

The nucleotide sequence of the infectious cloned DNA
components of potato yellow mosaic virus

R. H. A. Coutts, R. S. Coffin, E. J. F. Roberts and W. D. O. Hamilton

Presented by : Sankalya Ambagaspitiye

The nucleotide sequence of the infectious cloned DNA components of potato yellow mosaic virus

- Goal : The complete genome sequencing of a Venezuelan isolate of potato yellow mosaic virus (PYMV)
- Approach: Nucleotide sequencing using dideoxynucleotide chain termination procedure

Introduction

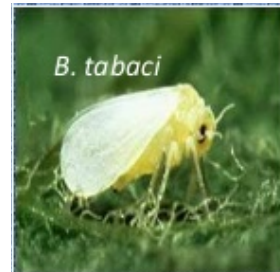
Life cycle of the PYMV (Potato Yellow Mosaic Virus)



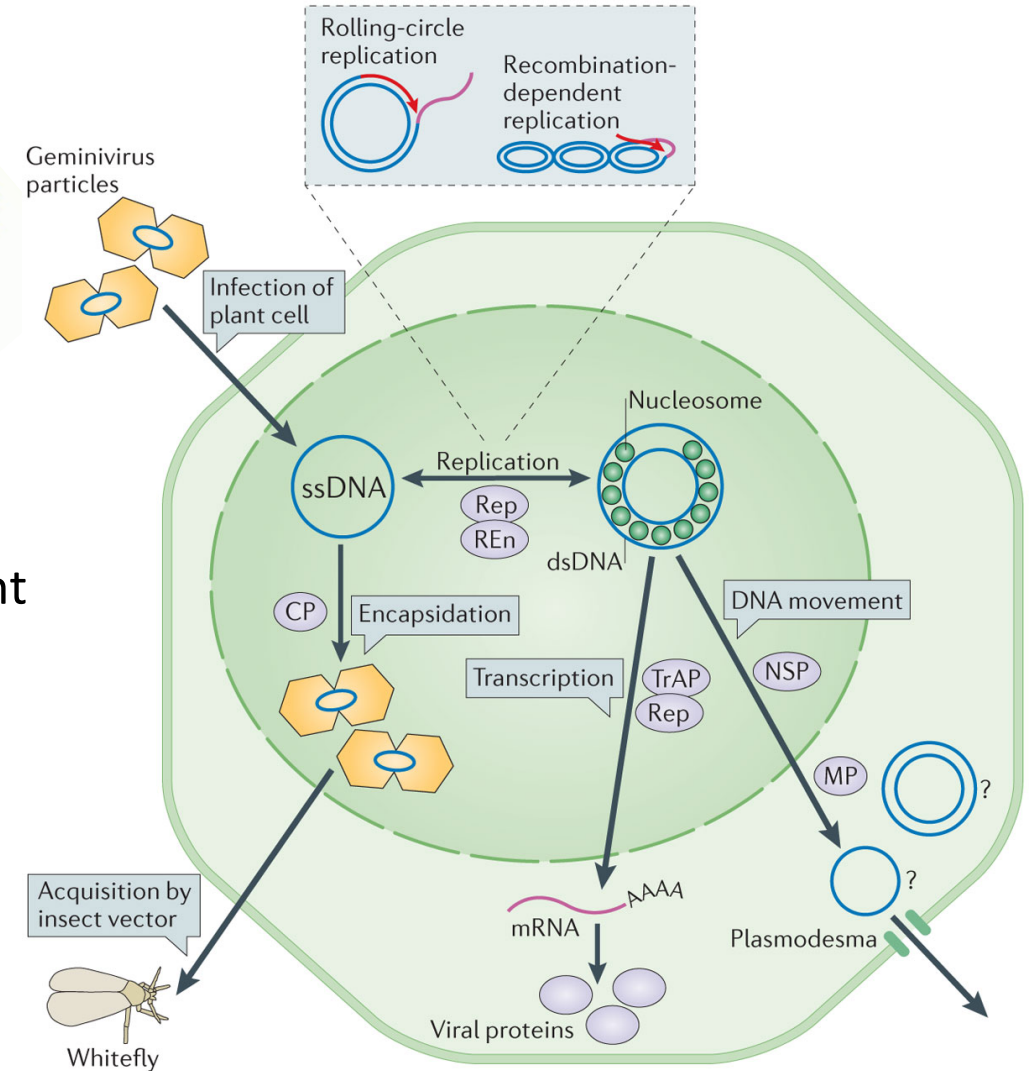
Infests new dicotyledonous plants



Infected potato plant



Bemisia tabaci



Nature Reviews | Microbiology

Potato Yellow Mosaic Virus

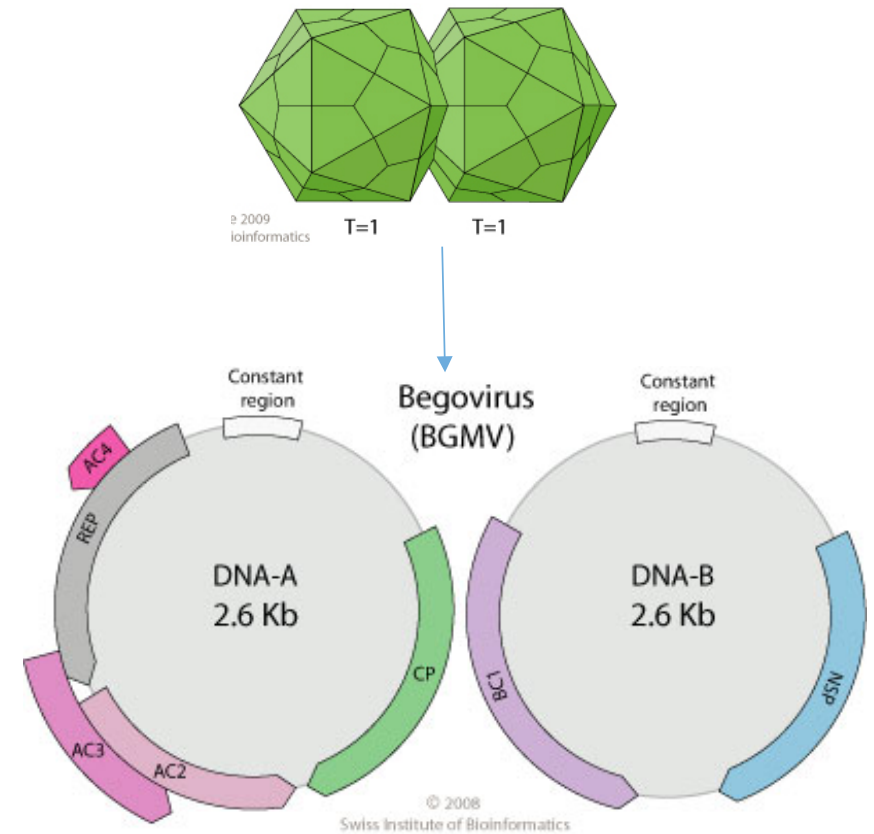
Taxonomy

Domain: Virus

Group: "ssDNA viruses"

Family: Geminiviridae

Genus: Begomovirus



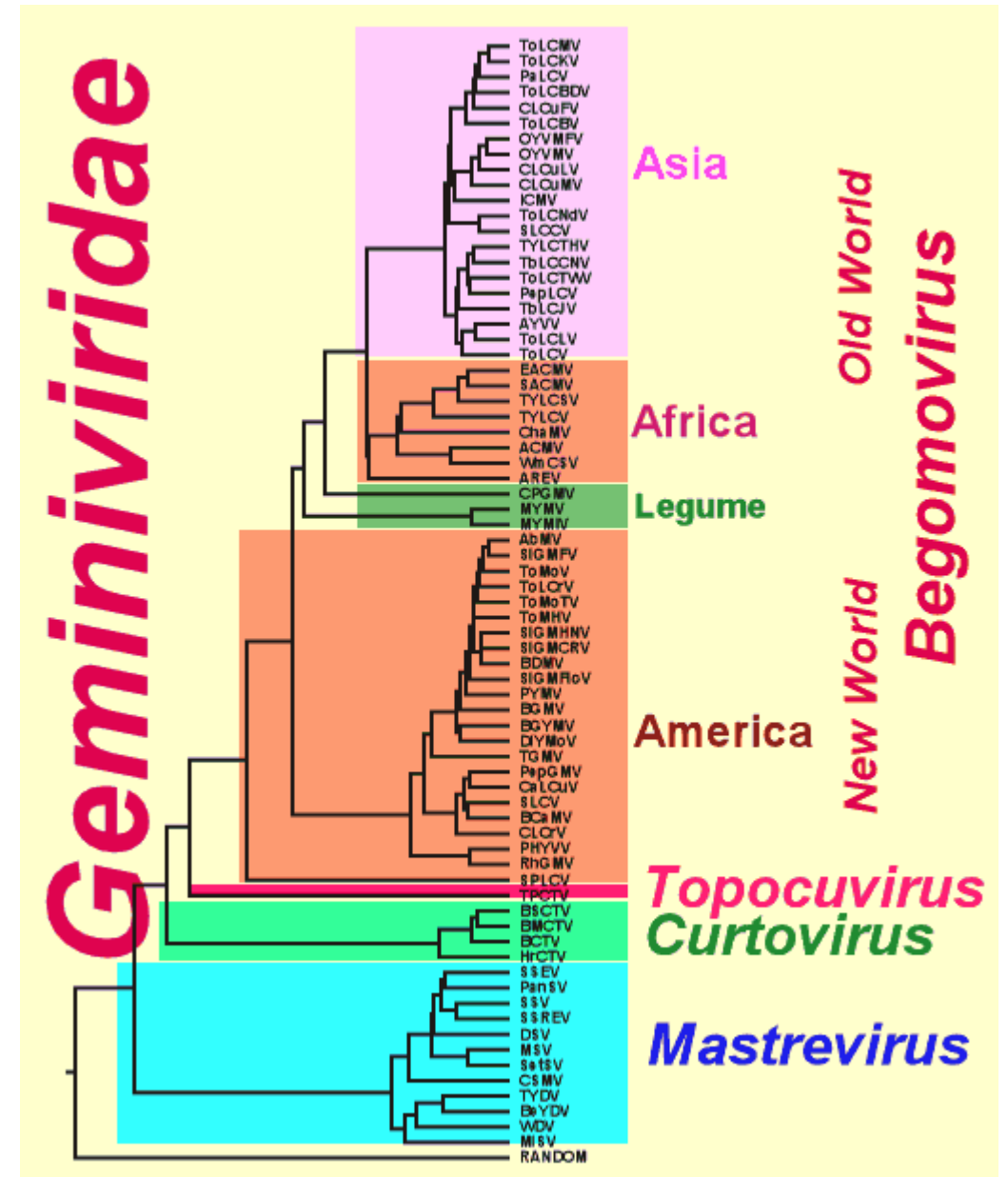
Circular bipartite ssDNA genome with DNA components A and B

