British Biotechnology Journal
12(3): 1-10, 2016, Article no.BBJ.24362
ISSN: 2231–2927, NLM ID: 101616695

SCIENCEDOMAIN *international*
www.sciencedomain.org

Molecular Diversity Analysis of Coat Protein Gene Encoded by Legume Begomoviruses and PCR Assay to Detect Yellow Mosaic Viruses Infecting Soybean in India

Shunmugiah V. Ramesh[1*], Bhagat S. Chouhan[2], Girish K. Gupta[1], Rajkumar Ramteke[1], Suresh Chand[2] and Syed M. Husain[1]

[1]ICAR- Indian Institute of Soybean Research (ICAR- IISR), Khandwa Road, Indore 452 001, Madhya Pradesh (MP), India.
[2]School of Life Sciences, Devi Ahilya Visva Vidhyalaya (DAVV), Indore, Madhya Pradesh, India.

Authors' contributions

This work was carried out in collaboration between all authors. Authors SVR, GKG and SMH designed the study. Authors GKG and RR contributed in disease symptomatology. Authors BSC and SC performed the bioinformatics analysis. Authors SVR and BSC performed experiments and drafted the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BBJ/2016/24362
Editor(s):
(1) Chung-Jen Chiang, Department of Medical Laboratory Science and Biotechnology, China Medical University, Taiwan.
Reviewers:
(1) Belal Hossain, Sher-e-Bangla Agricultural University, Dhaka, Bangladesh.
(2) Chen-Chin Chang, University of Kang Ning, Taiwan.
Complete Peer review History: http://sciencedomain.org/review-history/13464

Original Research Article

Received 16th January 2016
Accepted 15th February 2016
Published 26th February 2016

ABSTRACT

Aim: Coat protein *(CP)* genes encoded by Legume yellow mosaic viruses (LYMVs) were analysed to study molecular diversity and to devise effective PCR based assay to distinguish major *Begomovirus* species (*Mungbean yellow mosaic India virus* and *Mungbean yellow mosaic virus*) infecting soybean
Design of the Study: All the known coat protein gene sequences encoded by begomoviruses causing yellow mosaic disease (YMD) in legumes were obtained from GenBank. YMD infected soybean leaf samples were collected from different parts of India during Kharif 2012 and species of virus infections identified using *CP* gene based primers in a PCR assay.

**Corresponding author: E-mail: ramesh.sv@icar.gov.in, rameshsvbio@gmail.com;*

Methodology: DNA polymorphism, phylogeny, and test of theory of neutral evolution were studied to decode variability and molecular evolutionary lineage of LYMVs encoded *CP* gene. *CP* based primers have been designed and employed to differentiate MYMIV and MYMV using infected soybean samples collected across India.
Results: Nucleotide diversity and DNA polymorphism studies revealed relatively low levels of diversity in *CP* genes encoded by LYMV isolates. Test of neutral evolution and codon substitution analysis also reiterated the operation of purifying selection indicating deleterious mutations in *CP* gene are not tolerated in the LYMV population. Geographical confinement of species of yellow mosaic viruses infecting soybean is further validated as diseased soybean samples collected from Northern and Central India showed infection due to MYMIV and samples obtained from Southern and Western India were infected with MYMV.
Conclusion: Evolutionary genomic analysis revealed conserved nature of LYMV encoded *CP* gene however a variable region has been identified.PCR assay for differentiation of two major begomoviruses *viz.,*) MYMV and MYMIV infecting soybean in India has been standardised. This is the first report of population genetics in LYMVs and it's implications for yellow mosaic disease (YMD) resistance breeding in soybean are also discussed.

Keywords: Begomovirus; conserved domain; detection; population selection; soybean.

1. INTRODUCTION

Legume yellow mosaic viruses (LYMV) belong to genus *Begomovirus* of family *Geminiviridae* which are transmitted by whiteflies and cause severe yellow mosaic disease (YMD) of legumes in tropical and sub-tropical conditions [1]. Severity of yellow mosaic disease greatly hampers the productivity of the grain legumes including soybean [2]. YMD of legumes in SE Asia is attributed to four species of Begomoviruses such as, *Mungbean yellow mosaic virus* (*MYMV*), *Mungbean yellow mosaic India virus* (*MYMIV*), *Dolichos yellow mosaic virus* (*DoYMV*) and *Horsegram yellow mosaic virus* (*HgYMV*) [2-4]. Economic losses caused due to the infection of yellow mosaic virus accounts for 300 million US $ in all legumes, including Soybean, mungbean, urdbean, field bean etc. [1]. Begomoviruses are characterized with twinned icosahedral particles that encase single-stranded DNA (ssDNA) genomic components. Bipartite genome of legume infecting *Begomovirus* comprise two DNA genomic components namely DNA-A and DNA-B each approximately of 2800 nucleotides long [4-6]. DNA-A encodes for protein involved in virus encapsidation, intracellular replication, and transcription of viral genomic components whereas DNA-B encodes for movement proteins which help the virus in intracellular and intercellular movement within the host [7].

