Review of Literature
Being primary producers, plants occupy foremost position in the food chain. Directly or indirectly all other living organisms depend upon these “primary producers” for their daily needs. Such plants, growing in a normal environment, continuously fight for their existence with biotic and abiotic factors. Among these, biotic factors severely injure the plants and affect qualitative and quantitative production of plant products up to 31 – 42%. Such biological agents that cause diseases in plants are very similar to that of animals, and range from small particulate like viruses, bacteria, fungi, protozoa and even nematodes. Among these, viruses threaten and contribute majorly to crop loss. Such losses due to viral diseases impact heavily on crop production and are one of the major thrust for the detailed study of plant viruses. Crop loss due to virus infection on plants is estimated 47%, which is greater than the loss encountered by fungi, bacteria, phytoplasma or nematodes put together (Anderson PK. 2004). During past two decades, losses due to geminivirus infection are US$ 5 billion for cotton in Pakistan by cotton leaf curl virus (Briddon RW and Markham PG, 2001) in 1992 – 97, US$ 2.3 billion for cassava by African cassava mosaic virus (Thresh JM and Cooter RJ, 2005), US$ 300 million for grain legumes in India (Varma A and Malathi VG, 2003) and US$ 140 million in Florida, USA for Tomato (Moffat AS, 1999).

Chemically being nucleoprotein in nature, viruses’ genome code for utmost required protein and very smartly modifies host cellular machineries for their efficient replication, movement and transmission. Small genome size and ability to multiply within host cell have made virus particles extremely dynamic and diverse. On the other hand, possessing a small genome and restricted host range, make virus particles, model organisms to understand the concepts and principles of molecular biology.

2.1 History and importance of plant viruses

Historically, plant viruses were reported much later in written form then human or animal viruses. The earliest record of viruses like symptoms was found in writings of Japanese poet, Empress Koken, in 752 BC. The first known virus to infect plant, then named as, was Tobacco leaf curl virus. In 16th century, Western Europe experienced “Tulipomania”, in which, tulip blooms featuring striped patterns were prized highly and caused hyperinflation, due to extensive demand. In 1886, Adolf Mayer gave term “Mosaikkrankheit” for mosaic symptoms observed on
tobacco plant. Followed in 1892, Ivanowski proved that sap retained infectivity even after filtration, and hence infectious agent is subparticulate. In 1894, Hashimoto, for the first time provided evidence of vector transmission of Rice dwarf virus by leafhopper. Later on, it was found that such stripes are result of virus infection. Then after, Beijerinck (1898), finally proved the cause of tobacco mosaic disease, and gave term “*contagium vivum fluidum*”; virus. For his stone breaking findings, Beijerinck is considered as the father of plant virology. Much later in 1970, Harrison and group classified plant viruses into 16 genera and introduced term “Geminiviruses” and proved that they transmit by Whitefly and have ssDNA.

Geminiviruses derived their name from unique structure and geometry of virus particles that look like small balls stuck together. They induce diseases in food, fiber and ornamental crops and reduce the yields, sometimes up to 100%, depending upon severity of infection (Moffat AS, 1999; Fargette D *et. al.* 2006). In the decade of 1980s alone, they have caused tremendous loss in agricultural crops worldwide. They mutate rapidly, attack phloem system and thus systemically infect the plants, quickly adapting to diverse host range (Bedford ID *et. al.*, 1994; Bird J and Maramorosch K, 1978). The overall importance of WTGs is obvious as severe crop losses are continued to be reported from many parts of the world, but it is still very difficult to assess the relative and efficient measures to lessen or stop such infections (Oliveira MRV *et. al.*, 2001; Polston JE *et. al.*, 1997).