Objective

- Complete sequencing of PYMV DNAs A and B

Applications

- To determine the phylogenetic relationship of PYMV with other subgroup I geminivirus



Methods

Viral DNA is extracted from *Nicotiana* plants



Purified DNA cloned into plasmid and bacteriophage vectors



Infectious clones of A and B DNAs recloned to make 2 vectors
(pMAH2 and pMBBI)

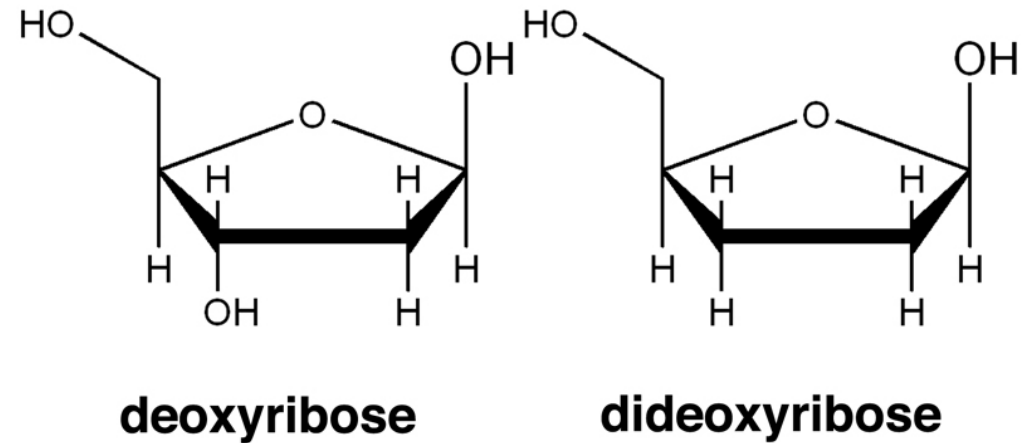


Sequence determined by Sanger's dideoxynucleotide chain termination procedure

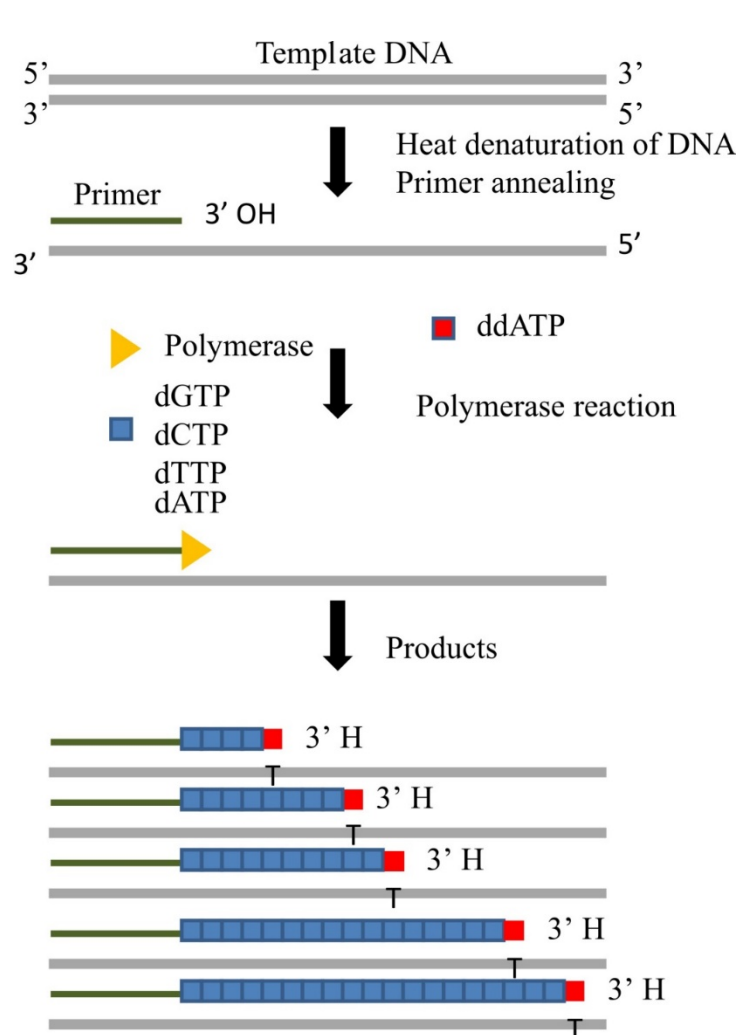
Sanger Dideoxy chain termination method

Components for the chain termination method:

