Genes Genom
DOI 10.1007/s13258-016-0422-6

RESEARCH ARTICLE

New insights into the evolutionary history of resistance gene candidates in coconut palms and their expression profiles in palms affected by lethal yellowing disease

Carlos Puch-Hau[1] · Carlos Oropeza[1] · Manuel Góngora-Paredes[1] · Iván Córdova[1] · José Tun-Suárez[2] · Luis Sáenz[1]

Received: 15 February 2016 / Accepted: 26 April 2016

Abstract The nucleotide binding site and leucine rich repeat (NBS–LRR) class of *R* genes is the most comprehensively studied in terms of sequence evolution; however, in coconut palm and, more generally, in the family of Arecaceae, our understanding of the evolution of these genes is rather limited. In this study, disease resistance gene candidates (RGCs) of the nucleotide binding site (NBS) type of coconut palm were used to investigate evolutionary relationships in *Arecaceae*, *Poaceae* and *Brassicaceae* species. The results indicate a species-specific evolution of RGCs in coconut palm. However, strikingly similar RGCs between species of Arecales indicate a 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. Finally the comparative analysis of the expression of four RGCs in healthy coconut palm and those affected with lethal yellowing disease revealed differences in their expression profiles. This study provides new insights for future efforts towards the improvement of disease resistance in coconut palm and other species of Arecaceae.

Keywords Resistance gene candidates · Coconut palm · Lethal yellowing · Evolution · *Cocos nucifera*

✉ Luis Sáenz
vyca@cicy.mx

[1] Biotechnology Unit, Centro de Investigación Científica de Yucatán, Calle 43 No. 130. Colonia Chuburna de Hidalgo, 97200 Mérida, YUC, Mexico

[2] Instituto Tecnológico de Conkal, Mérida, YUC, Mexico

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 legumes (Leguminosae). Among the palm crops, the top three out of the five major domesticated palm species in the world are the African oil palm (*Elaeis guineensis*), the coconut palm (*Cocos nucifera*) and the date palm (*Phoenix dactylifera*), and 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 12 million hectares. The coconut tree is highly versatile, with extensive applications in agriculture and industry; these applications include food, fiber, oil, soil fertilizers, spa ingredients, furniture, fashion accessories, garments, construction and building materials, oleochemicals and biofuels (Fan et al. 2013). Unfortunately, this palm is susceptible to several pathogens that hinder yields including viruses (Rohde et al. 1990), viroides (Hanold and Randles 1991), protozoa (Parthasarathy et al. 1978), nematodes (Griffith 1987), insect pests (Gitau et al. 2009) and mollicutes (Howard and Barrant 1989) such as the '*Candidatus* Phytoplasma palmae', which is responsible for the devastating disease known as lethal yellowing (LY) in the Americas (Harrison and Oropeza 2008). Currently genetic studies have focused on the identification of genes with biotechnological applications, such as those involved in defense mechanisms, which have become increasingly more relevant for this crop (Fan et al. 2013; Huang et al. 2014; Puch-Hau et al. 2015; Nejat et al. 2015).

Plant resistance to a range of pathogenic organisms is conferred by a diverse group of disease resistance

Published online: 03 May 2016

(*R*) genes (Martínez et al. 2003). These *R* genes encode proteins with cytoplasmic nucleotide-binding site and leucine-rich repeat (NBS–LRR) domains and have been associated with effector-triggered immunity (ETI) (Jones and Jones 1997). The NBS domain is involved in signaling and includes several highly conserved and strictly ordered motifs such as P-loop, kinase-2 and GLPL motifs (Belkhadir et al. 2004), which have 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 specificity (Ellis et al. 1999). 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 (Meyers et al. 1999; Richly et al. 2002). 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 been focused on this family of proteins because of their only known function in disease resistance (Meyers et al. 2005).

The past few years have seen remarkable progress in understanding the mechanisms of the *R* gene evolution in Arabidopsis, but notable discoveries have also emerged from studies in other plant species such as wheat, sorghum, maize, brachypodium, rice and cotton, among others (Li et al. 2010; Bouktila et al. 2015; Khan et al. 2015; Gu et al. 2015).

Comparative genomic analyses have 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 of *Carica papaya* and *Cucumis sativus,* contain approximately 50 NBS–LRR genes (Porter et al. 2009; Wan et al. 2013); this is in contrast to the 653 genes that have been identified in *Oryza sativa* (Shang et al. 2009). 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 evolution (Hulbert et al. 2001).

In coconut palm and the family Arecales in general, RGCs of the NBS type have been successfully cloned and characterized (Puch-Hau et al. 2015; Rajesh et al. 2015); however, their evolutionary process remains largely unknown. In fact, studies of NBS sequence evolution in coconut palms may shed some light on breeding strategies for disease control in coconut palm crops. Annotation and identification of RCGs in *Cocos nucifera* (Puch-Hau et al. 2015; Rajesh et al. 2015), *Phoenix dactylifera* (Al-Mssallem et al. 2013) and *Elaeis guineensis* (Foan et al. 2012) will permit comparisons of resistance gene candidates between species of the family Arecales and provide insight into their evolutionary relationships. The comparative analysis of these sequences among Arecales and Poales as monocot plants and Brassicales as dicot plants may allow us to find evolutionary relationships, common features or specific characteristics of *R* genes that have not been reported in this family.

Therefore, the objective of this study was to characterize the NBS sequences of coconut palms, to investigate their evolutionary relationships with NBS sequences of *Arecaceae, Poaceae and Brassicaceae* species, to estimate their evolutionary rates and study their gene expression in healthy palms and plants affected by LY disease in the field. 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 no resistance genes in this plant family have been reported to date.

Materials and methods

Study of NBS sequences

Identification of NBS sequences

The NBS sequences used in this study were retrieved from the transcriptome of the embryo (Accession No. GBGL00000000), endosperm (Accession No. GBGK00000000) and coconut leaves (Accession No. GBGM00000000) reported by Huang et al. (2014) as well as a previous study that we developed (Puch-Hau et al. 2015). All NBS sequences were downloaded from the NCBI database (National Center for Biotechnology information). 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).

Phylogenetic relationships and calculation of K_a/K_s ratios

The NBS domain was used for phylogenetic analysis. Sequences with less than 50 % of the full-length NBS were excluded. Multiple alignments of amino acid sequences were performed using ClustalW with default options (Larkin et al. 2007). Phylogenetic trees were constructed by the neighbor-joining (NJ) method (Saitou and Nei 1987) with Poisson correction using the NJ algorithm implemented with version 6.0 of the Molecular Evolutionary Genetic Analysis software (MEGA) (Tamura et al. 2013). 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. The analysis was calculated with the Nei–Gojobori model (Nei and Gojobori 1986) using version 6.0 of MEGA (Tamura et al. 2013). The ratio of non-synonymous to synonymous nucleotide substitution (K_a/K_s) was detected in the NBS region between sequences of the same clade and in sequences of different clades. It is widely accepted that a K_a/K_s ratio >1 indicates positive selection, a ratio of 1 indicates neutral evolution, and a ratio <1 indicates purifying selection.

