



RESEARCH ARTICLE

Evolution and selection patterns in begomoviruses infecting *Phaseolus vulgaris* L.

Chahat Slathia & Yogesh Kumar*

Department of Botany, Central University of Jammu, Jammu 181 143, Jammu and Kashmir, India

*Correspondence email - yogesh.bot@cuammu.ac.in; yogesh.ihbt@gmail.com

Received: 15 January 2026; Accepted: 12 June 2026; Available online: Version 1.0: 04 August 2026

Cite this article: Chahat S, Yogesh K. Evolution and selection patterns in begomoviruses infecting *Phaseolus vulgaris* L. Plant Science Today. 2026; 13(sp5): 1-10. <https://doi.org/10.14719/pst.13691>

Abstract

Begomoviruses affect *Phaseolus vulgaris* L. that considerably decrease yield in many bean growing areas. Their evolutionary dynamics, genetic diversity and selection patterns were investigated in this study. Complete genome sequences of begomovirus isolates from different parts of the world were retrieved from GenBank. The MEGA11 was employed to align the sequences and to select the best nucleotide substitution models. Phylogenetic clustering appeared broadly consistent with geographical origin in maximum-likelihood phylogenetic trees. Transition/transversion analysis indicated relatively higher substitution rates in AV1 and AC4 as compared with other genomic regions, most likely due to vector influence and host–virus interactions. Analysis was carried out using RDP4.1 revealed that recombination is one of the major factors for diversification of these viruses. Neutrality tests using DnaSP v6.12 indicated that purifying selection is a predominant force driving the evolution of these viruses. High nucleotide diversity was observed in both DNA-B and betasatellites. The role of geographical distribution, mutation, recombination and selection has been analysed and discussed in the present study. In addition, the observed genetic variability among isolates highlights the adaptive potential of begomoviruses under diverse environmental and host conditions. The findings provide valuable insights into the evolutionary mechanisms shaping *Begomovirus* populations and contribute to a better understanding of their emergence and spread. Such information may be useful for developing effective disease management strategies and supporting breeding programs aimed at improving resistance in common bean cultivars.

Keywords: begomoviruses; genetic diversity; *Phaseolus vulgaris*; phylogenetic analysis; recombination; selection pressure

Introduction

The genus *Phaseolus* (Leguminosae family) comprises legumes commonly known as beans, which are cultivated as food crops. The beans and other legumes are an important source of nutrients, particularly in the developing countries, where they are useful in supplementing the human diets. They are usually referred to as the poor man's meat because of being rich in protein. This is particularly true when it comes to those areas where the availability of animal products is scarce and very expensive (1). They are nutritious and, thus, can be effectively used as a dietary supplement to help fight protein-energy malnutrition, which is prevalent among low-income populations (2, 3). The species possesses a diploid number of chromosomes ($2n = 22$) (4). Despite their nutritional significance, common bean cultivars are characterised by a narrow genetic base. Small floral structures, drought susceptibility, limited cultivation and the existence of a wide but poorly used germplasm are causes of the low genetic diversity. Because most cultivars were produced with few donor lines, the genetic bottleneck has been intensified. Here, mutation breeding can be viewed as an important approach to increase genetic diversity and accelerate the improvement of desirable characteristics in *Phaseolus vulgaris* L. (5).

Plant viruses can cause significant global effects on crop quality and yield. The genus *Begomovirus* is the largest within the

family Geminiviridae. According to the most recent report of the International Committee on Taxonomy of Viruses (ICTV), the family Geminiviridae comprises 522 species across 15 genera. The genus *Begomovirus* is the largest genus in the family, with 455 different species that infect a large variety of dicots and with different types of illnesses. These viruses are circular single-stranded DNA (ssDNA) genomes within unique geminate particles. The whitefly *Bemisia tabaci* is the main vector that feeds on a range of host plants, enhancing its capacity to acquire and transmit begomoviruses to new plants. The begomovirus infection is a severe threat to common bean, with evidence pointing toward the occurrence of multiple species infecting the same plant (6). High mutation rates, rapid replication and large population sizes contribute to extensive genetic variability within viral populations. RNA viruses exhibit higher mutation rates because their RNA-dependent RNA polymerase lacks proofreading capacity. Although DNA viruses can be replicated more precisely with the involvement of DNA-dependent DNA polymerases, research has shown that some single-stranded DNA viruses evolve at rates comparable to those of RNA viruses. Over the past few years, it was found that begomoviral populations undergo frequent recombination (7). They reproduce through a rolling-circle mechanism, and they have distinct patterns of genetic exchange as revealed in both non-homologous and homologous recombination. Frequent recombination contributes to the emergence of novel viral

variants and, in certain instances, they eventually evolve into highly virulent strains that cause massive outbreaks of disease. While recombination plays a vital role in shaping their genetic diversity, it relies on existing sequences and cannot introduce entirely new mutations. Consequently, the constant production of mutations is an essential aspect of the evolution of these viruses (8).

Begomoviruses evolve through multiple mechanisms such as frequent mutations, genetic recombination, reassortment in case of bipartite viruses and interactions with their insect vectors. The rate of their evolution poses a significant challenge to the development of effective diagnostics and disease management strategies. Therefore, an in-depth analysis of the evolutionary behaviour and genetic variation of begomoviruses that affect beans is necessary.

Viral populations usually have many different lineages that change over time. Each lineage has its own genetic variation, is found in different places and is specific to certain hosts. There is a lot to be learned about the historical processes that influenced the evolution of the species by combining the information about their evolutionary relationship, the date when the specimen was collected, the place of origin and the host species. The duration in which these changes have been occurring can also be determined (9).

Materials and Methods

Sequence selection and alignment

A total of 203 DNA-A, 112 DNA-B and 9 betasatellite sequences associated with begomoviruses infecting *P. vulgaris* were retrieved from the National Centre for Biotechnology Information (NCBI) GenBank database (<https://www.ncbi.nlm.nih.gov/>) for the period between January 1, 2010 and July 31, 2025. From these, 42 complete DNA-A, 30 DNA-B and 8 betasatellite sequences were selected for analysis (Supplementary Table 1 and 2). Sequences were down selected to reduce redundancy. Full genome sequences with concise annotations that contained information on the host, place of collection, year of the sample and respective geographical location were only considered reliable and relevant. All the chosen sequences were downloaded in the FASTA format. The ClustalW, an algorithm of MEGA 11, was applied to align these sequences and six open reading frames (ORFs) of DNA-A (8, 10, 11).

