Computational Sequence analysis of DNA-A&B segments of Begomoviruses infecting vegetables.

Submitted By**:**

Khizar Hayat (Student ID: 14001231012)

Supervised By**:**

Dr. Muhammad Ali

Department of Life Sciences, School of Science

University of Management and Technology

Lahore, Pakistan

2020

Table of Contents

List of Tables

List of Figures

S. No.Title **Page no.**

List of Abbreviations

BGYMV	Bean Golden Yellow Mosaic Virus
BGMV	Bean Golden Mosaic Virus
CTV	Curly Top Virus
CIAT	Center Internacional de Agricultura Tropic
CLV	Cassava Latent Virus
CPT	Cycling Probe Technology
DNA	Deoxyribonucleic acid
MSV	Mosaic Streak Virus
PTGS	Post Transcriptional gene silencing
RDR	Recombination Dependent Replication
RNA	Ribonucleic Acid
PCR	Polymerase Chain Reaction
RCA	Rolling Circle Amplification
SsDNA	Single stranded DNA
TYLCD	Tomato Yellow Leaf Curl Disease

Abstract

Begomoviruses are the members of *Geminiviridae,* which mostly infects dicots. Vector which is the cause transmission of this virus are whiteflies *Bemisiatabaci*. They have been characterized into monopartite and bipartite genome associated with DNA A&B. including some devastating strains which are linked with monopartite genome. They infect a broad range of horticulture and vegetable crops resulting into a huge economic loss every year. It has been a nutshell for scientists since the twentieth century and till now there are regular updates in its strains. Total number of DNA-A sequences infecting in vegetables were 151. 30 Species including the extract from phylogenetic tree of maximum likelihood. Species from tree were 15 (Based on 91% identities Brown et al. 2015) Pair wise sequence comparison revealed 86% cutoff value and beyond this point new Species are predicted DNA -B consisted on total number of 17 sequences from which 13 Species were included (Based on 69% sequence identities) Pair wise sequence Comparison for DNA-B segment depicted 69%

Chapter 1

Introduction

These are plants infecting viruses with a very vast range of hosts infected, majorly dicotyledonous plants are susceptible by these viruses. Overall the globe they are notorious for devastating the economy by infecting important crops like tomatoes, beans, squash, cassavas and cotton crops. These crops are more susceptible to parasitic infections and are central component of agricultural economy. There are currently **322** species of Begomovirus are yet identified including ***Bean golden yellow mosaic virus*.** This genus possesses small circular single stranded DNA (ssDNA) with certain morphological adaptations(Zerbini*et al.,* 2017)

With ultimate morphological adaptations, these viruses are **twinned quasi-isometric virions**. The viral particles are non-encapsulated. They have nucleocapsid of 38nm in length and usually 15-22nm in diameter. These viruses have basic icosahedral symmetry, consisting on two incomplete icosahedra. They are so called incomplete, because of one vertex missing and are coupled together. There are 22 capsomeres constituting per nucleocapsid.

The vector for the transmission of genus begomovirus is whitefly***Bemisiatabaci***.Regarding to this vector it allows rapid growth and propagation of this viral genus due to its indiscriminate feeder adaption. The indigenous begomovirus first widely affected the noncultivated plants e.g. weeds, which serve as ancestors of crop-infecting viruses. Later on with rapid increase in population and increasedgeographical range of invasive variable B biotype of *B. tabaci*hasled to begomovirus infections in cultivated plants or crops.(Brown, et al.,2007).

The genus begomovirus consist of two genetically different Species, **Monopartite and Bipartite** genomes. Mostly begomovirus possess bipartite genome which means that their genomic class are divided into two separate parts, which are termed as segments. These are in the form of two circular small ssDNA termed as **DNA-A and DNA-B**. Both these segments aid in development of infectious symptoms. DNA-B on the other hand is totally dependent on DNA-A for its replication. Separate functions are associated with both of these DNAs, but DNA-A specifically encodes five or six proteins including replication protein Rep, coat protein and transport or regulatory proteins. This part of begomovirus is homologous to all other genomic categories of geminiviridae. The genes present on DNA-A play major role in different functions like replication, gene expression, host defense overcome, encapsidation and insect transmission. On the other hand DNA-B encodes two different movement proteins. The two components A and B are similar to a small portion of around 200 nucleotides, with 85% identity known as common region.Begomovirus genome has pathogenic adaptation of rapid evolution through common mutation, pseudo recombination, recombination and by adaptind new DNA components and satellites.

Pseudorecombinaton is a cascade also known as regular grafting, DNA-A donating its common region to the instantly occupied DNA-B. Proteins in this genus present either on the sense strand (positive orientation) or its complement (negative orientation)(Pita., 2001) The genome of these viruses consists of following genes as in **DNA-A,** V1 which is (R1) termed as positive orientation: Coat protein— with 29.7 (kDa) molecular weight , V2 the positive orientation: Movement protein precoat ORF with 12.8 kDaMW , C1 (L1) also known as negative orientation: Replication initiation protein (Rep) with 40.2 kDa MW, C2:as in L2 known as negative orientation: Transcription activator protein (TrAP) with 19.6 kDaMW, C3: L3 The negative orientation: Replication enhancer with 15.6 kDaMW,C4: the negative orientation: might determine symptom expression with 12.0 kDaMW. The other half;**DNA-B,** V1 R1 knowns as positive orientation: Nuclear shuttle protein with 33.1 kDa MW, C1 L1 also known as negative orientation: Movement protein with29.6 kDa as referred in (Sattar.et al, 2012).

The genomic structural variation of both of these mono and bipartite begomoviruses are also depicted in their symptoms.Monopartiteinfections are only limited to phloem tissues. The symptoms are induced stunting and leaf curl. These are not transmitted by Sap. Bipartite infect phloem and other tissues also. Symptoms include are leaf curling, crumpling and mosaic symptoms which are actually suspected to be transmitted through sap.

Recently identified monopartite begomovirus are associated with DNA satellites also known as betasatellites and nanovirus like DNA satellite molecules also known as alphasatellites**.** Betasatellites are a bit different in nature and are linked to helper begomovirus. They are half in size than the helper begomovirus. It contains the potential stem loop having nonanucleotide sequence TAATATTAC. Betasatellites has a specific region on 100 nucleotides whisch termed as satellite conserved region SCR. Satellites are entirely dependent on helper virusfor its replication. All functions of betasatellite are regulated by BC1. This protein regulates post transcriptional gene silencing and pathogenicity control as stated in .(Briddon *et al*., 2001)

They are not supposed to be true satellites because they can replicate themselve, that's why they are termed as satellite like molecules. They are half in size in genome than their helper begomovirus, but larger than betasatellites somehow. They are highly conserved in their structure consisting of only a single large gene, In the virion-sense orientation which encodes a Rep protein, an A-rich region and a predicted hairpin structure which has a nona-nucleotide sequence (TAGTATT/AC) Forming section of the loop. They can cause decrease in levels of viral DNA. Some absurd alphasatellites are identified which associate with reduction of accumulation on betasatellites as referred in. (Briddon et al, 2006)

Like all other viruses they don't have their own DNA polymerase, they depend on hosts DNA replication machinery present in the nuclei of infected plant cells. Begomoviruses replication is a combination of RCR and recombination-dependent replication RDR mechanism. The translated Rep protein is transferred to the plant nuclei to activate the viral DNA replication. The whole process consist of regular 3 phases initiation, elongation and finally termination (Pasumarthy et al., 2011).

The world is developing at very high pace in the domain of technology. On the other hand, competition in pathogens is also increasing with every passing day. Every year, new organism are reported and considered after thorough researches. Overall economy of world and its entire needs depend on agricultural economy. Crops are the main source of survival and earning of any country. But ever since, man has stepped into the agriculture field, he has been facing related issues from beginning. Massive efforts have been done every day to eradicate the ailments of fruitful plants and preservation from possible attacks. From insects and other microorganisms these crops are suspected to get infected. Round the world it is first and foremost priority of every country to protect its agricultural horizons from any possible harm. Major crops of this era are rice, wheat, tomatoes and potatoes etc. including these majority aid fall in earning through exporting goods. At the top of all the aim of this study is to know the classification of BGMV related infections, their types and subtypes, genomic composition, hosts, vectors and their brief lineages all around the world.

Plant viruses are classified into 15 families and there are about 47% plants diseases which are caused by plant viruses (Yadava et al., 2010). *Geminiviridae*is plant-pathogenic virus family. Geminivirus has single-stranded circular DNA (ssDNA) nucleic acid ranging from 2.5 to 5.2 kb genome (Lazarowitz et al., 1992). They are transmitted by insects as in (Davies et al., 1989).

Geminiviruses can cause diseases in multiple crops of both monocotyledons and dicotyledons plants round the world, that results remarkable yield reduction in economically among important crops (Zhou et al, 2013). The Family*Geminiviridae* is further classified in to nine genera (Silva et al., 2014). These genera are, *BegomovirusBecurtovirus, Curtovirus, Capulavirus, Eragovirus, Grablovirus, Mastrevirus, Topocuvirus and Turncurtovirus* (Varsani et al., 2014).

