Genetic variability of *Papaya ringspot virus* isolates in Norte de Santander - Colombia

Variabilidad genética del virus de la mancha anular de la papaya en Norte de Santander - Colombia

Giovanni Chaves-Bedoya[1] and Luz Yineth Ortiz-Rojas[1]

ABSTRACT

The *Papaya ringspot virus* (PRSV), a member of the potyvirus that is transmitted by aphids within the *Potyviridae* family, is the main limiting factor for papaya (*Carica papaya* L.) and Cucurbits worldwide and causes losses of up to 100%. In this study, we conducted research on the genetic diversity of PRSV isolates collected from two locations in the department of Norte de Santander, Colombia. The analysis was performed by comparing the nucleotide sequences of the region that encode the coat protein (CP) of nine PRSV isolates from the Villa del Rosario location and 12 isolates from the Campo Hermoso location. The analysis included three sequences of the CP of PRSV isolates reported in the Colombian departments of Arauca and Valle del Cauca. The bioinformatic analysis suggested that the PRSV isolates from the locations in Norte de Santander were different from each other, grouping into different phylogenetic groups. Anexistence of recombination events in the coat protein was observed. This is the first study of PRSV genetic variability that has been conducted at the local level in Colombia.

Key words: PRSV, coat protein, recombination, papaya, sequencing.

RESUMEN

El virus de la mancha anular de la papaya (PRSV), un virus miembro de los potyvirus transmitidos por áfidos dentro de la familia *Potyviridae*, es el principal factor limitante para los cultivos de papaya (*Carica papaya* L.) y cucurbitáceas en todo el mundo, llegando a ocasionar pérdidas hasta del 100% de la producción. En este trabajo se realizó el estudio de la diversidad genética de diferentes aislados de PRSV colectados en dos localidades del departamento de Norte de Santander, Colombia. El análisis se realizó comparando las secuencias nucleotídicas de la región que codifica para la proteína de la cubierta (CP) de nueve aislados de la localidad de Villa del Rosario y 12 aislados de la localidad de Campo Hermoso. Para comparación el análisis incluyó tres secuencias de la CP de PRSV de aislados reportados en los departamentos colombianos de Arauca y Valle del Cauca y de otras secuencias informadas en el GenBank. Los análisis bioinformáticos sugieren que los aislados de PRSV de las localidades en Norte de Santander son distintos entre sí agrupándose en grupos filogenéticos diferentes, y se demostró la existencia de eventos de recombinación. Este es el primer estudio de variabilidad genética de PRSV a nivel local en Colombia.

Palabras clave: PRSV, proteína de la cubierta, recombinación, papaya, secuenciación.

Introduction

The papaya (*Carica papaya* L.) is a species whose origin is registered as Tropical America, especially Central America and the west coast of South America. Currently, it is grown in all of the tropical and subtropical areas of the world, between 32 degrees north latitude and south of the Equator (Arango *et al*., 2010). It is an herbaceous plant with large leaves and rapid growth. Botanically, *Carica papaya* (L.) cannot be clearly described as a tree, plant or shrub. It reaches a height of 10 m and an age of 15 to 30 years. The size of the fruit can weigh between 100 g and 10 kg. Its shape is round-oval to oblong and the color of its flesh is yellowish-white, deep yellow, orange or reddish yellow. It is cultivated in land rich with organic matter and abundant moisture. The papaya was first described in 1,526 by the Spanish chronicler Oviedo, who found it on the coasts of Panama and Colombia (Cabrera *et al*., 2011).

A limiting factor for papaya yield worldwide is the disease caused by PRSV. Discovered in 1945 (Jensen, 1949), PRSV is the most widely distributed and harmful virus in papaya. Plants infected at the seedling stage or within two months after planting do not produce normal, mature fruits. The virus is transmitted by many species of aphids in a non-persistent manner (Suzuki *et al*., 2007). PRSV strains are divided into two biotypes: the papaya infecting type (PRSV-P), which affects both papaya and cucurbits,

Received for publication: 10 April, 2015. Accepted for publication: 30 June, 2015. Doi: 10.15446/agron.colomb.v33n2.50095

[1] Plantae Research Group, Faculty of Basic Sciences, Universidad Francisco de Paula Santander. Cucuta (Colombia). gchavesb@ufps.edu.co

and the cucurbit infecting type (PRSV-W) from Taiwan, which affects cucurbits but not papaya (Gonsalves *et al*., 2010). The virions of both biotypes cannot be distinguished in serological tests, but differ in their ability to infect papaya (Bateson *et al*., 2002). PRSV-P infects papaya (*Carica papaya* L.) naturally and is the main limiting factor in the production of papaya worldwide. PRSV-P was described for the first time in Hawaii and, since then, has been reported in several countries where it usually causes devastation to the production of papaya in the early years of the first infection (Gonsalves, 1998). The epidemiology of PRSV-P in papaya is similar to that of other non-persistent viruses and it is transmitted directly from papaya to papaya (Bateson *et al*., 2002).

PRSV has flexible, 760-780 x 12 nm filamentous, flexuous rod particles and consists of a monopartite single-stranded positive sense RNA, about 10,326 nucleotides long, excluding a poly-A tract found at its 3'end. PRSV encodes a single large protein which is subsequently cleaved into smaller proteins with several functions (Gonsalves *et al*., 2010). The central region of PRSV is highly conserved, while the N-terminal region is more variable (Abdalla and Ali, 2012).

The symptoms on papaya are somewhat similar to those on cucurbits. In papaya, the leaves develop prominent mosaic and chlorosis on the leaf lamina and water-soaked, oily streaks on the petioles and upper part of the trunk (Gonsalves *et al*., 2010).

Genetic engineering (Kung *et al*., 2009; Mangrauthia *et al*., 2010; Yu *et al*., 2011), and cross protection (Gonsalves *et al*., 2010) are options for managing viral diseases such as PRSV. Knowledge of the nucleotide sequence and genetic diversity is necessary to be able to create control strategies. There have been different studies on the genetic diversity of PRSV, particularly variant P, in different countries around the world, including Australia (Bateson *et al*., 1994), Brazil (Lima *et al*., 2002), India (Jain *et al*., 2004), Jamaica and Venezuela (Chin *et al*., 2007), Mexico (Noa-Carrazana *et al*., 2007), Taiwan (Wang and Yeh, 1997), Thailand, Vietnam and Philippines (Bateson *et al*., 2002), Venezuela (Fernández-Rodríguez *et al*., 2008) and the USA (Abdalla and Ali, 2012). In Colombia, at the sequence level, data on PRSV are scarce, with the report of a few sequences corresponding to the helper component proteinase (HC-Pro) and the coat protein (CP) obtained in the departments of Valle del Cauca (four sequences) and Arauca (one sequence) (Olarte Castillo *et al*., 2011).