Detection of DNA in LY phytoplasma

Plant material and DNA extraction

Three adult asymptomatic coconut palms and palms with symptoms of LY disease growing in field conditions on the northern coast of the Yucatan Peninsula (21°22′N, 89°00′W) were sampled to determine expression profiles of four NBS sequences in leaf, stem and root tissues. 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. (1983): nut fall (1) and appearance of necrotic inflorescences (2). Additionally, we used 12-month-old plantlets cultured in liquid medium Y3 (Eeuwens 1976) as the control of expression. The DNA for the diagnosis of LY phytoplasma was extracted following the protocol described by Harrison et al. (1994) with minor modifications.

Real time-PCR assay

For the detection of LY phytoplasma DNA from root tissues of asymptomatic and symptomatic coconut palms, the TaqMan LY16S assay reported by Córdova et al. (2014) was used. Reactions were performed in 20 µl of TaqMan Universal PCR master mix with AmpErase Uracil *N*-Glycosylase (UNG) (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 s 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).

Transcript expression analysis

RNA extraction and cDNA synthesis

Total RNA was extracted from coconut palms (leaf, stem and root) following the methodology described by Montero-Córtes (Montero-Cortes et al. 2010). The RNA samples were treated with DNase-free (Ambion, Foster city, CA, USA) to remove the genomic DNA. The RNA quality was checked on a ethidium bromide-stained formaldehyde agarose gel (1.2 %) by examining integrity of the rRNA bands followed by quantification using a 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 manufacturer's instructions.

Real-time PCR assay

Specific primers were designed for four randomly selected RGCs to study expression in healthy palms and those diseased with LY field (CnRGC121: FW 5́-ACTCCATCC TCTTCCCGCAGTTGT-3́ and RV 5́-AGAGAATGTCCC CACTACCCAGA-3́, for CnRGC46: 5′-TCCCACCTCTC CTCAGCTGTTTC-3′ and 5′-CCTTGACTGTTTACCAG AAGACCA-3′, for CnRGC22: 5′-TTGCTTCTCTCCTT CGACCTTAAAACC-3′ and 5′-TATTAGAGCGAAGAT CCACCAAGTC-3′ and for CnRGC42: 5′-CTTTCGGA CATTGGCATTTTAGTC-3′ and 5′-CTAGCAGTTGAAA GCCGTGTTCTGG-3′). We used the 18S rRNA gene as a normalizer of the expression (FW 5′-CGGCTACCACA TCCAAGGAA-3′ and RV 5′-GCTG GAATTACCGC GGCT-3′) in each sample.

Real-time RT-PCR was performed using a platinum SYBR Green qPCR superMix-UDG (Invitrogen, USA) and analyzed with an iCycler IQ real-time PCR detection system (Bio-Rad, USA) according to the manufacturer's instructions. The 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 $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001), with appropriate validation experiments performed beforehand (Real-Time PCR Applications Guide BIORAD). The results are expressed as relative expression levels. Data represent the mean ± standard deviation (SD) of each sample by triplicate.

Results

Identification of NBS sequences from transcriptomic sequences

To characterize the most RGCs of coconut palm, we identified 62 RGC sequences from embryo, endosperm and leaf tissue transcriptomes reported by Huang et al. (2014). Of the 62 RGC sequences, there were 53 open reading frames (ORFs) and 9 stop codons. Of the three tissues, the embryonic 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). The largest number of NBS genes with ORFs (17 sequences) was found in the leaf transcriptome compared with the embryo and endosperm (15 sequences). The total number of RGCs identified in the three coconut transcriptomes and others characterized in a previous study by our group (Puch-Hau et al. 2015) are described in Table 1.

The RGCs with an identified ORF in the three transcriptomes 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 (Fig. 1).

Analysis for the Conserved Motif Structures

To analyze the structural diversity of the conserved motifs in the RGCs, we used the MEME motif detection software. All previously characterized RGCs (Puch-Hau et al. 2015) and those identified in the three transcriptomes were used to identify the conserved motifs. We found 20 putative, conserved motifs among them. The detailed motif sequences are 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) groups (Meyers et al. 1999). 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 and RNBS-D-non TIR) identified previously, confirming that the NBS domain is 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 the LRR domain. In addition, some of the sequences presented unknown motifs (5, 6, 10, 11, 12, 13, 15, 17, 18, 20 and 29 motifs) in the CC and NBS domains.

Organization and phylogeny of NBS sequences among species of Arecaceae

One hundred forty one, fifty nine, and thirty four RGCs from coconut palm, date palm and oil palm, respectively, were used to clarify the phylogenetic relationships between Arecaceae species (Fig. 2). The generated tree showed nine major groups that are quite distinct from each other. Of the nine groups, seven had sequences of the three palm species and showed clear orthologous relationships between each clade, suggesting that these genes are conserved between Arecaceae species. Interestingly, we revealed two groups (VIII and IX) comprised of only coconut palm sequences, whose homologs were classified into a species-specific gene family.

The largest group of sequences (50 sequences) was found in clade III of the coconut palm, while clade I had the minority of sequences (1 sequence). For the date palm, most sequences were from group VII (15 sequences) and the minority in the group I (1 sequence). In the oil palm, the majority of sequences were in group VI (13 sequences); a minority were from groups IV and VII (2 sequences). The great variation in the number of NBS sequences in each

Table 1 Number of RGCs identified in coconut palm

Ecotype	Source	Total transcripts	Total of NBS–LRR sequences	# sequences with ORF	# sequences with stop codons	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