Yield constraint, tremendous losses and economical outbreaks leads researchers and policy makers gain interest in Geminiviruses. Sugar beet infection by beet curly top virus, cassava infection by African cassava mosaic virus, cotton infection by cotton leaf curl virus, bean golden mosaic virus of common bean, maize infection by maize streak virus and finally tomato infection by tomato leaf curl virus are past pandemics that causes huge loss in production of respective crops, one or the other parts of the world (Legg JP and Fauquet CM, 2004; Morales FJ and Anderson PK, 2001).

Small genome size and ability to multiply within host cell have made virus particles extremely dynamic and diverse. On the other hand, possessing a small genome and diverse host range, make virus particles, model organisms in understanding the concepts and principles of molecular biology.
Due to small genome size and ability to integrate with the host genome, geminiviruses are emerging as the potential vector system for delivery of foreign genes into the plant cells. Further, total dependence of these viruses on host cellular machineries for their efficient replication, attracts researchers to use them as model system for understanding the molecular interaction of virus infection in the host cells, including replication, cell cycle control and RNA silencing. Such findings, not only clears the understanding of virus replication into host cells, but also clears their evolution, lineage and adaptation for their efficient spread through vector into the host.

2.2 Structure of geminivirus virion particle

Structurally, plant viruses resemble the other viruses. Viral nucleic acid is encapsidated in closed shell or tube like structure, made of protein, termed as capsid. Geminiviruses constitute an important group of plant pathogens with genomes of ssDNA. Geminiviruses derived their name from unique structure and geometry of virus particles that look like “twin moon”, small balls are stuck together (Lazarowitz SG, 1987). They are non-enveloped, icosahedral virions that consist of a geminate capsid, made up of 22 capsomers each. The capsid is 30 nm long and has a diameter of 18-20 nm. A single molecule of covalently closed circular Single stranded viral sense DNA is encapsidated in the each paired particle (Ikegami M et. al., 1981). The ssDNA is often associated with host histone proteins during replication via rolling circle mechanism in the host nucleus to form geminiviral minichromosome. Transmission electron microscopic studies showed large number of virus particles and viral assembly, suggesting that nucleus also may be the site of virus assembly.

![Figure 2.1: Structure and composition of Geminivirus virion particle. (A) Electron micrograph of Bean dwarf mosaic virus (Adapted from Levy A. and Tzfira T. (2010) (B) Picture shows the]
geometrical model of the viron particle. One virion can accommodate only one DNA i.e. either DNA A or DNA B.

2.3 Genome Organisation

Geminiviruses genome is comprised of a closed circular ssDNA of 2.6 kb to 3.0 kb size (Goodman RM, 1977). Intergenic region (IR, common region CR, large intergenic region LIR) contains bidirectional promoters for RNA polymerase binding. (Fontes EP et. al., 1994; Frischmuth S et. al., 1991; Hanley-Bowdoin L et. al., 1990; Lazarowitz SG et. al., 1992; Morris-Krsinich BA et. al., 1985; Petty ITD et. al., 1988; Shivaprasad PV et. al., 2005; Sunter G and Bisaro DM, 1989; Sunter G et. al., 1993; Townsend R et. al., 1985). Intergenic region also houses signature nonanucleotide hairpin loop structure that is recognition and initiation site for viral DNA replication. (Heyraud-Nitschke F et. al., 1993a, b). The ORFs are bidirectional and overlapped. The termination codons of the ORFs are on respective direction and are independent for shared ORFs. (Accotto GP et. al., 1989; Frischmuth S et. al., 1991; Mullineaux PM et. al., 1993; Petty ITD et. al., 1988; Shivaprasad PV et. al., 2005; Townsend R et. al., 1985). Genes responsible for molecular events, viz. replication (Rep, REn), Transcription (TrAP) etc. house on right side; while genes for virion particle packing and movement are on left side with respect to the origin of the viral genome.