Norte de Santander is potentially an agricultural department due to its geographical conditions, presenting a diversity of crops. The papaya crop occupies the ninth position of importance for fruit crops (Tafur *et al*., 2006).

Given that there is not any information on PRSV in the department of Norte de Santander, Colombia, this study was conducted for a preliminary determination of their genetic variability, based on the sequence of the CP gene, using sequences of 21 PRSV isolates collected from symptomatic papaya plantsin two different locations in the department. Phylogenetic relationships between these isolates and PRSV isolates reported in other parts of the world were also evaluated. Studies on the molecular evolution of viruses are essential for understanding their biological characteristics, changes in virulence, geographical ranges and their emergence as new epidemics. This understanding is crucial for the design of strategies for viral control (Moreno *et al*., 2004; Tan *et al*., 2004).

Materials and methods

Field sampling. Symptomatic leaves of PRSV-infected papaya plants were collected in the months of March and September of 2014 from different plots in the Campo Hermoso (12 isolates) and Villa del Rosario (nine isolates) locations, for a total of 21 isolates. The harvested leaves were kept in polystyrene coolers with ice to keep them fresh, taken to the Plantae Research Group Laboratory at the Universidad Francisco de Paula Santander, and stored at -80°C until the extraction of total RNA from each sample.

RNA extraction and RT-PCR. The papaya leaves were used for total RNA extraction using the reagent TRIZOL (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. The RNA extracted from each plant was quantified in a spectrophotometer, Genesys 10S UV-VIS (Thermo Scientific, Waltham, MA), reading at a wavelength of 260 nm and 280 nm. The total RNA was quantified and preserved at -80°C in aliquots until use as a template for amplification by RT-PCR. The primers for the RT-PCR reactions were designed based on the consensus sequence of conserved regions reported in the GenBank (Benson *et al*., 2014). The annealing sites were the 3' region of the nuclear inclusion B (NIB) protein and the 5' region of the untranslated region (3UTR). The sequences of the primers were: forward 5'-GTCATGGGGATATGGGGAGTTAACACA-3' and reverse 5'-TTTTTTTTCTCTCATTCTAAGAGGCTCG-3'.

cDNA synthesis and cloning. Samples of the total RNA were used as a template to make the RT-PCR reactions to amplify the CP of the PRSV. For the reverse transcription,

we used a Tetro cDNA Synthesis Kit (Bioline, Tauton, MA). An 9.5 µL of ultrapure water were incubated with 2 µL of total RNA and 2 µL of the primer 3' (10 µM) at 70°C for 10 min and then placed in ice to stop the reaction. Each of the previous reactions were mixed with 6.5 µL of mixture containing 100 UMMLV reverse transcriptase, 4 µL of buffer 5X, 1 µL of ribosafe and 1 µL10 mM dNTP mix. The reactions were incubated at 37°C for 1 h.

Each of the resulting cDNA samples were amplified by PCR in a thermal cycler, ARKTIK (Thermo Scientific, Waltham, MA), using 5 uL of the corresponding reaction of RT, 2.5U of high fidelity Taq Accutaq polymerase (Bioline, Tauton, MA), 0.1 mM dNTPs and 0.5 µM uL^{-1} of each of the primers, reverse and forward, for a final volume of 50 uL. The conditions for the amplification of PCR were 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 56.3°C and 2 min at 72°C. Finally, there was one cycle at 72°C for 5 min.

The PCR reactions were loaded in 2% agarose gels and ran to 100 V in an Enduro Electrophoresis Chamber (Labnet, Edison, NJ) in 1% TAE buffer for a period of 60 min. The gels were stained with ethidium bromide and visualized in a UVP transilluminator at 302 nm. The PCR products with the expected size of 1,350 bp were cut with a sterile scalpel blade and purified using the Isolate II PCR and Gel Kit (Bioline, Tauton, MA). The purified DNA from each sample was cloned into the Topo vector (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. Briefly, 2 µL of each PCR was mixed with 1.5 µL of saline solution, 0.5 µL of the vector and 2 µL of water. This mixture was incubated for 10 min at room temperature. Subsequently, the vector was introduced into competent cells of *Escherichia coli* JM109 (Invitrogen, Carlsbad, CA), using heat shock. The plasmid DNA was isolated from bacterial cultures grown overnight using the alkaline extraction procedure (Bimboim and Doly, 1979). The plasmid DNA were analyzed with restriction enzymes and sequenced twice using M13 universal primers (Macrogen, Seoul, Korea).

Analysis of sequences. The DNA sequences obtained from each isolate were assembled using the SeqMan program. The edition and consensus sequence of nucleotides was made with the EditSeq program. Multiple alignment of the nucleotide and deduced amino acid sequences from two clones of each isolate were carried out using the MegAlign program. All of the programs are part of the DNAStar package software(Madison, WI). CP sequences were adjusted to the same length of 861 nucleotides for analysis.

A phylogenetic analysis was conducted initially with 21 new CP sequences of PRSV from Norte de Santander, including 54 isolates reported in 11 different countries obtained from the GenBank database. The accesion number, including for the new PRSV from Norte de Santander (Tab. 1). This analysis included the PRSV CP sequences reported in the Valle del Cauca (2 sequences) and Casanare (1 sequence). The phylogenetic tree was built with the MEGA program (Tamura *et al*., 2007), using maximum likelihood and the general time reversible (GTR) model with gamma distribution. The 1,000 replicates were performed to estimate the reliability of the branching patterns of the trees.

TABLE 1. Accession numbers and country of origin for the PRSV isolates used in this study. The PRSV sequences generated in this research are in bold.

