

Comparative *In silico* Analysis of Partial Coat Protein Gene Sequence of Zucchini Yellow Mosaic Virus Infecting Summer Squash (*Cucurbita pepo* L.) Isolated From India

Sharma Neha*, Bhardwaj Satya Vrat, Jarial Kumud, Thakur P D, Kaur Rajinder and Handa Anil

Dr Y S Parmar University of Horticulture and Forestry, Solan 173230, (HP)

Abstract

Zucchini yellow mosaic virus (ZYMV; Family: *Potyviridae*, Genus: *Potyvirus*) is a serious virus infecting summer squash causing severe damage to both crop and ornamental cucurbit crop production. In the present study, molecular characterization of ZYMV (at genomic and proteomic level) infecting summer squash was carried out and cDNA of approximately 700bp was amplified. The PCR amplified product was further sequenced and analyzed. The sequence of partial coat protein of 154 nucleotides of Indian isolate of zucchini yellow mosaic virus (ZYMV) was determined and translated to proteins. Later, the sequence was submitted to NCBI and has got accession no. GU144796 with protein id ACZ36948. In BLASTN analysis, nucleotide test sequence showed 91% homology with D13914 (sequence from USA), whereas, protein test sequence was 75.9% homologous in BLASTP analysis with a number of protein sequences present in the database. The alignment score of test sequence with 67 other isolates of ZYMV retrieved from NCBI database was highest for USA among varied countries and lowest for China in case of nucleotides and Korea in case of proteins. Phylogenetic analysis revealed similarity of the test virus sequence with a USA ZYMV CP (D13914) and similarity of the partial polyprotein sequence with that of Japan (BAE75935). Conserved domain of the test virus was found to show homology with the potyvirus coat protein domain alignment collection (pfam00767). Computational restriction digestion revealed that 22 different restriction enzymes restrict present isolate of ZYMV. Secondary structures for polyprotein of the test virus was predicted which inferred dominance of alpha (α) helix in the protein sequence.

Keywords: Cucurbitaceae; cDNA; Potyvirus; Sequence analysis; Squash; ZYMV

Introduction

Zucchini yellow mosaic virus (ZYMV) generally produces symptoms like mosaic, yellowing, shoe stringing of leaves, fruits and seed deformations and stunting of plants [1] and infection at early stage of the crop could cause as much as 94 per cent reduction of marketable fruits of summer squash. ZYMV In this study, the partial coat protein gene sequence of ZYMV of Indian isolate was determined and compared with 67 other isolates of ZYMV at both genomic and proteomic level because a better knowledge of the virus characterization and tracing its phylogeny provide a major contribution to understand the complexity and epidemiology of the pathogen and hence management of the disease.

Materials and Methods

Collection of samples and maintenance of the virus isolate

Tender leaves of summer squash plants showing symptoms of ZYMV were collected from the hill state of Himachal Pradesh located in the North Western Himalayan regions and the virus cultures were maintained on healthy seedlings of summer squash variety "Australian Dark Green" by mechanical sap inoculation under insect proof glass house conditions.

Serological Indexing of the virus

Serological detection and identification of virus was carried out by following enzyme linked immunosorbant assay (ELISA). Commercially available immunoreagents (BIOREBA – AG Switzerland) were used and protocols of suppliers of ELISA kits were used. The presence of potyvirus in experimental test plants was checked by DAC ELISA and the ZYMV presence was detected by DAS ELISA.

RNA extraction and cDNA synthesis

The RNA of the test virus was extracted from the experimental test plants using the RNA easy Plant Mini Kit (Qiagen). The RNA was used as a template for cDNA synthesis by using specific Oligonucleotide primer p9502 shown in Table 1.

Amplification of cDNA using specific primers

For further amplification of cDNA, PCR was carried out in a thermocycler (Applied Biosystem, USA) using specific primers p9502 (as reverse primer) and CPUP (as forward primer) shown in Table 1. PCR was run for 40 cycles and final elongation was carried out at 72°C for 10 minutes.