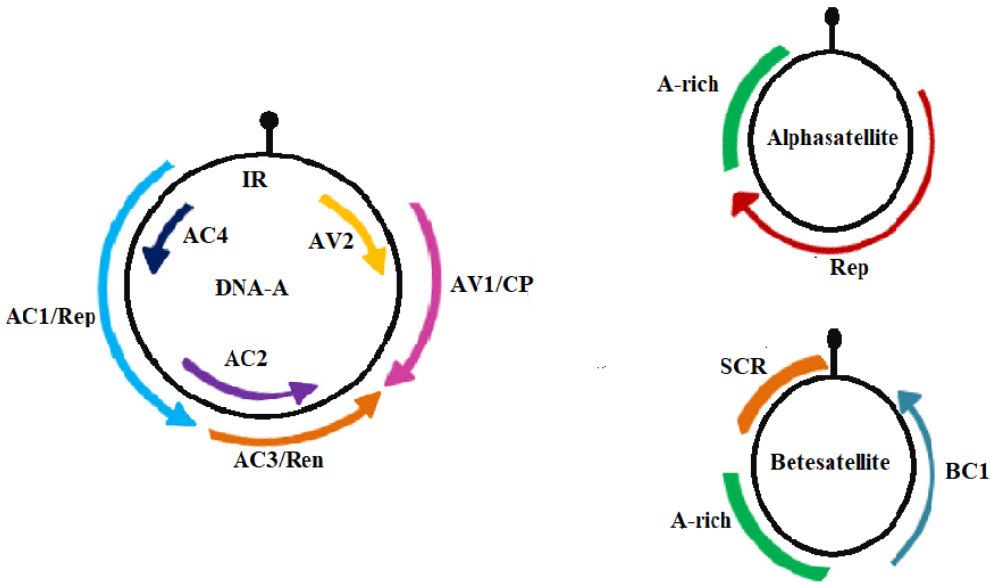

Figure 1.1The genomic component DNA-A of new world begomovirus show a set of five open reading frames ORFs, one is present in the virion sense strand AV1 or coat protein which encodes the coat protein. the AC1 or Rep; AC2 or TrA*P*; AC3 or REn*;* and AC4 that encode proteins p;ay a key role in multiple functions including plant cell cycle interference and DNA replication. They act as temporal regulation and other viral activities. DNA-B has only two components BV1 and NSP. BC1 and MP are involved in inter and intracellular movements. Bipartite BGVs has a 180bp sequence also known as common region CR. This section contains origin of replication which serve as stem loop in rolling circle amplification

Genomic organization of monopartite begomoviruses. DNA-A with four open reading frame in the complementary stand also known as C1, C2, C3 , C4 and on the other hand V1 and V2 are parts of viral stand. Whereas C1 serves as Replication initiation, C2 Transcription activator of right side, C3 as Replication enhancement, C4 as Viral replication and V1 Encapsidation,V2 consists on Cell to cell movement protein while on the other hand alpha and beta satellites are joined with DNA-A of monopartite begomoviruses. Alphasatellite has only Rep replication linkage protein and betasatellite has C1 Replication initiation as in (Hanley et al., 2013).

In genomic organization, bipartite begomoviruses are usually New World and the other monopartite are old world (OW) genome with some exceptions. Bipartite begomoviruses consist of two DNA genomic components (DNA-A and DNA-B) and monopartite mostly have DNA associated with satellites complexes, (alpha and beta) satellites. The genome size of DNA-A component of monopartite begomoviruses is similar to the begomoviral DNA-A component bipartite begomoviruses which is 2.5 to 5.2 kb in length(Zhao et al., 2018).

DNA-A has five or six proteins in the complementary stand .It has four open reading frames as following; (ORFs), C1, C2, C3 and C4 and in viral stand it has two ORPs,V1 and V2 are in viral stand. While C1 is Replication initiation, C2 Transcription activator, C3 Replication enhancer, C4 with variation and V1 coat protein, V2 pre coat protein gene as in (Nawaz et al., 2009).

DNA-B have two genes V1 known asNuclear shuttle protein and C1 termed as movement protein. Whereas one is responsible specifically for nuclear trafficking and the other is involved in pathogenicity determination and Cell to cell movement respectively. In few cases, the DNA-A of bipartite begomovirus may possess ability to spread disease without DNA-B (Briddon et al., 2010). It is not dependent and the purpose of Rep protein is strongl in the viral replication in Rolling circle mechanism as in (Arguello et al., 1994). Presence and absence of AV2 only affects viral motion. AV2 is majorly associated with symptom development and sometime determines remarkable yield of affected plants (Padidam et al., 1996)

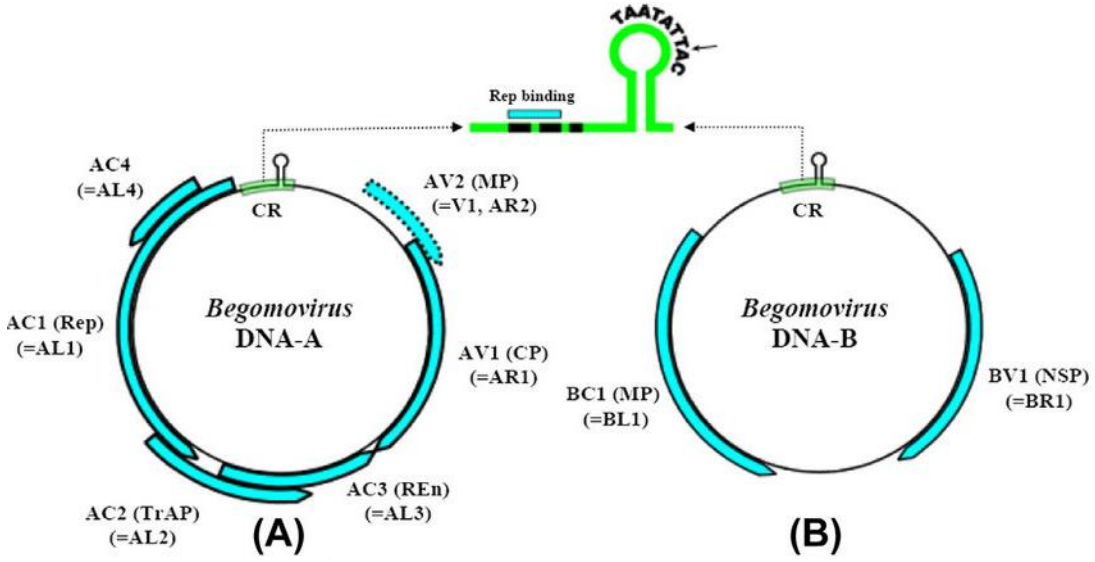

Figure 1.2 Bipartite: DNA A&B likewise in TYLCV, ToLCV and AYVV which has monopartite genome, the circular ssDNA of positive strand shares 2.6 bp length out of 5.2 kb. Poly A tract is not present on 3 terminus. They are present in both positive and complementary strand of virion.
genome replicates with double stranded mediators. Host DNA polymerase is used to start rolling circle amplification. Fixed nonanucleotide sequence serving as stem loop (TAATATTAC). The synthesis of ssDNA B completely depends in DNA A. The both components have 200 bp IR region. Temporal expression is controlled by two different promoters present on common region.

Genomic classification of bipartite begomovirus. DNA-A with four open reading frame in the complementary stand which are termed as C1, C2, C3 and C4 and V1 and V2 are present in viral stand. Whereas C1 acts as Replication initiation, C2 as Transcription activator of right side, C3 as Replication enhancement, C4 as Viral replication and V1 as Encapsidation,V2 contains Cell to cell movement protein while on the other hand DNA-B involves V1 and C1 which is responsible for pathogenicity, Whereas nuclear trafficking and determination of Cell to Cell movement is changed from. IR which is non coding intragenic region where conserved sequences are present in common region known as "iterons". Rep nicking which take place at origin of replication present in the IR. The nonanucleotide(TAATATT$^{\downarrow}$AC) is present the hairpin is the Rep binding site as in (Hanley et al., 2013).

Objective:

To analyze the available segments of DNA-A and DNA-B present in the databases in comparison with the prevailing classification system.

Chapter 2

Literature Review

2.1 Geminivirus

A wide range of plant pathological issues are related with geminivirus. CTV Curly Top Virus which is associated with major diseases caused in the crops of Western United States (Bennett, 1971). Paired particles were observed in semi purified particles from beans which had clear symptoms of bean golden mosaic virus BGMV which was later on confirmed by scientists of Colombia and in United States as in(Goodman et al., 1977), This virus was transmitted by the native whitefly of America known as Bemisiatabaci as shown in figure 2.which will be elaborated further in histological analysis of the vector specimen. Their investigations confirmed a single stranded DNA (ssDNA) with not even a hope or expectations of such results as they were seeking for RNA based genome (Goodman et al 1977). Other collaborated studies notified some notorious maize damages like maize streak disease caused by Mosaic Streak Virus MSV. At that time, remembering the radical efforts of H.H Storey and coworkers. The natural vectors of this virus are leafhoppers Cicadulinambila which are much similar in epidemiological terms of all other vectors of this strain (Bennett et al 1971). In Africa the variant Cassava Latent Virus CLV was detected(Bock et al., 1978),comparative study showed ssDNA presence in both variants(Harrison et al., 1977).Another variant similar to the recently confirmed series **CSMV** was confirmed in Australia(Franckil et al, 1979).

Figure 2.Whitefly also known as Bemisia tabaci.There is 24 different species varying. They are highly sensitive to cold temperature as it kills them. They can be clearly observed under the leaves of infected plants or plants showing clear symptoms. Their diversity varies from ecological locactions.

2.2 Taxonomy
It has enormous range of strains .

2.3Becurtovirus
This genus has overall two species Beet curly top Iran virus and Spinach curly top Arizona virus. Leafhopper are the vectors transmitting this disease to dicots as in .(Varsani et al, 2014)

2.4Capulavirus
Four species are included in this genus.Aphids are the primary vectors as mentioned in. (Varsani et al 2015)

2.5Curtovirus
Three species are included in this genus Beet curly top virus they are remarkable pathogens in North America and Iran. Leafhoppers are the vectors transmitting this disease to dicots as in. (Chen et al ,2010)

2.6Mastrevirus
Dicots and monocots both are infected by mastrevirus. They are extensively transmitted by leafhoppers 30 species are yet identified Maize streak virus and Wheat dwarf virus are among well studied as mentioned in. (Muhire et al, 2013)

2.7Eragrovirus
Only one species is included in this genus. Majorly Eragrostiscurvula streak virus as in. (Varsani et al, 2009)

2.8Grablovirus
Like above it also includes only one species. Species are Grapevine red blotch virus. .(Varsani et al, 2014)

2.9Topocuvirus
It has only one species, Tomato pseudocurly top virus, treehoppers are the vectors as mentioned in. (Briddon et al, 1996)

2.10Turncurtovirus
Turncurtovirus Turnip curly top virus is one spesies that is included. Every species included in this genus are yet identified and the vector for its transmission are leafhoppers *Brassica rapa* or *Raphanussativus* in Iran. As referred in (Briddon et al, 2010)

2.11 Replication cascade

According to series of researches carried out it was estimated that the replication of this viral components occurs in nucleus. Fibrillar rings of deoxynucleoprotein were observed with nuclear modifications (Kim *et al.*, 1978).