ISOLATE NUMBER	COUNTRY	ISOLATE NUMBER	Country
DQ008449	MEXICO	AF196839	USA
DQ008448	MEXICO	AF196838	USA
DQ008447	MEXICO	AF344650	BRAZIL
DQ008446	MEXICO	AF344645	BRAZIL
KC149502	INDIA	AF344640	BRAZIL
KC149501	INDIA	AF344639	BRAZIL
DQ088670	INDIA	DQ339581	VENEZUELA
JN979404	INDIA	DQ339580	VENEZUELA
HM778170	INDIA	DQ339579	VENEZUELA
HM626467	INDIA	DQ339578	VENEZUELA
HM626465	INDIA	DQ339576	VENEZUELA
JQ394693	INDIA	AF469066	CHINA
JN979406	INDIA	AF469065	CHINA
JQ394690	INDIA	HQ328798	COLOMBIA
JQ394694	BANGLADESH	HQ328795	COLOMBIA
JQ394687	BANGLADESH	HQ328794	COLOMBIA
JN979398	BANGLADESH	**KR075692 (CH1)**	**COLOMBIA**
JN979400	BANGLADESH	**KR075693 (CH11)**	**COLOMBIA**
DQ088671	THAILAND	**KR075694 (CH14)**	**COLOMBIA**
AY010718	THAILAND	**KR075695 (CH2)**	**COLOMBIA**
AY010717	THAILAND	**KR075696 (CH20)**	**COLOMBIA**
AY010715	THAILAND	**KR075697 (CH21)**	**COLOMBIA**
AY010712	THAILAND	**KR075698 (CH22)**	**COLOMBIA**
DQ104821	JAMAICA	**KR075699 (CH23)**	**COLOMBIA**
DQ104820	JA MAICA	**KR075700 (CH4)**	**COLOMBIA**
DQ104817	JAMAICA	**KR075701 (CH9)**	**COLOMBIA**
DQ104816	JAMAICA	**KR075702 (VR11)**	**COLOMBIA**
DQ104815	JAMAICA	**KR075703 (VR12)**	**COLOMBIA**
KC748224	CUBA	**KR075704 (VR2)**	**COLOMBIA**
KC748225	CUBA	**KR075705 (VR3)**	**COLOMBIA**
KC748226	CUBA	**KR075706 (VR5)**	**COLOMBIA**
KC748227	CUBA	**KR075707 (VR7)**	**COLOMBIA**
KC748228	CUBA	**KR075708 (VR8)**	**COLOMBIA**
FR696596	EGYPT	**KR075709 (VR9)**	**COLOMBIA**
FR696595	EGYPY	**KR075710 (VR10)**	**COLOMBIA**
FR696591	EGYPY	**KR075711 (CH26)**	**COLOMBIA**
FR696594	EGYPT	**KR075712 (CH24)**	**COLOMBIA**
FR696590	EGYPT		

Genomic distances. The genomic distances for each pair of isolates were estimated by the method of two parameters used by Kimura (Kimura, 1980). The genetic distances at the synonymous sites (dS) and non-synonymoussites (dn), as well as the nucleotide diversity, were estimated according to the Pamilo-Bianchi-Li method (Pamilo and Bianchi, 1993). The SimPlot program (Lole *et al*., 1999) was also used to analyze the nucleotide similarities between each isolate along the CP using a window of 200 pb.

Analysis of recombination. Three different strategies were used to determine the putative recombination events. In the first case, the aligned sequences in the FASTA format were uploaded to the remote server datamonkey (http://datamonkey.org/GARD/) to run the program GARD (Genetic Algorithm for Recombination Detection) (Kosakovsky Pond *et al*., 2006). The second strategy was the RDP3 (Recombination Detection Program) program (Martin *et al*., 2010). Finally, we analyzedthe haplotype map builtin the SNAP program (Suite of nucleotide Analysis Program) (Price and Carbone, 2005).

Results and discussion

Sequence analysis

The CP Sequences of the PRSV isolates from Villa del Rosario and Campo Hermoso in Norte de Santander were 861 nucleotides long and encoded for a protein of 287 amino acids. The average content of bases in the 21 isolates was A, 35.4018%; C, 15.9781%; G, 23.9201% and T, 24.7%. Table 2 shows the pairwise distance matrix between each PRSV isolate from Norte de Santander.

The nucleotide similarity between the PRSV isolates from Norte de Santander was between 95.7 and 99.9%. The isolates with greater divergence were the collected in Campo Hermoso, grouping into a different clade (Fig. 1).

The nucleotide similarity between the PRSV from Valle del Cauca and Casanare and the isolates from Villa del Rosario was between 92.5-93.7% and 88.3-90.5%, when compared with the isolates from Campo Hermoso. These similarity values between isolates from separate regions are consistent with the genetic variability found in other countries for PRSV (Chin *et al*., 2007; Akhtler *et al*., 2013).

The nucleotide diversity, defined as the probability that the base of a certain position differs between two individuals randomly selected from the population (García-Arenal *et al*., 2001), was 0.023 for the isolates from Campo Hermoso and 0.008 for the isolates from Villa del Rosario. Similar values of nucleotide diversity have been reported for various plant virus populations (García-Arenal *et al*., 2001). The lowest value of diversity for the PRSV isolates from Villa del Rosario could have been caused by one of the three bottleneck events that restrict the amount of genetic diversity in the viral populations of plants (Simmons *et al*., 2013).

TABLE 2. Pair distances of sequences of the CP of the 21 isolates of PRSV from Norte de Santander (Colombia). Low similarity values from Campo Hermoso isolates are highlighted.

	CH_1	CH_11	CH_14	CH_2	CH_20	CH_21	CH_22	CH_23	CH24	CH_26	CH_4	CH_9	VR_11	VR_12	VR_2	VR_3	VR_5	VR_7	VR_8	VR_9	VR10
CH _ 1	***	99.3	96.1	99.1	98.8	95.7	95.9	99.3	95.8	99.2	99.4	99.1	96.2	96.5	96.4	97.7	96.4	96.5	96.5	96.4	96.6
CH _ 11		***	96.5	99.5	99.3	96.2	96.4	99.5	96.3	99.4	99.9	99.5	96.6	97.0	96.9	98.1	96.9	97.0	97.0	96.9	97.1
CH _ 14			***	96.3	96.1	99.4	99.7	96.3	99.8	96.2	96.6	96.3	99.0	99.3	99.4	97.7	99.2	99.2	99.3	99.2	99.4
CH _ 2				***	99.1	95.9	96.2	99.3	96.1	99.2	99.7	99.3	96.4	96.7	96.6	98.4	96.6	96.7	96.7	96.6	96.9
CH _ 20					***	95.7	95.9	99.1	95.8	99.0	99.4	99.1	96.2	96.5	96.4	97.7	96.4	96.5	96.5	96.4	96.6
CH _ 21						***	99.3	95.9	99.7	95.8	96.3	95.9	98.4	98.7	98.8	97.1	98.6	98.6	98.7	98.6	98.8
CH _ 22							***	96.2	99.7	96.1	96.5	96.2	98.6	99.0	99.1	97.3	98.8	98.8	99.0	98.8	99.1
CH _ 23								***	96.1	99.4	99.7	99.3	96.4	96.7	96.6	97.9	96.6	96.7	96.7	96.6	96.9
CH _ 24									***	95.9	96.4	96.1	98.7	99.1	99.2	97.4	99.0	99.0	99.1	99.0	99.2
CH _ 26										***	99.5	99.2	96.3	96.6	96.5	97.8	96.7	96.6	96.6	96.5	96.7
CH _ 4											***	99.7	96.7	97.1	97.0	98.3	97.0	97.1	97.1	97.0	97.2
CH _ 9												***	96.4	96.7	96.6	97.9	96.6	96.7	96.7	96.6	96.9
VR _ 11													***	99.4	99.3	97.6	99.3	99.3	99.4	99.3	99.5
VR _ 12														***	99.7	97.9	99.7	99.7	99.8	99.7	99.9
VR _ 2															***	98.0	99.5	99.5	99.7	99.5	99.8
VR _ 3																***	97.8	97.9	97.9	97.8	98.0
VR _ 5																	***	99.5	99.7	99.5	99.8
VR _ 7																		***	99.7	99.5	99.8
VR _ 8																			***	99.7	99.9
VR _ 9																				***	99.8
VR _ 10																					***