Sequencing and sequence analysis

Sequencing using both reverse and forward primers was carried out [2] and the sequence so obtained was translated using EXPASY (Expert Protein Analysis System) tool. Nucleotide and protein sequence data was analyzed using BLAST program at NCBI website whereas Clustal W was performed for multiple sequence alignment of the test sequence with other 67 isolates of ZYMV available in NCBI database. Phylograms

*Corresponding author: Sharma Neha, Dr Y S Parmar University of Horticulture and Forestry, Solan 173230, (HP), India, E-mail: nehabtc@gmail.com

Received November 18, 2010; Accepted February 22, 2011; Published February 24, 2011

Citation: Neha S, Vrat BS, Kumud J, Thakur PD, Rajinder K, et al. (2011) Comparative *In silico* Analysis of Partial Coat Protein Gene Sequence of Zucchini Yellow Mosaic Virus Infecting Summer Squash (*Cucurbita pepo* L.) Isolated From India. J Proteomics Bioinform 4: 068-073. doi:10.4172/jpb.1000168

Copyright: © 2011 Neha S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and phylogenetic trees analysis was carried out using maximum likelihood (ML), maximum parsimony (MP), neighbor joining (NJ) and unweighted pair group mathematical averages (UPGMA) methods using phylip 3.68 and EXOME™. Restriction enzyme map was constructed (<http://tools.neb.com/NEBcutter2/index.php>) and conserved domain search (<http://www.ncbi.nlm.nih.gov/structure/cdd/wrpsb.cgi>) and secondary structure prediction of proteins was carried out (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA_server.html). Amino acid content calculation was performed to know more about the protein test sequence.

Name of primer	Sequence	No. of Bases	Designated as
P9502	5'- GCGGATCCTTTTTTTTTTTTTTTT - 3'	25	Reverse
CPUP	5'-TGAGGATCTGGTGYATHGARAAGG - 3'	25	Forward

Table 1: Primers used for PCR amplification and Sequencing.

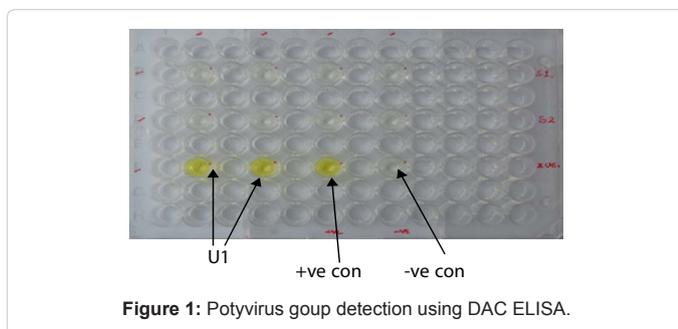


Figure 1: Potyvirus goup detection using DAC ELISA.

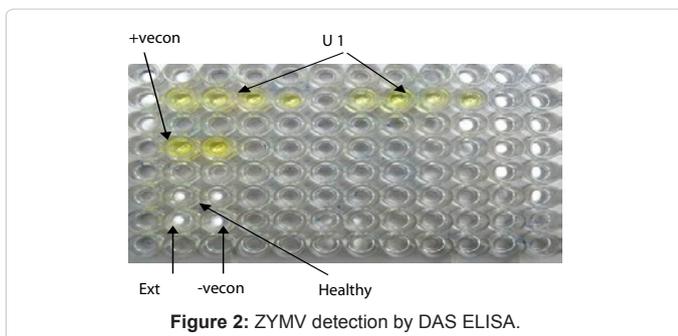


Figure 2: ZYMV detection by DAS ELISA.

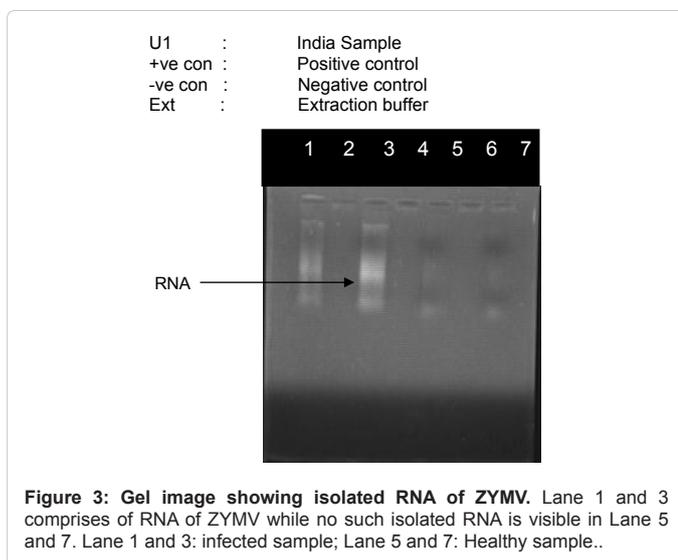


Figure 3: Gel image showing isolated RNA of ZYMV. Lane 1 and 3 comprises of RNA of ZYMV while no such isolated RNA is visible in Lane 2 and 4. Lane 1 and 3: infected sample; Lane 5 and 7: Healthy sample..