Phylogenetic relationship

To derive the most suitable nucleotide substitution model to be used on each dataset, a thorough model selection test was conducted with the help of MEGA software. The selection criteria were considered based on the Bayesian Information Criterion scores (BIC) and the corrected Akaike Information Criterion (AICc), which are both taken as common and widely accepted measures to balance the model fit in consideration of the complexity of the model. The sequences of DNA-A, DNA-B and betasatellites were constructed using maximum likelihood (ML) method in the construction of phylogenetic trees. The strength of every clade was tested with support from bootstrap analysis based on 1000 replicates (9, 12).

To understand the trends in the nucleotide substitution patterns, the transition/transversion bias (R), which specifies the relative frequency of two types of substitutions, was calculated. This was done consistently with all six ORFs, the entire sequences of the DNA-A and DNA-B components and the corresponding betasatellites (11, 13).

Recombination analysis

A detailed recombination analysis was done using recombination detection program version 4.1 (RDP4.1) to determine the existence of recombination in the selected sequences. In this analysis, seven independent detection methods were used, which are RDP, BOOTSCAN, MaxChi, GENECONV, SISCAN, CHIMERA and 3Seq. They were all executed on default parameter settings. The significance level was set to a maximum *p*-value of 0.05 to reduce multiple testing errors and was further corrected by the Bonferroni error. The incidences of recombination were considered statistically significant when a combination of three or more algorithms detected the incidences. This was to reduce the risk of false outcomes (14, 15).

Population genetics and structural analysis

A detailed analysis with the help of DnaSP v6.12 was performed to evaluate the genetic variability and the population structure of the viral populations (16). This was done on the entire DNA-A sequence, the 6 distinct ORFs of DNA-A, as well as on DNA-B and the respective betasatellites. Several molecular diversity measures were calculated to describe the amount and the type of variation among and within populations. These included the total number of segregating sites (*S*), which reflects the number of variable nucleotide positions and the total number of mutations (*n*) observed across sequences. Nucleotide diversity (π) was determined to provide a quantitative measure of the genetic diversity to measure the average proportion of nucleotide differences per site between any two sequences. Moreover, the mean number of nucleotide differences (*k*) was calculated to create an additional understanding of the extent of divergence within the sequences (17, 18).

Neutrality tests

Neutrality tests were conducted with DnaSP v6.12 to determine deviations from the neutral evolution and to estimate the selection pressures that might be applied to the viral populations. These tests included Tajima's *D*, Fu and Li's *D** and Fu and Li's *F** (16). Such analyses were applied systematically to the entire DNA-A, all six ORFs of DNA-A and DNA-B and betasatellite sequences to thoroughly examine selection throughout various genomic locations (16, 19).

Results

Phylogenetic analysis

The phylogenetic trees revealed several groups with clustering patterns that appear to be associated with *Begomovirus* species and their geographical origin. Most major nodes showed strong bootstrap value support, which implies that inferred groupings are reliable. Moreover, the strong support indicates low uncertainty in the topology, which confirms the suitability of the chosen substitution model and alignment. TN93+G+I, HKY+G+I and T92+G+I were the best models for DNA-A, DNA-B and betasatellite respectively.

Phylogenetic analysis of DNA-A sequences

Phylogenetic analysis of DNA-A sequences showed several well-supported clades corresponding broadly to New World and Old World begomovirus lineages (Fig. 1). Most of the isolates from the Americas grouped into one cluster, in accordance with previous observations indicating evolutionarily distant New World begomoviruses. Within these groups, additional regional clustering was observed among Caribbean, Central American and South American isolates, suggesting localised diversification patterns. The