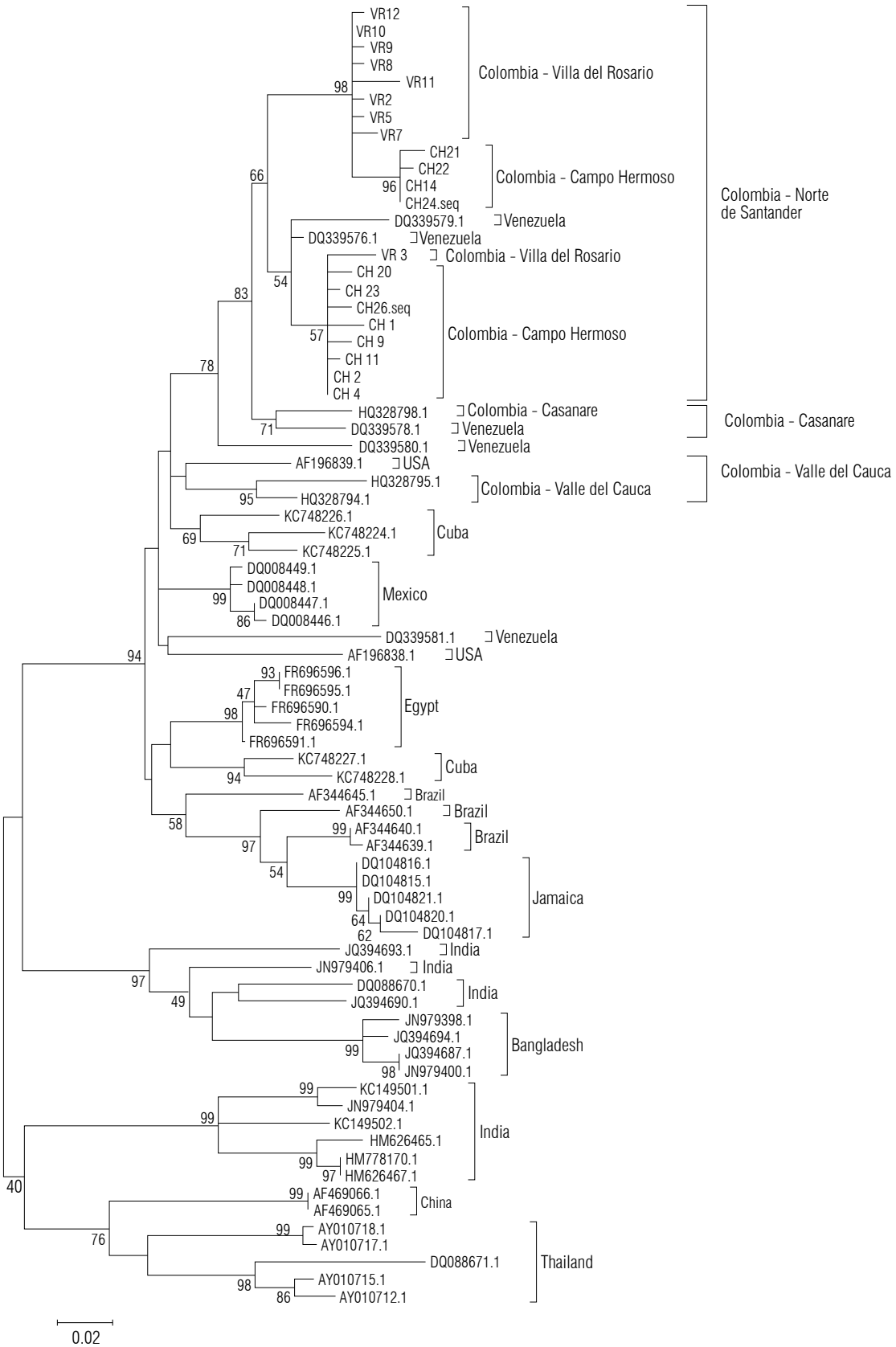

FIGURE 1. Phylogenetic relationship between the PRSV isolates from Campo Hermoso and Villa del Rosario in Norte de Santander. This treewas built with 71 sequences of the CP of PRSV from different geographical origins. On the top of the tree is the grouping of new isolates obtained in Norte de Santander, Colombia.

The number of non-synonymous substitutions (dN) was 0.012 for the isolates from Campo Hermoso and 0.006 for the isolates from Villa del Rosario. The number of synonymous substitutions (dS) was 0.066 for the PRSV isolates from Campo Hermoso and 0.017 for Villa del Rosario. These results suggest a greater evolutionary dynamic for the PRSV isolates from Campo Hermoso due to nucleotide substitutions. The dN/dS ratio of the CP from Campo Hermoso and Villa del Rosario was 0.17 and 0.35, respectively.

The low proportion of nucleotide substitution between non-synonymous and synonymous sites indicates a selection for the conservation of the amino acid sequence, or negative selection (Kryazhimskiy and Plotkin, 2008). In most cases the selection for virus populations is negative although, in some instances, a diversifying selection for some plant viruses has been found (Chaves-Bedoya *et al.*, 2013; Moury *et al.*, 2002).

At the time of preparation of this manuscript, there were only five sequences of the CP of PRSV in Colombia from the departments of Valle del Cauca and Casanare reported in the GenBank, the latter being phylogenetically closest to the PRSV isolates from Norte de Santander. This grouping could be due to the closeness between these two departments. In this way, the PRSV isolates reported in Casanare are phylogenetically more related to isolates from Venezuela (Fig. 1). However, the PRSV isolates from Valle del Cauca have greater similarity with PRSV isolates from the USA, suggesting a long-distance transport of infected material between countries (Fig. 1).

Potyviruses are able to infect different types of tissue in host plants, including seeds, although the function of the transmission by seed in the evolution and epidemiology of viral pathogens remains uncertain. The association of a plant virus with seeds provides a convenient means for virus survival and facilitates movement of the virus, including human intervention with harvesting and transporting seeds for subsequent cropping.