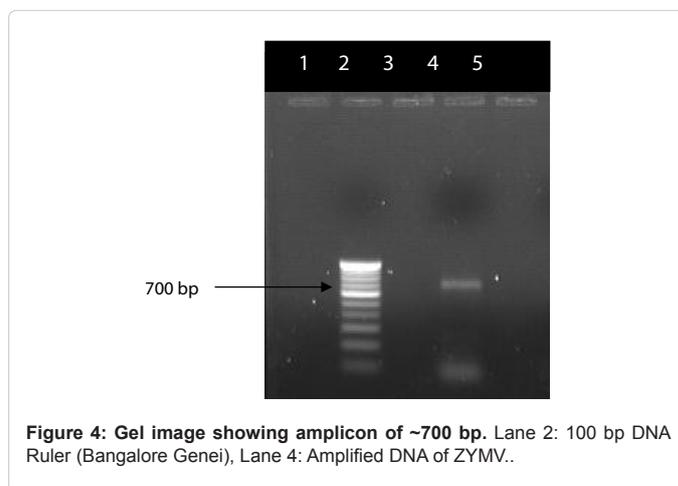


Figure 4: Gel image showing amplicon of ~700 bp. Lane 2: 100 bp DNA Ruler (Bangalore Genei), Lane 4: Amplified DNA of ZYMV..

Nitrogenous Base	Nucleotide Count	
	Total	Percentage (%)
Adenine (A)	41	26.62
Thymine (T)	35	22.73
Cytosine (C)	37	24.03
Guanine (G)	41	26.62
G+C	78	50.65
A+T	76	49.35

Table 2: Nucleotide base composition in the query sequence (Indian isolate of ZYMV).

Results

ZYMV isolate of India was collected from naturally infected summer squash plants shown in Figure 9 and the serological tests showed presence of potyvirus and ZYMV in particular using DAC and DAS ELISA respectively. Results of serological indexing indicated presence of the virus in Indian sample and the concentration of the virus was also high in the sample Figures 1 and 2.

Fresh leaves from infected and healthy plants were collected and then used for RNA isolation. Presence of RNA was checked by running it on 0.8% agarose gel under 100 V potential difference for 1.5 hours in TAE buffer (Figure 3). The RNA was reverse transcribed into cDNA (complementary DNA) with RT-PCR. This RT-PCR was followed by amplification of the cDNA with PCR. Amplified PCR product ~700bp obtained from the sample was subjected to agarose gel electrophoresis using 100 bp DNA ladder (Banglore Genei) as molecular weight marker (Figure 4).

After sequencing, the sequence so obtained was of 154 nucleotides, is as under:

```
1 gctacgaac ctacgggata gcagtctcac acttgacgct ttcgatttct atgaagtcaa
61 ttctacaact cctgaaagag cccgtgtagc ttagcgcag atgaaagcag cagctcttag
121 caatgtttct tcaaggcggg ttggcatagg tgat
```

It was further observed that the sequence consisted of 41, 35, 37 and 41 of A, C, G and T respectively (Table 2).

Nucleotide sequence accession number

The nucleotide sequence reported in this paper has been deposited in the GenBank database under accession number GU144796 at NCBI.

Translation of the test sequence was carried out

SLTLDAFDYEVNSTTPERARVAQMetKAAALSNVSSRRFGIGD

BLAST

The nucleotide and protein sequence of Indian isolate was analyzed using Blast program available at NCBI with test virus coat protein gene sequences from all around the world, available in the database of NCBI.

BLASTN and BLASTP analysis

In BLASTN analysis at NCBI, nucleotide test sequence showed 91% homology with D13914 (sequence from USA), and 90% homology with isolates ZYMV C-16 (DQ645729) and TW-NT1 (AF127933). Whereas, protein test sequence was 75.9% homologous in BLASTP analysis with number of protein sequences present in the database e.g. CAB63753, AAQ17214 (Korean isolate), ABL01532 (Israel isolate), AAO61299 (France isolate), AAA48511, BAE75934, BAA00596 (USA isolate), ABL01531 (Israel isolate), CAC87636 (China isolate), CAD12311 (Austria isolate), ACB47285, BAD74201 (Japan isolate) and 74.3% homologous with ABL09422 (Australian isolate) and AF513552 (China isolate) of Zucchini yellow mosaic virus.