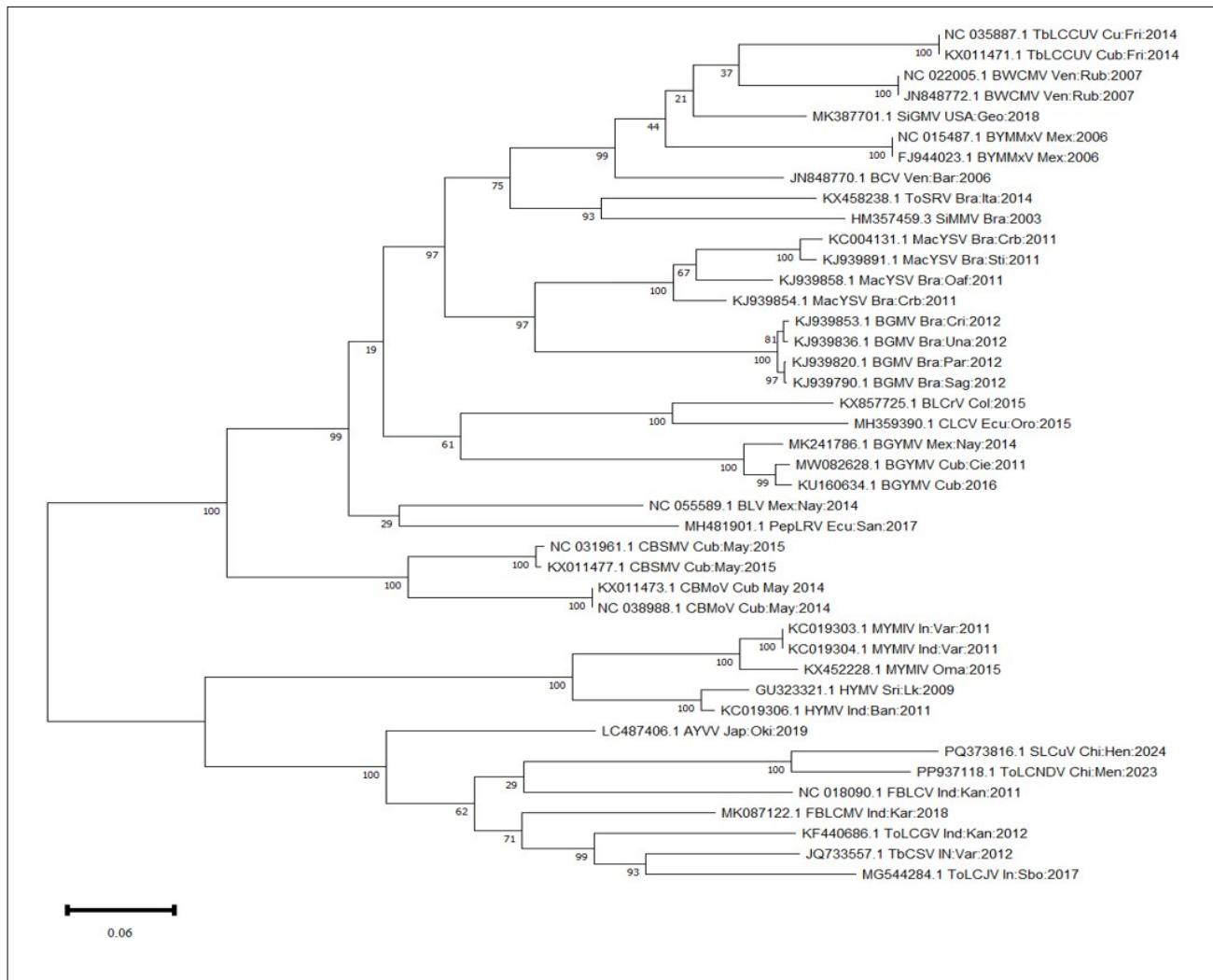


Fig. 1. Maximum likelihood phylogenetic tree of begomoviruses infecting *Phaseolus vulgaris* based on DNA-A sequences.

Asian isolates emerged as genetically distinct clades apart from the American strains, thereby confirming the known distinction between the Old World and the New World begomoviruses. Despite the presence of some regional clustering tendencies, these observations should be interpreted cautiously, as no formal population structure analysis was conducted in this study. This phylogenetic structure could, in part, be an indication of recombination history and sequence evolution between geographically isolated viral strains. Similar clustering trends have previously been reported in begomovirus evolutionary studies, where recombination, mutation, host adaptation and vector-mediated transmission collectively contribute to viral diversification. Overall, the phylogenetic relationships indicate substantial genetic diversity among begomoviruses infecting *P. vulgaris*, with lineage differentiation consistent with broader evolutionary patterns reported for global begomovirus populations.

Phylogenetic analysis of DNA-B sequences

Phylogenetic grouping of DNA-B sequences revealed numerous well-established groups representing geographically isolated begomovirus populations (Fig. 2). American isolates belonged primarily to the New World groups, while Asian isolates were grouped into their own genetically distinct clades in agreement with Old World begomovirus populations.

Within the American group, Caribbean, Central American and Northern South American isolates exhibited relatively close phylogenetic relationships, suggesting that the degree of similarity was higher in this region compared with other American regions. In

contrast Andean and Southern South American isolates formed comparatively distinct subclusters, reflecting increased genetic divergence within the New World lineage.

Asian isolates were grouped into a distinct clade that clearly differed from American isolates, which agrees with previous evidence of considerable evolutionary divergence between Old World and New World begomoviruses. While some degree of clustering could be observed for certain regional groups, such observations should be made carefully since no formal population structure was analysed in this study.

The observed DNA-B phylogenetic structure may reflect the combined effects of mutation, recombination and long-term evolutionary divergence among begomovirus populations infecting *P. vulgaris*. The occurrence of multiple recombination events identified within DNA-B sequences may have contributed to the diversification and phylogenetic placement of several viral lineages.

Phylogenetic analysis of betasatellite sequences

Betasatellite phylogenetic analysis demonstrated that there were many sequence variations and several different clusters (Fig. 3). Unlike the DNA-A and DNA-B sequences, betasatellites were not very well clustered because of their evolutionary nature and their tendency to be associated with different begomovirus groups.

Multiple Asian betasatellite isolates formed well-supported phylogenetic subclades with high sequence similarity, while others branched off to form a different cluster, showing independent evolution in local populations. This variability is consistent with the

