) al amarillamiento letal. Los análisis de segregación utilizando este marcador demostraron que existe una segregación de la resistencia al amarillamiento letal en una relación de 3:1, es decir, por cada 3 individuos resistentes hay un individuo susceptible, coincidiendo de esta manera con los resultados de los análisis realizados por Ashburner y Been (1997), sobre la alta heredabilidad de la resistencia al amarillamiento letal. En cocotero, el contar con marcadores moleculares que puedan ser capaces de discernir entre material de cocotero resistente o susceptible al ALC, podría ser de gran importancia para la selección de germoplasma cuya respuesta al amarillamiento letal es desconocida. Por otra parte, se han introducido 12 variedades de material vegetal de cocotero de otros países a México para su cultivo *in vitro*, pero su respuesta al ALC es desconocida, por lo tanto, el contar con estos

marcadores podría permitirnos conocer de manera temprana cual podría ser su posible comportamiento ante esta enfermedad.

4.2 MATERIALES Y METODOS

4.2.1 MATERIAL VEGETAL Y EXTRACCIÓN DEL ADN

Los ecotipos utilizados en este estudio para la identificación de los fragmentos polimórficos fueron 5 individuos del ecotipo Enano Malayo Amarillo (MYD, por sus siglas en inglés) resistente al ALC y 5 del ecotipo Alto del Oeste Africano (WAT, por sus siglas en inglés) susceptible al ALC (Zizumbo-Villareal *et al.* 2008). El ecotipo MYD está localizado en la localidad de Telchac, Yucatán, México, mientras que el ecotipo WAT está localizado en la plantación experimental de “Caridad de cobre”, en San Crisanto, Yucatán, México (Tabla 4.1).

Las bandas polimórficas obtenidas fueron validadas en 4 familias de plantas segregantes (n=81) ubicadas también en la plantación de “Caridad de cobre” en San Crisanto Yucatán, México, resultado de la cruce entre los ecotipos MYD (resistentes) y WAT (susceptibles) (Tabla 4.1). El ADN fue extraído a partir de 1 g de tejido de hoja espada siguiendo el protocolo descrito por Harrison *et al.* (1994) y fue cuantificado usando el equipo de Nanodrop 1000 (Thermo Scientific, USA).

Tabla 4.1. Descripción de genotipos de cocotero empleados para el desarrollo de marcadores moleculares *AFLP-RGC*.

# de palmas	Ecotipo	Reacción a la enfermedad del AL	Lugar de procedencia	Población segregante	# de Individuos	Lugar de procedencia
1	MYD	R	Telchac Puerto, Yucatán			
2	MYD	R	Telchac Puerto, Yucatán			
3	MYD	R	Telchac Puerto, Yucatán			
4	MYD	R	Telchac Puerto Yucatán	Familia 1	24	San Crisanto, Yucatán
5	MYD	R	Telcha Puerto, Yucatán	Familia 3	36	San Crisanto, Yucatán
			Telchac Puerto, Yucatán	Familia 4	12	San Crisanto, Yucatán
6	WAT	S	San Crisanto, Yucatán	Familia 6	9	San Crisanto, Yucatán
7	WAT	S	San Crisanto, Yucatán			
8	WAT	S	San Crisanto, Yucatán			
9	WAT	S	San Crisanto, Yucatán			
10	WAT	S	San Crisanto, Yucatán			

MYD: Enano Malayo Amarillo, WAT: Alto del Oeste Africano
R: Resistente, S: Susceptible

4.2.2. DESARROLLO DE LA TÉCNICA DE AFLP-RGC

La técnica de AFLP-RGC fue llevada a cabo tal como describe Hayes y Saghai-Marrof (2000) y Egea-Gilabert *et al.*, (2003). El ADN digerido y las ligaciones de los adaptadores fueron llevados a cabo siguiendo el protocolo estándar de AFLP (Vos *et al.* 1995) con ciertas modificaciones. Posteriormente el ADN digerido fue pre-amplificado usando iniciadores homólogos a los adaptadores EcoR1 y Mse1 con un nucleótido selectivo (EcoR1 + A y Mse1 + C). La pre-amplificación fue desarrollado en un volumen total de 20 μ l, conteniendo 2.0 μ l de ADN, 1 pmol de cada iniciador, 0.2 mM de dNTP's, 1.5 mM de MgCl₂ y 0.5 U de Taq ADN polimerasa (Invitrogen). La PCR fue llevada a cabo en un GeneAmp PCR System 9700 thermal cycler (Perkin-Elmer) usando las siguientes condiciones: un paso de desnaturalización a 94°C por 5 min, seguido por 20 ciclos de 94°C por 30 s, 56°C por 60 s y 72 °C por 60 s y un paso de elongación final de 72°C por 7 min.

Los productos de la pre-amplificación fueron diluidos (1:20) y 2 μ l fueron usados como templado para la amplificación selectiva. Como se muestra en la Tabla 2, las amplificaciones selectivas fueron desarrolladas con dos iniciadores, uno complementario al adaptador de EcoR1 o Mse1 con dos y tres nucleótidos selectivos y el otro con un iniciador degenerado correspondiente a alguno de los motivos conservados (P-loop, RNBS-D, GLPL) del dominio NBS de las proteínas de resistencia. EL protocolo de PCR touchdown descrito por Vos *et al.*, (1995) fue utilizado. Los productos de la amplificación selectiva fueron mezclados con 3 μ l de buffer de carga (95% formamide, 0.01 M NaOH, 0.05% xileno cianol y 0.05% azul de bromofenol). Las muestras fueron desnaturalizadas durante un periodo de incubación a 95°C por 3 min, puestas directamente en hielo y corridas en un gel de poliacrilamida al 6% que posteriormente fue teñido con nitrato de plata. El tamaño de los fragmentos amplificados fue estimado usando un marcador molecular de 10 pb (Invitrogen).

Tabla 4.2. Secuencia de los iniciadores empleados para la técnica de AFLP-RGCs.

	Nombre	Secuencia (5'-3')
Iniciadores de la técnica de AFLP		
Pre-amplificación	<i>EcoR1</i> + A	GACTGCGTACAATTCA
	<i>Mse1</i> + C	GATGAGTCCTGAGTAAC
Amplificación selectiva	<i>EcoR1</i> + AA	GACTGCGTACAATTCAA
	<i>EcoR1</i> + AAG	GACTGCGTACAATTCAAG
	<i>EcoR1</i> + AAC	GACTGCGTACAATTCAAC
	<i>Mse1</i> + CAT	GATGAGTCCTGAGTAACAT
	<i>Mse1</i> + CTG	GATGAGTCCTGAGTAACTG
Iniciadores degenerados para RGCs		
Amplificación selectiva	PIB	GGIATGCCIGGIIIIGGIAARACIAC
	RNBS-D	GGRAAIARISHRCARTAIVIRAARC
	P3A	AIITYIRIIRYIAGIGGYAAICC
	PIA	GGIATGCCIGGIIIIGGIAARACIAC

Códigos para las posiciones de degeneración: I = inosina; R = A/G; Y = C/T; M = A/C; K = G/T; W = A/T; S=C/G; B = C/G/T; D = A/G/T; H = A/C/T; V = A/C/G; N = A/C/G/T.

4.2.3 PURIFICACIÓN Y CLONACIÓN DE LOS FRAGMENTOS POLIMÓRFICOS

Las bandas polimórficas AFLP-RGC fueron cortadas desde el gel de poliacrilamida y purificadas usando el Kit QIAEX II Gel Extraction (Qiagen, Valencia, CA, USA). El ADN fue reamplificado en un volumen total de 25 µl conteniendo 1.5 mM de MgCl₂, 0.2 mM de cada dNTP, 50 ng de cada iniciador y 1 U de Taq DNA polimerasa (Invitrogen). La PCR fue desarrollada en una GeneAmp PCR System 97200 Thermal Cycler (Perkin-Elmer) siguiendo el protocolo de touchdown descrito por Vos *et al.* (1995). Los productos de PCR fueron separados por electroforesis en un gel de agarosa al 1% (Promega, Madison, Wiscosin, USA) teñido con bromuro de etidio, visualizado por un transiluminador de UV y fotografiado con un GelDoc 2000 (Bio-Rad, Hercules, CA, USA). Los productos de PCR fueron cortados del gel de agarosa y purificados usando el Kit de purificación QIAquick PCR purification (Qiagen, Valencia, CA, USA). El ADN purificado fue clonado dentro del plásmido p-GEM-T Easy (Promega Madison, Wiscosin, USA) y transformado dentro de las células competentes One Shot® TOP10 (Invitrogen). El plásmido recombinante fueron extraídos usando el Kit QIAprep Spin Miniprep (Qiagen, Valencia, CA, USA) y cuantificado en el NANODROP 2000c Spectrophotometer (Thermo SCIENTIFIC). El plásmido recombinante fue secuenciado por la compañía Davis Sequencing (CA, USA).