Further, Geminiviruses may be divided based on the number of genomic molecule/s they carry and require for efficient infection. Monopartite Geminiviruses carry one genomic component, termed as DNA A. While bipartite geminivirus possess two genomic components, DNA A and DNA B. In case of bipartite genome, both of genome components are essential for efficient disease transmission and systemic infection (Evans D and Jeske H, 1993). DNA A component encodes for major proteins for virus replication and multiplication inside the host cell while DNA B cares for intra- and intercellular movement of virus particles (Stanely J, 1983; von Arnim A and Stanely J, 1992). The movement protein genes of viruses with monopartite genomes are encoded in the single genome component.

Some of the monopartite Geminiviruses are also associated with additional circular ssDNA molecules, such as beta satellite or alpha satellite which are nearly half the size of DNA A. These
virus/satellite complexes are widespread, rapidly recombine and diverse which have wider host range. Beta satellites have been involved in pathogenicity by enhanced helper virus accumulation in host but alpha satellites have no known function and are certainly not involved in symptom induction (Mansoor S et. al., 1999). The process through which betasatellite regulates pathogenesis is unknown, but it might affect the replication of its helper virus by either facilitating its spread in host plants, or by suppressing host gene silencing (Saunders K et. al., 2000).

Despite of the smaller genome size, viruses are able to encode more proteins due to a) ability to transcribe and translate overlapping ORFs b) Availability of different ORFs on viral as well as complementary strand. The later ability of virus’s genome leads the researcher to provide an efficient nomenclature system for each viral ORFs. Two systems of nomenclature have been proposed by genes and gene products by number and have prefix of the DNA type (A in case of DNA A, B in case of DNA B etc.). One nomenclature system uses the word “V” or “C” to denote ORF of virus or complementary sense origin respectively. Other system uses the basis of orientation and location of ORFs on the viral map, as “R” (Right) for viral sense, clockwise and “L” (Left) for complementary sense, counter clockwise. Throughout this study, “V” and “C” nomenclature will be used for understanding.

2.4 Taxonomy of viruses and Nomenclature

Mutation, Recombination, Pseudorecombination, Synergism, Reassortment and Transcomplementation are common features of Geminiviruses (Garcia – Arenal F et. al., 2001). These viruses often adopt such “molecular arrangements” and are major source of genetic variation in plant viruses. During multiple infections in a single host, viruses can exchange or rearrange their genetic material through these mechanism/s which result in diverse host range (Rajeshwari R et. al., 2005). As increasing numbers of viral sequences become available, recombinant viruses are recognized to be frequent in nature and clear evidence is found for recombination to play a key role in virus evolution (Moonan F et. al., 2000; Padidam M et. al., 1999). Many of such mechanisms are traced back to the severe epidemic of plant diseases (Pita JS et. al., 2001). These molecular events are basis of enormous molecular diversity of plant viruses. Due to diverse host range and diversity in symptoms produced by same group of viruses
onto different plants, classification of plant viruses is a very thoughtful process (Fauquet CM and Stanely J, 2003).

*Geminiviridae* family is the second largest family, after *Potyviridae*, among plant viruses. Currently, International Committee on Taxonomy of Viruses (ICTV) is an apex body for creating a universal taxonomic system for all the viruses infecting animals (vertebrates, invertebrates and protozoa), plants (higher plants and algae), fungi, bacteria and Archaea. Currently, ICTV database (ICTVdb) also maintains primary metadata of individual viruses (Gibbs AJ, 2013)

Nature of genome (DNA or RNA), type of genome (Single stranded or double stranded; Sense or anti sense), envelope etc. are major criteria for classification. According to 9th ICTV report, approximately ~2284 species of viruses are divided into 349 genera, 87 families 19 subfamilies and 6 orders. For *Geminiviridae*, viral species with less than 89% sequence identity to previously reported species is considered as new species, while 89 - 93% sequence similarity as different strain of same species. Greater than 93% sequence similarity is considered as variant of the same species. For betasatellite less than 76% nucleotide sequence identity can be consider as new species (Brown JK et. al., 2012). Presently, plant viruses are divided in 20 families and are sub grouped into 90 genera, which include ~800 plant virus species. Plant viruses have a range between 17 nm and 2000 nm size with isometric or helical symmetry.