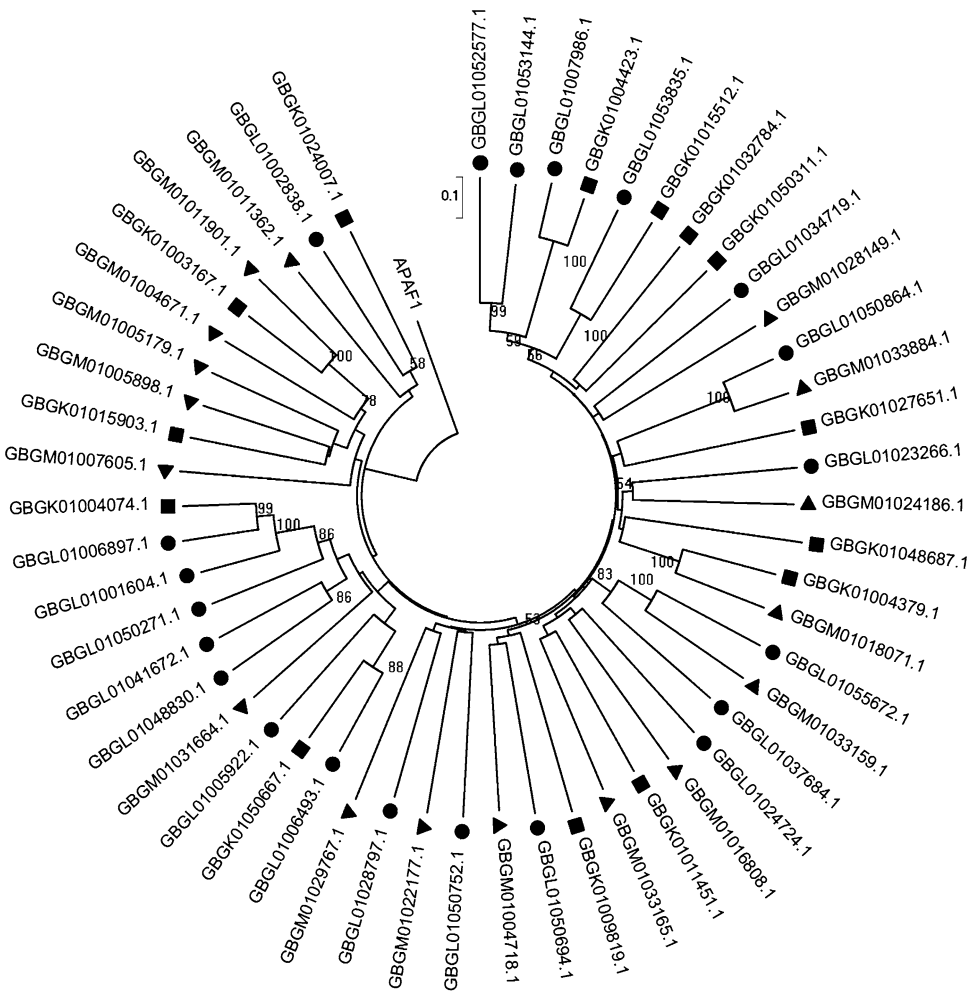


Fig. 1 Phylogenetic tree of RGAs obtained from three transcriptome of coconut palm (embryo, endosperm and leaf). The RGAs from embryo transcriptome are show as *solid black circles*, the endosperm in white squares and the leaf in *solid black triangles*. The tree was constructed by the neighbor-joining method with a statistical analysis of bootstrap with 1000 replicates. The values of bootstrap >50 % are shown on the branches. APAF-1 was used as an outgroup

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 diverse pathogens (Li et al. 2010).

Comparing evolutionary rates among NBS sequences of coconut palm

In genetics, the K_a/K_s ratio is used as an indicator of selective pressure acting on a protein-coding gene (Michelmore and Meyers 1998). The groups with less than three sequences were excluded (groups I and II) for this analysis. The average K_a/K_s 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 K_a/K_s ratio between NBS sequences of different clades. The results showed that most K_a/K_s 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 K_a/K_s 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 palms.

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	IIESVTGEQCQLTNMDAMQHELKEQLKGK	CC	–	29	2.8e−283
NB-ARC	Motif 4	GGGGKTTLAQEIYNH	NB-ARC	P-loop	15	1.3e−1136
	Motif 7	RRIDYHFHVRIWVCVSQDFDV	NB-ARC	RNBS-A-non-TIR	21	2.5e−964
	Motif 17	DLMRRLIKDLYNENRDILPGNIDAM PCDSLAEVLHGYLQQK	NB-ARC	–	41	1.0e−172
	Motif 12	IQQMIADRLGVPWKDNDSEILRAKMLLRA	NB-ARC	–	29	2.1e−419
	Motif 10	INLLKDIYDGTTGDLAGDQSKSSLEPKVE	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	ANGCKIIITTRNEQVCRQMGA	NB-ARC	RNBS-B	21	4.8e−1446
	Motif 18	VGIPRPNK	NB-ARC	–	8	5.9e−171
	Motif 6	LSEEDCWSLFCKKAF	NB-ARC	–	15	9.0e-834
	Motif 1	HQHLEDIGMEIVKKCHGLPLA	NB-ARC	GLPL	21	1.2e−1652
	Motif 11	AKVVGSAMRHETDPREWRHAL	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	PSHLKQCFAYCSMFPKDYHFDKDYL VQCWMAQGFIQPQGSN	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	PDSLGNLIHLRYLNMSYTKISVMPESIGNL RNLQFLMMSYC	LRR1	Motif1 LDL	41	3.9e−165

[a] Motifs are listed according to their rank derived from the program MEME analysis

[b] Consensus amino acid sequence derived from MEME program analysis

Phylogenetic relationships between species of Arecaceae, Poaceae and Brassicaceae

To study the evolutionary relationships among RGCs 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 did not show correct alignment were excluded from this analysis (Fig. 3).

To clarify the phylogenetic relationships, three types of clades were identified among the eight species: a species-specific clade, a family-specific clade and a complex gene clade. If the homologous 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 homologous genes of species from the same family. The remaining clades were categorized as complex gene clades, which include homologous 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 *Oriza sativa* and 6 in *Sorghum bicolor*) and only 1 in Brassicaceae (comprising 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 *Oriza 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 homologous genes of Arecales and Poales, 3 between Arecales and Brassicales and 2 between Poales and Brassicales, but none were composed of NBS sequences of the three families. These results indicate that the sequences of Arecaceae have more phylogenic relationships with the Poaceae family (9 complex clades) than with the Brassicaceae family (3 complex clades) (Table 4). This result was expected because both families (Arecaceae and Poaceae) belong to the class of monocots; however,