- DNA fragment for sequencing
- Primers
- dNTPs
- ddNTPs

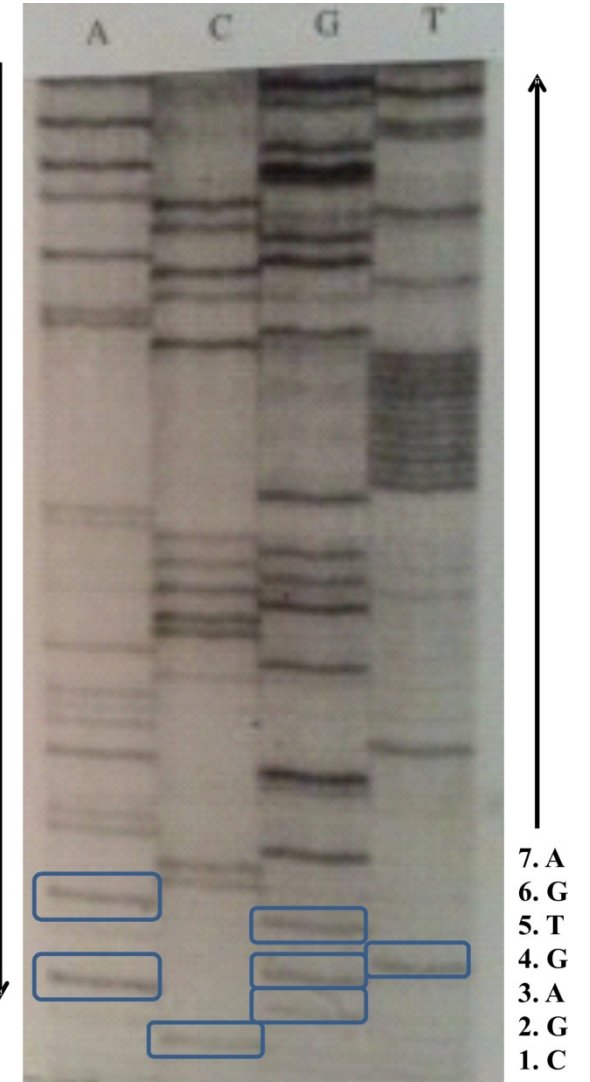


Sanger Dideoxy chain termination method



ddNTP in the reaction

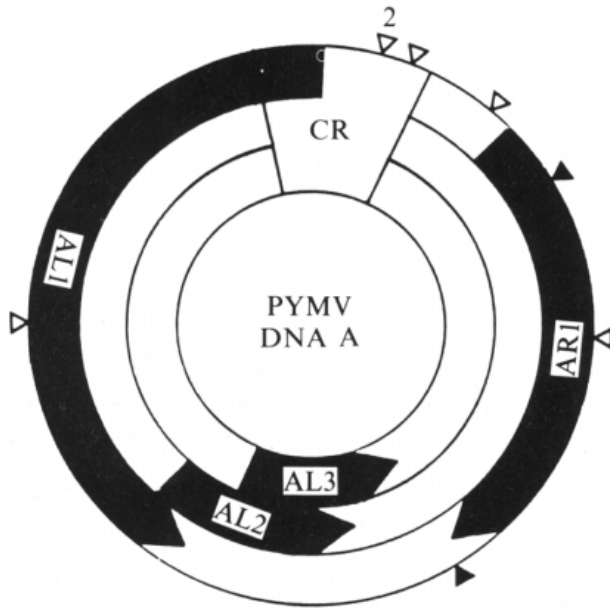
Direction of migration



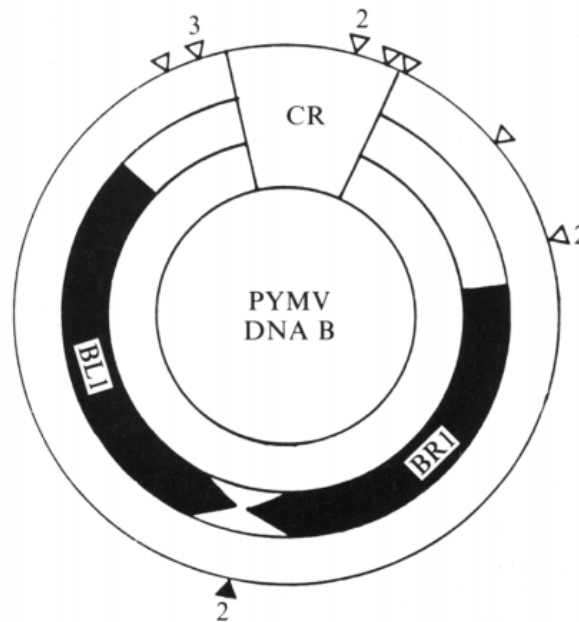
Results and Discussion

- The sequence of PYMV DNA was determined completely

DNA A
2593 bp



DNA B
2547 bp

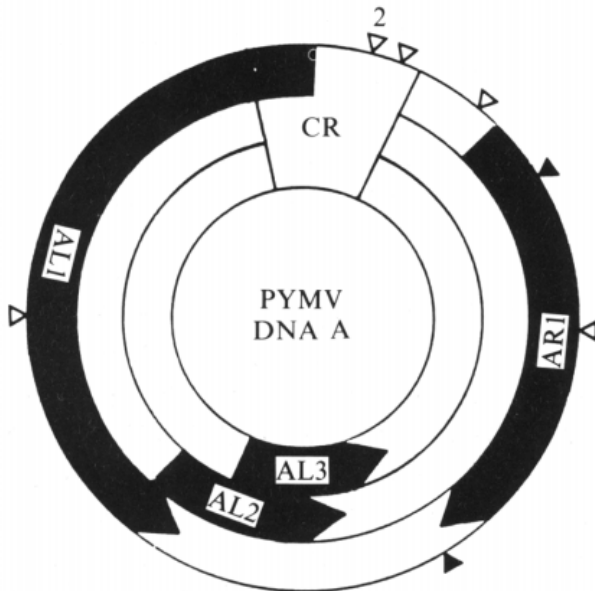


Important features

- 268 nt common region in A and B
- Similarity in amino acid sequences (41%) in ORFs AR1 and BL1
- Highly conserved inverted repeat TAATATTAC in CR, a candidate for origin of replication