### 2.5 Family Geminiviridae

Geminiviruses are characterized by circular single stranded DNA (ssDNA) genomes encapsidated in twinned quasi isometric particles of about 18 x 30 nm (Goodman RM, 1977; Howarth AJ and Vandemark GJ, 1989) The *Geminiviridae* family has been divided on basis of genome organization, host range and insect vector (Fauquet CM et. al., 2008; Padidam M et. al., 1995) into seven genera: *Mastrevirus, Curtovirus, Topocuvirus, Begomovirus, Becurtovirus, Eragrovirus* and *Turncurtovirus* (Fauquet CM et. al., 2008; Brown JK et. al., 2012; Adams MJ et. al., 2013).
Subgroup I: Genus *Mastrevirus* (Type species: Maize streak virus, MSV)

Subgroup I contains leafhopper transmitted monopartite viruses infects monocots. Well-characterized subgroup I pathogens include maize streak virus (MSV) and wheat dwarf virus (WDV). Two members of this genus, TYDV (Tomato yellow dwarf virus) and BeYDV (Bean yellow dwarf virus), also infect dicotyledonous species (Krabberger S *et. al.*, 2012) Infections are largely are confined to members of *Poaceae* family (Shepherd DN *et. al.*, 2010). This genus currently comprises 29 distinct virus species.

Subgroup II: Genus *Curtovirus* (Type species: Beet Curly Top Virus, BCTV)

Genus curtovirus has only one well-characterized virus, beet curly top virus (BCTV). Curtovirus infects dicotyledonous hosts and, like subgroup I, is transmitted persistently by a leafhopper but it differs significantly in its genome organization and is mostly monopartite. This genus currently comprises 4 distinct virus species (Baliji S *et. al.*, 2004).

Subgroup III: Genus *Begomovirus* (Type species: Bean Golden Mosaic Virus, BGMV)

Subgroup III is the largest genera of geminiviruses, includes white fly transmitted viruses with dicotyledons as their host range (Mansoor S *et. al.*, 2003). Currently, this group comprises of 288 distinct virus species. The members of this subgroup contain monopartite (~ 2.7 kb), monopartite with some subviral satellite molecule (~ 1.3 kb) or bipartite genome component (Figure 2.2). Begomoviruses are credited with most complex genome organization among the Geminiviruses. The genome majorly is bipartite with two ss circular DNA, of 2600 nucleotides each, is encapsidated independently and designated as DNA A and DNA B (Davis JW, 1987; Harrison BD *et. al.*, 1977). However, in few Begomoviruses, the genome is monopartite and satellite DNA, named betasatellite, is associated with it (Briddon RW *et. al.*, 2008, 2003; Navot N *et. al.*, 1991). In case of bipartite genome, both genome components are essential for efficient disease transmission and systemic infection (Evans D and Jeske H, 1993). DNA A component encodes for major proteins for virus replication and multiplication inside the host cell while DNA B cares for intra- and intercellular movement of virus particles (Stanely J, 1983; von Arnim A and Stanely J, 1992). The movement protein genes of viruses with monopartite genomes belonging to subgroups I, II and III are encoded on the single genome component.
Figure 2.2: Begomovirus genome organization and associated satellite molecules
Subgroup IV: Genus *Topocuvirus* (Type species: Tomato Pseudo-Curly Top Virus, TPCTV)

This genus Topocuvirus was split off from curtovirus (Pringle CR, 1999). They are transmitted by tree hopper, *Mirculatis malleifera*. Tomato Pseudo-Curly Top Virus, TPCTV, is a sole member of this genus. They are found in the South eastern United States and cause diseases in dicotyledonous plant. Topocuvirus encodes six proteins.