Soybean is cultivated as a major oilseed crop in India occupying ~11 million ha with annual production of ~ 12 million tonnes (Anonymous, ICAR-DSR Vision 2050). The crop, in the recent years, has been playing an important role in the oil economy as it contributes 25% of total edible oil production of the country and earns valuable foreign exchange due to export of soy meal. The potential of the crop is hindered by devastating viral infection caused by *Yellow mosaic virus* which may cause yield loss ranging from 5-90% [2]. The disease is also spreading to new areas of major soybean growing belt in the country which were hitherto free from its incidence.

Virus isolates causing the disease in soybean and other grain legumes are diverse despite no apparent differences in the visible symptoms. In Northern and Western parts of India *Mungbean yellow mosaic India virus* and *Mungbean yellow mosaic virus* cause the disease in soybean respectively [3,8]. In central India, where cultivation of soybean is a major *Kharif* crop, incidence of MYMIV was observed [9]. In addition, prevalence of other LYMVs such as Horsegram yellow mosaic virus and Dolichos yellow mosaic virus is also reported [10]. Incidence of LYMVs in wild species of legumes has also been observed [11]. Simple, reliable detection of yellow mosaic virus infection in legumes is imperative to identify causal organism of the disease. Mixed viral infections and enormous genome diversity of legumoviruses warrant effective and rapid detection. It aids in devising appropriate yellow mosaic disease (YMD) resistance breeding efforts in soybean. Hence it is imperative to differentiate and rapidly detect major *Begomovirus* (MYMV and MYMIV) infecting soybean in India in order to help soybean breeders to develop location and virus specific resistant cultivars. Thus the present

study was carried out to analyse coat protein gene encoded by legume begomoviruses and to devise simple and effective PCR based detection method to differentiate both the major species of *Begomovirus* infecting soybean in India.

2. MATERIALS AND METHODS

2.1 Nucleotide Diversity, DNA Polymorphism and Phylogeny Reconstruction

Complete coat protein *(CP)* gene sequences encoded by LYMVs were obtained from GenBank. DnaSP software [12] was used in determining the nucleotide diversity and DNA polymorphism in the *CP* gene encoded by legume begomoviruses. DNA polymorphism study also involves analysis of haplotypes and haplotype diversity. Phylogeny reconstruction of known complete *CP* gene sequences encoded by legumoviruses was done in MEGA 6 using ClustalW algorithm [13]. The evolutionary lineage analysis was carried out using Maximum Likelihood method in Tamura-Nei model [14] with default parameters and 1000 replicates in the bootstrap analysis.

2.2 Population Selection, Genetic Differentiation, Gene Flow Estimates and Neutrality Tests

All the codons of the *CP* gene in a multiple sequence alignment was analysed individually in SLAC (single like ancestor counting), FEL (fixed effects likelihood), and REL (random effects likelihood) methodologies using DATAMONKEY server (http://www.datamonkey.org). However it was ensured that the recombinant isolates identified from the RDP-4 detection were excluded from the codon substitution analysis. Codon substitution analysis involving mean rates of non synonymous (dN) and synonymous (dS) substitutions and dN/dS ratio were computed. In order to assess the neutral evolution of *CP* gene Tajimas's D [15], Fu & Li's D and Fu & Li's F [16,17] were also computed using DnaSP [12], Furthermore to study the extent of genetic differentiation among the species of LYMVs nucleotide test statistics such as Ks, Kst, Snn (Hudson's statistic of genetic differentiation) Hudson [18] and haplotype statistics such as Hs and Hst [19] were computed using DnaSP along with the estimation of gene flow parameters through statistic Fst [20].

2.3 Multiple Sequence Alignment (MSA) and Designing of Primers

Multiple sequence alignment studies were conducted using CLUSTALW interface available in BioEdit sequence alignment editor [21]. *CP* gene specific primers that distinguish *Begomovirus* species *Mungbean yellow mosaic India virus* (*MYMIV*) and *Mungbean yellow mosaic virus* (*MYMV*) were designed using Primer 3 plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/).

2.4 Virus Source, DNA Extraction and PCR Assay

Soybean leaves that showed typical yellow mosaic symptoms were surveyed and collected from various parts of India (Table 1). Infected leaves were collected based on the presence of characteristic yellowing and mosaic symptoms. Extraction of total DNA from infected leaves and healthy leaf samples (as negative control) were carried out as described previously [22]. In order to distinguish the *Begomovirus* species infecting soybean, specific primers were used in the PCR reaction with 50 ng of extracted DNA as template, 1 µl each of forward and reverse primers (10 µM), 1 µl of dNTPS (10 mM), 1 µl of $MgCl_2$, 1 U of *Taq* polymerase (Fermentas, Massachusetts, USA). The thermal cycler program for PCR amplification comprise 1 cycle of DNA denaturation at 94℃ for 5 min followed by 35 cycles each having a denaturation at 94℃ for 30 sec. annealing for 30 sec (annealing temperature is mentioned in Table 6) and a primer extension at 72℃ for 30 sec followed by final extension of 72℃ for 5 min. To confirm the reliability of the amplified products, amplicons were sequenced at Merck Biosciences (Bengaluru, India).