Many examples of long-distance spread and establishment of plant virus in new regions or countries can be attributed to seed-borne inoculum transported by the seed trade (Wilson, 2014). Transmission by seed can be significant in the spread of a virus (Simmons *et al.*, 2013). Although PRSV is a virus that typically is not seed transmitted (Tripathi *et al.*, 2008), there is a study showing the transmissibility of PRSV through papaya seeds (Bayot *et al.*, 1990).

Even low levels of seed transmission for PRSV can have a great impact since the disease can spread very quickly and can play an important role in the epidemiology and control of PRSV (Olarte Castillo *et al.*, 2011). Likewise, PRSV can be transmitted by seeds in alternative hosts, suggesting an important way to spread the virus (Laney *et al.*, 2012).

Phylogenetic analysis

The differences in the nucleotide similarity between the PRSV isolates from each location in Norte de Santander are reflected in the phylogenetic tree that groups the isolates into two different clades (Fig. 1). Nevertheless, the isolates CH14, CH21, CH22 and CH24 from Campo Hermoso grouped with the isolates from Villa del Rosario. Likewise, the VR3 isolate from Villa del Rosario grouped with the Campo Hermoso PRSV isolates. This phylogenetic relationship of the PRSV isolates in Norte de Santander suggests a possible transport of plant material and viral isolates between locations and possible recombination events.

The phylogenetic analysis also suggested that the PRSV isolated from Casanare was more related to the PRSV isolates from Campo Hermoso, while the isolates from Valle del Cauca were more related to the isolates from Casanare. The phylogetic trees indicated that Colombian papaya crops are affected by variable PRSV isolates that differentially group according to their procedence. These results of PRSV genetic variability coincide with recent reports that describe a growing genetic diversity and dispersion of the virus in spite of efforts to control this virus in different countries (Olarte Castillo *et al.*, 2011).

In order to understand the phylogenetic relationships of new Colombian PRSV isolates with isolates from other parts of the world, 51 sequences from Venezuela, Cuba, Mexico, Brazil, Jamaica, China, India, Egypt, Thailand, and Bangladesh were analyzed. The results suggest that Colombian PRSV isolates are more related to PRSV isolates reported in Venezuela. At the same time, isolates from Colombia and Venezuela have a greater relationship with isolates from Cuba and Mexico. In the phylogenetic tree, a large group is formed with PRSV isolates from America, including countries such as Colombia, Venezuela, Cuba, Mexico, Jamaica, Brazil and the United States. Interestingly, in this group, we found PRSV isolates from Egypt (Fig. 1). The possible reason for this type of cluster that includes isolates from geographically distant regions was discussed earlier.

Recombination analysis

With the aim of making a statistical validation of the possible recombination event observed from the groupings in the phylogenetic trees, especially for isolate VR3, we

analyzed the aligned sequences in a FASTA format with three different methodologies. In the first analysis using the GARD program, we found no evidence of recombination in the PRSV isolates from Norte de Santander. The second method was carried out using the RDP3 program. RDP3 is a computer program for the identification of statistical recombination events using non-parametric methods including BootScan, MaxChi, Chimaera, 3SEQ, GENECONV, Siscan, Phylpro and VIRSD (Martin *et al*., 2010). GENECONV (P-val= $1.316 \cdot 10^{-02}$) (Padidam *et al*., 1999), MAXCHI (P-val = $7.771 \cdot 10^{-05}$) (Smith, 1992), Chimaera (P-val=$5.841 \cdot 10^{-03}$) (Posada and Crandall, 2001), SiScan (P-val= $1.234 \cdot 10^{-05}$) (Gibbs *et al*., 2000), BootScan (P-val= $6.780 \cdot 10^{-03}$) (Martin *et al*., 2005) and 3 Seq (P-val= $1.538 \cdot 10^{-05}$) (Boni *et al*., 2007) detected the VR-3 isolated as an recombinant sequence. Putative breakpoints were located in positions 208 and 605. The putative major parent was the isolate VR_11 (100% similarity) and the putative minor parent was the isolate CH_20 (98.8% similarity).

Finally, the visual analysis of the haplotype map built with the SNAP program (Price and Carbone, 2005) allowed us to identify three types of PRSV haplotypes, two of them present in Campo Hermoso. The VR3 isolate shared haplotype segments with the PRSV isolates from Campo Hermoso and Villa del Rosario (Fig. 2).

The similarity plot (SimPlot) (Fig. 3) shows that the 5' region of the VR3 sequence was more similar to the isolates from Villa del Rosario; in the middle, there were