Subgroup V: Genus *Becurtovirus* (Type species: Beet Curly Top Iran Virus, BCTIV)

This genus was derived from the type species *Beet curly top virus* and contains two species Beet curly top Iran virus (BCTIV) and Spinach curly top Arizona virus (SpCTAV). Its biological properties resemble that of curtovirus (Varsani A et. al, 2014). Becurtovirus is monopartite, comprises closed circular ssDNA genome about 2850 nt in size. It comprises five ORFs. Three ORFs are on virion sense and codes for movement protein (MP), a capsid protein (CP) and ssDNA/dsDNA regulatory protein. Rests two are on complementary sense strands which encode replication associated protein (RepA) and a spliced transcript which encodes a full length Rep protein. IR contains nonanucleotide sequences TAAGATT/CC (Heydarnejad J et. al, 2013).

Subgroup VI: Genus *Eragrovirus* (Type species: Eragrostis curvularia streak virus, ECSV)

This genus is recently proposed and accepted in 9th ICTV taxonomy reports. It carries only one type species as its member. Biological properties of Eragrovirus resemble Mastrevirus (Varsani A et. al, 2014). It has four ORFs out of which two are on virion sense that encode movement protein (MP) and coat protein (CP). Other two on complementary sense strand rep protein (Rep) and transcription activator protein (TrAP) IR possesses nonanucleotide sequence TAAGATT/CC (Varsani A et. al., 2009).

Subgroup VII: Genus *Turncurtovirus* (Type species: Turnip curly top virus, TCTV)

This genus was also proposed in 9th ICTV taxonomy reports and carries type species *Turnip curly top virus* as sole member (Varsani A et. al, 2014). Genome codes for six ORF, four on complementary sense strand for replication associated protein (Rep), transcription activator protein (TrAP), replication enhancer protein (REn) and C4 protein. Two ORFs encodes
movement protein (MP) and coat protein (CP) are on virion sense strand (Razavinejad S et. al., 2013).

2.6 Gene functions

Gene functions in Geminiviridae family underwent positive selection in evolutionary history (Bradeen, JM et. al., 1997). Organization and gene functions in Geminiviruses are highly conserved. Following section describes gene functions of begomoviral genome and interaction of viral protein with host cellular factors.

2.6.1 Coat Protein/AV1/V1

Coat protein is encoded by AV1/V1 ORF in begomovirus and is composed of 250 – 270 amino acids. Coat protein (CP) forms the covering shield under which geminate particles of begomoviruses rest. Nucleotide and amino acid sequence analysis of N-terminal region (60-70 amino acids) of the coat protein is important in identification of geographical variation in begomoviruses (Harrison BD et. al., 2002). In monopartite begomovirus, CP also functions as movement protein. CP binds with viral ssDNA and provides protection and nucleo-cytoplasmic transport of ssDNA molecules. Intercellular systemic movement in plants is also facilitated by CP. CP also protects virion particles from degradation in vector white fly. It binds to GroEL, produced by bacterial endosymbiont (Morin S et. al., 2000). A conserved zinc finger motif at N-terminal of CP is essential for ssDNA binding (Kirithi N and Savithri HS, 2003)

![Figure 2.3](image.png)

**Figure 2.3:** Organization of CP. The CP contains three nuclear localization signals (NLS). The middle NLS overlaps with Nuclear Export Signal (NES) and cell wall targeting motif (CW). N-Terminal motif acts as a DNA binding site.
2.6.2 PreCoat Protein/AV2/V2