3. RESULTS AND DISCUSSION

Legumoviruses are serious threat to the cultivation of grain legumes in India [1]. It is also essential to devise effective detection methodology so as to differentiate the *Begomovirus* species infecting grain legumes in general and soybean in particular. Earlier reports indicate that two different species of Begomoviruses are infecting soybean in India [3,8,9]. PCR amplification and RFLP has been used not only to detect the virus species but also to differentiate multiple DNA-B genomic components associated with the YMV causing disease in soybean [3]. In this context it is

imperative to devise efficient and simple, PCR based Begomovirus detection assay so that use of sensitive restriction enzyme digestion of PCR amplicons could be precluded.

3.1 Genetic Diversity and Phylogeny

A total of 102 full length *CP* genes encoded by LYMVs have been analysed for genetic diversity and DNA polymorphism (Table 2). Among the four species of LYMVs, number of segregating sites was found to be increasing with number of isolates under study. However *CP* gene encoded by DoYMV exhibited 120 segregating sites (for 8 isolates under study) despite showing low haplotype diversity of 0.929 among the four species of LYMVs. Nucleotide diversity (π) values are relatively low for HgYMV (0.02073) and MYMIV (0.04180) when compared to MYMV (0.07219).

Molecular phylogeny of complete *CP* genes encoded by LYMVs was reconstructed in maximum likelihood method (Fig. 1). It revealed two major clusters with all the known DoYMV isolates forming a basal cluster to the rest of LYMVs. In the rest of all LYMVs cluster a major sub-clade was formed by MYMIV isolates and a minor cluster with two sub clusters were formed by HgYMV and MYMV isolates. All the virus isolates irrespective of the host legume have grouped based on the virus species. The inherent genetic differences in the *CP* genes of MYMIV and MYMV are further corroborated in the phylogeny reconstruction studies.

3.2 Test of Neutral Evolution and Codon Substitution Analysis

Coat protein *(CP)* gene of LYMVs was evaluated for neutral evolution by studying population selection statistics (Table 3). The parameters revealed that all the species of LYMVs showed negative Tajima's D indicating the operation of purifying selection and population expansion (Table 3). Similarly with other population statistic parameters like Fu & Li's D and Fu & Li's F the major LYMVs, (MYMIV, MYMV), and HgYMV also revealed negative values reiterating the operation of purifying selection and population expansion that could have played a role in the observed diversity. Nevertheless, DoYMV genotype showed positive values for Tajima's D, Fu & Li's D and Fu & Li's F indicating the operation of neutral selection in the population (Table 3). The combination of low negative Tajima's D and negative Fu & Li's D and Fu & Li's F values, indicate that with LYMVs population as a whole is under purifying selection. Purifying selection in *CP* gene thus suggests any mutations in the gene are not tolerated hence immediately purifying action revert the mutations to wild types. It thus elucidates overall low genetic diversity and conserved nature of the gene.

Population selection pressure on the *CP* gene encoded by LYMVs was determined by studying codon substitution analysis (Table 4). The mean ratio of rate of non synonymous substitutions and rate of synonymous substitutions (dN/dS) for

Table 1. Plant material collection details

S. No	Location	Code
1	Punjabrao Deshmukh Krishi Vidyapeeth, RRC, Amravati, Maharashtra	AM
2	Adilabad, Telengana	AD
3	University of Agricultural Sciences (UAS), Bengaluru, Karnataka	BG
4	Coimbatore, TamilNadu	CB
5	University of Agricultural Sciences (UAS), Dharwad, Karnataka	DW
6	Kota, Rajasthan	KT
7	ICAR-Directorate of Soybean Research, Indore, Madhya Pradesh	ID
8	Jhansi, Uttar Pradesh	JH
9	JN Krishi Vishwa Vidhyalaya (JNKVV), Jabalpur, Madhya Pradesh	JB
10	Kangra, Himachal Pradesh	KG
10	Punjab Agricultural University, Ludhiana, Punjab	LU
11	Indian Agricultural Research Institute (IARI), New Delhi	ND
12	GB Pant University of Agricultural and Technology (GBPAU&T), Pant Nagar, Uttarkhand	PN
13	Agharkar Research Institute (MACS) Pune, Maharashtra	PU
14	R&D Unit, Ugar Sugar Works, Ltd., Ugharkhurd, Karnataka	UG
15	ICAR Research complex for NEH region, Umiam, Meghalaya	UM
16	Healthy leaf sample from ICAR-DSR, MP (NC)	NC

under purifying selection. Another parameter, dN-dS computed from REL methodology showed values -0.806 and -0.807 for *CP* gene from MYMIV and MYMV respectively suggesting that all the codons are under the influence of purifying selection so that deleterious mutations are being removed from the population. MYMIV population showed four positively selected codon sites (22, 36, 116 and 184) along with three negatively selected codon positions (29, 79b, 28) whereas MYMV showed two positively selected codons (36 and 210) and three negatively selected codon positions (14, 69, and 137). The amino acid substitutions in the corresponding positions are mentioned in the Table 4. The other LYMVs also followed similar trend showing positive dN/dS values and negative dN-dS values implying codons are under purifying selection (Table 4). However HgYMV isolates showed no positively selected codon sites (Table 4).