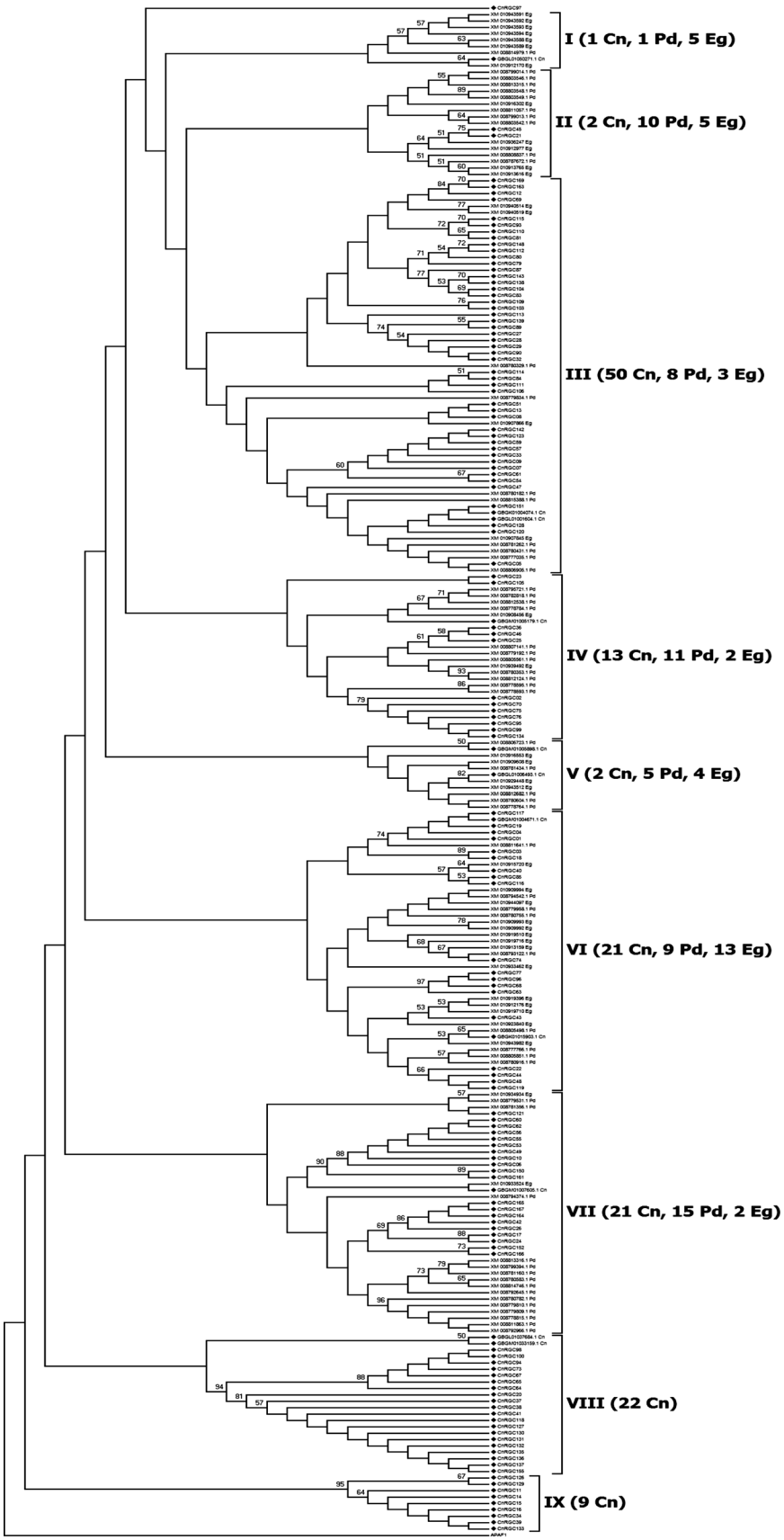


Fig. 2 Phylogenetic relationship of RGCs from coconut palm, date palm and oil palm. The 234 RGCs were grouped in nine main clades (I to IX), seven of which consisted of orthologous sequences of the three species and two of which consisted of species specific to coconut palm. The *black rhombuses* represent RGAs of coconut palm. The tree was constructed by the neighbour-joining method with a statistical analysis of bootstrap with 1000 replicates. The values of bootstrap >50% are shown on the branches. APAF-1 was used as an outgroup

Table 3 K_a/K_s ratios for pairwise comparisons among members of each clade and among RGCs of the different clades of coconut palm

	Clade name	K_a	K_s	K_a/K_s
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

homologous sequences were also identified between monocots and dicots (Arecaceae–Brassicaceae and Poaceae–Brassicaceae).

Expression analysis of four NBS sequences of coconut palms infected with LY-phytoplasma

For the investigation of gene expression in the field, four NBS sequences previously reported by our group (Puch-Hau et al. 2015) were randomly selected: CnRGC22, CnRGC42, CnRGC46 and CnRGC121. Symptomatic palms showing disease stages 1 and 2 and asymptomatic palms were collected in the locality of San Crisanto, Yucatan, Mexico (21°22′N, 89°00′W). To confirm whether these palms were infected, extracts from trunks were analyzed using an LY 16S TaqMan/Real-time PCR assay (Córdova et al. 2014). The three symptomatic palms tested positive for the LY-phytoplasma, and the asymptomatics tested negative (Table 5).

Because field conditions are highly variable, in vitro plantlets were used as controls for expression. The expression profiles of four genes were evaluated in the three different tissues (leaf, stem and root) of symptomatic and asymptomatic coconut palms. The CnRGC22 gene was highly expressed in the stem tissue of diseased coconut palms, while no changes were observed in the three tissues analyzed from healthy palms. The same pattern was observed in the three evaluated palms. CnRGC42 had the highest expression level in the root tissues of LY-affected coconut palms, while only a minor change in the expression was observed in stem tissues of healthy palms. CnRGC46 had an elevated expression level in the leaf tissues of LY-affected coconut palms compared to healthy palms. However, in healthy palms, CnRGC46 showed a similar pattern in all of the three analyzed tissues. Interestingly, CnRGC121 was highly expressed in roots of healthy palms, and the same pattern was observed in the three analyzed coconut palms. However, in LY-affected palms, CnRGC121 was highly expressed in leaf and stem tissues, but not in root tissues. This profile was observed in two coconut palms for leaf tissues and three for stem tissues. No changes were observed for in vitro plantlets in the expression of these four sequences (Fig. 4). These results show that the NBS sequences modify their expression during plant-phytoplasma interactions.