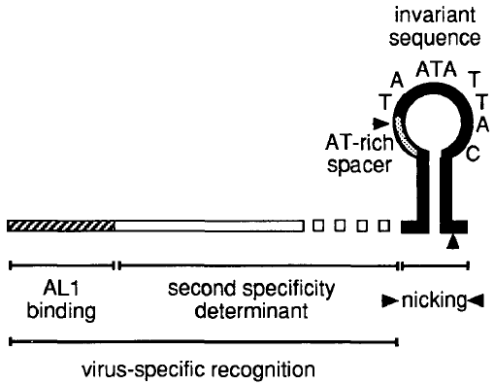

Fig 2.2:shows the replication origin of the virus class. The AL1 binding region is indicated. AT rich spaces are marked . DNA sequence that contains the possible nick site for initiation of rolling circle replication shown by (Etessami et al, 1989),. The process shows the entire phenomenon of initiation elongation and termination finally..

2.11 Diversity of Begomovirus infections

All over the world this viral strain infects hundreds and thousands of horticulture and edible crops with a high polymorphic degree. Ranging from tobacco its infections are down streaming to major economic crops like soya bean, tomatoes and potatoes etc.

2.12Infection in soybean

In north western region of Argentina which are important bean growing regions were reported to be infected by bean golden mosaic virus and bean dwarf mosaic virus (Morales *et al.*, 1990). Further studies on characterization of this virus revealed that virus was related to tomato golden mosaic virus (Rodríguez *et al.*, 1998). Their studies depicted that more than one viral strains could be involved in a particular infection.

2.13 Tomato yellow leaf curl disease

It is the most devastating disease of tomato crop in all over the tropical and subtropical regions of the world (Moriones and Navas-Castillo, 2000). With respect to tomato, some TYLCD-associated viruses also have been involved in infections of common bean (*Phaseolusvulgaris*L.) (Navas *et al.*, 1999) and pepper (*Capsicum annuum*L. and *C. chilensis*Jacq.) (Morillae et al., 2005) crops.

2.14 Genetic resistances and its inheritance

Most important genotype which is genetically resistant came from Middle American races (Morales et al, 1991). Green house breeding was carried out at Centro Internacional de AgriculturaTropic. Germplasm was analyzed in that study (CIAT, 1982).

2.15 BGMV invasion in phloem

Sap incubated infected plants were observed. Microscopic examination was conducted on 15-20 days incubated samples as in (Carr et al, 1983).(Honda et al, 1969). Clea symptoms and associated comples were observed.

2.16 Alternative hosts of BGMV

Possibility of various weeds serving as host for BGMV infections were predicted (Chagas et *al*, 1981). *Canavalia*sp., *Crotalaria* spp., *Macroptiliumlathyroides*, and *M. erythroloma*are the possible alternative hosts considered in this study. Using fingerprint diagnosis, confirmed thepresence of distinct sets of restriction fragments that specifically hybridized with BGMV DNA sequences in *M.*

*lathyroides*and*Malvastrumcoromandelianum*. None of these early studies, however, established infectivity cycles as proof to corroborate their findings (Haber et *al,* 1987).

The AL1/C1/Rep Protein:

Genomic organization of bipartite begomovirus. DNA-A have four open reading frame in the complementary stand which are known as C1, C2, C3 and C4 and V1 and V2 are in viral stand. Where C1 is Replication initiation, C2 Transcription activator of rightward, C3 Replication enhancement, C4 Viral replication and V1 Encapsidation,V2 consists of Cell to cell movement protein while DNA-B have V1 and C1 responsible for pathogenicity, Nuclear trafficking and determination of Cell to Cell movement Modified from. IR is non coding intragenic region where conserved sequences are present in common region called “iterons”. Rep nicking is take place at origin of replication present in the IR. The nonanucleotide(TAATATT$^{\downarrow}$AC) present in the hairpin is the Rep binding site (Hanley et al., 2013).

Trans-replication

The trans-replication is very important phenomenon in the bipartite begomoviralcognate and non-cognate genomic DNA components.The DNA-A have Rep protein involves in the trans-replication of viral DNA-B. The trans-replication occurs by the interaction between the Rep and iterons sequences. The Rep attached to thenonanucleotide sequences at origin of replication in the IR (Gladfelter et al., 1997). It has been studied that, iterons are very important for the trans-replication of begomoviral cognate and non-cognate DNA components. The Rep binds to the conserved iterons in the CR and viral replication takes place. They are important in infectivity causing in the plants. It has been studied that AV2 and Rep have key roles in the Trans-replication.

Intergenic Region (IR)

IR is a non-coding region present in both DNA-A and DNA-B of bipartite begomoviruses. This conserved sequence are located in the middle of start codon of rightward and leftward coding region. The CR is located within the IR and Rep protein binds to CR only while replication. The nucleotide sequence of IR is highly conserved in begomoviruses and studied that IR is necessary for replication (Laufs et al., 1995). The IR, with in a hairpin or stem-loop structure, contains the ori of virion strand DNA replication and promoter elements with the conserved nonanucleotide sequence (TAATATT↓AC) and also repeated upstream motifs called “iterons” (Chatterji et al., 2000).

Common Region (CR)

The split genome of begomoviral DNA-A and DNA-B are completely distinct from one other except a region of about 200 to 250 nt known as common region. Cognate DNA-A and DNA-B shares ≥ 85% nucleotide sequence identities in their CR (Hanley et al., 1999). The CR is located in the IR between the Rep-protein and V2 in DNA-A, and C1 and V1 in DNA-B (Zhao et al., 2018).

The CR contains (Ori)origin of replicationcovered the segment from tandem repeat of iteron to end of the stem loop and several sequential elements that are TATA box premotor for the viral and complementary sense ORFs, sequence repeats (iteron) and a nonanucleotide sequence (TAATATT↓AC) is present in the hairpin loop structure. It has been confermed know these conserved sequences may be the actual binding sites of Rep protein and replication initiation (Chatterji et al., 2000). This hairpin motif included an AT rich loop and GC rich stem. Rep protein has leading role in the replication and host defense mechanism. Both genomic DNA components of bipartite begomoviruses are necessary for infectivity causing (Stanley et al, 1983).

There are 2-3 iterons which are present upstream of TATA box of rep protein in the OW begomoviruses and their complementary iterons are present in their downstream while this kind ofiterons sequence absence in the NW begomoviruses(Arguello et al., 1994).TheDNA-B captures CR from DNA-A through a mechanism called ‘regulon grafting’. The phenomena in which DNA-A component gives its CR, by recombination process which is captured by DNA-B of a bipartite begomovirus, resulting a new dependent interactions between two DNA components called ‘regulon grafting’ (Saunders et al., 2002).

Rolling circular Mechanism (RCR)

The RCR is the predominant mode of replication in begomoviruses DNA. The DNA-A replicates independently in the absence of DNA-B, but this independence does not hold true for DNA-B. Begomoviruses depends upon the host machinery for replication because they encodes a few. These factors are DNA polymerase and repair polymerase to amplify their genome.

DNA replication is the two-step process.In first stage of replication the conversions of viral circular ssDNA into replicative forms (RF) or supercoiled dsDNA. The dsDNA used as template for the synthesis of ssDNA in the next stage which provides foundation for the viral assembly (Preiss et al., 2003).The Rep protein is the prime

viral protein necessary for the initiate and extend RCR process. Rep protein have a specific binding and nicking site in the CR region.

The single stranded DNA of bipartite begomoviruses is converted into double strand DNA by the help of host DNA polymerase the Rep binds and nicks the origin of replicationin the circular dsDNA and initiation of RCR replication with in nonanucleotides(TAATATT↓AC) of hairpin loop takes place (Laufs et al., 1995; Orozco et al., 1996). Rep nicking takes place between the 7^{th} and 8^{th} nucleotide in nonamer sequence (Preiss et al., 2003). The DNA polymerase extends the 3' site, whereas rep protein still bonded to the 5' site and travels along with template as a result of that a looped rolling circle is formed.

regeneration of origin of replication Ori site present on the prenatal site can takes place. Rep protein starts a new cycle of replication by parting the ori site and transferring it to the prenatal stand. Whereas rep protein initiates RCR and also aids in termination. It also supports the unwinding of double stranded dsDNA which is in joined with rep helicase as in (Baas et al., 1988).

RCR is literally two step process occurs in leading and lagging strand while DNA synthesis (Kornberg et al., 1992). Research reveals, that the leading strand is the single strand also called as plus (+)ve strand. It is used as template or virtual copy during viral DNA synthesis of minus (-)ve strand or lagging strand for the production of a double strand replicative form RF, whereas RF is called to be an intermediate stage of replication.

The plus strand synthesis can be done by using the RF as template for generating free ssDNA.The RCR has site-specific activity by nicking necessary for the synthesis of plus strand, whereas the synthesis of minus strand is initiated through DNA primase generated RNA molecule. The study shows, that the major difference between both of these strands is the plus strand is present in both ssDNA and dsDNA genomic components. While the minus strand is only studied in the double stranded dsDNA as in (Stanley et al, 1995).