Table 2. Genetic diversity of Coat protein *(CP)* gene of legume infecting begomoviruses

Virus species	N	*S*	π	Hd
MYMIV	62	254	0.04180	0.997
MYMV	25	230	0.07219	0.987
DoYMV	8	120	0.06233	0.929
HgYMV	7	46	0.02073	0.952
All	102	370	0.13505	0.997

N, number of isolates; S, number of polymorphic (segregating) sites; Hd, haplotype diversity; π, nucleotide diversity within species. Mungbean yellow mosaic India virus (MYMIV); Mungbean yellow mosaic virus (MYMV); Dolichos yellow mosaic virus (DoYMV); Horsegram yellow mosaic virus (HgYMV)

3.3 Genetic Differentiation and Gene Flow

The major LYMVs (MYMIV and MYMV) are genetically isolated as genetic differentiation estimates reveal higher S_{nn} values (values close to 1 indicates genetic differentiation) and K_{st} values are near zero reiterating inherent differences of *CP* gene at the genetic level. Further, frequency of genetic exchanges was also estimated using statistic F_{st} (F_{st} value close zero indicates frequent gene flow whereas value towards one indicates infrequent gene flow). Both MYMIV and MYMV reveal infrequent gene flow with other LYMVs (0.51893 and 0.52456) respectively and among themselves (0.68745) (Table 5).

Among the legumovirus genes *CP* was found to be suitable for differentiating the virus species infecting legumes in the Indian context [10]. Coat protein gene encoded by LYMVs were analysed for their conserved sequence features, and to test for neutral evolution. Nucleotide diversity and DNA polymorphism studies revealed relatively low diversity in *CP* genes encoded by LYMV isolates. Between MYMV and MYMIV the former showed little more diversity compared to the latter. Test of neutral evolution reiterated the operation of purifying selection thus indicating deleterious mutations in *CP* gene are not tolerated in the population. Phylogeny reconstruction also indicates clades are formed based only on the *Begomovirus* species irrespective of the crop species infected. Evidence of recombination among the isolates infecting diverse legumes is reminiscent of recombination between *Begomovirus* infecting grain legumes and other non-legumes as reported in phylogeographic analysis of begomoviruses infecting legumes and beta satellites [23]. Similar global population selection analysis based on *nucleocapsid protein gene (N gene*) of Iris yellow spot virus (IYSV) led to identification of and factors contributing to Tospovirus evolution [24]. Population selection analysis of Tomato spotted wilt virus (TSWV) identified purifying selection and population expansion influence its evolution [25] whereas neutral evolution of codons and genetic drift has been found to shape the evolution of Fig mosaic virus [26].

Table 3. Neutrality tests in *CP gene* encoded by begomoviruses infecting legumes to test theory of neutral evolution

Virus species	N	Tajimas's D	Fu &Li's D	Fu &Li's F
MYMIV	62	-1.98754	-3.70674	-3.62486
MYMV	25	-1.17959	-0.28995	-0.67846
DoYMV	8	-0.40757	0.12653	0.00147
HgYMV	7	-1.23529	-1.06903	-1.22454
All	102	-0.38242	-0.47610	-0.52119

(N-number of isolates, Tajimas's D, Fu &Li's D, and Fu &Li's F refer to the test statistic parameters)

Table 4. Codon-substitution analysis in coat protein *(CP)* of legume infecting begomoviruses

Genotype	Positively selected codons	Amino acid substitutions	Negatively selected codons	ω=*d*N/dS	dN-dS
MYMIV	22 [b]	Asn-Pro,Thr,Lys	29[a] 79[b]	0.223485	-0.806
	36 [b]	Thr-Val,Ser,Ala	28[c]		
	116 [b]	Tyr-Asp,Trp			
	184 [c]	Ser-Asn,Val,Gln			
MYMV	36 [a]	Val-Ala,Pro	14[a] 69 [b]	0.288406	-0.807
	210 [a]	Val-Arg,His	137[c]		
DoYMV	17 [c]	Val-Thr	9[a] 46 [b]	0.117865	-0.913
	25 [c]	Thr-Val,Ser Ser-Asn Val-Val Thr-Thr	6[c]		
	44 [c]	Pro-Ser,Gln,Ala			
	46 [c]	Thr-Gln			
HgYMV	-	-	2[a] 15 [b] 36[c]	0.0982273	-0.934