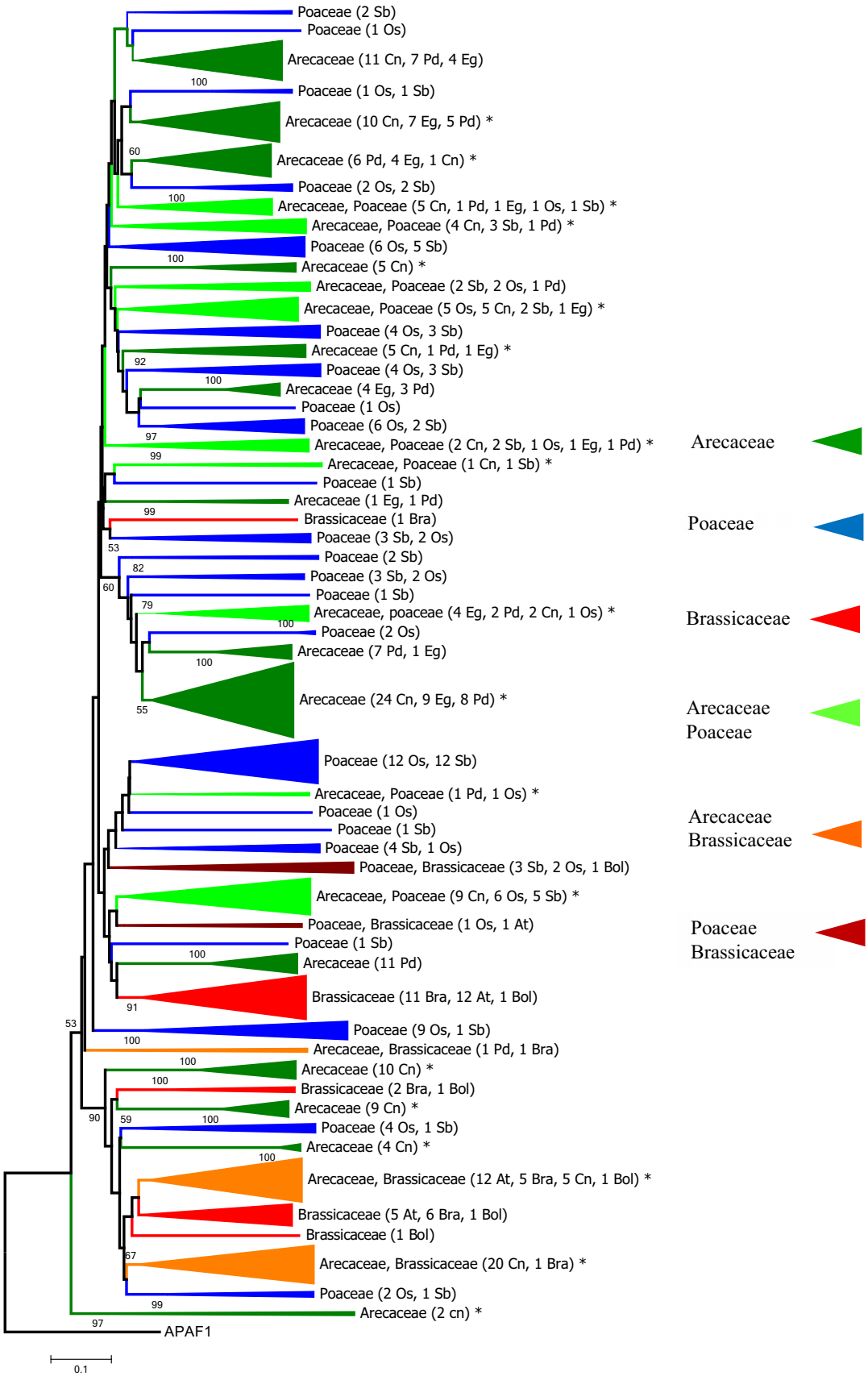
Discussion

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 (Meyers et al. 2003; Porter et al. 2009; Li et al. 2010; Lozano et al. 2015); however, only two studies of NBS sequences have been reported in coconut palms (Puch-Hau et al. 2015; Rajesh et al. 2015). These studies only focused on cloning and characterization, but studies of diversity, distribution and phylogenetic relationships of NBS sequences of coconut palms with species of the Arecales, Poales and Brassicales have not been reported.

The number of these genes always varies in different plants, including in plants within the same family or genus. For example, in the grass species, the numbers are quite different; they range from approximately 100 in maize to 500 in rice (Li et al. 2010; Luo et al. 2012), while the numbers in Brassicales range from 92 in *Brassica rapa* to 149 in *Arabidopsis thaliana*. In palms, of the two sequenced genomes, only 144 NBS genes have been reported in *Phoenix dactylifera,* and there are no precise data for these genes in the genome of *Elaeis guineensis.* In coconut palms, we identified a total of 231 NBS sequences, of which 169 were reported in a previous study (143 with ORFs, and 23 with stop codons) (Puch-Hau et al. 2015), and 62 were reported in this study (53 with ORFs, 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 size (Wan et al. 2012). Therefore, the likely number of NBS genes is higher in the genome of the coconut palm compared with date palm, but this could be due more to the difference in the genome sizes (~2.5 Gb for the coconut palm and ~671 Mb for the date palm) (Al-Mssallem et al. 2013; Alsaihati et al. 2014) than to the nature of the selective pressure that is imposed by their pathogens as well as to their very different life histories.

Transcriptome sequencing of coconut palms provides abundant sequence information for identifying the decisive role of these genes in individual tissues for conferring resistance against pathogens. We identified NBS transcripts from the transcriptome reported by Huang et al. (2014) in the embryo, endosperm and leaf tissues of healthy coconut palms (Table 1). All the transcripts were different in each tissue (Fig. 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 a 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* genes usually result from tandem duplications of paralogous genes and have been observed in NBS–LRR encoded in the Arabidopsis and rice genomes (Richly et al.