Multiple sequence alignment of selected nucleotide and protein sequences of zucchini yellow mosaic virus with that of Indian isolate was performed using CLUSTAL W program [3] available online at European Bioinformatics institute (EBI) (<http://www.ebi.ac.uk/>) and similarly, country wise CLUSTAL W alongwith query nucleotide and

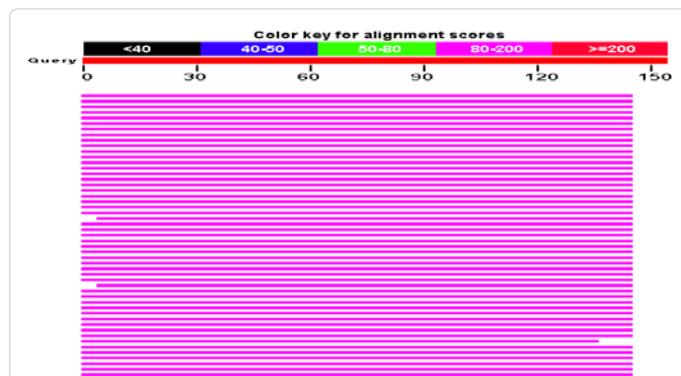


Figure 5: Color key for alignment score of nucleotide query sequence Indian isolate with other nucleotide sequences present in the database using BLASTN analysis (Pink lines show the similarity of the query sequence is more than 80%).

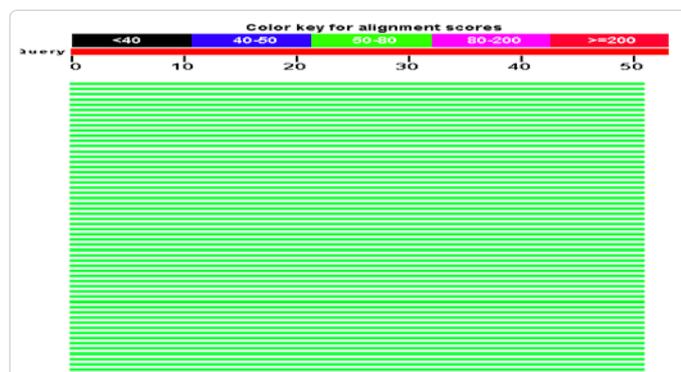


Figure 6: Color key for alignment score of protein query sequence Indian isolate with other protein sequences present in the database using BLASTP analysis (Green lines show the similarity of the query sequence is not more than 80%).

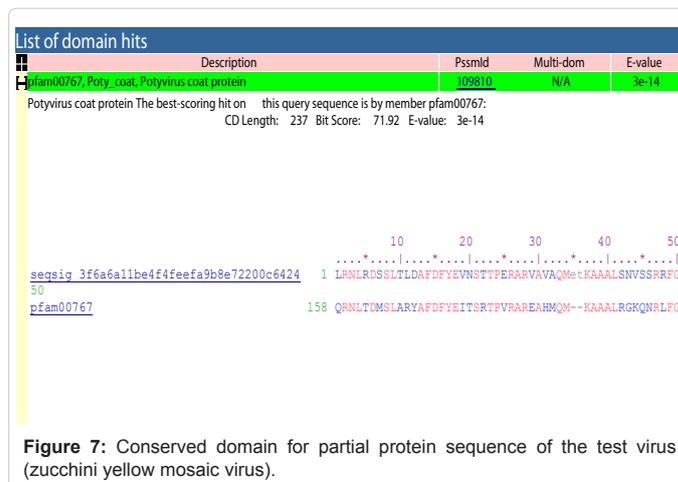


Figure 7: Conserved domain for partial protein sequence of the test virus (zucchini yellow mosaic virus).

protein sequence was also performed and these CLUSTAL W outputs were then used in (phylip 3.68 and EXOME™ software) bioinformatic tools for constructing phylograms and phylogenetic trees.

Phylograms (using phylip 3.68) and phylogenetic trees (using EXOME™) analysis was carried out using the test sequence (nucleotide and protein) with other isolates of ZYMV from eight different countries retrieved from NCBI. And these studies revealed similarity of the test nucleotide with D13914 ZYMV isolate of USA, while in case of proteins the test protein found to be most similar to Japanese sequence of BAE75935.