[a]Codon sites from SLAC at a cut off p-value 0.1; [b]-codon sites identified in FEL at a cut off p-value 0.1; [c]-codon sites from REL at a cut off Bayes factor value 50. dN, the number of non-synonymous substitutions per non-synonymous site; dS, the number of synonymous substitutions per synonymous site; ω - ratio of dN/dS from SLAC (single like ancestor counting) methodology ; dN-dS obtained from REL (random effects likelihood)

Table 5. Genetic differentiation and gene flow estimates

Virus species	H_s	H_{st}	χ2	P value	K_t	K_s*	K_{st}*	S_{nn}	Z*	F_{st}
MYMIV vs other *LYMVs*	0.99458	0.00271	99.902	0.1816	91.83207	3.58029	0.13288	0.94526	7.12289	0.51893
MYMV vs other *LYMVs*	0.99452	0.00277	99.298	0.1929	91.83207	3.71318	0.10069	0.94526	7.20555	0.52456
MYMIV vs *MYMV*	0.99401	0.00306	84.558	0.2347	69.39508	3.23554	0.15254	0.93582	6.80915	0.68745

H_s, H_{st}: Haplotype based statistic to estimate genetic differentiation; K_s, K_{st}, S_{nn}, Z: Nucleotide based test statistic to estimate the genetic differentiation (Kst value close to zero indicates no differentiation; Snn value close to one indicates differentiation); F_{st}: Statistic estimates the extent of gene flow between various genotypes (Value close to zero indicates free gene flow or panmixis value close to one indicates genotypic groups are closed to gene flow)

Table 6. Oligonucleotides employed in the differentiation of Begomoviruses infecting soybean

S. No	Primer	Sequence	Viral target region	Annealing temperature (°C)	Amplicon size
1	RUGEMF1 RUGEMR1	5' TGTGAGGGACCATGTAAAGTTC 3' 5' GCATGAGTACATGCCATATAC 3	Conserved *CP* (AV-1)	57℃	447bp
2	MYMV F YMV R	5' GTGTTAAGTCTATCTGGG 3' 5'CACAGGATTTGATGCATGAG 3'	MYMV *CP* region	50℃	391
3	MYMIV F	5' GCATCAAGTCCGTGTACATTAC 3'	MYMIV *CP* region	50℃	391

(Initially primers RUGEMF1 and RUGEMR1 have been used to detect legume infecting begomovirus then positive DNA samples from this detection were used as a template to differentiate begomovirus species using corresponding set of primers [MYMV F/ MYMIV F and YMV R]. Note that universal reverse (YMV R) primer was used to distinguish MYMIV and MYMV)

Thus the molecular evolutionary genomics analysis of LYMV encoded *CP* genes indicated that the gene is conserved across the virus species. Hence any PCR assay based on *CP* gene would be appropriate to differentiate the major *Begomovirus* infecting soybean.