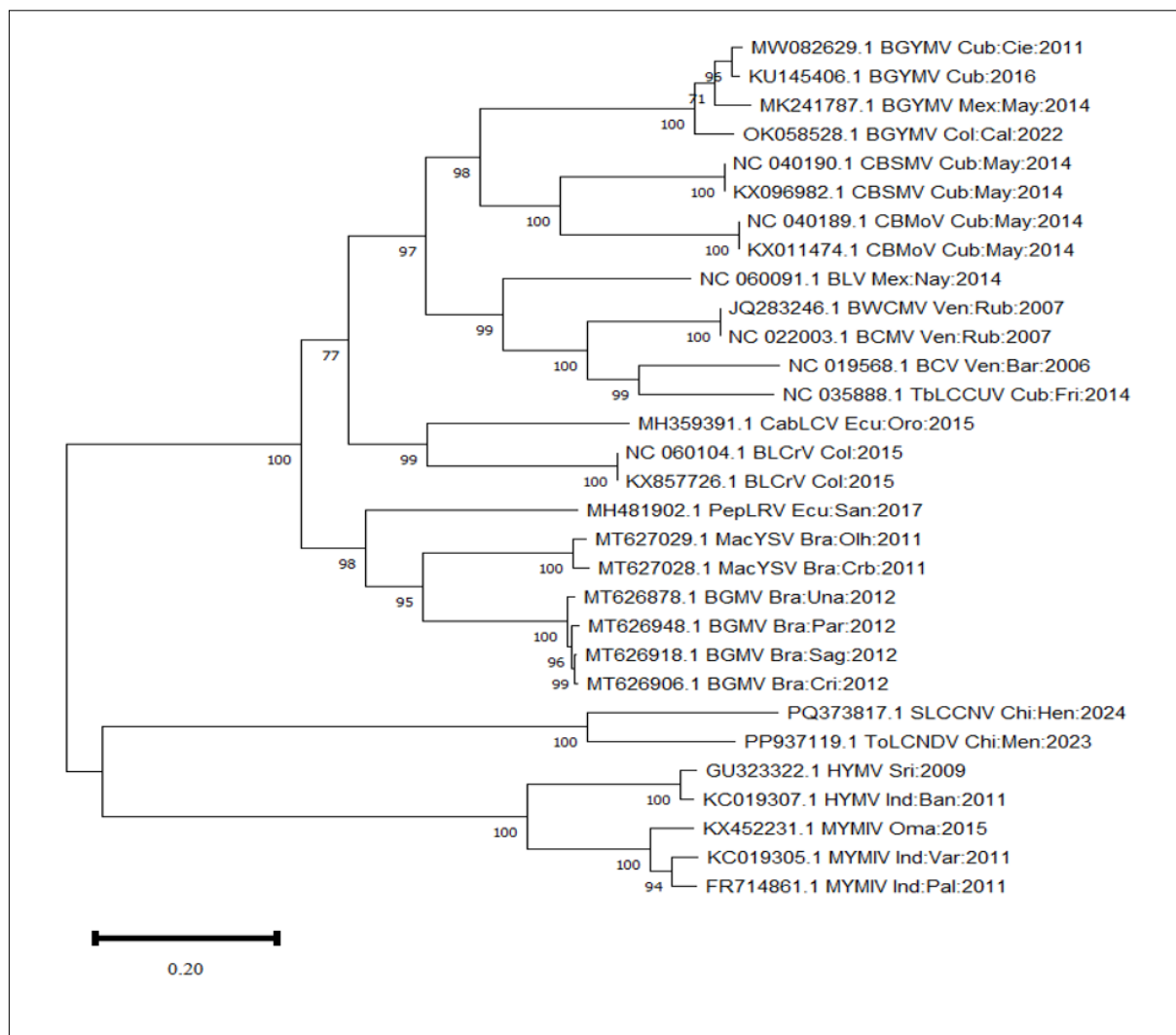


Fig. 2. Phylogenetic tree created using maximum likelihood method illustrating the phylogenetic relations among begomoviruses DNA-B sequences associated with *Phaseolus vulgaris*.

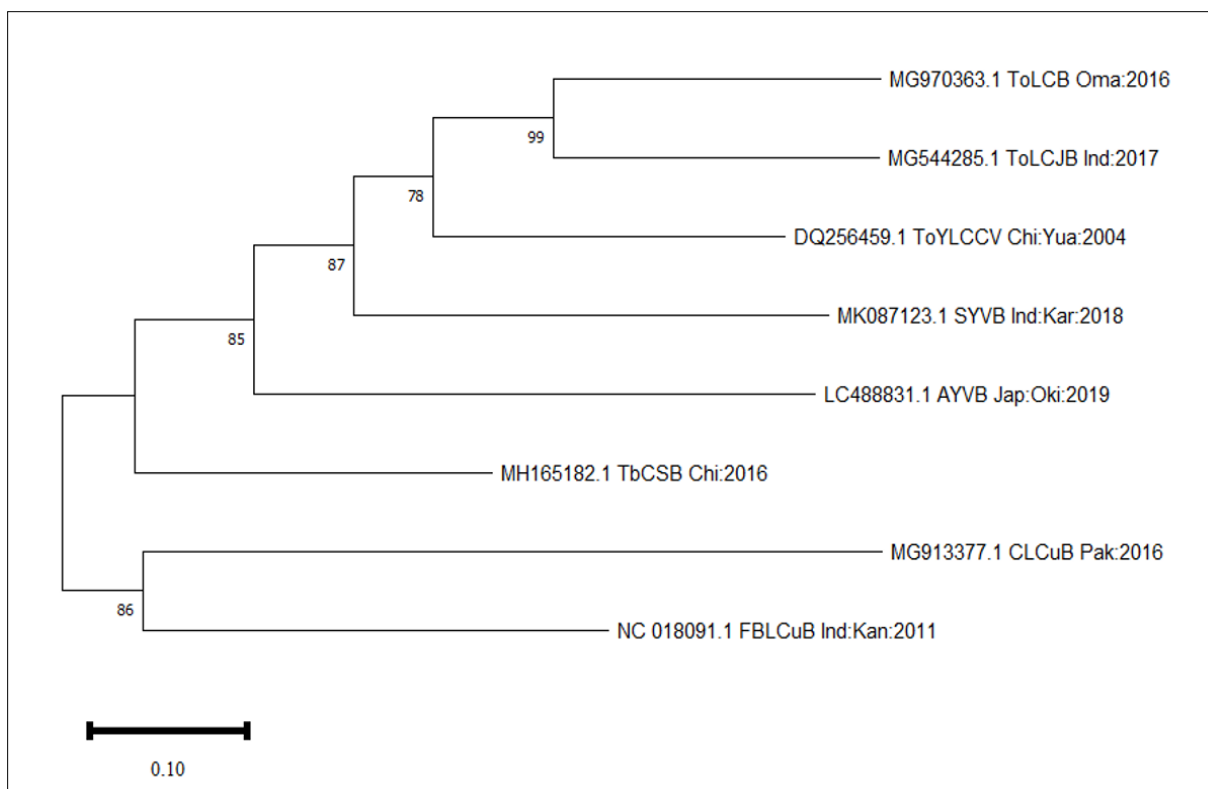


Fig. 3. Phylogenetic tree constructed using maximum likelihood method for betasatellites associated with begomoviruses infecting *Phaseolus vulgaris*.

elevated nucleotide diversity observed in betasatellites and may reflect their rapid evolutionary turnover.

Recombination signals observed in several betasatellite sequences additionally prove their nature as highly variable components of genomes, thus providing evidence for their participation in begomovirus evolution. In general, the results obtained from the study indicate that the betasatellites infecting *P. vulgaris* can undergo significant genetic variations.

Transition/transversion bias (R value)

Variation in transition and transversion frequencies was observed when nucleotide substitution rates were compared across several begomovirus genomic components. With transitional substitution rates ranging from 10.96 to 20.27 and 8.78 to 19.63 respectively, the AC4 and AV1 regions had the highest rates, suggesting that they had faster evolution than other genomic components. Both R values for AC4 (1.357) and AV1 (1.083) are larger than 1.0, which suggests a tendency towards a higher frequency of transitional changes in these sequences relative to the remaining parts of the genome, whose R values are less than 1.0 (Table 1). The increased transition bias in AC4 and AV1 may be a result of either relaxed selective constraints or positive selection affecting these genomic regions due to their function in the virus-host relationship.