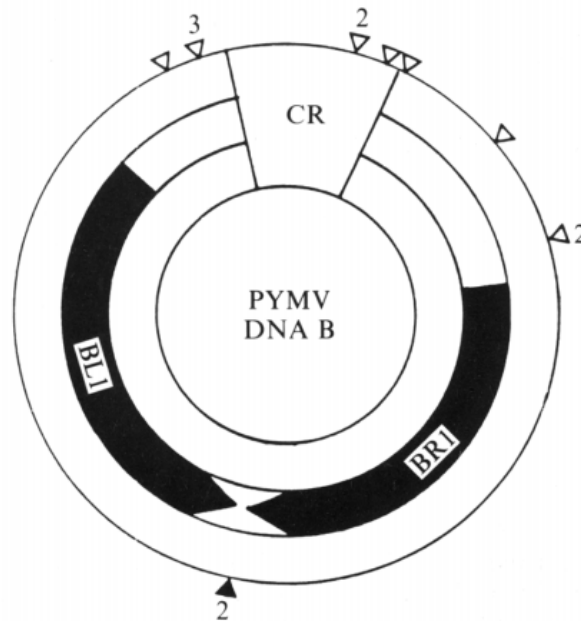
Both DNAs contained ORFs in the viral as well as the complementary strands

DNA A



ORF	M_r	nt
AR1	29266	321-1073
AL1	40849	11-1522
AL2	14735	1607-1221
AL3	15783	1472-1076

DNA B



ORF	M_r	nt
BR1	29743	575-1342
BL1	33279	2248-1370

- Six ORFs found
4 in DNA A
2 in DNA B
- Polyadenylation signals present

Comparison of PYMV DNA ORFs with other subgroup I geminiviruses

Table 1. *Amino acid sequence similarity between the most closely related subgroup I geminiviruses*

Pairwise comparison*	ORF					
	AR1	AL1	AL2	AL3	BR1	BL1
PYMV-BGMV	93†	81	83‡	87	78	91
PYMV-TGMV	94	86	81	85	75	89
PYMV-AbMV	97	83	90	88	77	77
BGMV-TGMV	95	80	81‡	86	80	91
BGMV-AbMV	93	84	80‡	89	75	77
TGMV-AbMV	94	79	84	88	72	78

- PYMV groups with subgroup I geminivirus from New world
- More closely related to Abutilon mosaic virus (AbMV)

Sequence similarities and evolution

1) ORF BL1 of AbMV 75% similar to TGMV, BGMV, PYMV

- Few changes in nucleotides raise the similarity to 90%
Explains how viruses evolved based on the environment

2) G to A transition in 360 nucleotide extends the ORF AR1 protein by 10 aa in BGMV and AbMV, similar to the length of PYMV

- Suggestion : This region of coat protein became unnecessary in the infectious cycle

Coat proteins are conserved in the subgroup

	0				49
BGMV		MAGTSKVSRS	ENYSPPGGME	SKSNKANAWV	NRPMYRKPRI
TGMV	MPKRDAPWRL	MAGTSKVSRS	ANYSPPGSLP	...KRDWV	NRPMYRKPRI
AbMV		MAGTSKVSRS	ANYSPPRIG	PRVTKASEWV	NRPMYRKPRI
PYMV	MPKRDAPWRS	MAGTSKVSRS	ANYSPPGIG	PRVTKASEWV	NRPMYRKPRI
	50				99
BGMV	YRMYSPPDVP	KGCEGPCKVQ	SYEQRHDIH	VGKVMCISDI	TRGNGITHRV
TGMV	YRSLRPPDVP	KGCEGPCKVQ	SYEQRHDIH	VGKVMCISDV	TRGNGITHRV
AbMV	YRTLRTADVP	KGCEGPCKVQ	SYEQRHDIH	VGKVMCISDV	TRGNGITHRV
PYMV	YRTLRTPDVP	KGCEGPCKVQ	SYEQRHDIH	VGKVMCISDV	TRGNGITHRV
	100				149
BGMV	GKRFCVKSIV	ILGKIWMEN	IKLKNHTNSV	MFVLVRDRRP	YGTDMDFGQV
TGMV	GKRFCVKSIV	ILGKIWMEN	IKLKNHTNSV	MFVLVRDRRP	YGTDMDFGQV
AbMV	GKRFCVKSIV	ILGKIWMEN	IKLKNHTNSV	MFVLVRDRRP	YGTDMDFGQV
PYMV	GKRFCVKSIV	ILGKIWMEN	IKLKNHTNSV	MFVLVRDRRP	YGTDMDFGQV
	150				199
BGMV	FNMPDNEPST	ATVKNDLRDR	YQVMHRYFAK	VTGGQYASNE	QALVRRFWKV
TGMV	FNMPDNEPST	ATVKNDLRDR	YQVMHRYFAK	VTGGQYASNE	QALVRRFWKV
AbMV	FNMPDNEPST	ATVKNDLRDR	YQVMHRYFAK	VTGGQYASNE	QALVRRFWKV
PYMV	FNMPDNEPST	ATVKNDLRDR	YQVMHRYFAK	VTGGQYASNE	QALVRRFWKV
	200				249
BGMV	NNHVVYNHQE	AGKYENHTEN	ALLYMACTH	ASNPVYATLK	IRIYFYDSIT
TGMV	NNHVVYNHQE	AGKYENHTEN	ALLYMACTH	ASNPVYATLK	IRIYFYDSIT
AbMV	NNHVVYNHQE	AGKYENHTEN	ALLYMACTH	ASNPVYATLK	IRIYFYDSIT
PYMV	NNHVVYNHQE	AGKYENHTEN	ALLYMACTH	ASNPVYATLK	IRIYFYDSIT
	250				
BGMV	N				
TGMV	N				
AbMV	N				
PYMV	N				