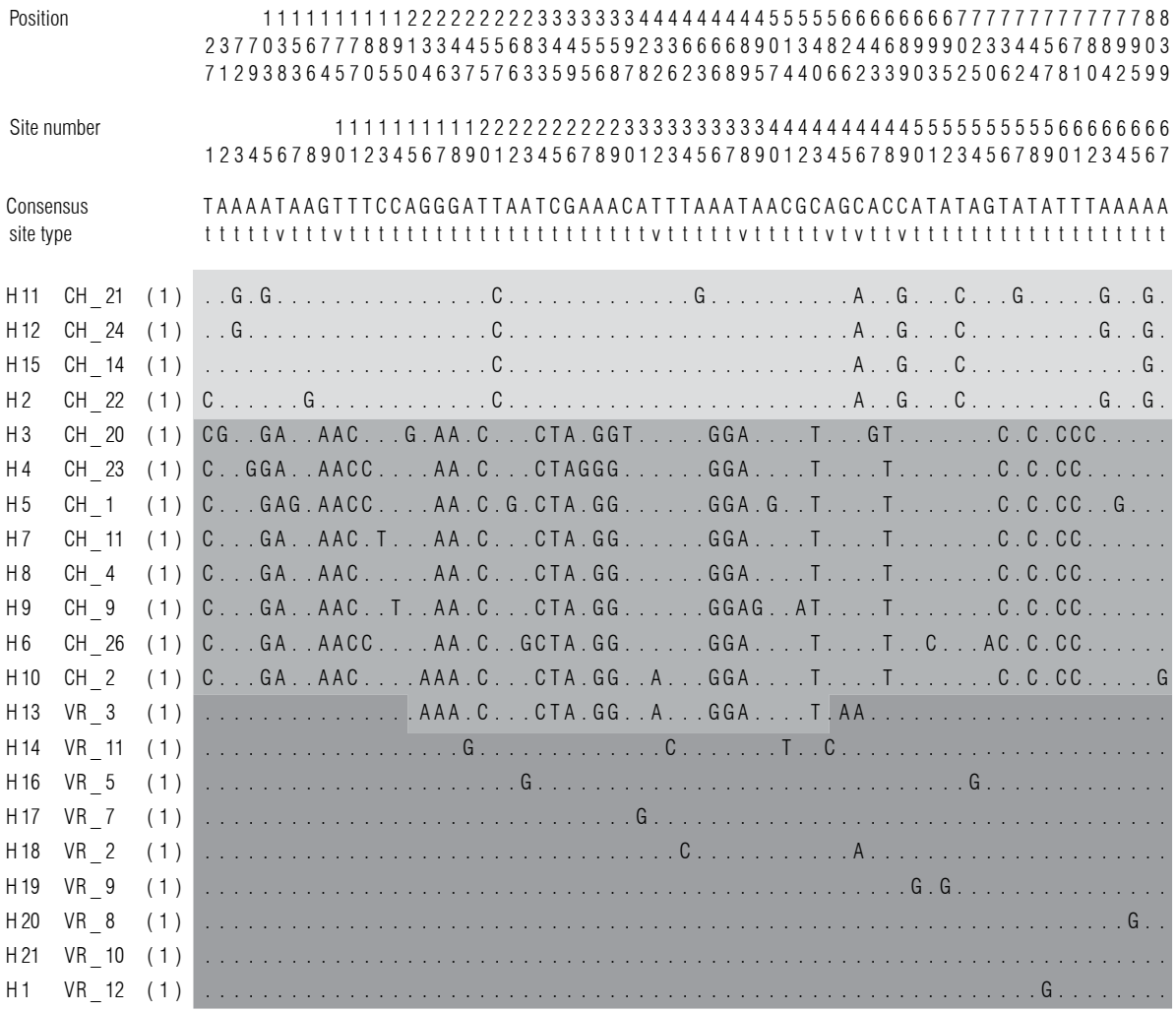

FIGURE 2. Haplotype map of 21 sequences of PRSV CPs from Norte de Santander. In the left, each haplotype (H11 to H1) and their corresponding PRSV isolate are indicated (CH21 to VR12). Haplotype H13 corresponds to recombinant VR_3 isolate. VR 3 shares the haplotype constitution of the Campo Hermoso and Villa del Rosario types. Reading vertically, the numbers at the top correspond to the position in the sequence. The numbers of below correspond to the position on the haplotype map. In positions where the RDP3 suggested recombination, i.e. 208 and 605, the constitution of the haplotype changed to VR3. The shadows correspond to the different constitutions of haplotypes of the PRSV isolates from Norte de Santander.

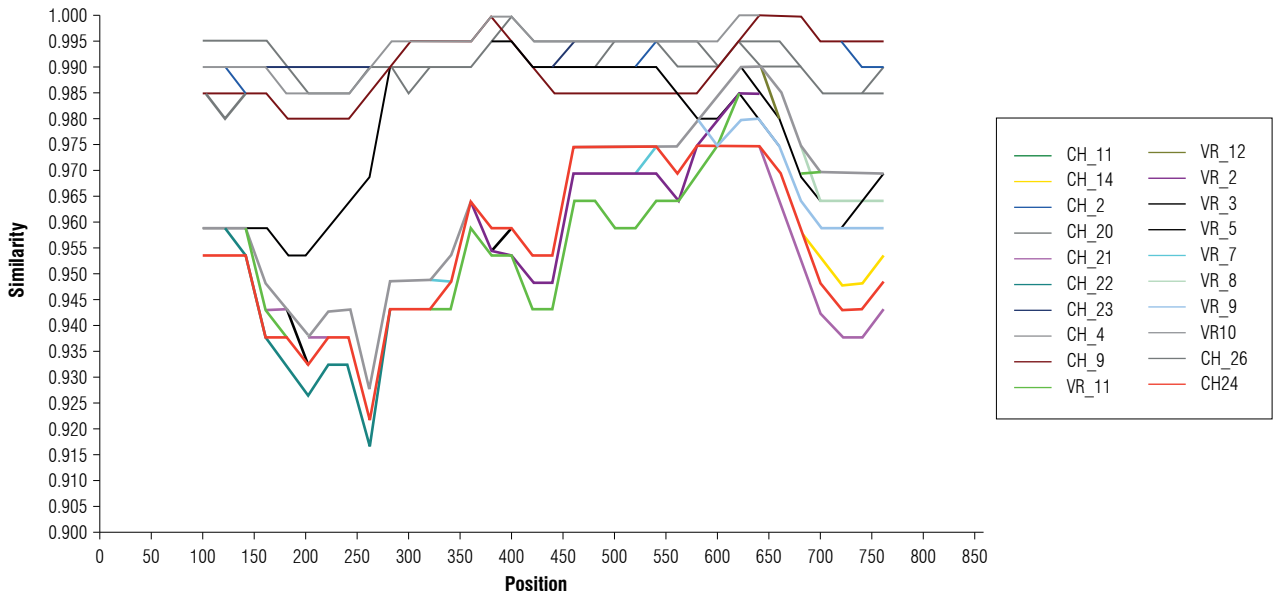

FIGURE 3. Similarity Plot along the CP of the PRSV. CH1 was used as the reference sequence. Each line of a different color represents a particular isolate indicated in the box on the right. The Y-axis shows the percentage of similarity within a sliding window of 200 base pairs.

more similar with isolates from Campo Hermoso and, in the 3' region, the similarity was higher again with the isolates from Villa del Rosario. This similarities in the different CP regions led to the conclusion that the VR3 isolate (black line in Fig. 3) was effectively a natural PRSV recombinant.

Recombinations, such as the one found in the VR3 isolate, occur in the majority of RNA viruses, are of great evolutionary importance and constitute one of the greatest forces that shape the virus genomes of plants (Sztuba-Solinska *et al.*, 2011). Recombinations have been reported in PRSV in almost all of its genome with the exception of the small region that encodes the protein 6K1. The hot spots of recombination in PRSV are concentrated in the region encoding the protein 1 (P1), protein 3 (P3), cytoplasmic inclusion (CI) and the Helper component proteinase (HC-Pro) (Mangrauthia *et al.*, 2008). Recombination events in the CP of PSRV appear to be less frequent than in other regions of the genome (Bateson *et al.*, 2002). The presence of recombinations in PRSV provides an understanding of its molecular evolution and help in the study of the characteristics such as the specificity of the host, geographical distribution and emergence as new epidemics (Mangrauthia *et al.*, 2008).

PRSV is one of the viruses that most limit factors in the production of papaya in Colombia and in tropical regions where the fruit is grown. In Colombia, as part of the management of the disease by PRSV, genetic control by plant selection has been considered for releasing resistant varieties. Nevertheless, in the field, the varieties were susceptible to the virus so the use of transgenic plants with the gene of the CP of PRSV opens new possibilities for solutions to the problem (Paéz, 2003).