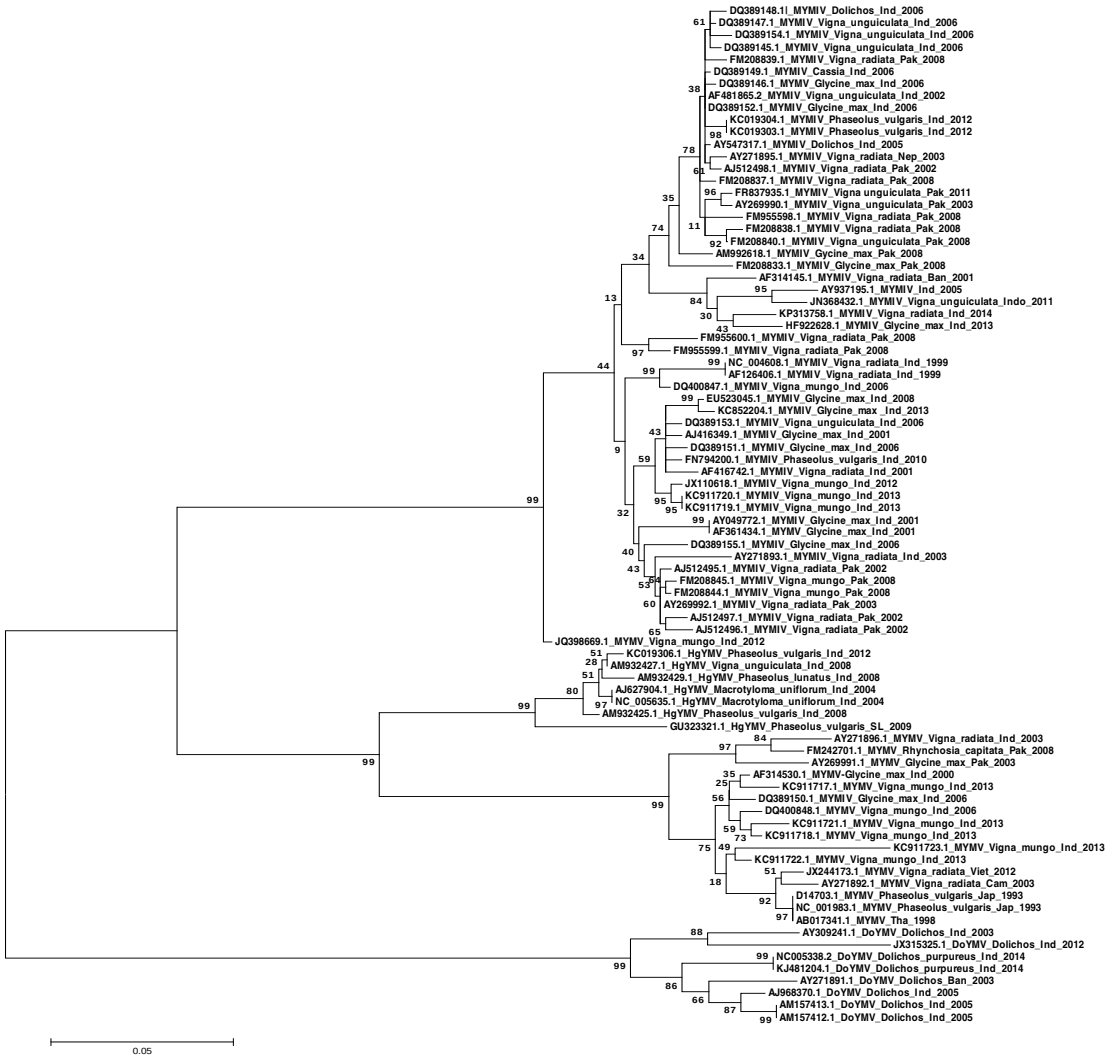

Fig. 1. Molecular phylogeny reconstruction of coat protein gene encoded by LYMVs

(Evolutionary history was inferred using the Maximum Likelihood method with 1000 bootstrap replicates based on the Tamura-Nei model. The tree with the highest log likelihood (-6168.2588) is depicted)

3.4 Multiple Sequence Alignment (MSA) and Oligonucleotides Designing

All known coat protein gene sequences of MYMIV and MYMV infecting legumes were analysed using CLUSTALW. Despite the conserved nature of *CP* gene sequence, a stretch of nucleotide sequences were identified that differentiates both the species of the Begomovirus infecting soybean. *In silico* analysis identified two different forward primers (MYMV F: GTGTTAAGTCTATCTGGG to detect MYMV and MYMIV F: GCATCAAGTCCGTGTACATTAC to detect MYMIV) in combination with a single reverse primer (YMV R: CACAGGATTTGATGCATGAG) to distinguish the virus species in infected plant samples (Table 6).

3.5 PCR Assay

In order to detect the presence of legume infecting yellow mosaic virus from the infected plant DNA samples, PCR amplification of conserved DNA-A was carried out using the primers RUGEMF1 5' TGTGAGGGACCATGTAAAGTTC 3' RUGEMR1 10: 5' GCATGAGTACATGCCATATAC 3' [3] (Table 6). All the samples except four (KT, JH, UG, UM) were found to be positive for LYMV implicating the presence of Begomovirus infection in those samples (Fig. 2). Despite the

apparent mosaic symptoms leaf samples from Hissar, Jhansi, Ugharkhurd and Umiam showed absence of legume yellow mosaic virus infections. DNA samples extracted from those positive leaves were further subjected to PCR amplification in order to detect *Begomovirus* species. Each DNA sample was used as template simultaneously in two PCR reactions each specific for detection of *Begomovirus* species infecting soybean. Soybean samples collected from Indore (ID), Jabalpur (JB), Kangra (KG), Ludhiana (LU), New Delhi (ND), and Pant Nagar (PN) showed positive amplification specific for MYMIV (Fig. 3) whereas soybean samples obtained from Amravati (AM), Adilabad (AD), Bengaluru (BG), Coimbatore (CB), Dharwad (DW) and Pune (PU) were positive for MYMV infection (Fig. 4). Thus soybean samples obtained from Northern and Central India showed presence of MYMIV infection whereas samples from Southern and Western region of the country showed MYMV infection.

Begomoviruses infecting soybean have been found to have geographical confinement too [8-9] further legumoviruses are genetically isolated [4]. The results of the present investigations also corroborate that MYMIV was found to be infecting in Northern, and Central India, whereas MYMV infection in soybean was observed in Southern and Western regions of the country. Breeding for YMD resistant cultivars in soybean requires easy and reliable identification of *Begomovirus* species prevalent in the region.