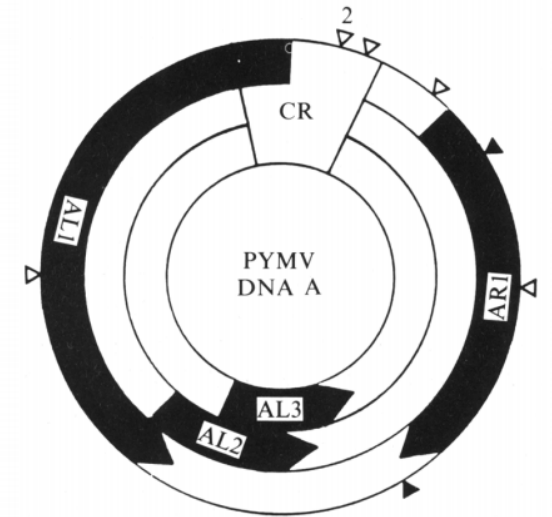
Comparison of ORFs AR1

Indicates similar geographical origins of these viruses

The common region of DNA A and DNA B

Similarities between PYMV and other class 1 geminivirus

- Possible stem-loop structure containing TAATATTAC
- Presence of inverted repeats between 3' end of the common regions and the start of ORF AR1



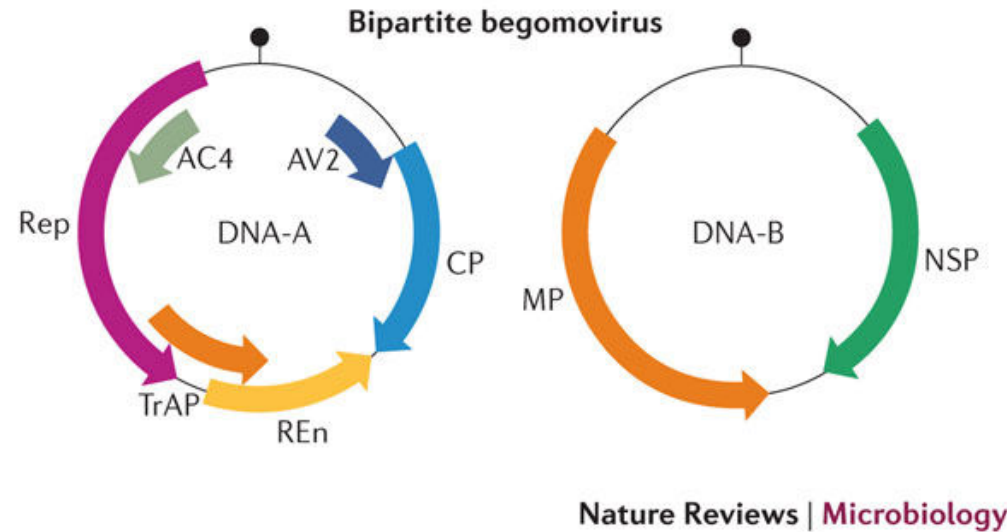
Differences

Other geminiviruses	PYMV common region
Common region range from 170-210 nt	Common region 268 nt
Extends only <15 nt into an ORF	Extends 100 nt into N-terminal in ORF AL1

Conclusions

- PYMV has a circular bipartite ssDNA genome and possesses bidirectionally oriented 6 ORFs
- PYMV is closely related to subgroup I geminiviruses isolated from the New World

Current knowledge on begomovirus genome



CP - Coat proteins

AV2 -anti-defence proteins

Rep - Replication initiator protein

TrAP- Transcriptional activator protein

Ren- Replication enhancer protein

NSP and MP : Movement proteins

Other references

- Linda Hanley-Bowdoin, Eduardo R. Bejarano, Dominique Robertson, Shahid Mansoor, 2013, Geminiviruses: masters at redirecting and reprogramming plant processes, *Nature Rev. Microbiol.*, 11:777-788
- <http://www.plantpath.wisc.edu/GeminivirusResistantTomatoes>
- http://education.expasy.org/images/Begomovirus_genome.jpg

Thank You

Nucleotide sequence of DNA [A](#)