64 phylograms (32 for nucleotide and 32 for protein) and 64 phylogenetic trees (32 for nucleotide and 32 for protein) were constructed using phylip 3.68 and EXOME™, respectively. On phylogenetic studies, in case of nucleotide sequences the test sequence showed common ancestry with DQ925447 (Australian), AJ420020 (Austrian), AB188115 and AB188116 (Japanese), AJ429071 (Korean), AF127933 (Taiwanese), D13914 (USA among varied countries) isolates of ZYMV whereas, in case of proteins the test protein sequence showed common ancestry with ABL0422 (Australian), CAD12315 (Austrian), CAD31036 and CAD31056 (Hungarian), BAE75935 and BAD74201 (Japanese) isolates of ZYMV.

Conserved domain search [4] of the test isolate was found with 3e-14E-value and a bit score of 71.92 (Figure 7). Further Conserved Domain Architecture Retrieval Tool (CDART) [5] was employed to perform similarity search for conserved domain in the test isolate with the NCBI Entrez Protein Database based on domain architecture (Figure 8).

Restriction enzyme maps for one cutter (Table 3 and 4), two cutter (Table 5) and three cutter (Table 6) restriction enzymes were constructed for the nucleotide query sequence using online web tool (NEB cutter 2.0). Secondary structure of protein was predicted with the help of various methods like PREDATOR, MLRC and SOPMA at Network Protein Sequence Analysis, (NPS@webserver) (Table 7).

Similarly amino acid content of the protein was calculated using Network Protein Sequence Analysis, (NPS@webserver). Total residue mass for protein of Indian isolate was 22.7kDa with 0.73 cm³.g⁻¹ mean specific volume and it was found that Alanine was in maximum percentage in the sequence (Table 8).

Discussion

In recent years, diseases caused by plant viruses have become a significant limiting factor in the sustainable production of vegetables,

ornamentals and fruit crops. Cucurbitaceous crops including summer squash (*Cucurbita pepo* L.) have been reported to be infected by more than 30 viruses under natural conditions in different parts of the world [6]. Crop failure due to debilitating viruses creates significant financial loss throughout the world, mainly in developing countries. Therefore, successful crop management strategies require improved scientific understanding about the viruses. For this purpose, sequencing of virus genome is prerequisite. Sequencing of the virus genome and its *in silico* analysis helps in generation of information about evolution of the virus and subsequently designing reliable management strategies against the virus.

In the present studies, partial CP gene sequence of Indian isolate of ZYMV compared with other 67 isolates of ZYMV at both genomic and proteomic level to see its evolutionary behavior.

Viral cultures under present investigations were selected on visual symptoms and the results of serological indexing of the samples collected from the hill state of Himachal Pradesh located in the North Western Himalayan regions indicated presence of zucchini yellow mosaic virus and these results are also in line with the other scientists having worked on the detection of the present virus [7-10].