Recombination analysis

Recombination analysis revealed evidence of genetic reshuffling within viral populations, suggesting the evolutionary plasticity of begomoviruses. The numerous and extensively validated recombination cues in many genomic elements suggest that similar recombination events had played a pivotal role in genome diversity and adaptation. The distribution of recombination breakpoints within specific genomic regions may indicate the existence of potential recombination hotspots, highlighting the dynamic aspect of these loci in producing new genetic variants. The relatively higher frequency of recombination in some of these coding areas may reflect the existence of functional constraints that perhaps favoured exchange of sequences to provide an advantage to fitness or host adaptability. Additionally, the identification of uncharacterised parental sequences was an indication of involvement of unsampled or cryptic viral lineages of viruses in the recombination process. Collectively, these findings suggest that recombination may act as a key driver of molecular evolution, contributing to the emergence of new viral strains with broadened host ranges and increased adaptability within the agroecosystem (Table 2).

Genetic analysis and DNA polymorphism

These evaluations of DNA polymorphism revealed a high degree of genomic variability across all the components of the virus,

highlighting the dynamic nature of their evolutionary processes underpinning the variation. The high level of nucleotide variation in specific regions suggest that there was an accumulation of mutations in them, thereby contributing to genetic differentiation and adaptation possibilities. The presence of numerous haplotype variations among various regions reflects extensive sequence heterogeneity among populations consistent with a lot of genetic exchange and diversification driven by selection pressures. In contrast, the relatively reduced diversity in specific coding regions points toward functional conservation, likely reflecting structural or replication constraints. The heterogenous distribution of polymorphism across genomic regions highlights the influence of evolutionary forces, resulting in variable rates of mutation and selection. The general patterns of polymorphism observed support the notion that the viral populations are shaped by complex interplay of mutation, recombination and selection, enhancing their adaptability and capacity to survive in varied host habitats (Table 3).

The findings of the neutrality test further suggested that demographic and selection pressures have contributed to shaping the evolutionary dynamics of the viral genome. The predominance of negative Tajima's D values across the DNA-A loci is consistent with the action of purifying selection or recent population expansion. However, the concurrent occurrence of positive Fu and Li's D* and F* values in several regions suggests the possible influence of balancing selection, maintaining genetic variation at specific loci. A similar pattern observed in DNA-B, implies coordinated evolutionary pressures acting on different genomic components. In contrast, the consistently low neutrality rates detected in betasatellites point towards strong purifying selection and this was likely to preserve important functional elements. All these findings indicate that the viral populations are shaped by a complex interplay of adaptive diversification and selective pressures, which collectively promote viral survival and ecological success (Table 4).

Discussion

Phylogenetic analysis revealed substantial genetic diversity among begomoviruses infecting *P. vulgaris*, with clustering patterns broadly consistent with previously described Old World and New World lineages (20–22). However, the phylogeny observed could be explained by the numerous recombination found in DNA-A, DNA-B and betasatellite segments. The detection of breakpoint sites at AC1 coding region indicates that genomic segment exchange has played a role in begomovirus evolution through the generation of novel viral lineages. Previous research has shown a similar evolution pattern whereby recombination was one of the mechanisms of begomovirus evolution and emergence (23,24).

Table 1. Substitution rates and transition/transversion bias among DNA-A, DNA-B, open reading frames and betasatellites

Virus component	Transitional substitution rate	Transversional substitution rate	Transition/transversion bias (R)
DNA-A	9.41–15.93	5.5–8.08	0.863
AC1	9.24–13.83	5.61–8.4	0.853
AC2	8.06–17.43	5.68–7.83	0.887
AC3	8.3–17.35	5.11–8.4	0.88
AC4	10.96–20.27	4.59–6.59	1.357
AV1	8.78–19.63	4.73–6.61	1.083
AV2	9.33–13.84	6.19–6.99	0.888
DNA-B	9.09–16.1	4.92–8.72	0.831
Betasatellite	8.55–14.79	4.98–8.66	0.804

Table 2. Recombination analysis for DNA-A, associated open reading frames, DNA-B and betasatellites determined using different algorithms, showing both major and minor parents along with recombinants and calculated *p*-value