Poaceae (2 Sb)
Poaceae (1 Os)
Arecaceae (11 Cn, 7 Pd, 4 Eg)
100
Poaceae (1 Os, 1 Sb)
Arecaceae (10 Cn, 7 Eg, 5 Pd) *
60
Arecaceae (6 Pd, 4 Eg, 1 Cn) *
Poaceae (2 Os, 2 Sb)
100
Arecaceae, Poaceae (5 Cn, 1 Pd, 1 Eg, 1 Os, 1 Sb) *
Arecaceae, Poaceae (4 Cn, 3 Sb, 1 Pd) *
Poaceae (6 Os, 5 Sb)
100
Arecaceae (5 Cn) *
Arecaceae, Poaceae (2 Sb, 2 Os, 1 Pd)
Arecaceae, Poaceae (5 Os, 5 Cn, 2 Sb, 1 Eg) *
Poaceae (4 Os, 3 Sb)
Arecaceae (5 Cn, 1 Pd, 1 Eg) *
92
Poaceae (4 Os, 3 Sb)
100
Arecaceae (4 Eg, 3 Pd)
Poaceae (1 Os)
Poaceae (6 Os, 2 Sb)
97
Arecaceae, Poaceae (2 Cn, 2 Sb, 1 Os, 1 Eg, 1 Pd) *
99
Arecaceae, Poaceae (1 Cn, 1 Sb) *
Poaceae (1 Sb)
Arecaceae (1 Eg, 1 Pd)
99
Brassicaceae (1 Bra)
Poaceae (3 Sb, 2 Os)
53
Poaceae (2 Sb)
82
Poaceae (3 Sb, 2 Os)
60
Poaceae (1 Sb)
79
Arecaceae, poaceae (4 Eg, 2 Pd, 2 Cn, 1 Os) *
100
Poaceae (2 Os)
Arecaceae (7 Pd, 1 Eg)
100
Arecaceae (24 Cn, 9 Eg, 8 Pd) *
55
Poaceae (12 Os, 12 Sb)
Arecaceae, Poaceae (1 Pd, 1 Os) *
Poaceae (1 Os)
Poaceae (1 Sb)
Poaceae (4 Sb, 1 Os)
Poaceae, Brassicaceae (3 Sb, 2 Os, 1 Bol)
Arecaceae, Poaceae (9 Cn, 6 Os, 5 Sb) *
Poaceae, Brassicaceae (1 Os, 1 At)
Poaceae (1 Sb)
100
Arecaceae (11 Pd)
Brassicaceae (11 Bra, 12 At, 1 Bol)
91
53
Poaceae (9 Os, 1 Sb)
100
Arecaceae, Brassicaceae (1 Pd, 1 Bra)
100
Arecaceae (10 Cn) *
100
Brassicaceae (2 Bra, 1 Bol)
Arecaceae (9 Cn) *
90
59
100
Poaceae (4 Os, 1 Sb)
Arecaceae (4 Cn) *
100
Arecaceae, Brassicaceae (12 At, 5 Bra, 5 Cn, 1 Bol) *
Brassicaceae (5 At, 6 Bra, 1 Bol)
Brassicaceae (1 Bol)
67
Arecaceae, Brassicaceae (20 Cn, 1 Bra) *
Poaceae (2 Os, 1 Sb)
99
Arecaceae (2 cn) *
97
APAF1
0.1
Arecaceae
Poaceae
Brassicaceae
Arecaceae
Poaceae
Arecaceae
Brassicaceae
Poaceae
Brassicaceae

◀**Fig. 3** Neighbor-joining tree showing the phylogenetic relationship among RGCs from different species. The complete tree was based on 434 sequences. The species-specific, family-specific and complex clades were collapsed into single branches and are show in different *colors*. The number on the branches indicates the percentages of 1000 bootstrap replications and only those with >50% support are shown. APAF-1 was used as an outgroup

2002; Michelmore and Meyers 1998; Zhu et al. 2002; Meyers et al. 2003; Baumgarten et al. 2003; Zhou et al. 2004). In addition, several clades include a mixture of NBS sequences from *Elaeis guineensis* and *Phoenix dactylifera,* indicating that similar resistance genes 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 (Seoighe and Gehring 2004; Demuth and Hahn 2009). The nature of selective pressure imposed by their pathogens is expected to be diverse; therefore, the selection is responsible for the drastic variations in the 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 K_a/K_s values among sequences of the same clade and sequences of different clades, we could identify >1 values, suggesting that positive selection is the force that directs the evolution of new R genes. However, values <1 were also obtained, indicating that they were under negative or diversifying selection. Our results suggest that NBS sequences of coconut palms are under selective constraints that provide further evidence consistent with a birth-and-death model of *R* gene evolution.

To infer 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 the coconut palm. These clusters frequently consist of tandem

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

Type of clade	Arecaceae			Poaceae		Brassicaceae			Total
	Cn	Pd	Eg	Os	Sb	At	Br	Bo	
Species-specific									
No. of clades	5	1	0	4	6	0	1	0	17
No. of genes	30	11	0	5	7	0	1	0	54

Type of clade	Arecaceae				Poaceae	Brassicaceae			Total
	Cn–Pd	Cn–Eg	Pd–Eg	Cn–Pd–EG	Os–Sb	At–Br	Br–Bo	At–Br–Bo	
Family-specific									
No. of clades	0	0	3	5	13	0	1	2	24
No. of genes	0	0	11 + 6	51 + 27 +25	55 + 41	0	2 + 1	17 + 17 + 2	255

Type of clades	Arecaceae–Poaceae	Arecaceae–Brassicaceae	Poaceae–Brassicaceae	Arecaceae–Poaceae–Brassicaceae	Total
Complex clades					
No. of clade	9	3	2	0	14
No. of genes	41 + 34	26 + 20	6 + 2	0	129

Cn Cocos nucifera, Sb Sorghum bicolor, Pd Phoenix dactylifera, At Arabidopsis thaliana, Eg Elaeis guineensis, Bn Brassica napus, Os Oriza sativa, Bo Brassica oleracea

Table 5 Detection of LY-phytoplasma DNA from leaf tissues of asymptomatic and symptomatic 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