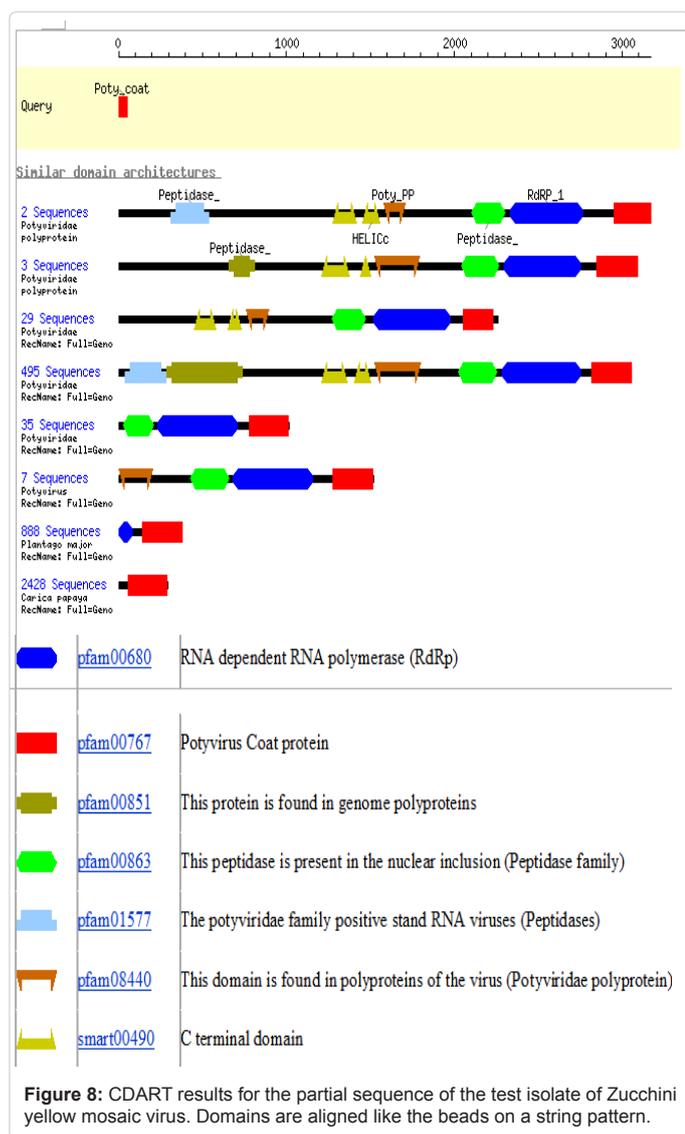


Figure 8: CDART results for the partial sequence of the test isolate of Zucchini yellow mosaic virus. Domains are aligned like the beads on a string pattern.

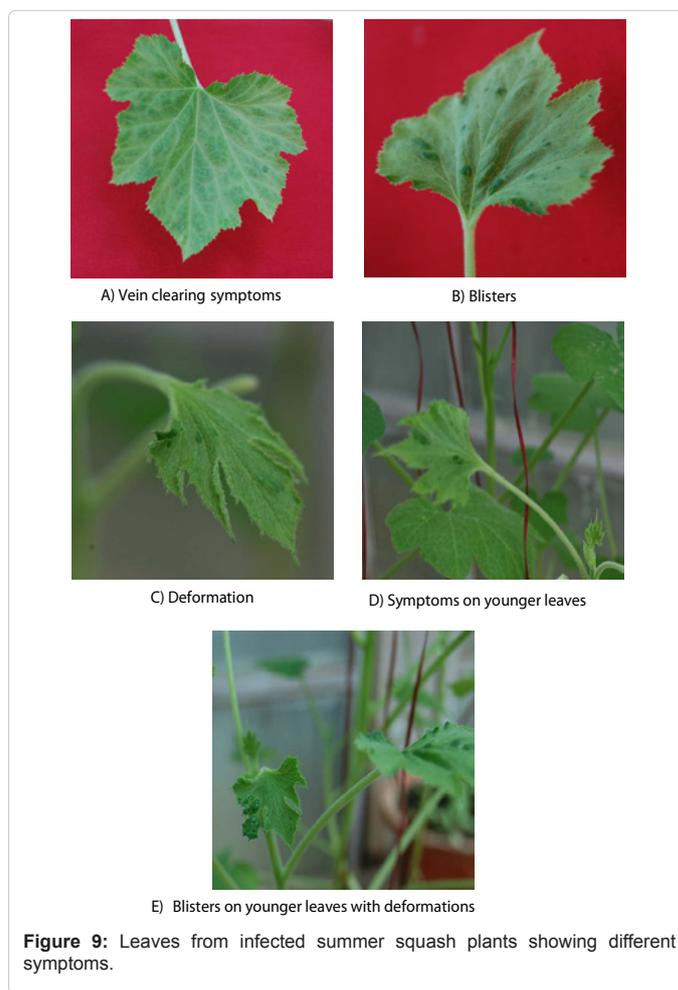


Figure 9: Leaves from infected summer squash plants showing different symptoms.

Country	Total number of sequences collected	Similarity score (nucleotides)	Similarity Score (proteins)
Australia	05	73-81%	75%
Austria	09	82-86%	77%
China	20	72-87%	75-77%
Hungary	04	82%	77%
Japan	06	82-87%	77%
Korea	05	74-84%	67-77%
Taiwan	08	84-87%	75-77%
Varied countries	10	74-88%	75-77%
Total	67		

Table 3: Nucleotide and protein sequences alignment data generated for different countries ZYMV isolates by Clustal W tool.