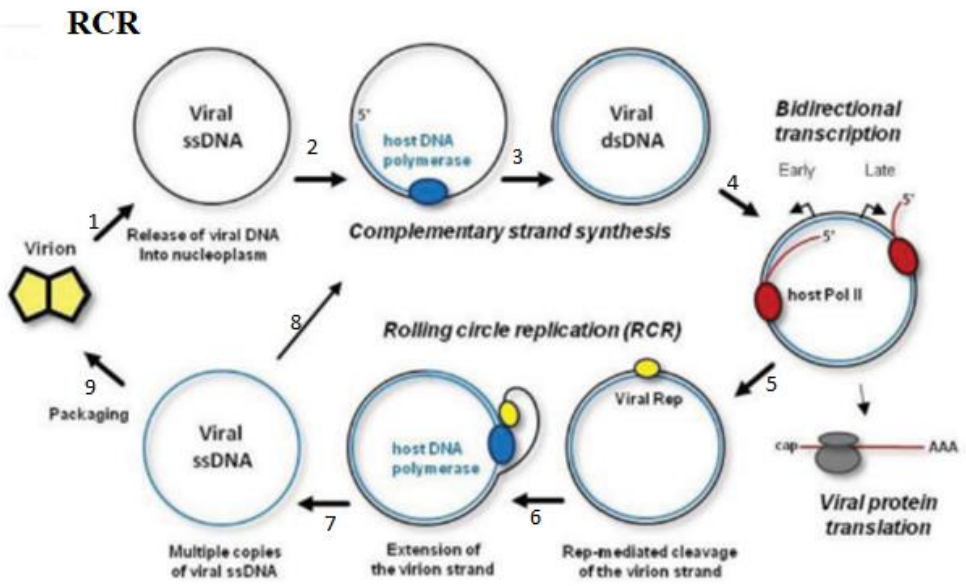

Figure 2.3: showing model of RCR Rolling circle replication**.** The ssDNA of bipartite begomovirus is injected into the nucleosome of host plant. Begomovirus which lacks its own DNA polymerase induce host DNA machinery for its DNA synthesis. In the next step, two directional DNA replication can occur by using dsDNA as template which is shown as coat protein on right side and Rep left side. The 5th step movement of Rep protein initiates RCR The Rep protein nicks at the particular site in the virion strand which is present in the origin of replication. The host DNA polymerase then extends the 3' end on complementary strand. This is also supported by the polymerase complex while Rep extends this process.

The polymerase complex covalently attached to the 5' end of virion strand and move the virionstrand.Then after some cycles of replication the termination can occur by nicking of Rep protein. In this process some copies of newly synthesis ssDNA are formed. In the 8th step of the RCR mechanism the copies of ssDNA are released restarting of replication round or get packed into virion strand (Pooggin et al, 2013).

Iterons

There are two main components, iterons and IRD (iteron related domain) necessary for the trans-replication.Iterons are shortcoserved repeat sequences of DNA located upstream stem-loop structure near Ori.Iterons are directly repeated DNA sequence and play a crucial in the regulation of plasmids. The iterons binds to the initiator protein as well as responsible for the negative control of initiation. Begomoviruses encode Rep protien which binds with iterated DNA motifs (Iterons) necessary for virus replication.

The Rep protein binds to iterons and are thought to be recognition sites for initiation (Abeles et al., 1995). Rep-protein has subdomains of about 8 to 10 residues. The core sequence of iterons is vary in begomoviruses starting from three direct repeats (I, II and III) with GGN1N2N3 core consensus sequence and one inverted (IV) repeat iteron located around TATA box (Arguello et al., 2001).

IRD(Iterons related domains)

Iterons related domains (IRD) are the amino acid sequence in Rep-protein. The Rep-IRD comprised of eight amino acids and form functional domains of Rep along with motif 1(Arguello-Astorga et al., 2001). Rep-IRD located at the N-terminal region of Rep protein with FX1X2X2 amino acid sequences (Argüello et al., 1994) and associated to core iterons sequences.

Pseudo-recombination

Pseudo-recombination occurs by the trans-replication genomic segments/components of multipartite viruses (Chakraborty et al., 2008) and competent capture is the phenomena in which a multipartite virus captures one or more components. Plants infected with pseudo-recombinant genomic DNA-A and DNA-B shows >96% similarity in IR region (Gilbertson et al., 1993). Study reveals that trans-replication is the basic need of a heterogenomic DNA-B component of a bipartite begomovirus for viable pseudo-recombination but it is not compulsory that trans-replication is resultant in a viable pseudo-recombination as per reference in (Hofer et al., 1997).

Materials and Methods

Sequence Retrieval

For the sake of Multiple sequence alignment of nucleotide sequences of the complete cloned fragment was carried out in order to know nucleotide basis and the product amino acid sequence of N-terminal region CP was performed through **Clustal W**. phylogenetic tree was then

generated by **MEGA X**. Method used for this study was **GTR+G+I with bootstrap 100.**.

Percentage sequence analysis:

With the help of excel sheets were created to attain a graph regarding cutoff value against under concerned species. Both segments DNA-A and DNA-B were separately analyzed.

Recombination detection:

RDP 4 was used to detect remarkable recombination among each segment. Each isolate with significant recombination score more than 0.4 and having at least 3 or more detectable programs especially RDP were honored to be noted down in the given tables. Throughout the table there is diversity in viral infections and also deviation from one species to another.

Table. Maximum Likelihood fits of 24 different nucleotide substitution models

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(G
GTR+G+I	307	206207.820	202854.745	-101120.141	0.07	0.94	1.06	0.263	0.298	0.203	0.236	0.073	0.043	0.106	0.065	0.120	0.063	0.056	0.176	0.053	0.119	0.0
TN93+G+I	304	206230.118	202909.804	-101150.676	0.07	0.93	1.11	0.263	0.298	0.203	0.236	0.070	0.048	0.104	0.062	0.126	0.055	0.062	0.184	0.055	0.117	0.0
GTR+G	306	206250.081	202907.926	-101147.734	n/a	0.75	1.06	0.263	0.298	0.203	0.236	0.073	0.043	0.106	0.065	0.121	0.063	0.056	0.176	0.053	0.118	0.0
TN93+G	303	206260.810	202951.417	-101172.484	n/a	0.75	1.06	0.263	0.298	0.203	0.236	0.071	0.049	0.105	0.063	0.121	0.057	0.063	0.176	0.057	0.118	0.0
HKY+G+I	303	206316.242	203006.849	-101200.200	0.07	0.94	1.06	0.263	0.298	0.203	0.236	0.072	0.049	0.122	0.063	0.106	0.057	0.063	0.154	0.057	0.137	0.0
HKY+G	302	206364.001	203065.529	-101230.541	n/a	0.75	1.06	0.263	0.298	0.203	0.236	0.072	0.049	0.122	0.063	0.106	0.057	0.063	0.154	0.057	0.137	0.0
T92+G+I	301	206512.004	203224.453	-101311.004	0.07	0.94	1.05	0.281	0.281	0.219	0.219	0.068	0.053	0.113	0.068	0.113	0.053	0.068	0.145	0.053	0.145	0.0
T92+G	300	206558.255	203281.624	-101340.592	n/a	0.75	1.05	0.281	0.281	0.219	0.219	0.068	0.053	0.113	0.068	0.113	0.053	0.068	0.145	0.053	0.145	0.0
K2+G+I	300	206933.924	203657.293	-101528.426	0.07	0.94	1.11	0.250	0.250	0.250	0.250	0.059	0.059	0.131	0.059	0.131	0.059	0.059	0.131	0.059	0.131	0.0
K2+G	299	206964.485	203698.775	-101550.168	n/a	0.75	1.04	0.250	0.250	0.250	0.250	0.061	0.061	0.128	0.061	0.128	0.061	0.061	0.128	0.061	0.128	0.0
JC+G+I	299	209038.618	205772.907	-102587.235	0.07	0.97	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.0
JC+G	298	209086.166	205831.376	-102617.471	n/a	0.77	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.0
GTR+I	306	215382.265	212040.111	-105713.826	0.14	n/a	0.99	0.263	0.298	0.203	0.236	0.077	0.044	0.105	0.068	0.115	0.066	0.057	0.168	0.055	0.117	0.0
TN93+I	303	215437.184	212127.791	-105760.670	0.14	n/a	1.03	0.263	0.298	0.203	0.236	0.072	0.049	0.104	0.064	0.119	0.057	0.064	0.174	0.057	0.116	0.0
HKY+I	302	215537.084	212238.612	-105817.082	0.14	n/a	0.99	0.263	0.298	0.203	0.236	0.074	0.051	0.118	0.066	0.102	0.059	0.066	0.149	0.059	0.132	0.0
T92+I	300	215674.823	212398.192	-105898.876	0.14	n/a	0.98	0.281	0.281	0.219	0.219	0.070	0.055	0.110	0.070	0.110	0.055	0.070	0.140	0.055	0.140	0.0
K2+I	299	216046.576	212780.866	-106091.214	0.14	n/a	1.03	0.250	0.250	0.250	0.250	0.062	0.062	0.127	0.062	0.127	0.062	0.062	0.127	0.062	0.127	0.0
JC+I	298	218014.716	214759.926	-107081.746	0.14	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.0
GTR	305	220161.640	216830.406	-108109.975	n/a	n/a	0.97	0.263	0.298	0.203	0.236	0.076	0.043	0.102	0.068	0.116	0.068	0.055	0.169	0.056	0.114	0.0
TN93	302	220238.244	216939.771	-108167.662	n/a	n/a	0.97	0.263	0.298	0.203	0.236	0.075	0.051	0.102	0.066	0.115	0.059	0.066	0.168	0.059	0.114	0.0
HKY	301	220299.206	217102.944	-108250.209	n/a	n/a	0.97	0.263	0.298	0.203	0.236	0.075	0.051	0.117	0.066	0.101	0.059	0.066	0.148	0.059	0.121	0.0

Figure: 3.1Maximum likelihood model retrieved by manual configuration for the sequences of DNA-A The model shows that the configuration for tree will be GTR+G+I