The results of this study constitute a contribution to the knowledge on the genomic sequences of the region coding for the CP of PRSV virus isolates collected from symptomatic papaya plants in the department of Norte de Santander and also present the first analysis of the genetic diversity of this virus in Colombia. The results may be of importance to implementing control strategies for PRSV in Colombia in the future through biotechnological approaches. The methods of pathogen-derived resistance as cross-protection (Wang and Yeh, 1997) and transgenic resistance (Gonsalves, 1998) are effective only for isolates that are genetically related (Tennant *et al.*, 1994). For this reason, it is necessary to gain knowledge on the nucleotide composition of viral isolates at the local level, which would provide a better addressing for the design of control strategies.

Acknowledgements

This work was conduced with the suport of the Fondo de Investigaciones Universitarias (FINU), from the Universidad Francisco de Paula Santander, Cucuta-Colombia. Project FINU-020-2013. We are grateful to Diego Armando Viviescas-López and Angie Maite Acevedo-Tarazona, students of Ingeniería Biotecnológica of the Universidad Francisco de Paula Santander, for technical assistance.

Literature cited

Abdalla, O.A. and A. Ali. 2012. Genetic diversity in the 3'-terminal region of papaya ringspot virus (PRSV-W) isolates from watermelon in Oklahoma. Arch. Virol. 157, 405-412. Doi: 10.1007/s00705-011-1184-8

Akhtler, M.S., Y.B. Basavaraj, A.-M. Akanda, B. Mandal, and R.K. Jain. 2013. High incidence of seed transmission of *Papaya ringspot virus* and *Watermelon mosaic virus*, two viruses newly identified in *Robinia pseudoacacia*. Indian J. Virol. 24, 70-73.

Arango, L., C. Román, C. Salamanca, E. Almansa, J. Bernal, G. León, V. Rey, M. Ariza, and P. Gómez. 2010. El cultivo de la papaya en Los Llanos Orientales de Colombia. Technical Assistance Manual No. 4. Corpoica; Sena; Asohofrucol, Bogota.

Bateson, M.F., J. Henderson, W. Chaleeprom, A.J. Gibbs, and J.L. Dale. 1994. Papaya ringspot potyvirus: isolate variability and the origin of PRSV type P (Australia). J. Gen. Virol. 75, 3547-3553. Doi: 10.1099/0022-1317-75-12-3547

Bateson, M.F., R.E. Lines, P. Revill, W. Chaleeprom, C.V. Ha, A.J. Gibbs, and J.L. Dale. 2002. On the evolution and molecular epidemiology of the potyvirus *Papaya ringspot virus*. J. Gen. Virol. 83, 2575-2585. Doi: 10.1099/0022-1317-83-10-2575

Bayot, R.G., V.N. Villegas, P.M. Magdalita, M.D. Jovellana, T.M. Espino, and S.B. Exconde. 1990. Seed transmissibility of papaya ringspot virus. Philipp. J. Crop Sci. 15, 107-111.

Benson, D.A., K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, and E.W. Sayers. 2014. GenBank. Nucleic Acids Res. 42, D32-D37. Doi: 10.1093/nar/gkt1030

Bimboim, H.C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513-1523. Doi: 10.1093/nar/7.6.1513

Boni, M.F., D. Posada, and M.W. Feldman. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics 176, 1035-1047. Doi: 10.1534/genetics.106.068874

Cabrera M., D., D. García H., M. Caballero Á., P. García M., and O. Portal V. 2011. Manejo de la mancha anular de la papaya mediante el uso de malla antiáfidos en viveros de *Carica papaya* L. var. Maradol roja. Fitosanidad 15, 241-244.

Chaves-Bedoya, G., M. Guzman-Barney, and L.Y. Ortiz-Rojas. 2013. Genetic structure and evidence of putative Darwinian diversifying selection in the *Potato yellow vein virus* (PYVV). Agron. Colomb. 31, 161-168.

Chin, M., Y. Rojas, J. Moret, G. Fermin, P. Tennant, and D. Gonsalves. 2007. Varying genetic diversity of *Papaya ringspot virus* isolates from two time-separated outbreaks in Jamaica and Venezuela. Arch. Virol. 152, 2101-2106. Doi: 10.1007/s00705-007-1035-9

Fernández-Rodríguez, T., L. Rubio, O. Carballo, and E. Marys. 2008. Genetic variation of papaya ringspot virus in Venezuela. Arch. Virol. 153, 343-349. Doi: 10.1007/s00705-007-1091-1

García-Arenal, F., A. Fraile, and J.M. Malpica. 2001. Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39, 157-186. Doi: 10.1146/annurev.phyto.39.1.157

Gibbs, M.J., J.S. Armstrong, and A.J. Gibbs. 2000. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16, 573-582. Doi: 10.1093/bioinformatics/16.7.573

Gonsalves, D. 1998. Control of papaya ringspot virus in papaya. Annu. Rev. Phytopathol. 36, 415-437. Doi: 10.1146/annurev.phyto.36.1.415

Gonsalves, D., S. Tripathi, J.B. Carr, and J.Y. Suzuki. 2010. Papaya ringspot virus. The Plant Health Instructor. The American Phytopathological Society (APS), St. Paul, MN. Doi: 10.1094/PHI-I-2010-1004-01

Jain, R.K., K.M. Nasiruddin, J. Sharma, R.P. Pant, and A. Varma. 2004. First report of occurrence of *Papaya ringspot virus* in Bangladesh. Plant Dis. 88, 221. Doi: 10.1094/PDIS.2004.88.2.221C

Jensen, D.D. 1949. Papaya virus diseases with special reference to papaya ringspot. Phytopathology 39, 191-211.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120. Doi: 10.1007/BF01731581

Kosakovsky Pond, S.L., D. Posada, M.B. Gravenor, C.H. Woelk, and S.D. Frost. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. 23, 1891-1901. Doi: 10.1093/molbev/msl051

Kryazhimskiy, S. and J.B. Plotkin. 2008. The population genetics of dN/dS. PLoS Genet. 4, e1000304. Doi: 10.1371/journal.pgen.1000304

Kung, Y.J., H.J. Bau, Y.L. Wu, C.H. Huang, T.M. Chen, and S.D. Yeh. 2009. Generation of transgenic papaya with double resistance to *Papaya ringspot virus* and *Papaya leaf-distortion mosaic virus*. Phytopathology 99, 1312-1320. Doi: 10.1094/PHYTO-99-11-1312

Laney, A., M.V. Avanzato, and I.E. Tzanetakis. 2012. High incidence of seed transmission of *Papaya ringspot virus* and *Watermelon mosaic virus*, two viruses newly identified in *Robinia pseudoacacia*. Eur. J. Plant Pathol. 134, 227-230. Doi: 10.1007/s10658-012-9985-5

Lima, R.C., J.R. Souza, G. Pio-Riberro, and J.A. Lima. 2002. Sequences of the coat protein gene from Brazilian isolates of *Papaya ringspot virus*. Fitopatol. Bras. 27, 174-180. Doi: 10.1590/S0100-41582002000200009

Lole, K.S., R.C. Bollinger, R.S. Paranjape, D. Gadkari, S.S. Kulkarni, N.G. Novak, R. Ingersoll, H.W. Sheppard, and S.C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73, 152-160.