Under present investigations a Reverse Transcription- Polymerase Chain Reaction assay using Potyvirus group specific primers for the test virus isolate amplified a product of ~700 bp in nucleic acid extracts (RNA) from the plants collected from District Una Himachal Pradesh. The Polymerase chain Reaction (PCR) and Reverse transcription- PCR are powerful tools for highly sensitive detection of plant viruses with DNA and RNA genomes. There have been many reports to detect plant viruses using RT-PCR. Lately, in 2007, detection of ZYMV using RT-

PCR was carried out in *C. sativus* L. and *Cucumis melo* L. in Poland [10]. Auger et al. [8] identified a strain of ZYMV on squash by means of DAS ELISA and PCR using ZYMV specific primers ZY-2 and ZY-3 and a segment of 1186 bp was amplified and sequenced.

As given in results, the partial CP sequence of ZYMV (Indian isolate) was determined, which was found to be 154 nucleotides long. Prieto had also sequenced a fragment of 395 bp in length from the 3' portion of CP gene of Chilean isolate of ZYMV. In the present case however only 154 nucleotide long DNA was amplified confirming only partial amplification and sequencing of the CP gene.

Conclusion

On sequence comparison, current studies revealed close relation of the test isolate of zucchini yellow mosaic virus with USA sequence of ZYMV coat protein. As the parent crop (zucchini) is of Central American origin [11] and the relationship of the present isolate with that of USA is not surprising. It is however indicating that the virus may

Enzyme	Specificity	Cut positions (blunt - 5' ext. - 3' ext.)
Acil	C▼CG▲C	*136/138
BanII	G▲RGCY▼C	82/78
BsmAI	GTCTCN▼NNNN▲	29/33
Bsp1286I	G▲DGCH▼C	82/78
BsrDI	GCAATG▲NN▼	#127/125
Ddel	C▼TNA▲G	116/119
EcoP15I	CAGCAG(N) ₂₅ ▼NN▲	138/140
HgaI	GACGC(N) ₅ ▼(N) ₅ ▲	*44/49
HhaI	G▲CG▼C	*97/95
HinP1I	G▼CG▲C	*95/97
Hpy188III	TC▼NN▲GA	71/73
MboII	GAAGA(N) ₇ ▲N▼	120/119
MwoI	GCNN▲NNN▼NNGC	86/83
Sfcl	C▼TRYA▲G	90/94
TaqI	T▼CG▲A	42/44
Tsp509I	▼AATT▲	58/62

Table 4: List of various one cutter restriction enzymes, their position and recognition site extracted using online tool NEBcutter, in partial nucleotide sequence of the test isolate of ZYMV.

Enzyme	Specificity	Cut positions (blunt - 5' ext. - 3' ext.)
AluI	AG▼▲CT	89, 113
ApeKI	G▼CWG▲C	107/110, 110/113
BbvI	GCAGC(N) ₈ ▼NNNN▲	119/123, 122/126
Fnu4HI	GC▼N▲GC	108/109, 111/112
TseI	G▼CWG▲C	107/110, 110/113

Table 5: List of various two cutter restriction enzymes, their position and recognition site extracted using online tool NEBcutter, in partial nucleotide sequence of the test isolate of ZYMV.

Enzyme	Specificity	Cut positions (blunt - 5' ext. - 3' ext.)
CviKI-1	RG▼▲CY	80, 89, 113

Table 6: List of various three cutter restriction enzymes, their position and recognition site extracted using online tool NEBcutter, in partial nucleotide sequence of Indian isolate of ZYMV.

Secondary Structures	MLRC		PREDATOR		SOPMA	
	No.	Per cent	No.	Per cent	No.	Per cent
Alpha helix (Hh)	17	33.3%	17	33.3%	31	60.8%
Extended strand (Ee)	8	15.7%	-	-	9	17.65%
Random Coil (Cc)	26	50.9%	34	66.67%	11	21.57
Ambiguous state (?)	-	-	-	-	-	-

Table 7: Different secondary structures predicted by PREDATOR, SOPMA and MLRC.