4.2.4 DISEÑO DE CEBADORES SCAR Y ANÁLISIS DE PCR

Basados en la información de la secuencia de los fragmentos clonados, iniciadores SCAR (Regiones amplificadas, caracterizadas y secuenciadas) fueron diseñados mediante el programa Primer3 (<http://frodo.wi.mit.edu/primer3>). Los iniciadores SCAR fueron validados en palmas resistentes y susceptibles al ALC para determinar tres principales características: a) la amplificación, b) su polimorfismo y c) su reproducibilidad. Las reacciones de PCR fueron llevados a cabo en un volumen de 25 µl conteniendo 200 ng de ADN genómico, 30 pM de cada iniciador, 200 µM de cada dNTP's (Invitrogen), 5 x Mango Taq Colored Reaction Buffer y 1.5 U de Mango Taq DNA polimerasa (BIOLINE). Las condiciones fueron: un paso de desnaturalización inicial a 94°C por 5 min, seguido por 30 ciclos a 95 °C por 20 seg, 58°C por 20 seg y 72°C por 20 seg, con una elongación final a 72°C por 5 min. Los productos amplificados fueron corridos en un gel de agarosa al 1.2 %, teñido con bromuro de etidio y fotografiado con un GelDoc 2000 (Bio-Rad, Hercules, CA, USA).

4.2.5 VALIDACIÓN DEL MARCADOR AFLP-RGC 1-05

El marcador *AFLP-RGC 1-05* (FW 5'-CACCTCATCTTCCAAATCCTC-3' y RV 5'-CCACTCTTCCTTTCACCTCAT-3') fue validado en 10 palmas resistentes (MYD) ubicadas en la localidad de Telchac Puerto, Yucatán, México y 10 susceptibles (WAT) localizadas en la plantación experimental de "Caridad del Cobre" en San Crisanto, Yucatán, México. Esto fue realizado con la finalidad de confirmar su asociación con palmas susceptibles a dicha enfermedad. Este marcador también fue evaluado en una plantación segregante F2 (n=81) resultado de la cruce entre ecotipos resistentes y susceptibles al ALC (MYD x WAT).

Para poder determinar si nuestros datos de segregación estuvieron de acuerdo con las proporciones Mendelianas (3:1) se realizó la prueba estadística de chi-cuadrado (X^2), usando la siguiente formula:

$$X^2 = (O-E)^2 / E$$

En donde O es el valor observado y E el valor esperado. Debido a que son dos clases genéticas para la resistencia al amarillamiento letal (resistentes y susceptibles), los datos se ajustan a un grado de libertad (Gl: n-1=2-1=1), si el valor de Chi-cuadrada calculada es

mayor que el correspondiente al de la probabilidad del 5% ($P \leq 0.05$) se rechaza la hipótesis y por lo tanto los datos no se ajustan a una distribución 3:1.

Por otra parte, también fue validado en 12 individuos de una plantación de Calkiní, Campeche, México, en donde el amarillamiento letal ha devastado aproximadamente el 80 % de la plantación. De los 12 individuos, 10 de ellos no presentaban algún síntoma de la enfermedad a diferencia de las otras 2 quienes presentaron los primeros síntomas del amarillamiento letal. La presencia del fitoplasma del ALC fue confirmada por la técnica de PCR en tiempo real usando la sonda TaqMan 16S (Cordova et al. 2014). Para este análisis solo fueron muestreados 12 individuos debido a que las demás palmas de la plantación estaban muy altas, otras ya se habían muerto y las demás ya habían sido eliminadas. En este análisis no se realizó la prueba de Chi-cuadrado debido a que no son palmas segregantes. Se realizó principalmente con la finalidad de confirmar la asociación del marcador *AFLP-RGC 1-05* con palmas enfermas de ALC.

4.3 RESULTADOS

4.3.1 DESARROLLO DE LA TÉCNICA DE AFLP-RGC

Para desarrollar la técnica de AFLP-RGC, nosotros evaluamos un total de 24 combinaciones de iniciadores. Las amplificaciones selectivas fueron desarrolladas en combinación con el iniciador degenerado para los RGA y cinco iniciadores de AFLP (*EcoR1* + AA, *EcoR1* + AAG, *EcoR1* + AAC, *Mse1* + CAT and *Mse1* + CTG) seleccionados en base al mayor número de polimorfismos obtenidos por Soriano *et al.* (2005) en Durazno (*Prunus armeniaca*) De las 24 combinaciones de iniciadores, 3 de ellos resultaron en 17 bandas polimórficas, de las cuales 9 fueron identificadas en ecotipos resistentes (*AFLP-RGC 1-02*, *AFLP-RGC 1-06*, *AFLP-RGC 1-08*, *AFLP-RGC 1-09*, *AFLP-RGC 1-10*, *AFLP-RGC 1-11*, *AFLP-RGC 1-12*, *AFLP-RGC 1-13* and *AFLP-RGC 1-17*) y 8 en susceptibles (*AFLP-RGC 1-01*, *AFLP-RGC 1-03*, *AFLP-RGC 1-04*, *AFLP-RGC 1-05*, *AFLP-RGC 1-07*, *AFLP-RGC 1-14*, *AFLP-RGC 1-15* and *AFLP-RGC 1-16*) (Tabla 4.3). En la figura 4.1, se muestra un perfil de AFLP-RGC generado con dos combinaciones de iniciadores.

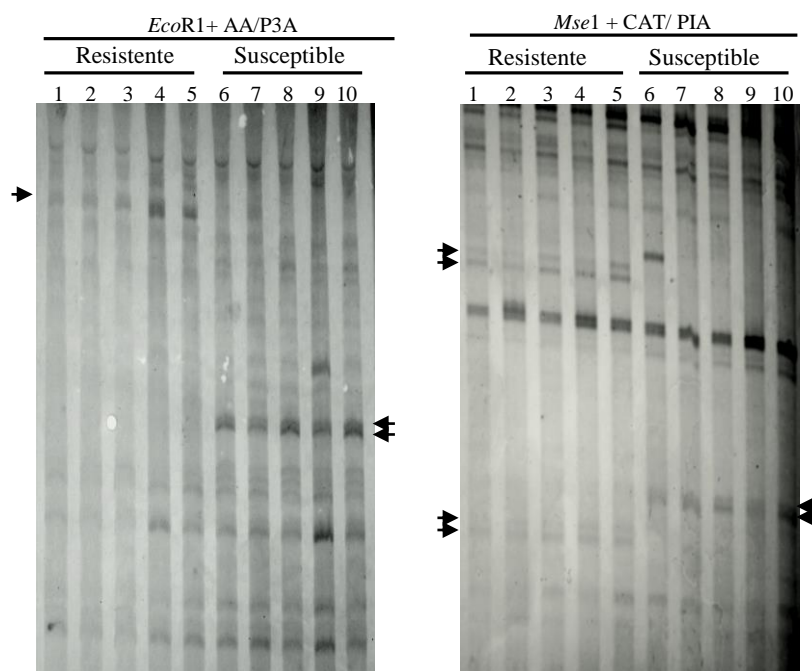


Figura 4.1. Perfiles de la técnica de AFLP-RGC entre ecotipos de cocotero resistentes y susceptibles la enfermedad del amarillamiento letal en gel de poliacrilamida al 6% usando las combinaciones de iniciadores *EcoR1*+ AA/P3A y *Mse1* + CAT/PIA. Carriles del 1 al 5 ecotipos resistentes y carriles del 6 al 10 ecotipos susceptibles. Las flechas señalan las bandas polimórficas entre ambos ecotipos de cocotero.