Table. Maximum Likelihood fits of 24 different nucleotide substitution models

Model	Parameters	BIC	AICc	*lnL*	(+*I*)	(+*G*)	*R*	*f*(A)	*f*(T)	*f*(C)	*f*(G)	*r*(AT)	*r*(AC)	*r*(AG)	*r*(TA)	*r*(TC)	*r*(TG)	*r*(CA)	*r*(CT)	*r*(CG)	*r*(GA)	*r*(GT)
TN93+G	37	65296.065	64974.375	-32450.155	n/a	0.92	1.12	0.276	0.313	0.194	0.217	0.072	0.045	0.105	0.064	0.115	0.050	0.064	0.185	0.050	0.133	0.072
TN93+G+I	38	65298.134	64967.751	-32445.842	0.05	1.17	1.11	0.276	0.313	0.194	0.217	0.073	0.045	0.105	0.064	0.114	0.050	0.064	0.183	0.050	0.133	0.073
HKY+G	36	65298.414	64985.417	-32456.678	n/a	0.92	1.12	0.276	0.313	0.194	0.217	0.073	0.045	0.116	0.064	0.104	0.050	0.064	0.168	0.050	0.148	0.073
HKY+G+I	37	65299.741	64978.051	-32451.994	0.05	1.18	1.11	0.276	0.313	0.194	0.217	0.073	0.045	0.116	0.064	0.104	0.050	0.064	0.167	0.050	0.147	0.073
T92+G	34	65304.359	65008.747	-32470.347	n/a	0.92	1.11	0.295	0.295	0.205	0.205	0.068	0.048	0.110	0.068	0.110	0.048	0.068	0.158	0.048	0.158	0.068
T92+G+I	35	65305.725	65001.421	-32465.682	0.05	1.18	1.10	0.295	0.295	0.205	0.205	0.069	0.048	0.109	0.069	0.109	0.048	0.069	0.157	0.048	0.157	0.069
GTR+G	40	65323.129	64975.361	-32447.644	n/a	0.91	1.12	0.276	0.313	0.194	0.217	0.073	0.040	0.105	0.064	0.114	0.054	0.057	0.184	0.050	0.134	0.078
GTR+G+I	41	65325.209	64968.749	-32443.336	0.05	1.16	1.11	0.276	0.313	0.194	0.217	0.074	0.041	0.105	0.065	0.114	0.054	0.058	0.183	0.050	0.134	0.078
K2+G+I	34	65673.243	65377.631	-32654.789	0.06	1.29	1.06	0.250	0.250	0.250	0.250	0.061	0.061	0.128	0.061	0.128	0.061	0.061	0.128	0.061	0.128	0.061
K2+G	33	65676.270	65389.351	-32661.650	n/a	0.96	1.07	0.250	0.250	0.250	0.250	0.060	0.060	0.129	0.060	0.129	0.060	0.060	0.129	0.060	0.129	0.060
HKY+I	36	66027.934	65714.937	-32821.438	0.14	n/a	0.96	0.276	0.313	0.194	0.217	0.078	0.049	0.108	0.069	0.097	0.054	0.069	0.156	0.054	0.138	0.078
TN93+I	37	66028.599	65706.909	-32816.422	0.14	n/a	0.96	0.276	0.313	0.194	0.217	0.078	0.049	0.101	0.069	0.103	0.054	0.069	0.166	0.054	0.129	0.078
T92+I	34	66036.033	65740.421	-32836.183	0.14	n/a	0.96	0.295	0.295	0.205	0.205	0.074	0.052	0.102	0.074	0.102	0.052	0.074	0.147	0.052	0.147	0.074
GTR+I	40	66053.547	65705.780	-32812.853	0.14	n/a	0.96	0.276	0.313	0.194	0.217	0.084	0.045	0.101	0.074	0.104	0.055	0.064	0.167	0.052	0.129	0.080
JC+G+I	33	66167.752	65880.833	-32907.391	0.07	1.41	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G	32	66175.076	65896.851	-32916.401	n/a	1.01	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2+I	33	66331.134	66044.215	-32989.082	0.15	n/a	0.94	0.250	0.250	0.250	0.250	0.064	0.064	0.121	0.064	0.121	0.064	0.064	0.121	0.064	0.121	0.064
JC+I	32	66772.808	66494.582	-33215.267	0.14	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
TN93	36	67155.505	66842.507	-33385.224	n/a	n/a	0.92	0.276	0.313	0.194	0.217	0.080	0.050	0.096	0.071	0.103	0.055	0.071	0.166	0.055	0.123	0.080
HKY	35	67164.175	66859.870	-33394.906	n/a	n/a	0.92	0.276	0.313	0.194	0.217	0.080	0.050	0.106	0.071	0.095	0.056	0.071	0.153	0.056	0.135	0.080
T92	33	67172.655	66885.736	-33400.842	n/a	n/a	0.92	0.295	0.295	0.205	0.205	0.076	0.053	0.100	0.076	0.100	0.053	0.076	0.142	0.053	0.142	0.076

Figure 3.2Maximum likelihood model obtained by manual configuration for the sequences of DNA-BThe model shows that the configuration for tree will be TN93+G

Results and Discussion:

Total number of DNA-A sequences infecting in vegetables were 151. 30 Species including extract from phylogenetic tree of maximum likelihood. Species from tree were 15 (Based on 91% identities Brown et al. 2015) Pair wise sequence comparison revealed 86% cutoff value and beyond this point new Species are predicted. (Based on our research) DNA -B consisted on total number of 17 sequences from which 13 Species were included (Based on 69% sequence identities) Pair wise sequence Comparison for DNA-B segment depicted 69% (Based on our research)

As far as results are concerned the total number of species under observation was 30. 151 isolates from all over the globe. The results given above are based on phylogenetic analysis and percentage sequence identities.

According to phylogenetic analysis 15 species from DNA-A with 86% cutoff value which is based on reference of 91% identities Brown et al. 2015. On the other hand 17 sequences from DNA-B with overall 69% cutoff as in reference to our studies.

Mostly begomoviruses don't have DNA-B segments. A total of 151 sequences of begomovirus particularly infecting vegetables were downloaded from the NCBI-GenBankdatabase.The list viruses included merely infecting vegetables were 30 according to the previously accepted 89% cutoff values.!5 identifications from DNA-A segments and 13 components from DNA-B segments affecting vegetables..For pair wise sequence comparison two closest species were cross checked using SDT. Muscle aligned file from clustal W were run through SDT to get excel sheets for percentage analysis of each segment respectively.Mega X was used to predict phylogenetic tree for each segment with manually configured protocol . The Model kep for DNA-A segment was GTR+G+I and on the other hand for DNA-B segment was TN93+I respectively. The bootstrap branching set strictly to 100.

pairwise sequence comparisons, using different cutoff values with rounded to the nearest full percentile 69% and 86% respectively to define potential species so as to determine sequences. The resultant taxonomic framework provided in the description of a small number of outliers in the list. Out of nowhere the results of begomovirus infection and pathogenicity keeps on evolving.To clearly elaborate the issue;it is still a nutshell for scientist.

isolates from a particular species, which might shares\91 % identity with the major number of isolates in thatspecies. Still, isolate having 94% identities can share genetic information with less percentage identities.they have been misunderstood with artificial infections and their biological nature is ignored . This is a misunderstanding.-

biological properties are intercepted with taxonomic classification. Difference in their biological properties can be clearly predicted with them.

For example in Latin America bean crops can be discussed. Two different diseases bean golden mosaic virus and bean golden yellow mosaic virus. First disease is majorly present in Brazil and Argentina whereas later is found in Caribbean and North America. The symptoms are very trolling but the causative vector is same for both which is whitefly. Both of the disease vary in disease symptoms and viral complexes. With the help of this research modern procedures for viral spread and threats prediction are depicted. The outcomes of this research demonstrate that for example if BGMV has 86 % cutoff rate that is the threshold value. Species which fall in higher percentage like 87 and 88% those are newly predicted species. Further this tool was used to elaborate more with the help of recombination program. Each segment was passed through cross sectional recombination scan with the help of RDP 4. Mfft aligned files were run through RDP4 with query of all possible recombination scan. The criteria was kept narrow with each segment having recombination score more than 0.4 and must show more than three to four confirmation scans like RDP Maxchill etc. Each segment with its extracted values were ploted on the table in word file to delineate cross recombination in species. Results reveal that one specie can cause disease in other species of plants too in same lineage. Begomovirus affectin tomato can also diverge to chilli, pepper and bean. This is also significant in probable threat anticipation and vector control. This research can also aid in Integrated vector management system to control the spread of vectors.

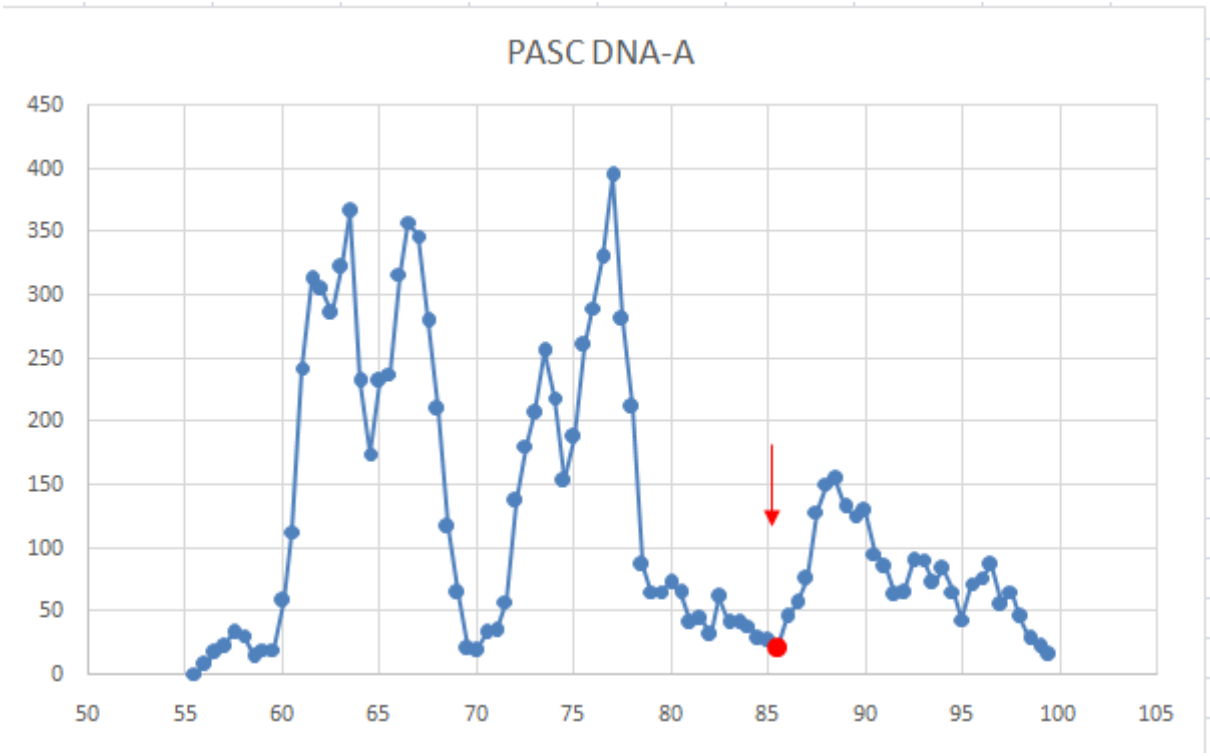

Figure4.1According to this graph 86 is the cut off value and beyond this point new species are predicted as per research. The criteria were followed with reference to (Brown et al, 2015). pairwise sequence comparisons, using different cutoff values which were rounded off to least . they described potential species so as to determine each sequences. The resultant taxonomic structures resulted in the description of a small number of outlying values. 86% is the cutoff value which depicts that beyond this point new species are predicted For example if BGYMV has 86% score it is the threshold. Species next to it starting from 87% are newly predicted.