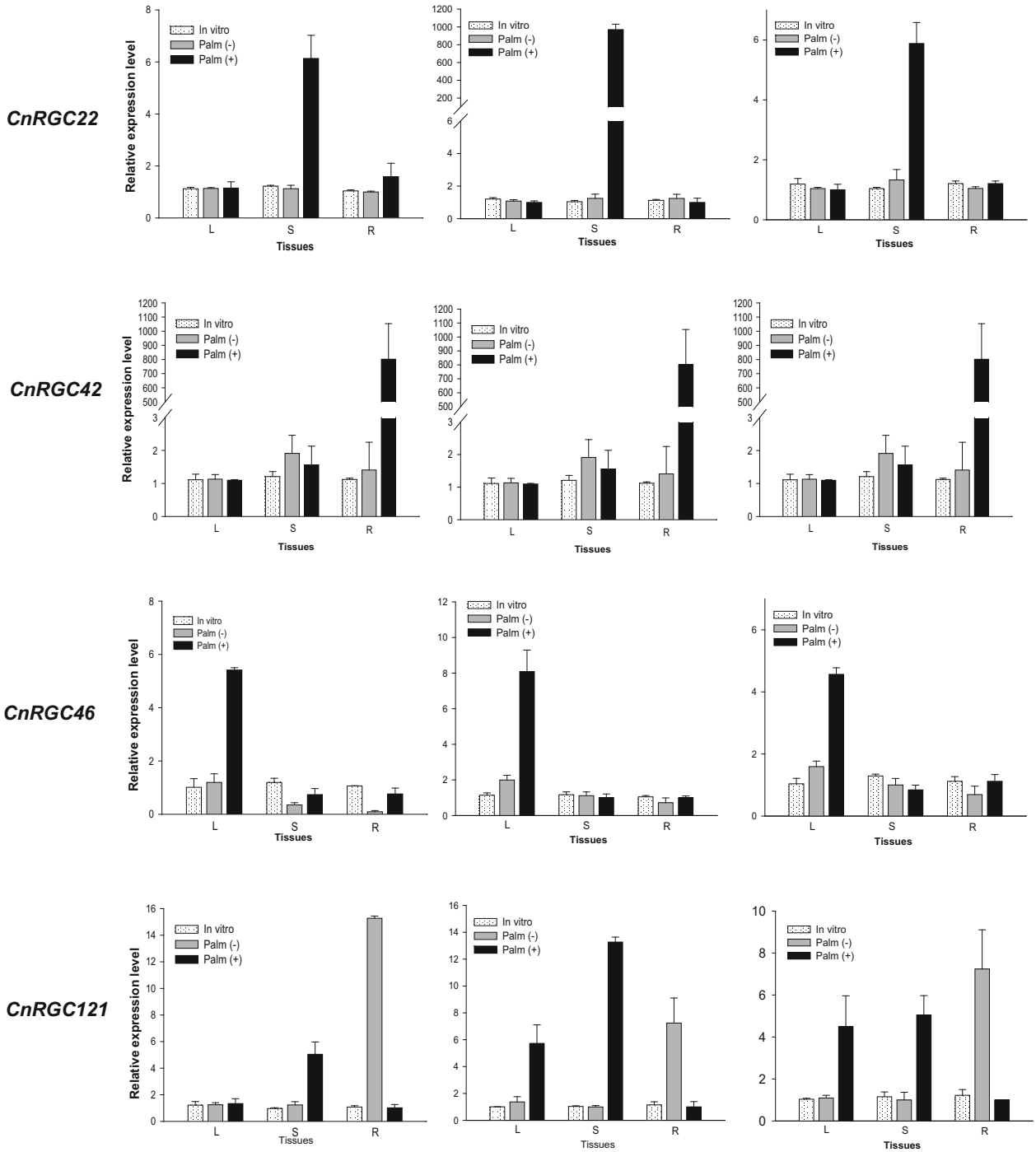


Fig. 4 Relative transcript levels of four CnRGCs obtained from quantitative real-time PCR experiments in different tissues of healthy coconut palms and those affected with lethal yellowing disease. *Black bars*: LY- coconut palms, gray bars: healthy coconut palm and *white bars*: in vitro coconut palm 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 (each sample was taken in triplicate)

duplications of the same species (Leister 2004; Yang et al. 2008). Family-specific clades were also identified. Five were 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 other species of Poaceae and Rosaceae (Yang et al. 2008; Gu et al. 2015). Family-specific NBS sequences may also exhibit restricted taxonomic functionality, as shown by the cloned *R* gene *Bs2* from pepper and *RPS2* from *Arabidopsis*, which function in related species within the same family but not in species outside of the family (Tai et al. 1999).

Heterogeneous clusters (complex clades) in which sequences belong to different phylogenetic lineages are also present. The major phylogenetic relationships were identified among monocot species (9 clades between Arecales and Poales). However, a phylogenetic relationship among NBS sequences of monocot and dicot species were also found (3 clades between Arecaceae and Brassicaceae and 2 clades between Poaceae and Brassicaceae). This result supports the proposal by Cannon et al. (2002) that there are several ancient clades of NBS sequences and that some of these predate the split between monocots and dicots.

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 (Dangl and Jones 2001). However, to our knowledge, there are no reports of *R* genes conferring resistance to phytoplasma even when phytoplasma effectors in plants have been characterized (Sugio et al. 2011). To evaluate the possible role of these genes against LY-phytoplasma, we selected four random 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 the tissue level (CnRGC22 in stems, CnRGC42 in roots, CnRGC46 in leaves, and CnRGC121 in both leaves and stems) compared to healthy palms. Moreover, the tissue-specific expression profile could be associated with pathogen distribution, as has been reported by Córdova et al. (2014). Interestingly CnRGC121 was found to be differentially expressed in the root tissue of healthy coconut palms compared to the diseased palms.

In conclusion, this paper examined the NBS sequences of coconut palms and other Arecaceae species to reveal the evolutionary history of these genes. Phylogenetic reconstruction in the Arecaceae 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 palms, date palms, and oil palms, thus highlighting a monophyletic origin of these three genera and the high conservation of some RGC in these palms. The phylogenetic relationship with the Brassicaceae family suggests that the RGC genes originated before monocotyledonous and dicotyledonous classes split. The expression profile of four sequences in healthy palms and those infected with LY-phytoplasma provides the first evidence of the role of NBS genes during plant-phytoplasma interactions. Importantly, pathogen-responsive NBS–LRR genes identified in the present study may be used as candidate genes for engineering pathogen resistance in coconut palm as well as other related species.

Acknowledgments Funding of the research reported here was provided by CONACyT, México (Grant No. CB 129717). Carlos Puch-Hau and Manuel Gongora thank CONACyT for the scholarship awarded No. 47245 and 265354, respectively.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Al-Mssallem IS, Hu S, Zhang X, Lin Q, Liu W, Tan J, Yu X, Liu J, Pan L et al (2013) Genome sequence of the date palm *Phoenix dactylifera* L. Nat Commun 4:2274

Alsaihati B, Lin Q, Al-Mssallem IS (2014) Coconut Genome de novo Sequencing. Plant & animal genome XXII. Conference San Diego

Baumgarten A, Cannon S, Spangler R, May G (2003) Genome-level evolution of resistance genes in *Arabidopsis thaliana*. Genetics 165:309–319

Belkhadir Y, Nimchuk Z, Hubert D, Mackey D, Dangl J (2004) 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 16:2822–2835

Bouktila D, Khalfallah Y, Habachi-Houimli Y, Mezghani-Khemakhem M, Makni M, Makni H (2015) Full-genome identification and characterization of NBS-encoding disease resistance genes in wheat. Mol Genet Genomics 290:257–271

Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, Young ND (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and Non-TIR NBS–LRR resistance gene subfamilies. J Mol Evol 54:548–562

Córdova I, Oropeza C, Puch-Hau C, Harrison N, Collí-Rodríguez A, Narvaez M, Nic-Matos G, Reyes C, 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. J Plant Pathol 96:343–352

Dangl L, Jones J (2001) Plant pathogens and integrated defense responses to infection. Nature 411:826–833

Demuth J, Hahn M (2009) The life and death of gene families. BioEssays 31:29–39

Eeuwens C (1976) 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

Ellis J, Lawrence G, Luck J, Dodds P (1999) Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. Plant Cell 11:495–506

Fan H, Xiao Y, Yang Y, Xia W, Mason AS et al (2013) RNA-Seq analysis of *Cocos nucifera*: transcriptome sequencing and subsequent functional genomics approaches. PLoS One 8:e59997

Foan C, Lee Y, Tan J, Alwee S (2012) Amplification and sequencing of partial-length disease resistance gene homologues coding for NBS–LRR proteins in oil palm (*Elaeis guineensis*). Asia Pac J Mol Biol Biotechnol 20:25–31