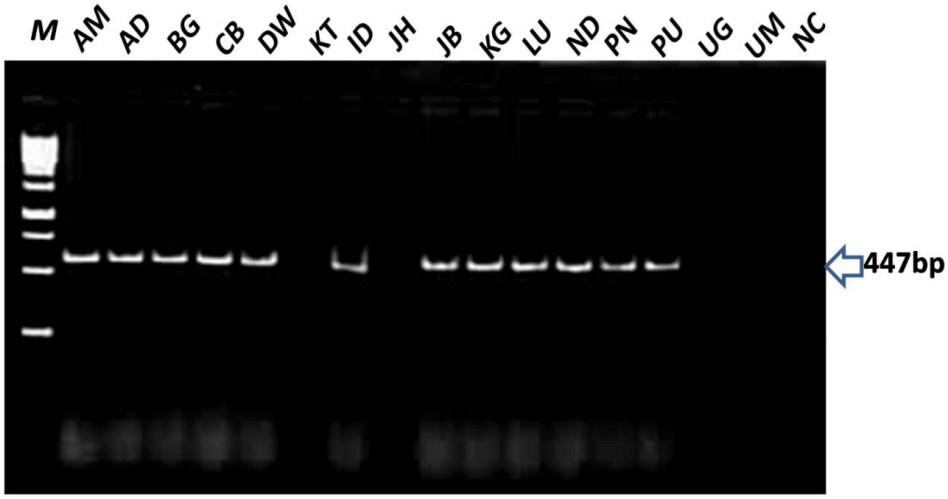

Fig. 2. PCR amplification of conserved coat protein (AV1) region using RUGEMF1 and RUGEMR1 primers to confirm the presence of legume infecting begomoviruses

(Lane M: Marker (Gene Ruler 1kb DNA ladder Thermo Fisher Scientific); Codes for lanes and locations where from infected samples were collected are presented in Table 1)

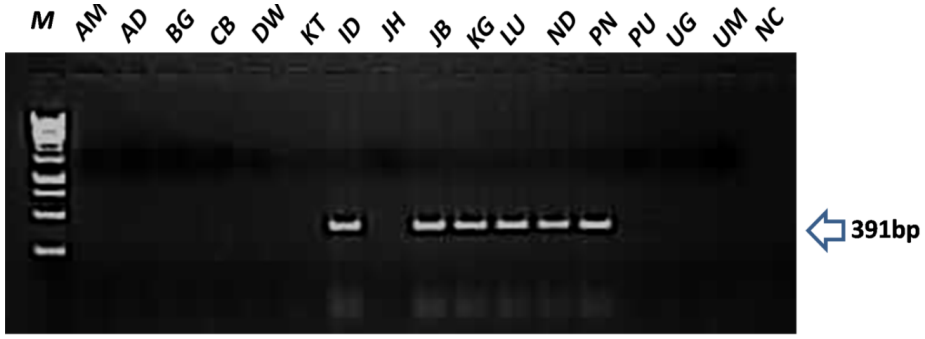

Fig. 3. PCR amplification and validation of primers to amplify only MYMIV encoded coat protein *(AV1)* region using MYMIV F and YMV R primers using DNA template extracted from infected soybean

(Lane M: Marker (Gene Ruler 1kb DNA ladder Thermo Fisher Scientific); Codes for lanes and locations where from infected samples were collected are presented in Table 1)

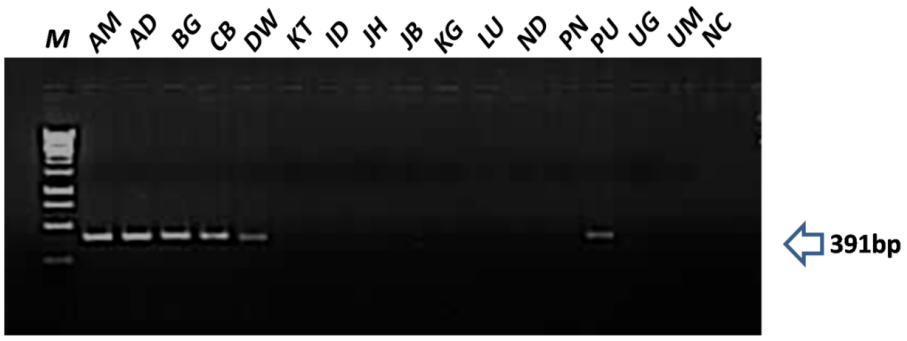

Fig. 4. PCR amplification and validation of primers designed to amplify only MYMV encoded coat protein *(AV1)* region using MYMV F and YMV R using DNA template extracted from infected soybean

(Lane M: Marker(Gene Ruler 1kb DNA ladder Thermo Fisher Scientific); Codes for lanes and locations where from infected samples were collected are presented in Table 1)

It thus helps the breeders to screen soybean genetic stocks against specific YMV infection and to identify any resistant differentials in order to use those genotypes in developing YMV resistant cultivars. Genetics of YMD resistance in soybean has not been linked to any particular species of *Begomovirus*. However the revelation of two different species of *Begomovirus* causing disease and their geographical confinement requires breeders to develop region specific YMV resistant soybean cultivars. In this regard the present study has a bearing on the detection of *Begomovirus* species and development of resistant cultivars.