Las bandas polimórficas fueron secuenciados y se identificó que todos los fragmentos contenían los iniciadores de la técnica de AFLP y el específico para los RGAs, por lo tanto, bandas no específicas de la técnica de AFLP-RGA no fueron detectadas. El tamaño de los fragmentos obtenidos estuvieron dentro de un rango entre 1026 a 114 pb (Tabla 3). La búsqueda de identidad usando el programa BLASTX (Altschuld *et al.*, 1997) mostró que 7 de los marcadores presentaron identidad con alguna de las accesiones del GenBank, mientras que 10 de ellos no presentaron similitud. De los 7 marcadores, *AFLP-RGC 1-01* presentó homología con la proteína de Resistencia tipo RPM1 de *Prunus mume* (accesión XP_008235888.1) involucrado en conferir resistencia contra *Pseudomonas syringae* en *Arabidopsis thaliana* (Grant *et al.*, 1995). *AFLP-RGC1-02* presentó homología a una chaperona de la clase DnaJ de *Phoenix dactylifera* (accesión XP_008786524.1), proteínas que juegan un importante papel en el plegamiento de proteínas, translocación, degradación y respuesta a diversos estreses ambientales (Parsell *et al.* 1994; Zhichang *et al.*, 2010). *AFLP-RGC 1-03* mostro identidad con una Glicosiltransferasa de la familia A (GT-A) de *Phoenix dactylifera* (accesión XP_008777793.1); *AFLP 1-04* a una polyproteina de *Oriza sativa* (accesión AAT07561.1); *AFLP-RGC 1-06* con el gen DEF1 de *Elaeis guineensis* (KF142646.1) y los otros marcadores *AFLP-RGC 1-05* y *1-06* con proteínas no caracterizadas de *Phoenix dactylifera* (XP_008787211.1 y XR_606152.1).

Tabla 4.3. Bandas polimórficas identificadas usando la técnica AFLP-RGCs, indicando la combinación de iniciadores, el tamaño del marcador, el ecotipo donde fue amplificado y las accesiones en el GenBank a la que mostraron identidad.

Nombre	Combinación de iniciadores	Tamaño (pb)	Ecotipo	Accesiones a las que presentaron identidad	Valor de E	Porcentaje de identidad	ORF
<i>AFLP-RGC I-01</i>	<i>EcoRI-AA + P3A-R</i>	311	S	Disease resistance protein RPM1-like (<i>Prunus mume</i>)	1 ^{e-12}	40	Si
<i>AFLP-RGC I-02</i>	<i>MseI-CTG + PIB-F</i>	133	R	DnaJ-class molecular chaperone (<i>Phoenix dactylifera</i>)	0.33	94	Si
<i>AFLP-RGC I-03</i>	<i>EcoRI-AA + P3A-R</i>	308	S	Glycosyltransferase family A (<i>Phoenix dactylifera</i>)	4 ^{e-29}	57	No
<i>AFLP-RGC I-04</i>	<i>EcoRI-AA + PIA-F</i>	160	S	Putative polyprotein (<i>Oryza sativa</i>)	4 ^{e-09}	63	Si
<i>AFLP-RGC I-05</i>	<i>MseI-CAT + PIA-F</i>	792	S	Uncharacterized protein (<i>Phoenix dactylifera</i>)	2 ^{e-04}	88	Si
<i>AFLP-RGC I-06</i>	<i>EcoRI-AA + P3A-R</i>	433	R	Disrupted DEF1 gene (<i>Elaeis guineensis</i>)	6e-36	72	No
<i>AFLP-RGC I-07</i>	<i>EcoRI-AA + PIA-F</i>	155	S	Uncharacterized protein (<i>Phoenix dactylifera</i>)	4e-19	94	No
<i>AFLP-RGC I-08</i>	<i>MseI-CAT + PIA-F</i>	192	R	NS	-	-	No
<i>AFLP-RGC I-09</i>	<i>MseI-CAT + PIA-F</i>	133	R	NS	-	-	No
<i>AFLP-RGC I-10</i>	<i>MseI-CAT + PIA-F</i>	136	R	NS	-	-	No
<i>AFLP-RGC I-11</i>	<i>MseI-CAT + PIA-F</i>	176	R	NS	-	-	No
<i>AFLP-RGC I-12</i>	<i>MseI-CAT + PIA-F</i>	192	R	NS	-	-	No
<i>AFLP-RGC I-13</i>	<i>MseI-CAT + PIA-F</i>	130	R	NS	-	-	No
<i>AFLP-RGC I-14</i>	<i>MseI-CAT + PIA-F</i>	141	S	NS	-	-	No
<i>AFLP-RGC I-15</i>	<i>MseI-CAT + PIA-F</i>	142	S	NS	-	-	No
<i>AFLP-RGC I-16</i>	<i>MseI-CAT + PIA-F</i>	142	S	NS	-	-	No
<i>AFLP-RGC I-17</i>	<i>MseI-CTG + PIA-F</i>	114	R	NS	-	-	No

R, Resistente (MYD)

S, Susceptible (WAT)

NS, no se encontró similitud

4.3.2 VALIDACIÓN DE LOS FRAGMENTOS POLIMORFISMOS

A partir de cada banda polimórfica secuenciada se diseñaron marcadores SCAR los cuales fueron validados en los ecotipos a partir de los cuales fueron generados, esto fue con la finalidad de determinar si en realidad son polimórficos y reproducibles. De los 17 marcadores SCAR, 6 fueron validados (*AFLP-RGC 1-01*, *AFLP-RGC 1-03*, *AFLP-RGC 1-05*, *AFLP-RGC 1-06*, *AFLP-RGC 1-08*, *AFLP-RGC 1-12*) en palmas resistentes y susceptibles al ALC y 11 (*AFLP-RGC 1-02*, *AFLP-RGC 1-04*, *AFLP-RGC 1-07*, *AFLP-RGC 1-09*, *AFLP-RGC 1-10*, *AFLP-RGC 1-11*, *AFLP-RGC 1-13*, *AFLP-RGC 1-14*, *AFLP-RGC 1-15*, *AFLP-RGC 1-16* and *AFLP-RGC 1-17*) no fueron considerados para este análisis debido a que no se pudieron diseñar iniciadores específicos por ser fragmentos pequeños. Entre los 6 SCAR evaluados, 5 de ellos (*AFLP-RGC 1-01*, *AFLP-RGC 1-03*, *AFLP-RGC 1-06*, *AFLP-RGC 1-08* y *AFLP-RGC 1-12*) resultaron no estar ligados a la característica de resistencia o susceptibilidad del ecotipo a partir del cual fue generado, es decir, resultaron no ser polimórficos, por lo tanto no fueron considerados para ser mapeados en la plantación segregante. El marcador *AFLP-RGC 1-05*, resulto ser el único marcador SCAR que fue polimórfico y reproducible para el ecotipo susceptible al amarillamiento letal (Tabla 4.4).

Tabla 4.4. Características de los marcadores SCAR validados

Nombre	Amplificación	Polimorfo	Reproducible	Ecotipo en el que amplifica	
				Resistente	Susceptible
<i>AFLP-RGC 1-01</i>	Si	No	Si	+	+
<i>AFLP-RGC 1-03</i>	Si	No	Si	+	+
<i>AFLP-RGC 1-05</i>	Si	Si	Si	-	+
<i>AFLP-RGC 1-06</i>	Si	No	Si	+	+
<i>AFLP-RGC 1-08</i>	Si	No	No	+	+
<i>AFLP-RGC 1-12</i>	Si	No	Si	+	+

4.3.3 VALIDACIÓN DEL MARCADOR SCAR *AFLP-RGC 1-05*

El marcador *AFLP-RGC 1-05* fue posteriormente validado en 10 palmas de cocotero resistentes y 10 susceptibles a la enfermedad del amarillamiento letal (3 de las palmas susceptibles fueron positivas para el amarillamiento letal). En la Figura 4.2A, se puede observar la presencia de un fragmento de ~790 pb solamente en las palmas susceptibles al amarillamiento letal pero ausente en las resistentes. Cuando se evaluó su presencia en palmas enfermas de amarillamiento letal el marcador *AFLP-RGC 1-05* también fue identificado (Figura 4.2B). Los resultados de la validación para las 30 palmas (10, resistentes, 10 susceptibles y 10 enfermas de ALC) se describen la tabla 4.5.