(a)

```
1   TTTGGTGGCATTTTGTAAATATGAATGTCTCCCAATATGTTCCCOCTATGCTCTGGCTCTCAAAACTCTTA7GAA7TGGGGGAAC7GGGGGAAC7TA7ATAGTAGAAGTTCCTAAAG
121 GCAGATCAACACGTGGGCCATCCGTATAAATATTACCGGATGGCCGGCTTTTTTTTTTAA7TGGGCTTA7TTGTA7TGGGCTTTA7TTGGCCCAATCAACGT7TAGTCTGACACCTA
241 GATAAGTGGAA7TACTTTGTGCTAAGTTGTGATTTTGTCTATAAA7TAAAGCCTCTTGCCCACTATCTTTAACTCAAAA7GCTAAGCCGATGCCCCATGGCGTTCAATGGCGGAA
361 CCTCAAAGGTTAGCCGCAACGCTAA7TACTCTCTCGTTCAAGAA7TGGGCCAAGAA7TAA7CAAGGCGCTGAATGGGTTAA7TAGGCCATGTACCGGAAGCCCAAGGATCTATCGGACGC
481 TGAAGACGCCCTGATGTCTCTAGAGGCTGTGAAGGCCCTGTAAAGGTCCAGTCTTTGAGCAGGACACGA7ATCTTACACACTGGCAAGGTAATGTGCATATCTGACGTTACTGGCGTA
601 ACGGTATTACTCACCGTGTGGGAAGCGGTTCTGTGTTAAGTCCGTTTACATATTAGGTAAAGATATGATGGAAGAGAA7ATCAAGCTGAAGAACACACCAACAGCGTCATGTTCTGGC
721 TGGTCAGAGACCGAAGACCGTATGGAAACGCTATGGHTTTGGACAAAGTGTTCACATGTTGACAA7TAGOCTAGCACCGCCACTGTTAAGAACGATCTTGTGATCGTTATCAGGTCA
841 TGCATAGGTTCTATGGCAAGGTGACAGGTGGACAGTATGCCAGCAAGGACATGCTATTGTGAGGGGTTTCTGGAAGGTCAACAAACGTTGTTTACAACCAAGGAAAGCTGGCAAA7
961 ACGAGAATCATACTGAGAACGCCCTATTATGTACATGGCATGTACTCATGCTCAAATCCGTGTATATGCAACACTTAAGATTGGGATCTATTTTTATGATTCGATCTTAAATTAATAAA
1081 ATTATATTTTATTGAATGATTTTCCAGTACATAA7TACATATGATCTATCTGTTGCAAAAACGAACAGATCTAAT7TACATTGTTAATTGAAATGACACCTAACCGGTCTAAGTACAGAA
1201 GAACTAATTGCTAAATCTAGCTAAATAATTGACCCAGAAAGCTGTCA7CGATGCGTCCAGACTTGGAAAGTTCAGGTATGCTTTGTGGAGATCCAA7GCTCTCCTGAGGTTGTGGTTGA
1321 ACCGGA7TTGGACGCTGTATA7CCTGGTCTGTGTAT7TGCATGTCTCTACTTGGTTTATCTTGAAT7TAGAGGGGATTTTCTATCTCCCAGATAAACACGCCATTCCTCTGCTGACGTG
1441 CAGTGATGAGCTCCOCTGTGGTGAATCCATGTCTGGCAGTCTATGTGGAAGTATATGAGCAACCGCAGTCTAAGTCAATGGGCTCTCTCTGATTCGCCCTCTCTTTGCTTGCCTG
1561 TGTGCTTCTTGATAGAGGGGGCTGTGATGGTGA7GAAGACCGCATTCCTTAGAGTCCAGTTT7TGAGAGAAGCA7TTTGTCTTTGTGAGGAAAGCTTTA7AGCTGGAACCTCACCC
1681 AGGATTCGAGGACAGATGATGGATACCGCCTTAA7TTGAACAGGCTTCCG7ACTTACAA7TTGAT7GOCATCCOCTTTGGGCCCCAAGCAGTCTTTCAGTGTCTTAGCTTTAG
1801 ATATTGGGTTGGACGTCATCAATGACGTTA7ACTCCA7TTCATTGGAATAGAACCTGGGA7TGAAT7CAAGATGACCGCTCAAGTAAT7ATGTGGGCTTAAGACTGGGGCCACATGCT
1921 CTTGCTCTTGGAGAA7CACCTTCAATTATAA7ACTAA7AGGTCTTTCTGGGGCGCGGCACTACTCTTTCCGAAA7AACCATCTGCCACTCTTGCA7CTGCTGGGAAGCTTAGTGAA
2041 AGAGGAGAGTGAACCGGAGGCCCATGTCTCTGGAGCCTTCATGAAATCCATCGAGGTTACAGGATAGGTTATGATCTGAAAAAGAACTTTCCGGCAACTCTCTTTAATGAT
2161 TTTCA7TGGCTGCTTCC7TTGTTCCAGAGTTTAGTGCTCGGCMGCTGGCTGTTAACTGTCTGCTGGCCCCCTCGAGCACTTCTTCGGTCAATCTGGAACAAACCCCATTCGATGGTGTG
2281 TCGCTCTTCTCGACATACGACTTGACATCGGAGCTCGATTAGCTCCCTGAATGT7TGGATGGAATGTGTTGA7CTGGTTGGGGACACCAAGATOGAACAGTCTGTTATTTGTGAGTT
2401 GTATTTGCTCTGAACTGGA7AAGAACGTGGAGATGAGGTTCCOCCATTCCTGTGAAGTCTCTGCAAA7CTTGATGAATTTCTGTTGACTGGGATAGTTAGGTTTGAAGTTGGGAAAG
2521 TGCTTCTTCTTTGCTAATAGAGCATTGAGGATACGTAAGGAAATAATTTTGGCTTTAATAGAGAA7GAAACCC
```

...