Mangrauthia, S.K., B. Parameswari, R.K. Jain, and S. Praveen. 2008. Role of genetic recombination in the molecular architecture of *Papaya ringspot virus*. Biochem. Genet. 46, 835-846. Doi: 10.1007/s10528-008-9198-y

Mangrauthia, S.K., P. Singh, and S. Praveen. 2010. Genomics of helper component proteinase reveals effective strategy for papaya ringspot virus resistance. Mol. Biotechnol. 44, 22-29. Doi: 10.1007/s12033-009-9205-5

Martin, D.P., P. Lemey, M. Lott, V. Moulton, D. Posada, and P. Lefeuvre. 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462-2463. Doi: 10.1093/bioinformatics/btq467

Martin, D.P., D. Posada, K.A. Crandall, and C. Williamson. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res. Hum. Retroviruses. 21, 98-102. Doi: 10.1089/aid.2005.21.98

Moreno, I.M., J.M. Malpica, J.A. Diaz-Pendon, E. Moriones, A. Fraile, and F. García-Arenal. 2004. Variability and genetic structure of the population of watermelon mosaic virus infecting melon in Spain. Virology 318, 451-460. Doi: 10.1016/j.virol.2003.10.002

Moury, B., C. Morel, E. Johansen, and M. Jacquemond. 2002. Evidence for diversifying selection in *Potato virus Y* and in the coat protein of other potyviruses. J. Gen. Virol. 83, 2563-2573.

Noa-Carrazana, J.C., D. Gonzalez-de-Leon, and L. Silva-Rosales. 2007. Molecular characterization of a severe isolate of papaya ringspot virus in Mexico and its relationship with other isolates. Virus Genes 35, 109-117. Doi: 10.1007/s11262-006-0039-y; 10.1007/s11262-007-0099-7

Olarte Castillo, X.A., G. Fermin, J. Tabima, Y. Rojas, P.F. Tennant, M. Fuchs, R. Sierra, A.J. Bernal, and S. Restrepo. 2011. Phylogeography and molecular epidemiology of *Papaya ringspot virus*. Virus Res. 159, 132-140. Doi: 10.1016/j.virusres.2011.04.011

Padidam, M., S. Sawyer, and C.M. Fauquet. 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218-225. Doi: 10.1006/viro.1999.0056

Paéz, A.R. 2003. Manejo del virus de la mancha anular de la papaya en la región Caribe colombiana. Corpoica, Valledupar, Colombia.

Pamilo, P. and N.O. Bianchi. 1993. Evolution of the Zfx and Zfy genes: rates and interdependence between the genes. Mol. Biol. Evol. 10, 271-281.

Posada, D. and K.A. Crandall. 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA 98, 13757-13762. Doi: 10.1073/pnas.241370698

Price, E.W. and I. Carbone. 2005. SNAP: workbench management tool for evolutionary population genetic analysis. Bioinformatics 21, 402-404. Doi: 10.1093/bioinformatics/bti003

Simmons, H.E., J.P. Dunham, K.E. Zinn, G.P. Munkvold, E.C. Holmes, and A.G. Stephenson. 2013. *Zucchini yellow mosaic virus* (ZYMV, Potyvirus): vertical transmission, seed infection and cryptic infections. Virus Res. 176, 259-264. Doi: 10.1016/j.virusres.2013.06.016

Smith, J.M. 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. 34, 126-129. Doi: 10.1007/BF00182389

Suzuki, J.Y., S. Tripathi, and D. Gonsalves. 2007. Virus-resistant transgenic papaya: Commercial development and regulatory and environmental issues. In: Punja, Z., S.H. De Boer, and H. Sanfacon (eds.). Biotechnology and plant disease management. CABI, Oxfordshire, UK. Doi: 10.1079/9781845932886.0436

Sztuba-Solinska, J., A. Urbanowicz, M. Figlerowicz, and J.J. Bujarski. 2011. RNA-RNA recombination in plant virus replication and evolution. Annu. Rev. Phytopathol. 49, 415-443. Doi: 10.1146/annurev-phyto-072910-095351

Tafur R., R., J.C. Toro, and E. Garcia. 2006. Desarrollo de la fruticultura en el Norte de Santander. Ministerio de Agricultura y Desarrollo Rural; Gobernación de Norte de Santander, Cucuta, Colombia.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599. Doi: 10.1093/molbev/msm092

Tan, Z., Y. Wada, J. Chen, and K. Ohshima. 2004. Inter- and intralineage recombinants are common in natural populations of *Turnip mosaic virus*. J. Gen. Virol. 85, 2683-2696. Doi: 10.1099/vir.0.80124-0

Tennant, P.F., C. Gonsalves, K. Ling, M. Fitch, R. Manshardt, J.L. Slightom, and D. Gonsalves. 1994. Differential protection against papaya ringspot virus isolates in coat protein gene transgenic papaya and classically cross-protected papaya. Phytopathology 84, 1359-1366. Doi: 10.1094/Phyto-84-1359

Tripathi, S., J.Y. Suzuki, S.A. Ferreira, and D. Gonsalves. 2008. *Papaya ringspot virus*-P: characteristics, pathogenicity, sequence variability and control. Mol. Plant Pathol. 9, 269-280. Doi: 10.1111/j.1364-3703.2008.00467.x

Wang, C. and S.D. Yeh. 1997. Divergence and conservation of the genomic RNAs of Taiwan and Hawaii strains of papaya ringspot potyvirus. Arch. Virol. 142, 271-285. Doi: 10.1007/s007050050076

Wilson, C.R. 2014. Applied plant virology. Gutenberg Press, Tarxien, Malta.

Yu, T.A., C.H. Chiang, H.W. Wu, C.M. Li, C.F. Yang, J.H. Chen, Y.W. Chen, and S.D. Yeh. 2011. Generation of transgenic watermelon resistant to *Zucchini yellow mosaic virus* and *Papaya ringspot virus* type W. Plant Cell Rep. 30, 359-371. Doi: 10.1007/s00299-010-0951-4