4. CONCLUSION

Molecular evolutionary genomics analysis of LYMV encoded *CP* genes indicated conserved nature of the gene. Multiple sequence alignment studies identified a set of forward primers in the variable *CP* region to distinguish MYMIV and MYMV. Furthermore, to circumvent the use of restriction enzyme digestion, here we devised an easy to use PCR detection method wherein using virus specific forward primer along with a common reverse primer both the virus species could be differentiated effectively. The results of the present investigations also corroborate that MYMIV was found to be infecting in Northern, and Central India, whereas MYMV infection in soybean was observed in Southern and Western regions of the country

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Varma A, Malathi VG. Emerging Geminivirus problems A serious threat to crop production. Ann Appl Biology. 2003;142:145-164.
2. Gupta GK, Chauhan GS. Symptoms, identification and management of soybean diseases. National Research centre for Soybean (ICAR), Indore, India, 2005;92.
3. Girish KR, Usha R. Molecular characterization of two soybean-infecting begomoviruses from India and evidence for recombination among legume- infecting Begomoviruses from South-East Asia. Virus Res. 2005;108:167–176.
4. Qazi J, Ilyas M, Mansoor S, Briddon RW. Legume yellow mosaic viruses: Genetically isolated begomoviruses. Mol Plant Pathology. 2007;8:343-348.
5. Briddon RW, Patil BL, Bagewadi B, Nawaz-ul-Rehman MS, Fauquet CM. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evolutionary Biology. 2010;10:97.
6. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva Jose CF, et al. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Arch Virology. 2015;160: 1593–1619.
7. Fauquet C, Briddon RW, Brown J, Moriones E, Stanley J, Zerbini M, et al. Geminivirus strain demarcation and nomenclature. Archives of Virology. 2008; 153:783–821.

8. Usharani KS, Surendranath B, Haq QMR, Malathi VG. Yellow mosaic virus infecting soybean in northern India is distinct from the species infecting soybean in southern and western India. Current Science. 2004;86:845–850.
9. Ramesh SV, Bhaskale R, Admane N, Gupta GK, Husain SM. Multiply Primed Rolling Circle Amplification (MPRCA) of *Yellow mosaic virus* genome from infected soybean in central Indian region divulges it as *Mungbean yellow mosaic Indian virus*-[sb] and its implications for RNAi mediated virus resistance. World Soybean Research Conference-IX (WSRC-IX); 2013. Durban South Africa.
10. Maheshwari R, Panigrahi G, Angappan K. Molecular characterization of distinct YMV (*Yellow mosaic virus*) isolates affecting pulses in India with the aid of coat protein gene as a marker for identification. Molecualr Biology Reports. 2014;41:2635-2644.
11. Naimuddin MA, Pratap A. First report of natural infection of Mungbean yellow mosaic India virus in two wild species of *Vigna*. New Disease Reports. 2011;23.
12. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451–1452.
13. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mole Bio Evol. 2013;30:2725-2729.
14. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mole Bio Evol. 1993;10:512-526.
15. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–595.
16. Fu Y. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147:915–925.
17. Fu Y, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993;133:693–709.
18. Hudson RR. A new statistic for detecting genetic differentiation. Genetics. 2000;155: 2011–2014.
19. Hudson RR, Boos DD, Kaplan NL. A statistical test for detecting geographic subdivision. Mole Bio Evol. 1992a;9:138–151.
20. Hudson RR, Slatkin M, Maddison WP. Estimations of levels of gene flow from DNA sequence data. Genetics. 1992b;132: 583–589.
21. Hall TA. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 1999;41:95–98.
22. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–15.
23. Ilyas M, Qazi J. Mansoor S, Briddon RW. Genetic diversity and phylogeography of begomoviruses infecting legumes in Pakistan. J Gen Virolology. 2010;91:2091-2101.
24. Iftikhar R, Ramesh SV, Bag S, Ashfaq M, Pappu HR. Global analysis of population structure, spatial and temporal dynamics of genetic diversity, and evolutionary lineages of Iris yellow spot virus (*Tospovirus*: *Bunyaviridae*). Gene. 2014;547:111-118.
25. Sundaraj S, Srinivasan R, Culbreath AK, Riley DG, Pappu HR. Host plant resistance against tomato spotted wilt virus in peanut (*Arachis hypogaea*) and its impact on susceptibility to the virus, virus population genetics, and vector feeding behavior and Survival. Phytopathology. 2014;104:202-210.
26. Walia JJ, Willemsen A, Elci E, Caglayan K, Falk BW, Rubio L. Genetic variation and possible mechanisms driving the evolution of worldwide Fig mosaic virus isolates. Phytopathology. 2014;104.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/13464