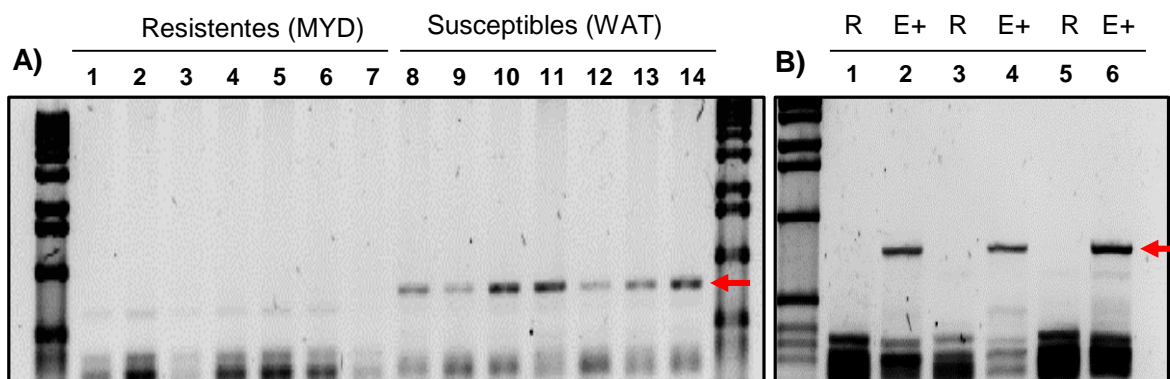


Figura 4.2. Imagen representativa de la validación del marcador *AFLP-RGC 1-05* en palmas resistentes, susceptibles y enfermas de amarillamiento letal. A) carriles 1 al 7; palmas resistentes (MYD), carriles del 8 al 14; palmas susceptibles (WAT) y B) 1, 3 y 5; palmas resistentes (MYD) y 2, 4 y 6; palmas enfermas de amarillamiento letal (WAT). R: resistente y E: enferma. La banda amplificada que corresponde al marcador *AFLP-RGC 1-05* es señalado con una flecha roja.

4.3.4 EVALUACIÓN DEL MARCADOR *AFLP-RGC 1-05* EN SEGREGANTES Y EN UNA PLANTACIÓN

El marcador *AFLP-RGC 1-05* fue mapeada en 81 individuos de 4 familias segregantes F2, resultado de la cruce entre ecotipos resistentes (MYD) y susceptibles (WAT) al ALC, con la finalidad de determinar el número de plantas que segregan dicho marcador y la proporción de plantas resistentes con respecto a las susceptibles en una población segregante. Los 81 individuos fueron definidos en 4 familias (Tabla 4.1) y como se puede observar en la tabla 4.5, de los 81 individuos solamente 13 de ellos presentaron el marcador, es decir, el 16 % de la plantación. En la familia 9, se pudo identificar el mayor porcentaje (33.3%) de palmas positivas para el marcador *AFLP-RGC 1-05* mientras que en la familia 3 el menor número (11.1%). En la figura 4.3 se puede observar la presencia del marcador de ~790 pd en palmas segregantes.

La resistencia al amarillamiento letal, es considerado un carácter dominante, y de acuerdo a las leyes de Mendel, en la F1 (híbridos) la dominancia se manifiesta sobre la recesiva. Sin embargo, en la segunda ley de Mendel conocida como la separación o disyunción de los alelos, en la F2 (resultado de la autopolinización de la F1) existe una segregación en una proporción 3:1. La segregación observada en esta población (n=81) fue de 68:13 (resistentes: susceptibles) y el valor de Chi-cuadrado fue de 3.46. Debido a que el valor calculado fue menor que el valor encontrado en la tabla de Chi-cuadrado (3.84) se acepta la hipótesis de que los datos se ajustan a una distribución Mendeliana de 3:1.

Tabla 4.5. Validación del marcador *AFLP-RGC 1-05* en plantas segregantes.

Familias	Individuos	Presencia del marcador	Individuos	Presencia del marcador	Porcentaje observado
1	1-1-1	-	1-3-1	-	12.5
	1-1-2	-	1-3-2	-	
	1-1-3	-	1-3-3	-	
	1-1-4	-	1-4-1	-	
	1-1-5	-	1-4-2	-	
	1-1-6	-	1-4-3	-	
	1-1-7	+	1-4-4	-	
	1-1-8	-	1-4-5	-	
	1-1-9	-	1-4-6	-	
	1-1-10	-	1-6-2	+	
	1-1-11	-	1-6-3	+	
	1-2-2	-	1-7-1	-	
3	3-1-1	-	3-5-6		11.1
	3-2-1	-	3-5-7	+	
	3-2-2	-	3-6-2		
	3-2-3	-	3-6-3		
	3-2-4	-	3-6-4		
	3-2-5	-	3-6-5		
	3-2-6	-	3-6-6		
	3-3-1	-	3-7-1		
	3-3-2	-	3-7-2		
	3-4-1	-	3-8-1		
	3-4-2	-	3-8-2		
	3-4-3	-	3-9-1		
	3-5-1	-	3-9-2		
	3-5-2	+	3-9-3		
	3-5-3	-	3-9-4		
	3-5-4	+	3-10-1	+	
3-5-5	-	4-7-4			
3-6-1	-	4-8-1			
4	4-2-1		4-5.2	+	25
	4-3-1	+	4-5-3	-	
	4-3-2	+	4-7-1	-	
	4-3-3	-	4-7-2	-	
	4-3-4	-	4-7-3	-	
	4-5-1	-	4-8-2	-	
6	6-1-1	+	6-2-3	-	33.3
	6-1-2	+	6-2-6	+	
	6-1-3	-	6-3-1	-	
	6-1-4	-	6-3-2	-	
	6-2-2	-			
Total= 81 individuos					

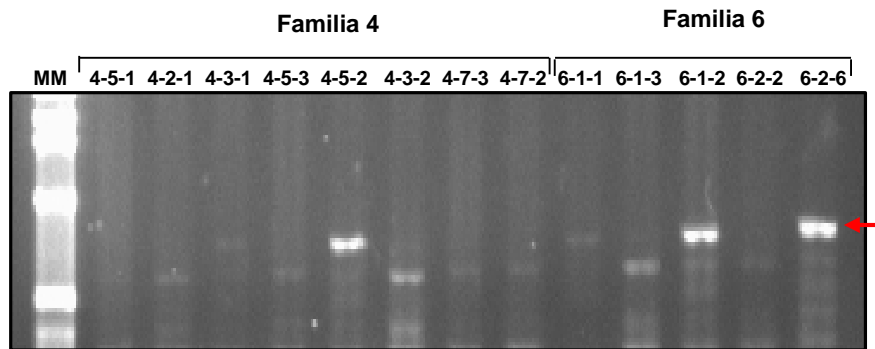


Figura 4.3. Validación del marcador *AFLP-RGC 1-05* en plantas segregantes, resultado de la cruce entre ecotipos resistentes y susceptibles al amarillamiento letal. Carril 1, marcador molecular, carriles del 2 al 9, familia 4 y carriles del 10 al 14, familia 6. La flecha en color rojo indica la banda del tamaño esperado para el marcador *AFLP-RGC 1-05*.

El marcador *AFLP-RGC 1-05* también fue validado en 12 palmas de una plantación de Calkiní, Campeche, México, en donde la enfermedad del ALC ha devastado aproximadamente el 80 % de la plantación (figura 4.4) y en la cual se desconoce el genotipo de las plantas que ahí se encuentran plantadas. Durante el muestreo de las palmas, solamente identificamos a 2 de ellas que presentaron los síntomas del amarillamiento letal y que fue confirmada por la técnica de PCR en tiempo real usando la sonda TaqMan (Cordova *et al.* 2014) (Figura 4.5). Las otras 10 palmas fueron muestreadas debido a que desde que fueron plantadas no se han infectado y presentan un fenotipo diferente a todas las que se han muerto de ALC. Los resultados de la validación indicaron la presencia del marcador solamente en las 2 palmas que fueron diagnosticadas como positivas al ALC, pero en las otras 10 palmas las cuales no se han enfermado de amarillamiento letal el marcador no fue identificado (Figura 4.5). Estos resultados nos proporcionan datos importantes sobre el posible uso de este marcador para la identificación de germoplasma susceptible al amarillamiento letal, sin embargo, estudios complementarios son necesarios para afirmar esta asociación. Las 10 palmas que no presentaron el marcador las hemos observado en diferentes periodos de tiempo y hasta el momento siguen libres de la enfermedad.