DNA-A						
Events	Breakpoints		Recombinant	Parents		Method
	Begin	End		Major	Minor	
1	2342	2672	KC004131.1_Mac	KJ939854.1_Mac	NC_018090.1_FB	R,G,B,M,C,S,T
2	2325	2725	NC_018090.1_FB	Unknown	JN848770.1_BCV	R,G,B,M,C,S,T
3	2324	2714	KX857725.1_BLC	JN848770.1_BCV	Unknown	R,G,B,M,C,S,T
4	2218	2412	HM357459.3_SiM	JN848770.1_BCV	KX458238.1_ToS	R,G,B,M,C,S,T
5	526	1128	KF440686.1_ToLCGV	MG544284.1_ToLCJV	MW082628.1_BGY	R,G,B,M,C,S,T
6	2154	2341	KJ939858.1_Mac	KC004131.1_Mac	Unknown	R,G,B,M,C,S,T
7	688	995	MG544284.1_ToLCJV	NC_018090.1_FB	Unknown	R,G,B,M,C,T
8	2430	2708	KJ939853.1_BGM	KJ939854.1_Mac	Unknown	R,G,B,M,C,T
9	2286	2710	MK087122.1_FBL	MG544284.1_ToLCJV	JN848770.1_BCV	R,B,M,C,S,T
10	2205	201	NC_055589.1_BLV	Unknown	JN848772.1_BWC	R,G,B,M,C,S
11	2346	2613	LC487406.1_AYV	JQ733557.1_TbC	JN848770.1_BCV	R,G,B,M,C,T
12	2254	2542	MH481901.1_Pep	KX458238.1_ToS	MG544284.1_ToLCJV	R,G,M,C,S,T
13	1068	2018	JN848770.1_BCV	MK387701.1_SiG	Unknown	R,B,M,C,S,T
14	319	2172	MW082628.1_BGY	NC_038988.1_CB	KX857725.1_BLC	R,B,M,C,S,T
15	1482	2024	NC_031961.1_CB	MK241786.1_BGY	MK087122.1_FBL	R,B,M,C,S,T
16	2633	2712	KJ939858.1_Mac	KJ939854.1_Mac	Unknown	R,M,C,S,T
17	2170	2304	MK387701.1_SiG	JN848772.1_BWC	Unknown	R,G,B,M,C,S,T
18	2445	2718	PP937118.1_ToLNV	Unknown	JN848772.1_BWC	R,M,C,T
19	1270	1408	MH359390.1_CLC	KX857725.1_BLC	NC_038988.1_CB	R,G,B,M,C,T
20	2258	2554	KX011477.1_CBS	KX458238.1_ToS	JN848722.1_BWC	R,G,B,M,C,S,T
21	603	1177	KX452228.1_MYM	JQ733557.1_TbC	KJ939858.1_Mac	B,M,C,S,T
22	496	1168	LC487406.1_AYV	NC_018090.1_FB	Unknown	M,C,S
23	2446	2734	JQ733557.1_TbC	KC019304.1_MYM	JN848772.1_BWC	R,M,C
25	2346	2376	KJ939858.1_Mac	KJ939854.1_Mac	Unknown	R,G,T
26	2206	2254	KX458238.1_ToS	KJ939854.1_Mac	KF440686.1_ToLCGV	R,G,M,C,T
28	2084	2526	KJ939854.1_Mac	Unknown	KX011471.1_TbL	R,B,M,C,S,T
29	2238	2674	NC_038988.1_CB	Unknown	JN848772.1_BWC	R,B,M,C,T
30	2463	2502	KC019306.1_HYM	KX452228.1_MYM	MW082628.1_BGY	R,G,T
31	1620	2113	KX011477.1_CBS	MH481901.1_Pep	Unknown	R,B,M,C,T
32	750	1161	HM357459.3_SiM	Unknown	FJ944023.1_BYM	R,M,C,T
33	1654	333	NC_055589.1_BLV	GU32321.1_HYM	Unknown	G,M,C,S
34	2198	2444	JQ733557.1_TbC	PP937118.1_ToLNV	GU32321.1_HYM	R,M,C,T
35	2852	515	KF440686.1_ToLCGV	MK087122.1_FBL	JQ733557.1_TbC	R,G,B,M,C
36	83	518	MG544284.1_ToLCJV	NC_018090.1_FB	PQ373816.1_SLC	M,C,T
DNA-B						
1	710	2396	MH359391.1_CLC	KX857726.1_BLC	KX096982.1_CBS	B,M,C,S,T
2	570	2357	PP937119.1_ToL	Unknown	MT627028.1_Mac	R,B,M,C,S,T
3	2396	706	KX452231.1_MYM	Unknown	MK241787.1_BGY	M,C,T
4	504	1406	KX011474.1_CBM	KX096982.1_CBS	KU145406.1_BGY	M,C,T
5	159	436	NC_035888.1_TbL	MT627028.1_Mac	Unknown	R,G,M,C,S,T
6	2404	2742	KC019307.1_HYM	Unknown	NC_022003.1_BC	R,M,T
7	752	1926	KX857726.1_BLC	MT627029.1_Mac	KU145406.1_BGY	M,C,T
Betasatellite						
1	1000	1286	MG913377.1_CLC	Unknown	NC_018091.1_FB	R,B,M,C,S,T
2	1356	622	MH165182.1_TbC	Unknown	MG544284.1_ToLCJV	R,B,M,C,S,T
3	88	208	NC_018091.1_FB	MG970383.1_ToL	MG913377.1_CLC	R,B,M,C,S,T
4	1182	1336	MK087123.1_SYV	Unknown	MH165182.1_TbC	R,B,M,C,S,T
5	1372	1456	MG913377.1_CLC	Unknown	MG970363.1_ToLB	R,G,B,M,C
6	366	610	MH165182.1_TbC	NC_018091.1_FB	MG970363.1_ToLB	M,C,S

AC1						
1	624	1102	MH359391.1_CLC	JN848770.1_BCV	Unknown	R,G,B,M,C,S, T
2	560	1091	KJ939858.1_Mac	KC004131.1_Mac	NC_031961.1_CB	R,B, M ,C,S,T
3	718	1116	KJ939853.1_BGM	KC004131.1_Mac	LC487406.1_AYV	R,G, M ,C,S,T
4	716	1112	NC_055589.1_BLV	KJ939891.1_Mac	KX458238.1_ToS	R,B,M,C,S, T
5	778	1074	LC487406.1_AYV	Unknown	JN848772.1_BWC	R,G,M,C,S, T
6	666	1010	JN848772.1_BWC	KJ939891.1_Mac	NC_031961.1_CB	R,G,B,M,C,S, T
7	722	1046	NC_018090.1_FB	JQ733557.1_TbC	JN848770.1_BCV	R,G,B,M,C,S, T
8	582	918	MH481901.1_Pep	KJ939853.1_BGM	MG544284.1_ToLCJV	R,G,B,M,C,S, T
9	828	882	MW082628.1_BGY	KJ939891.1_Mac	KC019306.1_HYM	R, G ,M,C,T
10	378	1057	JN848770.1_BCV	Unknown	MK387701.1_SiG	R,G,M,C,S, T
11	555	1030	MK387701.1_SiG	KC004131.1_Mac	KX458238.1_ToS	R,G,B,M,C,S, T
12	618	826	MK087122.1_FBL	JQ733557.1_TbC	Unknown	R, M ,C, T
13	604	1064	FJ944023.1_BYM	KJ939790.1_BGM	Unknown	B, M ,C,S,T
14	902	380	KX458238.1_ToS	MK241786.1_BGY	MG544284.1_ToLCJV	M, C ,S
15	828	884	KX458238.1_ToS	Unknown	GU323321.1_HYM	R, G ,M,C,S
16	1116	400	MG544284.1_ToLCJV	Unknown	KF440686.1_ToLCGV	M ,C,T
17	1116	236	JQ733557.1_TbC	Unknown	MK087122.1_FBL	R,G, M ,C,T
18	636	777	JQ733557.1_TbC	LC487406.1_AYV	KC019304.1_MYM	R,M,C,S, T
19	484	665	KJ939891.1_Mac	KJ939854.1_Mac	JN848772.1_BWC	M ,C,S
20	648	777	KF440686.1_ToLCGV	LC487406.1_AYV	GU323321.1_HYM	R, M ,C,T
21	489	576	NC_031961.1_CB	MH481901.1_Pep	Unknown	M,C, T
AC2						
1	212	292	MG544284.1_ToLCJV	LC487406.1_AYV	JN848772.1_BWC	R, M ,C,S,T
2	490	108	MH359390.1_CLC	KX857725.1_BLC	NC_038988.1_CB	R ,G,B,M,C,T
3	340	490	JN848770.1_BCV	KX011471.1_TbL	HM357459.3_SiM	R,G,M, C
4	344	490	NC_031961.1_CB	FJ944023.1_BYM	MG544284.1_ToLCJV	M,C, T
AC3						
1	123	258	MH359390.1_CLC	KX857725.1_BLC	NC_038988.1_CB	R,G, M ,C,T
2	350	64	PQ373816.1_SLC	LC487406.1_AYV	KC019304.1_MYM	M, C ,S
AC4						
1	474	318	KC004131.1_Mac	Unknown	KJ939854.1_Mac	M ,C,T
2	434	300	MK087122.1_FBL	NC_055589.1_BLV	Unknown	M ,C,T
3	366	477	MH481901.1_Pep	Unknown	NC_018090.1_FB	G, M ,C,S