ORF AV2 encodes Pre-coat protein (Pre CP) and mostly found absent in new world begomoviruses (Padidam M et al, 1996; Briddon RW et al., 2010). It regulates levels of ssDNA and dsDNA inside plants. GFP fusion studies reveal that Pre CP play important role in pathogenicity determinant and involves in movement of viral DNA in plants. Padidam M et al. (1996) had shown that loss of function mutation in AV2 ORF resulted in very mild symptoms and low levels of both single-stranded (ss) and double-stranded (ds) viral DNA. On the other hand, protoplasts inoculated with mutated ORF AV2 accumulated ss and dsDNA to wild-type levels, confirming role of AV2 in efficient viral movement (Lacatus G and Sunter G, 2008). Previous studies also demonstrate that Pre CP differently modulates CP expression in tissue specific manner. Pre CP/AC2 induces CP expression by activation of the CP promoter in mesophyll cells, while in phloem tissue it deactivates CP expression by derepression of the CP promoter (Sunter G and Bisaro DM, 2003). Pre CP also interacts with host protein and involved in suppression of host RNA silencing (Yadava P et al., 2010). Recently it has been identified that they are involved in perinuclear distribution of begomoviruses in association with endoplasmic reticulum and cytoplasmic strands (Sharma P et al., 2011).

2.6.3 Replication initiator protein (Rep)/AC1/C1

Replication initiator protein (Rep/AC1/C1) is multifunctional and essential for virus replication and hence disease transmission in the host cells. Localized into nucleus, rep, along with host cellular factors, recruits replication materials to the viral strand for replication (Settlage SB et al., 1996; Kong LJ, 2002). Computational homology studies shows that rep functions similarly as other replication initiator proteins of bacteriophages and eukaryotes plasmid. Sequence analysis of rep proteins from different Geminiviruses revealed three conserved motifs, located in half of amino terminus. Among three, tyrosine residue in motif III participates in cleavage and binds to exposed 5’ P group. Motif II contains histidine residues that may bind to Mg++ or Mn++, so as to inactivate cellular DNase. Ligation activity also resides into the amino terminus of the rep. Further, presence of NTP binding domain suggests that rep also function as ATPase and helicase. Mutation of conserved NTP binding domain drastically reduces replication in protoplast (Arguello-Astorga Get al., 2004). Rep represses its own transcription, by unknown
mechanism, and initiates virion sense expression of some proteins. In resting G0 phase, cell is divisionally inactive and replication machineries are available at minimal level. Retinoblastoma (Rb) protein, along with other factors, binds to transcription factor E2F, hence E2F will not bind to its recognition site. As a result, transcription and translation of proteins necessary for G0/G1 to S phase transition are not available. Upon geminiviral infection, rep protein displaces E2F factor, makes E2F free to bind to its site and initiate transcription. As a consequence, host cell enters into S phase, making replication machineries available for viral replication (Gutierrez C 2000; 2004; Kong LJ et. al., 2000). Further, rep modulates plant cell cycle by interacting with Proliferating Cell Nuclear Antigen (PCNA). PCNA acts as a processivity factor for DNA Polymerase δ during replication (Bagewadi B et. al., 2004). Geminivirus infection induces the expression of PCNA in G0 resting cells. Rep interacts with host replication protein A (RPA) (Singh DK et. al., 2007; 2008), a major protein for formation of the replication apparatus, such as DNA polymerase a, RFC and PCNA (Loor G et. al., 1997).

![Diagram of Geminivirus Rep and RepA proteins](image)

**Figure 2.4:** Geminiviral replication-associated protein (Rep) and replication-associated protein A (RepA) proteins. (a) Motif I DNA binding and Motif II (divalent cation binding) overlap the DNA binding and cleavage/ligation domains. Geminivirus Rep Sequence (GRS) and motif III also are part of DNA binding motif. N- termini of Rep interacts with various viral and host proteins, viz., replication enhancer protein (REn), retinoblastoma-related protein (RBR), proliferating cell nuclear antigen (PCNA), GRIK (Geminivirus Rep-Interacting Kinase), small
ubiquitin-related modifier (SUMO)-conjugating enzyme (SUMO) and ATP. (b) The RepA, C-terminus contains the binding sites for RBR, geminivirus RepA-binding transcription factors (GRAB), transactivation domain (TRA).