Figura 4.4. Plantación ubicada en la localidad de Calkiní Campeche, México, en donde la enfermedad del amarillamiento letal ha devastado cerca del 80 % de la plantación y en donde el marcador *AFLP-RGC 1-05* fue validado en 12 individuos.

Palmas	MM	1	2	3	4	5	6	7	8	9	10	11	12
PCR tiempo real		ND	ND	ND	38.0	ND	ND	ND	ND	39.2	ND	20.8	22.32
		(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(+)	(+)

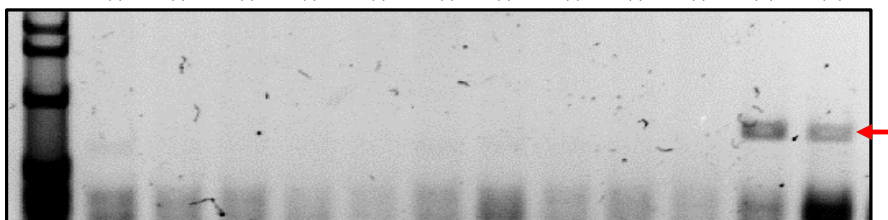


Figura 4.5. Validación del marcador *AFLP-RGC 1-05* en 12 palmas de la plantación de Calkiní Campeche, México. Las palmas de la 1 a la 10 no presentaban ningún síntoma del ALC, el diagnóstico por PCR tiempo real fue negativo y no se observó la presencia del marcador. Palmas 11 y 12 presentaban síntomas de ALC, el diagnóstico fue positivo y fueron las únicas en las que se observa la presencia del marcador *AFLP-RGC 1-05*. La flecha roja indica en tamaño del fragmento esperado.

4.4 DISCUSIÓN

El cultivo del cocotero, es afectado por una gran diversidad de patógenos entre las que se encuentra el fitoplasma causante de la enfermedad del amarillamiento letal, que afecta también a muchas otras especies de palmas. El desarrollo de un programa de mejoramiento genético asistido por marcadores moleculares en cocotero, representa una importante herramienta para poder identificar germoplasma resistente o susceptible al amarillamiento letal. En este estudio, empleando la técnica de AFLP-RGCs, logramos identificar 17 bandas polimórficas entre los ecotipos de cocotero que fueron evaluados, 9 de ellos estuvieron presentes en las palmas resistentes, mientras que los otros 8 en las susceptibles. El bajo número de polimorfismos obtenidos ya han sido reportados en estudios anteriores (Egea-Gilabert *et al.* 2003; Soriano *et al.* 2005), sin embargo, en cocotero, es posible que evaluando más combinaciones de iniciadores sea posible incrementar número de fragmentos polimórficos.

Todas las bandas polimórficas identificadas en este estudio fueron clonados y secuenciados. Los resultados de la secuenciación nos permitieron identificar 17 secuencias diferentes, de las cuales 7 presentaron identidad con alguna de las accesiones del GenBank, mientras que en 10 de ellas no se encontró alguna similitud. Todos los fragmentos polimórficos presentaron en su secuencia la presencia de los adaptadores de la técnica de AFLP-RGC, lo que indica que fragmentos inespecíficos de esta técnica no fueron generados. Entre las secuencias que presentaron identidad con las accesiones del GenBank solamente una de ellas presentó identidad con el gen de resistencia RPM1 de *Prunus mume*. Esto sugiere que los iniciadores degenerados para los RGCs empleados en este estudio no tuvieron tanta afinidad hacia esta familia de genes, sin embargo, las secuencias con identidad a otros genes, tales como poliproteínas y glicosiltransferasas, ya han sido reportados en otros estudios empleando esta misma técnica (Soriano *et al.* 2005).

A partir de cada secuencia nosotros tratamos de diseñar iniciadores específicos con la finalidad de convertirlos en marcadores SCAR, sin embargo, en 11 de ellas no fue posible debido a que fueron secuencias pequeñas (menores a 180 pb), lo que dificultó el diseño de los iniciadores, pero en 6 de ellas si fue posible. Cuando los iniciadores específicos fueron validados en palmas resistentes y susceptibles al amarillamiento letal, la mayoría de ellos no mostraron ser polimórficos y tampoco ser reproducibles, con la única excepción del

marcador *AFLP-RGC 1-05* quien resulto ser específico para el ecotipo susceptible al amarillamiento letal. La presencia de este marcador también fue identificado en palmas enfermas con ALC.

Los análisis de segregación de la resistencia al amarillamiento letal muestran que las plantas F2 segregan como es esperado en una proporción Mendeliana de 3:1. Es decir, por cada 4 palmas, 3 son resistentes y 1 susceptible. Estos datos sugieren que el carácter de resistencia al ALC es dominante con respecto a la susceptibilidad, tal como ha sido documentado (poner cita) y existe una alta herabilidad en las poblaciones F2 analizadas, lo cual también concuerda con los valores altos de heredabilidad obtenidos por Ashburner y Been (1997). En un principio se propuso que la resistencia al amarillamiento letal es una respuesta monogenica (Whitehead, 1968), sin embargo, otros estudios concluyeron en bases a sus resultados que la resistencia es poligenica (Been, 1981), en este trabajo aún no se puede concretar esta afirmación, sin embargo, la forma de segregación Mendeliana que se observa en este estudio, parece indicar que puede ser regida por un solo gen, debido a que cuando las características de segregación son de manera no Mendeliana, es cuando se trata de características fenotípicas gobernadas por diferentes genes, es decir, poligenica. Estas hipótesis aún necesitan ser comprobadas, sin embargo los estudios transcriptomicos durante la interacción molecular planta-fitoplasma cada vez generan más evidencias sobre los posibles genes encargados de generar dicha resistencia (Hren *et al.* 2009; Mardi *et al.* 2015).

El contar con un marcador molecular que pueda ser empleado para la selección de palmas con características de interés representa un gran reto. Rajesh *et al.*, (2013), después de evaluar un total de 200 iniciadores para la técnica de RAPD, solamente una combinación de iniciadores genero un fragmento polimórfico reproducible que estuvo presente exclusivamente en accesiones altos y ausente en cultivares enanos. En nuestro grupo de trabajo, el éxito de obtener un marcador molecular que pueda discernir entre ecotipos susceptibles con respecto a los resistentes o viceversa, representaría una excelente herramienta para la identificación y selección de germoplasma de cocotero cuya respuesta al amarillamiento letal es desconocida. El marcador *AFLP-RGC 1-05* también fue validado en una plantación segregante resultado de la crusa entre palmas resistentes y susceptibles, los resultados mostraron la presencia de este marcador en el 16 % de esta plantación, sin embargo, los resultados en cuanto este análisis de segregación necesitan repetirse debido

a que muchas de las plantas son muy altas y solo se pudo alcanzar muestrear las hojas más viejas y no la hoja espada, la más ideal para este estudio.

En este trabajo se logró identificar un posible marcador molecular para la identificación de germoplasma susceptible al amarillamiento letal. El contar con este tipo de marcadores moleculares podría permitirnos seleccionar plantas con características de interés desde la etapa de plántula, con alta precisión y con reducciones en su costo. Para poder identificar otros posibles marcadores moleculares asociados a la resistencia o susceptibilidad a las principales enfermedades del cocotero sería necesario evaluar un mayor número de combinaciones de iniciadores para la técnica de AFL-RGC, así como también emplear otras técnicas como los RAPD, AFLP, SSR y MSAP.