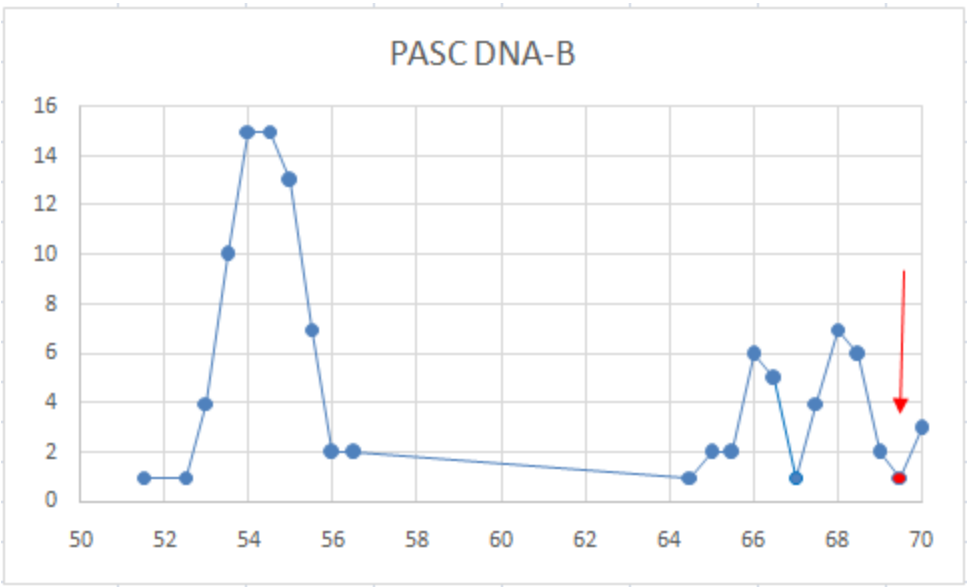

Figure 4.2: According to graph 69% is the cutoff point and beyond this point are new species predicted as per reference to this research. Pair wise sequence comparisons, using different cutoff values which were rounded off to least . they described potential species so as to determine each sequences. The resultant taxonomic structures resulted in the description of a small number of outlying values. 69% is the cutoff value which depicts that beyond this point new species are predicted For example if CabLCJV has 69% score it is the threshold. Species next to it starting from 70% are newly predicted.

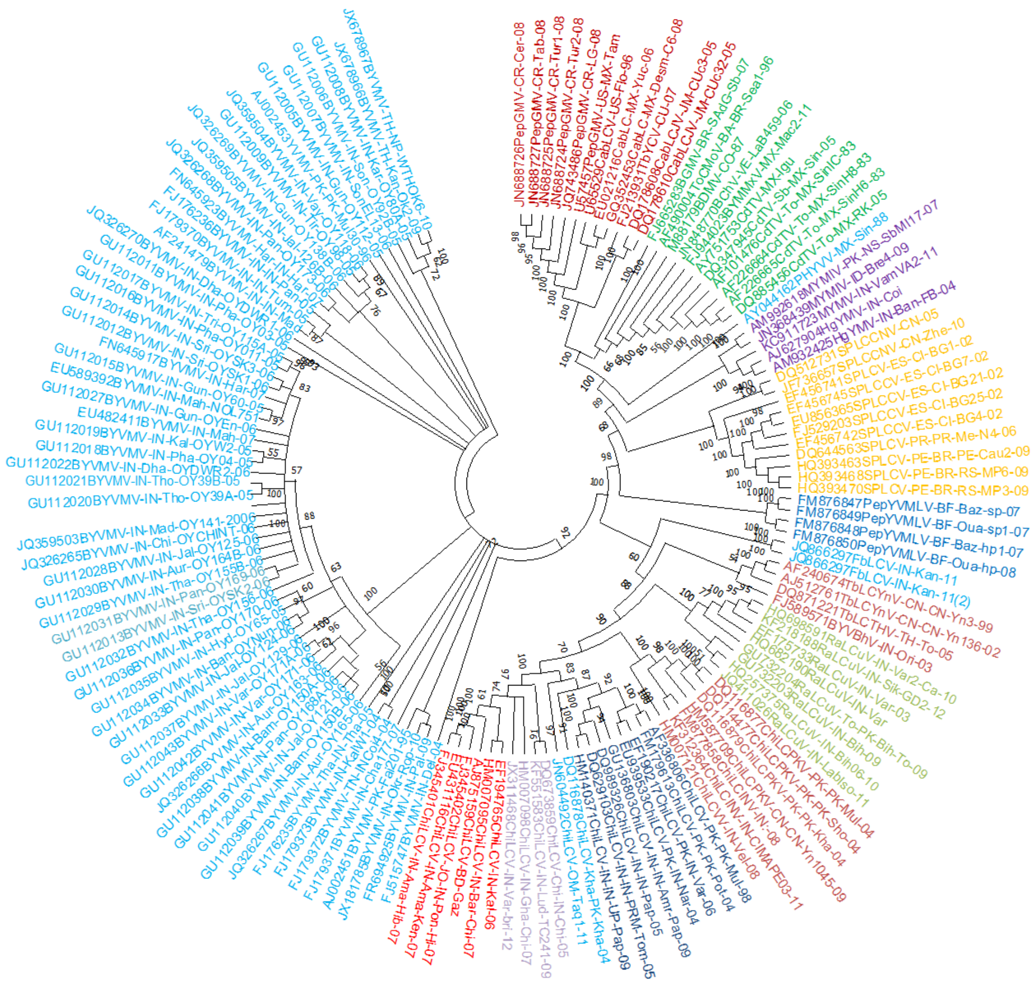

Figure 5.1 Maximum likelihood tree obtained manually configured values in MEGAX. Each color represents a separate class of species . This technique was employed to divide species which gave a total extract of 30 Species accounting DNA-A&B segment.15 species exclusively in DNA-A segments the whole program was manually configured with bootstrap scaled 100. Two closest identities were cross checked in pair wise sequence comparison to check 86% matches

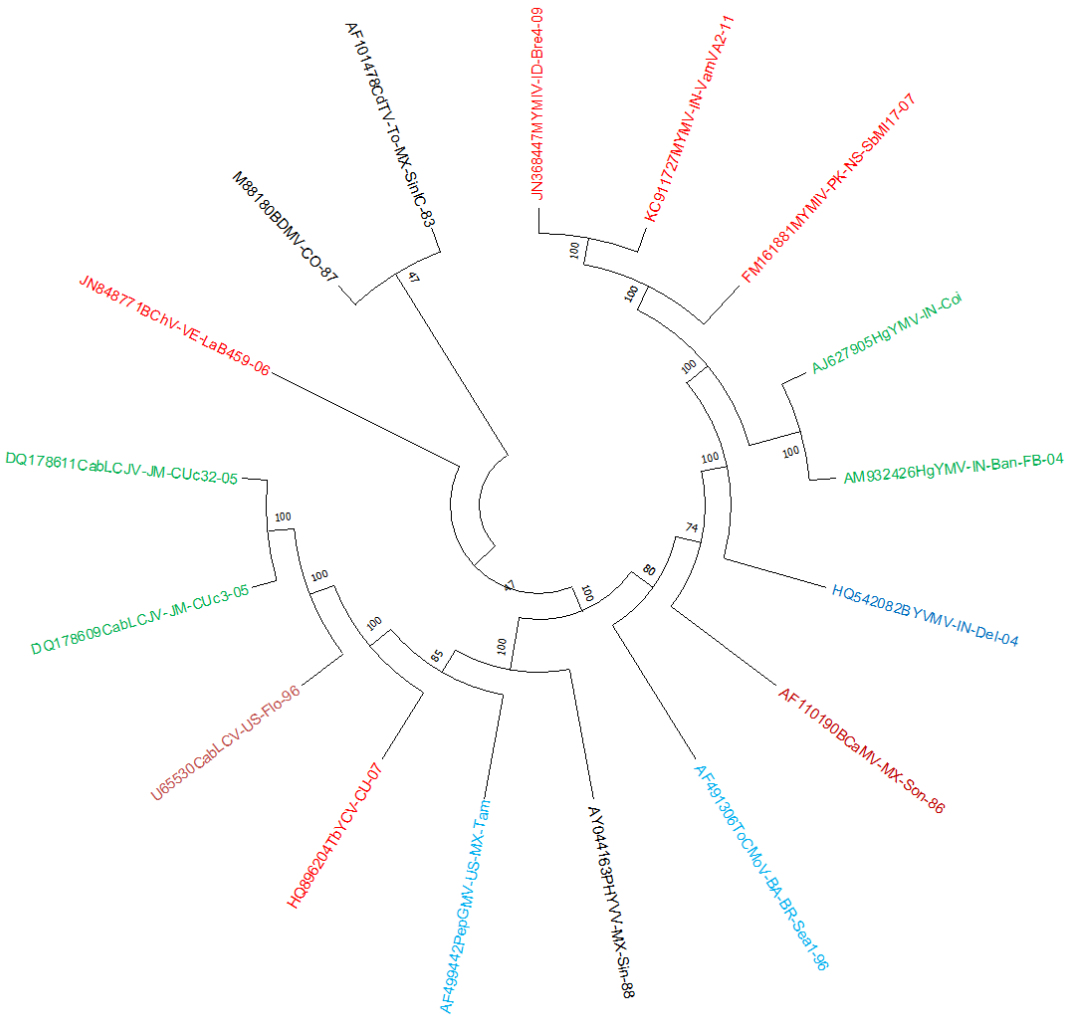

Figure 5.2 Maximum likelihood tree obtained manually configured values in MEGAX. Each color represents a separate class of species . This technique was employed to divide species which gave a total extract of 30 Species accounting DNA-A&B segment.13 species exclusively in DNA-B segments the whole program was manually configured with bootstrap scaled 100. Two closest identities were cross checked in pair wise sequence comparison to check 69% matches

1.1Recombination detection in DNA-A:

Table1.1 shown below have criteria strictly kept to check cross recombination among different species of begomovirus. As shown the table is deployed with recombinant including isolate its taxa, host, recombination score, probable detection programs and break points in columns of left hand side. On right hand side Minor and Major parents of the deviating species are deployed. Each isolate with significant recombination score more than 0.4 and having at least 3 or more detectable programs especially RDP were honored to be noted down in the given tables. Throughout the table there is diversity in viral infections and also deviation from one species to another.