¹Macrotium yellow spot virus²French bean leaf curl virus³Bean chlorosis virus⁴Bean leaf crumple virus⁵Sida micrantha mosaic virus⁶Tomato severe rugose virus⁷Tomato leaf curl Gujarat virus⁸Tomato leaf curl Joydebpur virus⁹Bean golden yellow mosaic virus¹⁰Bean golden mosaic virus¹¹French bean leaf curl Madikeri virus¹²Bean latent virus¹³Bean white chlorosis mosaic virus¹⁴Ageratum yellow vein virus¹⁵Tobacco curly shoot virus¹⁶Pepper leafroll virus¹⁷Sida golden mosaic virus¹⁸Common bean mottle virus¹⁹Tomato leaf curl New Delhi virus²⁰Cabbage leaf curl virus²¹Common bean severe mosaic virus²²Mungbean yellow mosaic India virus²³Tobacco leaf curl Cuba virus²⁴Horsegram yellow mosaic virus²⁵Squash leaf curl China virus²⁶Synedrella yellow vein betasatellite²⁷Tomato leaf curl betasatellite

Table 3. An overview of genetic composition of the begomoviruses associated with *Phaseolus vulgaris*, highlighting key elements like DNA-A, DNA-B, the 6 open reading frames and the associated betasatellite sequences

Virus component	S	η	π	k	θ-η	h
DNA-A	1558	2881	0.28034	621.519	0.17528	37
AC1	1119	1233	0.33335	243.346	0.39477	35
AC2	284	518	0.30023	114.086	0.3168	36
AC3	297	537	0.29605	115.46	0.32	36
AC4	120	215	0.25019	43.784	0.2943	29
AV1	528	882	0.26424	166.997	0.3281	33
AV2	246	413	0.39278	111.55	0.43825	14
DNA-B	1570	3246	0.39137	734.936	0.43629	26
Betasatellite	872	1531	0.4152	469.179	721.875	8

Table 4. Results of various neutrality tests performed on DNA-A, DNA-B, open reading frames and the betasatellites

Virus component	Neutrality test		
	Tajima's D	Fu & Li's D	Fu & Li's F
DNA-A	-0.26837	0.74873	0.44742
AC1	-0.58307	1.25879	0.69949
AC2	-0.19478	0.82618	0.53982
AC3	-0.2787	0.75559	0.44964
AC4	-0.56305	0.96743	0.50758
AV1	-0.73105	1.09302	0.5144
AV2	-0.45414	0.61857	0.36033
DNA-B	-0.40343	1.02344	0.64034
Betasatellite	-1.13021	-0.37528	-0.61744

A comparative substitution analysis revealed relatively high transition rates in the AC4 and AV1 regions, suggesting differential evolutionary pressures across genomic components. Because these genes are associated with pathogenicity, host-virus interactions and viral movement, elevated sequence variability within these regions may provide adaptive advantages under changing host and vector environments. Previous studies have similarly reported that genes involved in host interaction and pathogenicity frequently exhibit higher sequence variability than more conserved replication-associated regions (25, 26).

The higher level of nucleotide diversity identified for DNA-B and betasatellites further supports the notion that begomovirus-associated genomes are characterised by dynamic evolutionary processes. The significant sequence variability found for betasatellites along with multiple recombination events emphasises their crucial contribution towards genomic diversity and virus diversification (22, 27).

Although recombination and nucleotide variation were detected across all genomic regions analysed, neutrality tests revealed negative Tajima's D scores for almost all genomic loci, indicating the occurrence of purifying selection. This observation suggests that despite ongoing begomovirus diversification by recombination and mutations, purifying selection acts to conserve functional genomic loci essential for the processes of replication and transmission. Positive values observed for some loci analysed according to Fu and Li's test may suggest the maintenance of genetic variation at specific genomic loci (7, 28, 29).

Together, the results suggest that the evolution of begomoviruses is governed by the interactions of mutational, recombinational and selective processes in various segments of the genome. However, it is important to acknowledge that these

observations are based solely on phylogenetic tree topology and bootstrap support, and no formal statistical analysis of population structure was performed to confirm that geographic origin is a significant driver of this clustering. Phylogenetic co-clustering of geographically proximate isolates may also reflect shared host range, vector population structure, or historical dispersal routes rather than geographic isolation alone. Future studies incorporating larger and more geographically representative datasets, combined with formal population structure analyses, will be necessary to rigorously determine the extent to which geographic barriers shape the genetic diversity of begomoviruses infecting *P. vulgaris*.

Conclusion

The study highlights that begomovirus evolution is influenced by geographical distribution, host adaptation and vector dynamics, resulting in phylogenetic groupings that broadly reflect geographic distribution. Adaptive variation was observed in certain genomic regions, particularly AC4 and AV1, which exhibited relatively higher mutation rates. Frequent recombination events, especially within DNA-A and the AC1 region, contributed substantially to genetic variability. Despite this variability, purifying selection appears to conserve essential viral functions required for replication and transmission. Overall, the combined effects of mutation, recombination and selection shape begomovirus diversity, emphasising the importance of continuous molecular surveillance to better manage their impact on legume production. Future studies incorporating larger datasets and population genetic analyses will further improve understanding of begomovirus evolution and diversification.

Acknowledgements

The authors sincerely express their gratitude to the Vice Chancellor of the Central University of Jammu, Jammu and Kashmir, India for providing essential facilities and unwavering support. The authors are also thankful to Anusandhan National Research Foundation (ANRF), Govt. of India, for financial support (Project No. ANRF/IRG/2024/000372/LS).