4.5 CONCLUSIONES

1. Empleando la técnica de AFLP-RGCs logramos identificar 17 fragmentos polimórficos, 9 presentes en ecotipos resistentes y 8 en susceptibles.
2. Se logró desarrollar un marcador SCAR a partir de uno de los polimorfismos obtenidos. Dicho marcador denominado *AFLP-RGC 1-05* fue específico y reproducible para el ecotipo altamente susceptible a la enfermedad del amarillamiento letal.
3. Empleando el marcador *AFLP-RGC 1-05* asociado a plantas susceptibles pudimos identificar que la segregación de la resistencia al amarillamiento letal es dominante con respecto a la susceptibilidad, observándose una segregación Mendeliana de 3:1.
4. Para confirmar la asociación del marcador *AFLP-RGC 1-05* en palmas susceptibles al AL, muestreamos una plantación altamente afectada por esta enfermedad y solamente estuvo presente en palmas con ALC y no en las palmas aparentemente sanas.

4.6 REFERENCIAS BIBLIOGRAFICAS

- Ahsburner GR y Been BO. (1997). Characterization of Resistance to Lethal Yellowing in *Cocos nucifera* and implications for Genetic Improvement of this Species in the Caribbean Region. Resistance Screening & Rehabilitation. Pp. 173-183.
- Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. y Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
- Been BO. (1981). Observations of field resistance to lethal yellowing in coconut varieties and hybrids in Jamaica. *Oleagineux*, 36:9-12.
- Calenge F., Van der Linden C.G., Van de Weg E., Schouten H.J., Van Arkel G., Denance C. y Durel C.E. (2005). Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. *Theor Appl Genet.* 110:660–668.
- Cardeña R., Ashburner G.R. y Oropeza C. (2003). Identification of RAPDs associated with resistance to lethal yellowing of the coconut (*Cocos nucifera* L.) palm. *Scientia Horticulture.* 98; 257-263.
- Córdova I., Oropeza C., Puch-Hau C., Harrison N., Collí-Rodríguez A., Narvaez M., Nic-Matos G., Reyes C y Sáenz L. (2014). A Real-Time PCR assay for detection of coconut lethal yellowing phytoplasmas of group 16SrIV subgroups A, D and E found in the Americas. *Journal of Plant Pathology.* 96 (2): 343-352.
- Egea-Gilabert C., Dickinson M.J., Bilotti G. y Candela, M.E. (2003). Isolation of resistance gene analogs in pepper using modified AFLPs. *Biol Plant.* 47:27–32.
- Foale M: An introduction to the coconut palm. In: *Coconut Genetic Resources*. Edited by Batugal P, Ramanatha V, Rao GP, Oliver J. (2005). International Plant Genetic Resources Institute – Regional Office for Asia, the Pacific and Oceania (IPGRI-APO), Serdang, Selangor DE, Malaysia;1-8.

- Grant M., Godard L., Straube E., Ashfield T., Leward J., Sattler A., Innes R. y Dangl J. (1995). Structure of the Arabidopsis RPM1 enabling dual specificity disease resistance. *Science* 269:843–846.
- Griffith, R. (1987). Redring disease of coconut palm. *Plant Dis.* 71: 193-196.
- Hanold D y Randles JW. (1991). Coconut cadang-cadang disease and its viroid agent. *Plant Dis.* 75:330–5.
- Harrison NA, Oropeza C. Coconut lethal yellowing. In: *Characterization, Diagnosis and Management of Phytoplasmas*. Edited by Harrison NA, Rao GP, Marcone C. Houston TX, USA: Studium Press LLC; 2008: 219-248.
- Harrison N.A., Richardson P.A., Kramer J.B. y Tsai J.H. (1994). Detection of the mycoplasma-like organism associated with lethal yellowing disease of palms in Florida by polymerase chain reaction. *Plant Pathology.* 43: 998-1008.
- Hayes y Saghai-Marrof. (2000). Targeted resistance gene mapping in soybean using modified AFLPs. *Theor Appl Genet.* 100:1279–1283.
- Hren M, Nikolic P, Rotter A, Blejec A, Terrier N, Ravnikar M, Dermastia M. y Gruden K. (2009). Bois noir phytoplasma induces significant reprogramming of the leaf transcriptome in the field grow grapevine. *BMC genomics*; 10:460.
- Joseph T. y Radha K. (1975). Role of *Phytophthora palmivora* in bud rot of coconut. *Plant Dis Repr.* 59:1014-7.
- Mardi M., Karimi Farsad L., Gharechahi J. y Salekdeh G.H. (2015). In-Depth Transcriptome Sequencing of Mexican Lime Trees Infected with Candidatus *Phytoplasma aurantifolia*. *PLoS ONE* 10(7): e0130425. doi:10.1371/journal.pone.0130425.
- Olson E.L., Rouse M.N., Pumphrey M.O., Bowden R.L., Gill B.S. y Poland J.A. (2013). Introgression of stem rust resistance genes *SrTA10187* and *SrTA10171* from *Aegilops tauschii* to wheat. *Theor Appl Genet.* 126:2477–2484.

- Parsell D.A., Lindquist S., Morimoto R.I., Tissieres A. y Georgopoulos C. (1994). Heat shock proteins and stress tolerance. In *The biology of heat shock proteins and molecular chaperones*, eds Morimoto R.I., Tissieres A., Georgopoulos C. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp 457–494.
- Parthasarathy MV, van Slobbe WG. y Soudant C. (1978). Hartrot or fatal wilt of palms. 1. Coconuts (*Cocos nucifera*). *Principes*. 22:3–14.
- Rajasha M.K., Jerard B.A., Preethi P., Thomas R.J., Fayas T.P., Rachana K.E. y Karun A. (2013). Development of a RAPD-derived SCAR marker associated with tall-type palm trait in coconut. *Scientia Horticulturae*. 150; 312–316.
- Rohde W, Randles JW, Langridge P. y Hanold D. (1990). Nucleotide sequence of a circular single stranded DNA associated with coconut foliar decay virus. *Virology*. 176:648–51.
- Singh A., Singh V.K., Singh S.P., Pandian R.T.P, Ellur R.K., Singh D., Bhowmick P.K., Gopala Krishnan S., Nagarajan M., Vinod K.K., Singh U.D., Prabhu K.V., Sharma T.R., Mohapatra T. y Singh A.K. (2012). Molecular breeding for the development of multiple disease resistance in Basmati rice. *AoB PLANTS*. pls029; doi:10.1093/aobpla/pls029.
- Soriano J.M., Vilanova S., Romero C., Llácer G. y Badenes M.L. (2005). Characterization and mapping of NBS-LRR resistance gene analogs in apricot (*Prunus armeniaca* L.). *Theor Appl Genet*. 110: 980–989.
- Miedaner T. y Korzun V. (2012). Marker-Assisted Selection for Disease Resistance in Wheat and Barley Breeding. *Phytopathology*. 102(6).
- Vos P., Hogers V., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters J., Pelman J., Kuiper M. y Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids. Res.* 23:4407–4414.
- Zhichang Z., Wanrong Z., Jinping Y., Jianjun Z., Xufeng L.Z. y Yang Y. (2010). Over-expression of Arabidopsis DnaJ (Hsp40) contributes to NaCl-stress tolerance. *African Journal of Biotechnology*. 9(7): 972-978.

Zizumbo-Villareal D., Ruiz-Rodriguez M., Harries H. y Colunga-Garcia M. (2006). Population genetics, lethal yellowing disease, and relationships among Mexican and imported Coconut Ecotypes. *Crop Sci.* 46: 2509-2516.

Zizumbo-Villareal D., Colunga-García M., Fernández-Barrera M., Torres Hernández N. y Oropeza-Salín C. (2008). Mortality of Mexican coconut germplasm due to lethal yellowing. *Plant Genetic Resources Newsletter.* 156: 23-33.

CAPÍTULO V

5.1 DISCUSIÓN GENERAL

Las plantas continuamente están expuestas al ataque de bacterias, virus, hongos, invertebrados e incluso de otras plantas. Debido a su incapacidad de moverse y poder escapar, cada célula de la planta ha desarrollado tanto una capacidad de defensa preformada o inducible. El estudio de estos mecanismos de defensa en plantas puede permitirnos implementar soluciones prácticas sustentables para el control de las principales enfermedades en los cultivos agrícolas, así como también pueden ayudar a elucidar los mecanismos de señalización por el cual las células de las plantas responden a alguna situación de estrés. Los estudios moleculares enfocados a la identificación de genes con aplicaciones biotecnológicas, entre ellos, los genes involucrados en los mecanismos de defensa cada vez van tomando mayor relevancia. Actualmente los estudios de genómica y transcriptómica han acelerado este proceso y han permitido analizar grandes volúmenes de información respecto a la biología molecular, principalmente en cultivos de gran importancia comercial. En cocotero, su genoma aún no ha sido reportado, sin embargo, en años recientes ya empiezan a surgir los primeros estudios en el área de la transcriptómica (Fan *et al.*, 2013, Huang *et al.*, 2014, Nejat *et al.* 2015) con la finalidad de identificar genes de utilidad para el mejoramiento genético de este cultivo, sobre todo los involucrados en la resistencia a las principales enfermedades que lo afectan, tal es el caso del amarillamiento letal.