Symbol	Residue	Number	Percentage	Residue mass	Specific volume
A	Ala	7	13.73	90.1018	0.74
D	Asp	4	7.84	134.1116	0.60
E	Glu	2	3.92	148.1387	0.66
F	Phe	3	5.88	166.2003	0.77
G	Gly	2	3.92	76.0748	0.64
I	Ile	1	1.96	132.1829	0.90
K	Lys	1	1.96	147.1976	0.82
L	Leu	5	9.80	132.1829	0.90
M	Met	1	1.96	150.2225	0.75
N	Asn	3	5.88	133.1270	0.62
P	Pro	1	1.96	116.1399	0.76
Q	Gln	1	1.96	147.1540	0.67
R	Arg	6	11.76	175.2111	0.70
S	Ser	6	11.76	106.1011	0.63
T	Thr	3	5.88	120.1282	0.70
V	Val	4	7.84	118.1559	0.86
Y	Tyr	1	1.96	182.1997	0.71

Table 8: Estimation of the amino acid residues in the partial polyprotein sequence of the test isolate of ZYMV.

have been imported into India along with the crop long ago and evolved slowly into the present isolate.

Impact of the Study

This study will provide better knowledge of the virus and its phylogenetic relationship aimed at management of the disease caused by Zucchini yellow mosaic virus.

This study will give a flip to management of the viral disease.

References

1. Brunt AA, Crabtree K, Dallwitz MJ, Gibbs AJ, Watson L (1996) Viruses of plants: Description and lists from the VIDE Database. University Press Cambridge 1484.
2. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467.
3. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
4. Marchler-Bauer A, Bryant S H (2004) CD-search: protein domain annotations on the fly. *Nucleic Acids Res* 32: 327-331.
5. Geer LY, Domrachev M, Lipman DJ, Bryant SH (2002) CDART: protein homology by domain architecture. *Genome Research* 12: 1619-1623.
6. Massumi H, Samli A, Pour AH, Shaabamian M, Rahimian H (2007) Occurrence, distribution and relative incidence of seven viruses infecting greenhouse-grown cucurbits in Iran. *Plant Dis* 91: 159-163.
7. Chalam VC, Singh B, Rathi YPS, khetarapal RK (2003) Occurrence of zucchini yellow mosaic Dis. on summer squash (*Cucurbita pepo* L.) a new record from India. *J Mycol Plant Pathol* 33: 284-285.
8. Auger J, Martinez C, Esterio M, Prieto H (2004) Identification and transmission of strains of 24 zucchini yellow mosaic virus (ZYMV) causal agents of fruit deformation in squash (*Cucurbita maxima*) in Chile. *Fitopatologia* 39: 133-143.
9. Malik AH, Mansoor S, Iram S, Briddon RW, Zafar Y (2006) A severe outbreak of melon yellow mosaic Dis. caused by Zucchini yellow mosaic virus in Pakistan. *Plant Pathol* 55: 285.

Citation: Neha S, Vrat BS, Kumud J, Thakur PD, Rajinder K, et al. (2011) Comparative *In silico* Analysis of Partial Coat Protein Gene Sequence of Zucchini Yellow Mosaic Virus Infecting Summer Squash (*Cucurbita pepo* L.) Isolated From India. J Proteomics Bioinform 4: 068-073. doi:10.4172/jpb.1000168

10. Pospieczny H, Hasiow B, Borodynko N (2007) A polish isolate of zucchini yellow mosaic virus from zucchini is distinct from other European isolates. Plant Dis 91: 639.
11. Timothy J (1993) New opportunities in the Cucurbitaceae. In: New Crops. Janick J and Simon JE Wiley New York 538-546.
12. Carter W (1962) Insects in Relation to plant diseases. John Willy. New York. 1st ed. 705.
13. Hull R (2002) Matthew's Plant Virology, 4th ed. San Diego; Academic Press. USA.
14. Riechmann JL, Lain S, Garcia JA (1992) Highlights and prospects of potyvirus molecular biology. J Gen Virol 73: 1-16.
15. Rubatzky VE, Yamaguchi M (1997) World Vegetables. Chapman and Hall, New York, 843.
16. Siaw MFE, Shahababuddin M, Ballad S, Shaw JG, Roads RE (1985) Identification of a protein covalently linked to the 5' terminus of Tobacco vein mottling virus RNA. Virology 142: 134-143.
17. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carsten EB, Estates MK (2000) Virus Taxonomy: Classification and Nomenclature of Viruses: Seventh Report of the International Committee on Taxonomy of Viruses. San Diego, Academic Press.
18. Vander Vlugt RAA, Steffens P, Cuperus C, Barg E, Lasemann DE, et al. (1999) Further evidence of that shallot yellow stripe virus (SYSV) is a distinct potyvirus and reidentification of welsh onion stripe virus as a SYSV strain. Phytopathol 89: 148-155.