2.6.4 Transcription Activator protein (TrAP)/AC2/C2:

Geminiviral lifecycle includes DNA synthesis, encapsidation and movement, which require tight regulation for formation of complete virus particle and efficient disease transmission into the host cells. From the molecular perspective, in the viral life cycle, ssDNA have three different fates. First, it may either act as a template for complementary sense strand synthesis to form replicative form, which can initiate RCR. Second it may combine with complementary strand to form transcriptionally active double stranded DNA or third, it may encapsidated into capsid protein to form a virus particle. Such multifunctional task of ssDNA is tightly regulated, at least in Subgroup III geminiviruses, by protein called TrAP (Transcriptional Activator Protein), encoded by AC2 ORF (Sunter G and Bisaro DM, 1997). TGMV protoplast studies with AC2 mutants revealed that TrAP is required for active expression of capsid protein. TrAP mutants of TGMV accumulate ssDNA in lower amount and are unable for systemic spread of disease in the host (Sunter G. and Bisaro DM., 2003). Later experiments, describes that TrAP, transcriptionally regulates expression of nuclear shuttle protein (NSP) encoded by BV1 gene. NSP localized inside nucleus, binds with ssDNA, and is required for systemic spread (Haley A et. al. 1992). Since, both capsid protein and NSP levels are regulated by TrAP, TrAP mutant are, obviously, unable to spread disease systemically (Hartitz M et. al., 1999). Also, acidic c-terminus of TrAP interacts with phloem specific Geminiviral conserved late element (CLE) through N- terminal zinc finger motif and transactivates host and viral genes require for encapsidation and movement of virus particles (Cazzonelli CI et al., 2005).

**Figure 2.5:** Modular organization of TrAP protein. The N-terminal basic region houses nuclear localization signal (NLS). Central region contains zinc finger motif interacts with CLE and
domain that regulate programmed cell death (PCD). Acidic C-terminal is required for activation of transcription.

2.6.5 Replication enhancer protein (REn)/AC3/C3:

The AC3 ORF of geminivirus codes for 15 Kda protein. Expression of AC3 protein greatly enhances, up to 50 fold accumulation of viral DNA, hence often referred as Replication enhancer (Ren) (Morris BK et. al., 1991; Elmer JS et. al., 1988). Ren is also located in the nuclei of plant cells at the level similar to rep, and may facilitate stability to rep-DNA binding. Expression of Ren restores recognition capacity of rep where binding site is mutated in TGMV. Rep binding site is located upstream to the cleavage site (Hairpin Loop) in subgroup II and III geminiviruses. In such viruses, ren may also direct rep to cleavage site to facilitate cleavage (Settlage SB et. al., 2005). REn interact with Rep and upregulate Rep mediated ATPase activity in Tomato leaf curl Kerala Virus (Pasumarthy KK et. al., 2010).

Figure 2.6: Modular organization of REn. N – and C- terminal have Retinoblastoma – Related protein domain (RBR). Three hydrophobic clusters (HD), Rep interactive domain, proliferating cell nuclear antigen (PCNA) and SINAC1 (Solanum lycopersicum Non Apical Meristem 1) binding domains have been identified in REn.

2.6.6 AC4/C4:

ORF AC4/C4 completely integrates on ORF AC1/C1 (Rep protein) in different reading frame. AC4/C4 protein is least conserved with diverse function among all geminiviral proteins (Ho ES et. al., 2014; Sunitha S. et. al., 2013). Virus DNA levels were significantly reduced when infected with disrupted AC4/C4 ORF, in case of ToLCV (Rigden JE et. al., 1994) and TYLCV.
(Jupin I et. al., 1994). Symptom development in such plants was also found milder, suggesting role of AC4/C4 ORF in symptom determinant and viral movement. Peculiar vein swelling phenotype observed upon BCTV infection is resultant of AC4/C4 ORF (Mills-Lujan and Deom, 2010; Stanley J et. al., 1986). Transgenic expression of AC4/C4 protein in Arabidopsis resulted into characteristic viral infection symptoms, while in N. benthamiana, it induced ectopic cell division. In two independent study BSCTV C4 protein also found interacting with host RKP (RING finger E3 ligase) and Arabidopsis thaliana homeobox; ATHB7 and ATHB12 (Lai J et. al. 2009; Park J et. al. 2011). All these factors are involved in host cell cycle progression and regulation, indicating role of AC4/C4 protein in host cell development and differentiation.