Gitau CW, Gurr GM, Dewhurst CF, Fletcher MJ, Mitchell A (2009) Insect pests and insect-vectored diseases of palms. Aust J Entomol 48:328–342

Griffith R (1987) Red ring disease of coconut palm. Plant Dis 71:193–196

Gu L, Si W, Zhao L, Yang S, Zhang X (2015) Dynamic evolution of NBS–LRR genes in bread wheat and its progenitors. Mol Genet Genomics 290:727–738

Hanold D, Randles J (1991) Coconut cadang-cadang disease and its viroid agent. Plant Dis 75:330–335

Harrison N, Oropeza C (2008) Coconut lethal yellowing. In: Harrison NA, Rao GP, Marcone C (eds) Characterization, diagnosis and management of phytoplasmas. Studium Press LLC, Houston, pp 219–248

Harrison N, Richardson P, Kramer J, Tsai JH (1994) Detection of the mycoplasma like organism associated with lethal yellowing disease of palms in Florida by polymerase chain reaction. Plant Pathol 43:998–1008

Howard FW, Barrant CI (1989) Questions and answers about lethal yellowing disease. Principes 33:163–171

Huang Y, Lee CP, Fu FL, Chang B, 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 Genom Genet 4:2147–2157

Hulbert S, Webb C, Smith S, Sun Q (2001) Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol 39:285–312

Jones D, Jones J (1997) The role of leucine-rich repeat proteins in plant defenses. Adv Bot Res 24:89–167

Khan A, Khan A, Azhar M, Amrao L, Cheema H (2015) Comparative analysis of resistance gene analogues encoding NBS–LRR domains in cotton. J Sci Food Agric 96:530–538

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez A, Thompson J, Gibson T, Higgins D (2007) ClustalW and Clustal X version 2.0. Bioinformatics 23:2947–2948

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet 20:116–122

Li J, Ding J, Zhang W, Zhang Y, Tang P, Chen JQ, Tian D, Yang S (2010) Unique evolutionary pattern of numbers of gramineous NBS–LRR genes. Mol Genet Genomics 283:427–438

Livak K, Schmittgen T (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25:402–408

Lozano R, Hamblin M, Prochnik S, Jannink J (2015) Identification and distribution of the NBS–LRR gene family in the Cassava genome. BMC Genomics 16:360

Luo S, Zhang Y, Hu Q, Chen J, Li K, Lu C, Liu H, Wang W, Kuang H (2012) Dynamic nucleotide-binding site and leucine-rich repeat-encoding genes in the grass family. Plant Physiol 159:197–210

Martínez M, López-Solanilla E, Rodríguez-Palenzuela P, Carbonero P, Díaz I (2003) Inhibition of plant-pathogenic fungi by the barley cystatin Hv-CPI (gene Icy) is not associated with its cysteine-proteinase inhibitory properties. Mol Plant Microbe Interact 16:876–883

McCoy R, Howard F, Tsai J, Donselman H, Thomas D, Basham H, Atilano R, Eskafi F, Britt L, Collins M (1983) Lethal yellowing of palms. Florida Agricultural Experiment Station Bulletin 834, Gainesville

Meyers B, Dickerman A, Michelmore R, Sivaramakrishnan S, Sobral B, 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

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

Meyers B, Kaushik S, Nandety R (2005) Evolving disease resistance genes. Curr Opin Plant Biol 8:129–134

Michelmore R, Meyers B (1998) Clusters of resistance genes in plants evolve by divergent selection and birth-and-death process. Genome Res 8:1113–1130

Montero-Cortes M, Saenz L, Cordova I, Quiroz A, Verdeil J, Oropeza C (2010) GA_3 stimulate the formation and germination of somatic embryos and the expression of a KNOTTED-like homeobox gene of *Cocos nucifera* (L.). Plant Cell Rep 29:1049–1059

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nejat N, Cahill DM, Vadamalai G, Ziemann M, Rookes J, 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 290:1899–1910

Parthasarathy MV, van Slobbe WG, Soudant C (1978) Hartrot or fatal wilt of palms. Coconuts (*Cocos nucifera*). Principes 22:3–14

Porter B, Paidi M, Ming R, Alam M, Nishijima W, Zhu Y (2009) Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. Mol Genet Genomics 281:609–626

Puch-Hau C, Oropeza-Salín C, Peraza-Echeverría S, Gongora-Paredes M, Córdova-Lara I, Narvaez-Cab M, Zizumbo-Villareal D, Sáenz-Carbonell L (2015) Molecular cloning and characterization of disease-resistance gene candidates of the nucleotide binding site (NBS) type from *Cocos nucifera* L. Physiol Mol Plant Pathol 89:87–96

Rajesh M, Rachana K, Naganeeswaran S, Shafeeq R, Thomas R, Shareefa M, Merin B, Anitha K (2015) Identification of expressed resistance gene analog sequences in coconut leaf transcriptome and their evolutionary analysis. Turk J Agric For 39:1–14

Richly E, Kurth J, Leister D (2002) Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. Mol Biol Evol 19:76–84

Rohde W, Randles JW, Langridge P, Hanold D (1990) Nucleotide sequence of a circular single stranded DNA associated with coconut foliar decay virus. Virol 176:648–651

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Seoighe C, Gehring C (2004) Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. Trends Genet 20:461–464

Shang J, Tao Y, Chen X, Zou Y, Lei C, Wang J, Li X, Zhao X, Zhang M, Lu Z, Xu J, Cheng Z, Wan J, Zhu L (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

Sugio A, MacLean A, Kingdom H, Grieve V, Manimekalai R, Hogenhout S (2011) Diverse targets of phytoplasma effectors: from plant development to defense against insects. Annu Rev Phytopathol 49:175–195

Tai T, Dahlbeck D, Clark E, Gajiwala P, Pasion R, Whalen M, Stall R, Staskawicz B (1999) Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. Proc Natl Acad Sci USA 96:14153–14158

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729

Wan H, Yuan W, Ye Q, Wang R, Ruan M, Li Z, Zhou G, Yao Z, Zhao J, Liu S, Yang Y (2012) Analysis of TIR- and non-TIR-NBS LRR disease resistance gene analogous in pepper: characterization, genetic variation, functional divergence and expression patters. BMC Genom 13:1–15

Wan H, Yuan W, Bo K, Shen J, Pang X, Chen J (2013) Genome-wide analysis of NBS-encoding disease resistance in *Cucumis sativus* and phylogenetic study of NBS-encoding genes in *Cucurbitaceae* crops. BMC Genom 14:109

Yang S, Zhang X, Yue J, Tian D, Chen J (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. Mol Genet Genomics 280:187–198

Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D (2004) 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

Zhu H, Cannon S, Young N, Cook D (2002) 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