Nucleotide sequence of DNA B

(b)

```
1  TTTCGGGGCAITTTGATAAATAATTAGTGTCCCGGATATGTTCCCCCAATTGCTCTGGCCCTCAAAACTCTTATGAATTGGGGGAACCTGGGGGAACCTATATAGTAGAAGTTCCTAAAG
121 GCAGATCAACACGTGGGCCATCCGTTAATAATTACCGGATGGCCGGAGATTTTTTATACTGGGCTTCTACTAAGGCCCAAGTATGATTGGACTCTAGGCCCATTTTTIAGGGTCC
241 CTCITTTTCTCGTCTTTCTCTGCTGTTTCTTTTACTTTTAAATTTAAATCACTCTTAAATTTAAATATGTTGTGATCAATTTCAACAGCTGTCCCGCATATATATGTAAGTGAAGTGG
361 GGTAAATTCGACCAAGTTACAGTGTAAAGTCGAAATGATTTTAAAGTCTCTCCATTTATGAGCTCATGGTCTATTAAGCATTTACATTACAGTGTACCATTTAAATTTGATGTTGG
481 AGTCAATTTAAATTAATGTATACCTGATATGCTATATATTCGCTGACTAGGAATGAGCTCATTATTACGTGGTACAAAGACATTACATGATAATGTATCTAAATAGGCACAGGGGTGC
601 TTCTTTTGTAGCCAGCCACGTACTTACCCAGTAAATAGTTTGATTAGACAGCTCATTATTCAAGCGTAATGTAGCAACGACGACCATTTCAAACCGTGAAGATGGTTGATGACTC
721 CATGATGAAGGCACACGTATTATGAGAATCAATACGGTCCAGATTTTTCACGTGGCCCAATATACAGCGCTCTACATTTTAAAGTTACCCGTGATTTGCTAAGTCTCTGCCCAATAG
841 AACCAGGTCAATATTAAGCTAAACGACTTCGGTTCAAGGTATTTGTGAAGGTGGAACGTGTACATGTAGAGGTAAACATGGACTGTTCTGTGCTTAAGACCGAAGGAGTTTCTCTTT
961 GGTATTTGTAGTGGATCGTAAACCTCACTGGACCTCTGGGGGACTGCTACATTTGACGAGCTATTTGGCGCTAGGATCCACAGTCATGGTAACTTGGCAATAGTTCCATCTCTGAA
1081 GGATCGGTCTACATACGCCATGTAAGTGAAGGTGTGATATCAGTCGAGAAGGACACCATGATGGTGGACATAGAAGGTGTTGTAGCCCTTTCTAGCAGACGTTTAAATTTGTTGGGCTGG
1201 TTTTAAGGACCTTGACATAGAGTCCCGAAAGGGTGTATATGATAACATTAATAAGAAAGCCCTGTAGTTTATTTATTTGTTGGATGTGGATACAGTATCCAAAGCATCCACATTTGTATC
1321 GTTGATCTTGATTATATTTGATGATTGTATTCACTAATACAAATTTTATTTAAAAATTTTGGCTGAGAATTTACAAATTAATCTTTATACATTTCTGGACAGTTGCTTACAAAGTT
1441 CATTTAACTGGGCTACGGACATGGTTATGTTCGATTGGGCCATGTTAGCACCACTATAGATGCTGACTCTCCTGGATCTAGGACGCTGGCTCCTAGTCTTTGCAAACTCGATACGTAT
1561 GTAGCTGGTTCTCTATCTCTGACTGGCCCTCTGATTTGACCCACTCTCTATTTGTACTCTCGGAAGCCATGATTCACAGGCCCTTATTTCAATTTGGGCTCTGAGCCCACTCTGGACATTC
1681 ATGCGCATCTAATGGGCTTCTCTCCATCTCCCGTAGTCGAGTGGCGAGAAATCCACGTCTTATCTGTGAACGTGTTGGACAGGATCTTGACCGTTGGTCTCTGAACGGAATGTCTA
1801 CGGAATGTTTGGCGTCGACAGTTTCAGCTTCCCTTTGAACTTGGCAAGGTGTGTCGGTTGATGAACGTTCGGTGTGGAAACTCTGTAATAGAGTTTCCATGGGATAGGATCTTTTAGGG
1921 AGAAAAAGGAAGCTGAGAAATAGTGGAGATCTATGTTGCATCTTAACGGGAAGTCCATGACGCTCGCAGAGATTGTTGTCTGTATCTCTTATCATGAAATCCCACTATTACAGACC
2041 CGGTGGCGTTGATGGGTACCTGTGCTGTACTCAATTAACGAGTGGTCTATCTCATACAGCTGGACGTGGCCGTGGGTTAACTGGCGCGCGTCGACGGAATTTGACGATTTATCT
2161 CAGTTAGGTCATGGAAAGCTGATATTATCCGTTGAGACTCTATGTAATTGAAGCATGAGGAGATTAACTAACTGAGAATTCATTTATGAAATTAGGCGCGCAGCGGCACCGC
2281 TTCCAGAAATATGAACAGGAAAGAGGATGATAGGGTTTCTCGAACAGGCAAGCATGATCTTGAAGAGAGACAGAGAAATATAATAGTCTATGGAAGACCTGTAATATGTTAATGTT
2401 CAAGGATTAGGTAGTTTATATAAAGAAAGTTGTTACTGCTGAGGTACCAATTAAGAAATTTGAAGTTGGGAAAGTGATCTTCTTTGCTAAGATAGTATTGAGGATTCGTAAGGAAATAAT
2521 TTTTGGCTTAAATAGAGAAAGACCTT
```