En este estudio, nosotros caracterizamos a una superfamilia de secuencias del tipo NBS-LRR, las cuales en otras especies han servido de fuente para la identificación de genes de resistencia contra patógenos. La caracterización fue realizada a partir del ADN genómico de tres ecotipos de cocotero, así como también en tres transcriptomas reportados por Huang *et al.*, (2014). En total identificamos más de 200 secuencias, todas ellas presentaron los motivos característicos de las proteínas de resistencia del tipo NBS-LRR y estuvieron relacionadas filogenéticamente con genes *R* de otras especies cuya resistencia ante patógenos ya es conocida. Los estudios de evolución demostraron secuencias especie-específicas de cocotero, lo cual indica que posiblemente sean genes *R* específicos para patógenos que afectan solamente a esta especie. Por otro lado, también pudimos identificar

secuencias familia-específica, entre especies de palmas, lo que sugiere que se originaron a partir de un ancestro en común y que posiblemente reconozcan a patógenos específicos de esta familia de plantas. Por último, algunas de las secuencias de cocotero, estuvieron relacionadas filogenéticamente con secuencias de especies de dicotiledóneas, indicando que posiblemente se originaron antes de diversificarse entre monocotiledóneas y dicotiledóneas.

A pesar de que existen y que siguen surgiendo un gran número de estudios de genes NBS-LRR en muchas especies de plantas, estudios de expresión de estos genes que nos permitan conocer la forma en cómo se regulan y cómo responden ante el ataque de patógenos son muy limitados. En este estudio, nosotros identificamos que las secuencias NBS-LRR modifican su expresión en respuesta a la aplicación exógena del ácido salicílico y que cada secuencia responde de manera diferente a esta fitohormona. Por otra parte, en plántulas no tratadas con AS, los perfiles de expresión fueron constitutivos y en bajos niveles, lo cual coincide con la hipótesis de que los mecanismos de defensa se encuentran preformados antes de que ocurra el ataque de un patógeno (Van Der Biezen y Jones 1998).

En cocotero, la falta de una estrategia para infectar palmas con el fitoplasma del ALC bajo condiciones controladas impiden realizar muchos estudios de interacción planta-patógeno, sin embargo, los estudios de expresión en campo podrían darnos un acercamiento sobre algunos de los genes que pudieran estar involucrados en esta interacción. En este estudio nosotros también evaluamos la expresión de 4 RGCs en palmas de campo sanas y enfermas de ALC y pudimos observar que durante la interacción palma-fitoplasma ocurren cambios en los perfiles de expresión de estas secuencias. Estos resultados sugieren que los genes de resistencia del tipo NBS-LRR son inducibles durante una infección por fitoplasma, sin embargo, aún no puede asegurarse el papel de estos genes como responsables de la resistencia en cocotero.

Dentro de los programas de mejoramiento genético del cocotero el interés por contar con marcadores moleculares que puedan permitirnos identificar ecotipos de cocotero resistentes o susceptibles al amarillamiento letal, cada vez toma más relevancia. En este estudio nosotros empleamos la técnica de AFLP-RGC para identificar polimorfismos entre ecotipos resistentes y susceptibles al ALC. Con esta técnica logramos identificar al marcador *AFLP-RGC 1-05* el cual resultó ser específico y reproducible para el ecotipo Alto

del Oeste Africano (WAT), altamente susceptible al amarillamiento letal. La presencia de este marcador también fue observado en palmas enfermas con amarillamiento letal. Estos resultados sugieren que este marcador posiblemente está asociado con palmas susceptibles al ALC, sin embargo, es necesario realizar más validaciones en campo para confirmar este resultado. Algunas litaciones que fueron encontradas para la validación de este marcador, fue que el ADN extraído de otros tejidos que no fuera de la hoja espada, tales como hojas viejas o tejido de tronco, la validación de este marcador no fue reproducible debido a la cantidad e impureza del ADN que se extrae en estos tejidos. Nosotros recomendamos realizar algunas pruebas de purificación del ADN de hoja adulta, tallo e incluso raíz para probar la amplificación de este marcador. Esto ayudaría al análisis de palmas muy altas en donde la hoja espada es difícil de alcanzar.

Por otra parte, es importante seguir con la búsqueda de marcadores moleculares no solamente asociados a resistencia o susceptibilidad, más bien a otras características de interés agronómica, tales como marcadores moleculares asociados con producción, con la calidad de la copra, la calidad y cantidad de agua, el tamaño de la planta y de la nuez, entre muchos otros. Para lograr esto, es necesario aplicar y desarrollar nuevas técnicas para la búsqueda de marcadores moleculares e inclusive realizar estudios de omica, tales como, secuenciación de genomas, transcriptomas y proteomas, estos estudios aceleraría todo este proceso.

Finalmente el contar con un paquete biotecnológico de marcadores moleculares que puedan ser capaces de discernir entre ecotipos resistentes o susceptibles a los principales patógenos del cocotero, principalmente al ALC, podría tener grandes aplicaciones en la agricultura de este cultivo, sobre todo para la selección de germoplasma cuya respuesta al ALC es desconocido. Actualmente, contamos con 12 variedades de cocotero, que fueron introducidas a México, pero aún no sabemos como puedan responder al ALC, el contar con el marcador *AFLP-RGC 1-05* el cual se ha visto que puede discernir entre ecotipos de cocotero resistentes y susceptibles al amarillamiento letal podría ahorrar mucho tiempo y recursos. Por otra parte, también podría ser utilizado para evaluar otras variedades antes de ser introducidas al país o para caracterizar el germoplasma en otros países en las cuales se desconoce cual podría ser su respuesta al amarillamiento letal. En conclusión, este trabajo, proporciona una gran fuente de secuencias para la caracterización de genes de

resistencia en este cultivo y representa las bases para el inicio de un programa de mejoramiento genético asistido por marcadores moleculares.

5.2 CONCLUSIONES GENERALES

1. En este estudio se caracterizaron más 200 secuencias del tipo NBS-LRR. Todas las secuencias presentaron los motivos característicos de esta familia de genes y estuvieron filogenéticamente relacionados con genes de resistencia conocidos en otras especies de plantas.
2. Se observaron cambios en los niveles y perfiles de expresión de algunas de estas secuencias cuando plántulas de cocotero fueron expuestas al ácido salicílico (hormona involucrada en los mecanismos de defensa). Estos resultados sustentan el posible papel de estas secuencias dentro de los mecanismos de defensa del cocotero.
3. Los estudios de expresión de 4 RGCs en palmas de campo, revelaron cambios en sus niveles de expresión en palmas sanas y enfermas de amarillamiento letal. La expresión observada en los diferentes tejidos de cocotero, indican que existe una expresión tejido específico.
4. Los estudios evolutivos de los RGCs de cocotero con otras especies de palmas sugieren secuencias específicas para *Cocos nucifera*, así como también secuencias compartidas entre especies de la familia de arecáceas. La relación filogenética de algunos RGCs de cocotero con brasicáceas sugieren la existencia de estas secuencias mucho antes de su diversificación entre mono y dicotiledóneas.
5. Empleado la técnica de AFLP-RGC identificamos diferentes fragmentos polimórficos entre ecotipos de cocotero resistentes y susceptibles al ALC y a partir de ellos, se logró generar el marcador *AFLP-RGC 1-05* el cual fue específico y reproducible para el ecotipo susceptible al amarillamiento letal.

5.3 PERSPECTIVAS

En el presente estudio se caracterizó una fuente importante de secuencias candidatas a genes de resistencia, el cual abre la posibilidad para desarrollar diferentes trabajos que contribuyan al entendimiento de los mecanismos de defensa en cocotero y que permitan implementar estrategias para el control de las principales enfermedades en este cultivo.