SN	Recombinant			Recombination			Minor Parent			Major Parent		
	Isolate (Acc. No.)	Taxa*	Host	R. score	Detection Programs	Break points (Start-End)	Isolate	Taxa*	Host	Isolate	Taxa*	Host
1.	DQ116879	ChilCPKV	Chili	0.532	RGBMCS	2209-111	AJ271370	ChilCPKV	chili	X14307	ChilCPKV	Chili
2.	KF312364	ChilCPKV	chili	0.590	RBMCS	3468-4571	AF382829	ChilCPKV	chili	U27200	ChilCPKV	Chili
3.	FM877858	ChilCPKV	chili	0.479	RGBM	1489-2129	AF382829	ChilCPKV	Chili	U27200	ChilCPKV	Chili
4.	JQ866297	ChiLCV	Chili	0.572	RGMC	2250-95	JQ866298	ChiLCV	Chili	HJ334766	ChiLCV	Chili
5.	FM876847	ChiLCV	Chili	0.458	RMGP	2294-10	FM87888	ChiLCV	Chili	RF76012	ChiLCV	Chili
6.	FM876847	ChiLCV	Chili	0.588	RMP	2633-2712	JQ45544	ChiLCV	Chili	KF002398	ChiLCV	Chili
7.	GU112036	ChiLCV	Chili	0.622	RGBMC	1784-83	GU112032	ChiLCV	Chili	AJ002451	ChiLCV	Chili
8.	JQ326267	ChiLCV	Chili	0.699	RGMC	1377-87	MN90111	ChiLCV	Chili	JQ66676	ChiLCV	Chili
9.	JQ326267	ChiLCV	Chili	0.594	RGMC	1921-2525	FJ665283	ChiLCV	Chili	DQ44545	ChiLCV	Chili
10.	FR694925	ChiLCV	Chili	0.409	RGBM	155-688	GU112015	ChiLCV	Chili	GU112031	ChiLCV	Chili
11.	FR694925	BYVMV	Bhindi	0.653	RGMP	1582-1661	GU112025	BYVMV	Bhindi	GU114574	BYVMV	Bhindi
12.	FR694925	BYVMV	Bhindi	0.403	RGBM	2130-2720	GU112028	BYVMV	Bhindi	GU222074	BYVMV	Bhindi
13.	FM876848	BYVMV	Bhindi	0.458	RCS	2294-108	JQ326263	BYVMV	Bhindi	GU552074	BYVMV	Bhindi
14.	KF218188	BYVMV	Bhindi	0.524	RBMC	559-1456	JQ326269	BYVMV	mungbean	GU44074	BYVMV	Bhindi
15.	KF218188	BYVMV	Bhindi	0.508	RGMC	2156-2441	JQ326269	BYVMV	Bhindi	GU12374	BYVMV	Bhindi
16.	FM876849	BYVMV	Bhindi	0.458	RCS	2294-108	JQ326269	BYVMV	Bhindi	GU176764	BYVMV	Bhindi
17.	GH334553	BYVMV	Bhindi	0.432	RMP	888-231	SW467886	BYVMV	Bhindi	YD76876	BYVMV	Bhindi
18.	DQ178610	BYVMV	Bhindi	0.576	RGCC	765-98	JQ326270	BYVMV	tomato	GU44074	BYVMV	Pepper
19.	JQ326270	BYVMV	Bhindi	0.666	RMPG	324-987	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	Bhindi
20.	JQ326270	BYVMV	Bhindi	0.432	RMP	76-987	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	Bhindi
21.	JQ326270	BYVMV	Bhindi	0.677	RPGC	765-897	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	Bhindi
22.	JQ326270	BYVMV	Bhindi	0.432	RPGC	457-565	JQ326270	BYVMV	chil	GU44074	BYVMV	Bhindi

23.	JQ326270	BYVMV	Bhindi	0.432	RPGC	657-33	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	Bhindi
24.	JQ326270	BYVMV	Bhindi	0.409	RPGC	757-878	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	tomato
25.	JQ326270	BYVMV	Bhindi	0.477	RPGC	357-5455	JQ326270	BYVMV	Bhindi	GU44074	BYVMV	Bhindi
26.	HM357461	CdTAV	Tomato	0.407	RPGC	43-544	HM357461	CdTAV	Tomato	HM35746	CdTAV	Tomato
27.	HM357461	CdTAV	Tomato	0.532	RPGC	333-323	HM357461	CdTAV	Tomato	HM357461	CdTAV	Tomato
28.	HM357461	CdTAV	Tomato	0.434	RPBC	665-323	HM357461	CdTAV	chilli	HM357461	CdTAV	bhindi
29.	HM357461	CdTAV	Tomato	0.4343	RPGC	276-111	HM357461	CdTAV	Tomato	HM357461	CdTAV	Tomato
30.	HM357461	CdTAV	Tomato	0.502	RPBC	632-222	HM357461	CdTAV	pepper	HM357461	CdTAV	Tomato
31.	HM357461	CdTAV	Tomato	0.400	RCC	765-444	HM357461	CdTAV	Tomato	HM357461	CdTAV	Tomato
32.	HM357461	CdTAV	Tomato	0.599	RGGC	177-433	HM357461	CdTAV	chilli	HM357461	CdTAV	Tomato
33.	HM357461	CdTAV	Tomato	0.455	RMGC	4-222	HM357461	CdTAV	Tomato	HM357461	CdTAV	Tomato
34.	HM357461	CdTAV	Tomato	0.436	RPC	567-888	HM357461	CdTAV	Tomato	HM357461	CdTAV	pepper
35.	HM357461	CdTAV	Tomato	0.732	RBMGC	177-464	HM357461	CdTAV	Tomato	HM357461	CdTAV	Tomato
36.	AM932425	HgYMV	Horsegram	0.634	RPGC	057-464	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
37.	AM932425	HgYMV	Horsegram	0.431	RPC	355-222	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
38.	AM932425	HgYMV	Horsegram	0.435	RPBC	876-5756	AM932425	HgYMV	Horsegram	AM932425	HgYMV	mungbean
39.	AM932425	HgYMV	Horsegram	0.466	RPBC	323-777	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
40.	AM932425	HgYMV	Horsegram	0.466	RCB	454-876	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
41.	AM932425	HgYMV	Horsegram	0.433	RPGC	222-545	AM932425	HgYMV	chilli	AM932425	HgYMV	Horsegram
42.	AM932425	HgYMV	Horsegram	0.454	RBGC	733-544	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
43.	AM932425	HgYMV	Horsegram	0.567	RMCC	943-555	AM932425	HgYMV	Horsegram	AM932425	HgYMV	Horsegram
44.	AF416742	MYMIV	Mungbean	0.543	RPGGC	633-544	AF416742	MYMIV	Mungbean	AF416742	MYMIV	Mungbean
45.	AF416742	MYMIV	Mungbean	0.433	RPBG	743-43	AF416742	MYMIV	Mungbean	AF416742	MYMIV	Mungbean
46.	AF416742	MYMIV	Mungbean	0.455	RPGC	255-322	AF416742	MYMIV	Mungbean	AF416742	MYMIV	pepper
47.	AF416742	MYMIV	Mungbean	0.767	RPBG	332-121	AF416742	MYMIV	chilli	AF416742	MYMIV	Mungbean
48.	AF416742	MYMIV	Mungbean	0.432	RSC	444-222	AF416742	MYMIV	Mungbean	AF416742	MYMIV	Mungbean
49.	AJ416349	MYMIV	Mungbean	0.476	RPGC	522-5454	AJ416349	MYMIV	Mungbean	AJ416349	MYMIV	bhindi
50.	AJ416349	MYMIV	Mungbean	0.432	RBPC	844-987	AJ416349	MYMIV	Mungbean	AJ416349	MYMIV	Mungbean
51.	AJ416349	MYMIV	Mungbean	0.532	RGCC	845-777	AJ416349	MYMIV	Mungbean	AJ416349	MYMIV	Mungbean
52.	AJ416349	MYMIV	Mungbean	0.467	RPGC	055-23	AJ416349	MYMIV	Mungbean	AJ416349	MYMIV	Mungbean
53.	AJ416349	MYMIV	Mungbean	0.787	RPBC	365-877	AJ416349	MYMIV	tomato	AJ416349	MYMIV	Mungbean
54.	AJ416349	MYMIV	Mungbean	0.432	RPGC	542-222	AJ416349	MYMIV	Mungbean	AJ416349	MYMIV	Mungbean
55.	AJ416349	MYMIV	Mungbean	0.476	RPGC	7347-756	AJ416349	MYMIV	tomato	AJ416349	MYMIV	Mungbean
56.	AJ416349	MYMIV	Mungbean	0.455	RPGC	6457-555	AY049772	MYMIV	Mungbean	AY049772	MYMIV	pepper
57.	AY049772	MYMIV	Mungbean	0.466	RPGC	65-43	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
58.	AY049772	MYMIV	Mungbean	0.443	RC	655-22	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
59.	AY049772	MYMIV	Mungbean	0.644	RPGC	355-666	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
60.	AY049772	MYMIV	Mungbean	0.556	RPGC	466-343	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
61.	AY049772	MYMIV	Mungbean	0.401	RPC	532-87	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
62.	AY049772	MYMIV	Mungbean	0.465	RPG	656-444	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
63.	AY049772	MYMIV	Mungbean	0.400	RPBC	133-877	AY049772	MYMIV	chilli	AY049772	MYMIV	Mungbean
64.	AY049772	MYMIV	Mungbean	0.476	RBC	555-21	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
65.	AY049772	MYMIV	Mungbean	0.743	RPBBC	57-634	AY049772	MYMIV	Mungbean	AY049772	MYMIV	bhindi
66.	AY049772	MYMIV	Mungbean	0.432	RPGB	6666-737	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
67.	AY049772	MYMIV	Mungbean	0.432	RPBC	756-939	AY049772	MYMIV	Mungbean	AY049772	MYMIV	Mungbean
68.	KF156759	SPLCSiV	Sweet potato	0.487	RPGB	732-233	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
69.	KF156759	SPLCSiV	Sweet potato	0.543	RPG	822-211	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
70.	KF156759	SPLCSiV	Sweet potato	0.656	RPGC	933-434	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	pepper