Authors' contributions

CS conducted the research, performed data analysis and wrote the manuscript. YS conceptualised and supervised the study, provided

guidance throughout the research process and critically reviewed the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

Declaration of generative AI and AI-assisted technologies in the writing process

Generative AI or AI-assisted technologies were used solely for grammar checking. These tools were not used for data analysis, interpretation, or the generation of scientific content. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- Gepts P. Crop domestication as a long-term selection experiment. In: Plant Breed Rev. 2003. p. 1–44. <https://doi.org/10.1002/9780470650288.ch1>
- Blair MW, González LF, Kimani PM, Butare L. Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. Theor Appl Genet. 2010;121:237–48. <https://doi.org/10.1007/s00122-010-1305-x>
- Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. Beans (*Phaseolus* spp.)-model food legumes. Plant Soil. 2003;252:55–128. <https://doi.org/10.1023/A:1024146710611>
- Winge T. Seed legislation in Europe and crop genetic diversity. In: Lichtfouse E, editor. Sustainable Agriculture Reviews. Vol. 15. Cham: Springer; 2015. p. 1–64. https://doi.org/10.1007/978-3-319-09132-7_1
- Bhagat S, Kumar S, Dogra S. Identification of induced mutants for morpho-physiological and biochemical traits of rajmash (*Phaseolus vulgaris*). Indian J Agric Sci. 2023;93:1338–43. <https://doi.org/10.56093/ijas.v93i12.135595>
- Chang-Sidorchuk L, González-Alvarez H, Martínez-Zubiaur Y. Begomoviruses infecting common beans (*Phaseolus vulgaris* L.) in production areas in Cuba. Span J Agric Res. 2018;16. <https://doi.org/10.5424/sjar/2018162-11735>
- Lima ATM, Silva JCF, Silva FN, Castillo-Urquiza GP, Silva FF, Seah YM, et al. The diversification of begomovirus populations is predominantly driven by mutational dynamics. Virus Evol. 2017;3. <https://doi.org/10.1093/ve/vex005>
- Duffy S, Holmes EC. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. J Virol. 2008;82:957–65. <https://doi.org/10.1128/JVI.01929-07>
- Saitou N, Imanishi T. Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. Mol Biol Evol. 1989;6(5):514–25. <https://doi.org/10.1093/oxfordjournals.molbev.a040572>
- Kumar Chanchal D, Alok S, Kumar M, Kumar Bijauliya R, Rashi S, Gupta S. A brief review on *Abelmoschus esculentus* Linn. okra. Int J Pharm Sci Res. 2018;9:58–66.
- Posada D, Crandall KA. Selecting the best-fit model of nucleotide substitution. Syst Biol. 2001;50:580–601. <https://doi.org/10.1080/106351501750435121>
- Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Biol Evol. 2021;38:3022–7. <https://doi.org/10.1093/molbev/msab120>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547–9. <https://doi.org/10.1093/molbev/msy096>
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015;1. <https://doi.org/10.1093/ve/vev003>
- Bland JM, Altman DG. Statistics notes: Multiple significance tests: The Bonferroni method. BMJ. 1995;310:170. <https://doi.org/10.1136/bmj.310.6973.170>
- Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302. <https://doi.org/10.1093/molbev/msx248>
- Tajima F. Evolutionary relationship of DNA sequences in finite populations. Genetics. 1983;105:437–60. <https://doi.org/10.1093/genetics/105.2.437>
- Watterson GA. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 1975;7:256–76. [https://doi.org/10.1016/0040-5809\(75\)90020-9](https://doi.org/10.1016/0040-5809(75)90020-9)
- Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–95. <https://doi.org/10.1093/genetics/123.3.585>
- Lefeuve P, Lett JM, Reynaud B, Martin DP. Avoidance of protein fold disruption in natural virus recombinants. PLoS Pathog. 2007;3(11). <https://doi.org/10.1371/journal.ppat.0030181>
- Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, et al. Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol. 2015;160:1593–619. <https://doi.org/10.1007/s00705-015-2398-y>
- Mondal D, Mandal S, Shil S, Sahana N, Pandit GK, Choudhury A. Genome wide molecular evolution analysis of begomoviruses reveals unique diversification pattern in coat protein gene of Old World and New World viruses. Virusdissease. 2019;30:74–83. <https://doi.org/10.1007/s13337-019-00524-7>
- Fiallo-Olivé E, Navas-Castillo J. Begomoviruses: What is the secret(s) of their success? Trends Plant Sci. 2023;28:715–27. <https://doi.org/10.1016/j.tplants.2023.01.012>
- Crespo-Bellido A, Hoyer JS, Dubey D, Jeannot RB, Duffy S. Interspecies recombination has driven the macroevolution of cassava mosaic begomoviruses. J Virol. 2021;95. <https://doi.org/10.1128/JVI.00541-21>
- Naveed H, Islam W, Jafir M, Andoh V, Chen L, Chen K. A review of interactions between plants and whitefly-transmitted begomoviruses. Plants. 2023;12:3677. <https://doi.org/10.3390/plants12213677>
- Breves SS, Silva FA, Euclydes NC, Saia TFF, Jean-Baptiste J, Andrade Neto ER, et al. Begomovirus-host interactions: Viral proteins orchestrating intra- and intercellular transport of viral DNA while suppressing host defense mechanisms. Viruses. 2023;15:1593. <https://doi.org/10.3390/v15071593>
- Vinoth Kumar R, Singh D, Singh AK, Chakraborty S. Molecular diversity, recombination and population structure of alphasatellites associated with begomovirus disease complexes. Infect Genet Evol. 2017;49:39–47. <https://doi.org/10.1016/j.meegid.2017.01.001>
- Nigam D. Genomic variation and diversification in begomovirus genome in implication to host and vector adaptation. Plants. 2021;10:1706. <https://doi.org/10.3390/plants10081706>
- Chahat S, Akshu J, Rahul K, Yogesh K. Evolutionary dynamics and recombination patterns in begomoviruses infecting *Abelmoschus esculentus*: A phylogenetic and population structure analysis. Plant Sci Today. 2026;13. <https://doi.org/10.14719/pst.11190>

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonpublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonpublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). 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 (<https://creativecommons.org/licenses/by/4.0/>)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.