- Es importante realizar estudios de expresión complementarios en plántulas de cocotero tratadas con otras fitohormonas involucradas en los mecanismos de defensa, tales como el ácido jasmónico y el etileno. Esto podría ayudar a entender cuáles son las principales vías por las cuales se regula la expresión de estos genes. Por otro lado, para conocer más a detalle el papel de estos genes durante la interacción palma-fitoplasma, sería importante evaluar sus perfiles de expresión durante los diferentes estadios de la enfermedad del amarillamiento letal. Los transcritos que muestren alguna asociación con esta enfermedad serían excelentes candidatos para el aislamiento del gen completo por la técnica RACE y realizar estudios posteriores sobre transformación genética de especies modelos y del cocotero.
- La respuesta de estos genes ante los diferentes patógenos que afectan al cocotero es prácticamente desconocida, sin embargo, el estudiar su expresión en palmas afectadas por estos patógenos podría permitirnos identificar algún polimorfismo transcripcional posiblemente asociado con la resistencia a alguno de estos agentes.
- La búsqueda de marcadores moleculares en cocotero asociados a diferentes características de interés agronómica siguen siendo pilares para el establecimiento de un programa de mejoramiento genético en este cultivo, por lo tanto, es importante seguir trabajando en este objetivo, pero evaluando otras técnicas moleculares tales como, MSAP (Polimorfismos de amplificación sensitivos a metilación), SNP (Polimorfismos de un solo nucleótido), SSR (simples secuencias repetidas), entre otros. El mejoramiento genético del cocotero asistido por marcadores moleculares acortaría el tiempo para la obtención de variedades con características agronómicas deseables.

- Por último, los estudios de genómica y transcriptómica en la interacción palma-patógeno generaría una gran cantidad de información que podría ser aplicado para el mejoramiento genético del cocotero, tales como:
 - ✓ Identificación de genes con potencial biotecnológico
 - ✓ Elucidación de rutas activadas durante esta interacción
 - ✓ Estudios de evolución de las diferentes familias de genes
 - ✓ Análisis de polimorfismos a nivel genoma
 - ✓ Identificación de genes que pudieran ser blancos para el control de los patógenos.

5.4. REFERENCIAS BIBLIOGRÁFICAS

- Fan H., Xiao Y., Yang Y., Xia W., Mason A.S., *et al.* (2013). RNA-Seq Analysis of *Cocos nucifera*: Transcriptome Sequencing and Subsequent Functional Genomics Approaches. PLoS ONE. 8(3): e59997
- Huang Y., Lee C.P., Fu F.L., Chang B. y Matzke A. (2014). De Novo Transcriptome Sequence Assembly from Coconut Leaves and Seeds with a Focus on Factors Involved in RNA-Directed DNA Methylation. G3: Genes, Genome, Genetics 4: 2147-2157.
- Nejat N., Cahill D.M., Vadamalai G., Ziemann M., Rookes J. y Naderali N. (2015). Transcriptomics-based analysis using RNA-Seq of the coconut (*Cocos nucifera*) leaf in response to yellow decline phytoplasma Infection. Mol Genet Genomics.
- Van Der Biezen E. y Jones JDG. (1998). Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci. 23:454–456.

ANEXOS
Anexo 1. Iniciadores específicos diseñados a partir de la secuencias de los fragmentos polimórficos identificados por la técnica de AFLP-RGC.

AFLP-RGC 1-01	Inicio F1: GCGCATCTTCTCTTTTGTGG Inicio F2: AGTGGCTGCGCATCTTCTCT Termino R1: GGAATGTGGACATCTTCACTTGG Termino R2: GAGCAACTCATCACCTTCTCTG
AFLP-RGC 1-02	Inicio F1: AGTGTGAGGCAGCTGATTTG Inicio F2: TGTGAGGCAGCTGATTTGTCT Termino R1: GCCAAAACTGAAAACGAGA Termino R2: AACTGAAAACGAGAAGAAAG
AFLP-RGC 1-03	Inicio F1: GATTGAAGCCAACCCGAAGA Inicio F2: GCCTTTAGGGTTGCCTCAGAA Termino R1: AATTTCTTGAGGATTTGGAAGAAGG Termino R2: CGAACACCTTTTGACATTCTG
AFLP-RGC 1-04	Inicio F1: GCGGTCATCAGACAGAGACATTT Inicio F2: CTCTAAGCGGTCATCAGACAGAGA Termino R1: TGTAATCTCTTTGAACATCGCTGAC Termino R2: CGCGGTTGTAATCTCTTTGAAC
AFLP-RGC 1-05	Inicio F1: CACCCTCATCTTCCAAATCCTC Inicio F2: CCACTCTTCCTTTACCCTCAT Termino R1: CTGCTTGGGCGATTAGAAGG Termino R2: AGCCTGCTTGGGCGATTA
AFLP-RGC 1-06	Inicio F1: ATCCATCTCGTCCCCATACCTT Inicio F2: CTCGATATTTGGCCTCCACCTA Termino R1: ACCGACTTCAGATGGAAAGCTG Termino R2: GCGTACCAATTCAACCGACTTC
AFLP-RGC 1-07	Inicio F1: GAATGCGTATCGGCCTAGAT Inicio F2: TGAATGCGTATCGGCCTAGAT Termino R1: AGGCAATCTTGGATGGTCTG Termino R2: CGAGGCAATCTTGGATGGT
AFLP-RGC 1-08	Inicio F1: ATTGGGGGTGCGTTAGGAT Inicio F2: TTATTGGGGGTGCGTTAGGA Termino R1: TGGTCCTAACATGCCCCCTACA Termino R2: AATGGTCCTAACATGCCCCCTAC
AFLP-RGC 1-09	Inicio F1: AAAGGCTAGCCCCACCCCAT Inicio F2: GGCTAGCCCCACCCCATGCA Termino R1: CCGGTACACCCCATACTGAA Termino R2: ACACCCCATACTGAACCGGA

AFLP-RGC 1-10	Inicio F1: TTATTGCAGTTGGGTCACCA Inicio F2: GATCATTTATTGCAGTTGGGTCAC Termino R1: GACCTTGGCTCGAAGATCCT Termino R2: TGACCTTGGCTCGAAGATCC
AFLP-RGC 1-11	Inicio F1: CCTCACAGACCAAAGACATGGA Inicio F2: GTCAGAACCTCACAGACCAAAGAC Termino R1: CATTTGATGTAGTCAGGTTGTCGTG Termino R2: AGGCGGGACTGTCTAATATGGT
AFLP-RGC 1-12	Inicio F1: GAAGCTTGACATCTGAGGAGCA Inicio F2: TCATCTAGCTTTCACCACCATCC Termino R1: GAGGCTCAGATGTTGGTAGAGTAGATG Termino R2: GGAGGCTCAGATGTTGGTAGAGTAG
AFLP-RGC 1-13	Inicio F1: AGCAACAACCTATAGCGATGG Inicio F2: AACAACTATAGCGATGGTAA Termino R1: CTAAGGTGGGGGAAAGTGGG Termino R2: GGTGGGGGAAAGTGGGGGTG
AFLP-RGC 1-14	Inicio F1: GATCATGATGGCATCGAAAGG Inicio F2: GGATCATGATGGCATCGAAAG Termino R1: GGTGGATGGAATATGCGCTAGTA Termino R2: GCGGAGGTGGATGGAATATG
AFLP-RGC 1-15	Inicio F1: CAAGCATGCATATATATATA Inicio F2: GCATGCATATATATATATAC Termino R1: AGCACAATTTCCCTCCCTAGA Termino R2: ACAATTTCCCTCCCTAGATGG
AFLP-RGC 1-16	Inicio F1: GTCACTAGAGGACTTGAAGCTTTTCG Inicio F2: CCATGTCACTAGAGGACTTGAAGC Termino R1: TTGAGGACCGCCTTGTT Termino R2: AAACCTTGAGGACCGCCTTG
AFLP-RGC 1-17	Inicio F1: GTTGGTTGTTTGC GCGGTGT Inicio F2: GGTTGTTTGC GCGGTGTGGT Termino R1: GAAGGGGACGATAAGGCTGG Termino R2: GGGGACGATAAGGCTGGAAA