71.	KF156759	SPLCSiV	Sweet potato	0.554	RPGC	421-433	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
72.	KF156759	SPLCSiV	Sweet potato	0.466	RPGC	733-443	KF156759	SPLCSiV	chilli	KF156759	SPLCSiV	Sweet potato
73.	KF156759	SPLCSiV	Sweet potato	0.443	RPGC	84-111	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
74.	KF156759	SPLCSiV	Sweet potato	0.408	RPGC	944-111	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
75.	KF156759	SPLCSiV	Sweet potato	0.465	RPGC	344-32	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
76.	KF156759	SPLCSiV	Sweet potato	0.443	RPGC	155-433	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
77.	KF156759	SPLCSiV	Sweet potato	0.433	RPGC	254-29	KF156759	SPLCSiV	Sweet potato	KF156759	SPLCSiV	Sweet potato
78.	U38239	ToLCKaV	Tomato	0.421	RPGC	666-72	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
79.	U38239	ToLCKaV	Tomato	0.411	RPGC	222-82	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
80.	U38239	ToLCKaV	Tomato	0.431	RPGC	876-43	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
81.	U38239	ToLCKaV	Tomato	0.634	RBSC	987-8676	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
82.	U38239	ToLCKaV	Tomato	0.543	RPCBS	112-876	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
83.	U38239	ToLCKaV	Tomato	0.609	RPBC	432-43	U38239	ToLCKaV	Tomato	U38239	ToLCKaV	Tomato
84.	FJ660431	ToLCPalV	Tomato	0.888	RPGC	798-433	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
85.	FJ660431	ToLCPalV	Tomato	0.555	RPBG	965-43	FJ660431	ToLCPalV	chilli	FJ660431	ToLCPalV	Tomato
86.	FJ660431	ToLCPalV	Tomato	0.654	RPBD	157-232	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
87.	FJ660431	ToLCPalV	Tomato	0.454	RSB	576-65	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
88.	FJ660431	ToLCPalV	Tomato	0.454	RBSSC	298-655	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
89.	FJ660431	ToLCPalV	Tomato	0.466	RPSSB	121-432	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Mungbean
90.	FJ660431	ToLCPalV	Tomato	0.476	RPGSC	15-876	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
91.	FJ660431	ToLCPalV	Tomato	0.443	RPSG	667-965	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
92.	FJ660431	ToLCPalV	Tomato	0.443	RBSG	677-868	FJ660431	ToLCPalV	pepper	FJ660431	ToLCPalV	Tomato
93.	FJ660431	ToLCPalV	Tomato	0.454	RPGB	243-545	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
94.	FJ660431	ToLCPalV	Tomato	0.554	RBCC	365-323	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
95.	FJ660431	ToLCPalV	Tomato	0.466	RMBSC	255-232	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
96.	FJ660431	ToLCPalV	Tomato	0.655	RPBC	35-9876	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Chilli
97.	FJ660431	ToLCPalV	Tomato	0.454	RG	232-656	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
98.	FJ660431	ToLCPalV	Tomato	0.644	RGC	153-55	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
99.	FJ660431	ToLCPalV	Tomato	0.411	RPGC	621-867	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
100.	FJ660431	ToLCPalV	Tomato	0.543	RGC	887-7676	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato
10	FJ660431	ToLCPalV	Tomato	0.443	RPGC	232-544	FJ660431	ToLCPalV	Tomato	FJ660431	ToLCPalV	Tomato

1.2 Recombination detection in DNA-B:

Resultant table1.2 shows fewer values as described above in introduction due to less associated DNA-B segments

SN	Recombinant			Recombination			Minor Parent			Major Parent		
	Isolate (Acc. No.)	**Taxa***	**Host**	**R. score**	**Detection Programs**	**Break points** (Start-End)	**Isolate**	**Taxa***	**Host**	**Isolate**	**Taxa***	**Host**
1.	KC911727	MYMV	Mungbean	0.523	RGP	033-2322	FM161881	MYMV	mungbean	JN368447	MYMV	mungbean
2.	AF499442	PepGMV	pepper	0.469	RGPS	987-4344	Unknown	CdTV	mungbean	JN368447	MYMIV	mungbean
3.	AY044163	PHYVV	pepper	0.521	RBCS	76-433	AF101478	Cdtv	pepper	KC911727	MYMV	mungbean

Conclusions:

Pairwise sequence comparison (PASC) analysis showed the DNA-A Species cutoff as 86% & DNA-B as 69% nucleotide sequence identities. RDP analysis showed frequent interspecies recombination in DNA-A.

References

Brown, J. K., Idris, A. M., Alteri, C., andStenger, D. C. (2002).Emergence of a new cucurbit-infecting *Begomovirus* species capable of forming viable reassortants with related viruses in the *Squash leaf curl virus* cluster. *Phytopathology*, *92*(7), 734-742.

Bracero, V., Rivera, L. I., and Beaver, J. S. (2003). DNA analysis confirms macroptiliumlathyroides as alternative host of *Bean golden yellow mosaic virus*. *Plant disease*, *87*(9), 1022-1025.

Brown, J. K. (2007). The Bemisiatabaci complex: genetic and phenotypic variability drives *Begomovirus* spread and virus diversification. *Plant Dis*, *1*, 25-56.

Carr, R. J., and Kim, K. S. (1983).Evidence that *Bean golden mosaic virus* invades non-phloem tissue in double infections with *Tobacco mosaic virus*. *Journal of general virology*, *64*(11), 2489-2492.

Fontes, E. P., Gladfelter, H. J., Schaffer, R. L., Petty, I. T., and Hanley-Bowdoin, L. (1994).*Geminivirus* replication origins have a modular organization. *The Plant Cell*, *6*(3), 405-416.

Goodman, R. M. (1981). *Geminiviruses*. *Journal of General Virology*, *54*(1), 9-21.

Moriones, E., andNavas-Castillo, J. (2008). Rapid evolution of the population *of Begomoviruses* associated with the tomato yellow leaf curl disease after invasion of a new ecological niche: a review. *Spanish Journal of Agricultural Research*, *6*(S1), 147-159.

Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S., andFauquet, C. M. (2001). Recombination, pseudorecombination and synergism of *Geminiviruses* are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology*, *82*(3), 655-665.

Rodríguez-Pardina, P. E., Zerbini, F. M., andDucasse, D. A. (2006).Genetic diversity of *Begomoviruses* infecting soybean, bean and associated weeds in northwestern Argentina. *FitopatologiaBrasileira*, *31*(4), 342-348.

Rojas, M. R., Gilbertson, R. L., and Maxwell, D. P. (1993).Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted *Geminiviruses*. *Plant disease*, *77*(4), 340-347.

Sattar, M. N. (2012). Diversity and interactions of *Begomoviruses* and their associated DNA-satellites. (Vol. 2012, No. 23).

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., andVarsani, A. (2017). ICTV Virus Taxonomy Profile: Geminiviridae. *Journal of General Virology*, *98*(2), 131-133.

Brown, J. K., Zerbini, F. M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C., ... & Malathi, V. G. (2015). Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of virology*, *160*(6), 1593-1619.

Muhire, B. M., Varsani, A., & Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PloS one*, *9*(9), e108277.

Zhou, X. (2013). Advances in understanding begomovirus satellites. *Annual review of phytopathology*, *51*.

Varsani, A., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Brown, J. K., ... & Martin, D. P. (2014). Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. *Archives of virology*, *159*(8), 2193-2203.

Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annual review of phytopathology*, *49*, 219-248.

Sattar, M. N., Kvarnheden, A., Saeed, M., & Briddon, R. W. (2013). Cotton leaf curl disease–an emerging threat to cotton production worldwide. *Journal of General Virology*, *94*(4), 695-710.

Melgarejo, T. A., Kon, T., Rojas, M. R., Paz-Carrasco, L., Zerbini, F. M., & Gilbertson, R. L. (2013). Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. *Journal of Virology*, *87*(10), 5397-5413.

Rocha, C. S., Castillo-Urquiza, G. P., Lima, A. T., Silva, F. N., Xavier, C. A., Hora-Júnior, B. T., ... & Alfenas-Zerbini, P. (2013). Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. *Journal of Virology*, *87*(10), 5784-5799.

Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., ... & Varsani, A. (2013). A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of virology*, *158*(6), 1411-1424.

Lefeuvre, P., Martin, D. P., Harkins, G., Lemey, P., Gray, A. J., Meredith, S., ... & Heydarnejad, J. (2010). The spread of tomato yellow leaf curl virus from the Middle East to the world. *PLoS Pathog*, *6*(10), e1001164.

Valadez-Moctezuma, E., Samah, S., Zelaya-Molina, L. X., & Díaz-Rivera, J. B. (2020). Molecular characterization of a new bipartite begomovirus that infects okra plants in guerrero, mexico. *Journal of Plant Diseases and Protection*, *127*(6), 753-762.

Duffy, S., & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. *Journal of virology*, *82*(2), 957-965.

Ramesh, S. V., Chouhan, B. S., Gupta, G. K., Husain, S. M., & Chand, S. (2017). Genomic sequence characterization of Begomovirus infecting soybean and molecular evolutionary genomics of Legume yellow mosaic viruses (LYMVs). *Plant Omics*, *10*(2), 88.

Claverie, S., Bernardo, P., Kraberger, S., Hartnady, P., Lefeuvre, P., Lett, J. M., ... & Martin, D. P. (2018). From spatial metagenomics to molecular characterization of plant viruses: A geminivirus case study. In *Advances in virus research* (Vol. 101, pp. 55-83). Academic Press.