Several reports also shown that AC4/C4 protein also involved in suppression of RNA silencing in several bipartite and monopartite Geminiviruses by interacting with microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Fondong VN et. al., 2007; Vanitharani R et. al., 2004; Chellappan P et. al., 2004). Dogra et. al. (2009) demonstrated that ToLCV – Au C4 protein binds with shaggy-like kinase (SISK) at C- terminus and overcome silencing.

2.6.7 Nuclear shuttle protein (NSP)/BV1:

In bipartite Geminiviruses, Other molecule, named DNA B, carry function of virus movement. BV1 is one of two ORFs encoded by DNA B component. BV1 encodes Nuclear Shuttle Protein (NSP) and carries out trafficking of ssDNA between nucleus and cytoplasm. NSP localized into nucleolar region of host cell. NSP recognize and bind to DNA in sequence non-specific manner. In cytoplasm, NSP-DNA complex is encountered by movement protein (MP) for virus cell-cell and long distance movement. Some NSP-DNA complex return to nucleus to serve as template for new DNA replication cycle (Lazarowitz SG and Beachy RN, 1999; Ward BM et. al., 1997). As discussed before, CP is homologous and functions similarly as NSP. This led to conclusion that NSP and CP shared common ancestral genes in the evolutionary history (Zhou YC et. al., 2007; 2011). CaLCV NSP interacts with AtNSI (Arabidopsis thaliana nuclear shuttle protein interactor), resulting in acetylation of NSP (McGarry RC et. al., 2003). NSP also interacts with cytoplasmic GTPases and carries nuclear export of growing viral DNA from the nucleus to the cytoplasm by a leucine-rich NES (Carvalho CM et. al., 2008). In addition, NSP is also shown to
be avirulence determinant in case of BDMV, where it induced hypersensitive response in *Phaseolus vulgaris* L. (Zhou YC et. al., 2007).

**Figure 2.7:** The nuclear shuttle protein (NSP). N-terminal contains two nuclear localization signals (NLS-A and NLS-B). DNA binding domain is a part of NLS – b, while sequences required for the induction of the hypersensitive response are part of NLS – A. The C-terminus contains the nuclear export signal (NES), movement protein (MP) binding domain and AtNSI interactive domain.

### 2.6.8 Movement Protein (MP)/BC1:

BC1 ORF of bipartite geminivirus encodes movement protein (MP) that binds to DNA in sequence nonspecific manner and is required for the long distance movement of the viruses. Thus, MP plays a key role in systemic distribution of virion particles in host plant and determinant of infection. MP binds to NSP-DNA complex in cytoplasmic phase and enables transfer of novice viral particles to adjacent cells. Plant viruses use phloem system for systemic movement and for long-distance transport. MP increase size exclusion limit (SEL) of plasmodesmata (PD), thus facilitating entry of virus particles between cells (Wolf S et. al., 1989; Su S et. al., 2010). CaLCV MP interacts with synaptotagmin (SYTA) and recruits MP-DNA complex on endosome. Through this endosome, MP-DNA complex channels from PD to neighboring cells (Lewis JD and Lazarowitz SG, 2010). Hydrophobic domain of MP interacts with NSP-DNA complex, allows spreading of NSP-DNA complexes at the surface of cellular membranes, which could be suitable for lining of the MP-DNA complexes through plasmodesmata.
Figure 2.8: Organization of MP. The BC1 encoded MP contains three distinct domains: Left N-terminal domain is pilot domain which is required for recruitment of MP-DNA complex at cell periphery. Central domain is NSP interaction domain, by which MP relay over NSP-DNA complex to MP-DNA complex in cytoplasm. C-terminal right domain facilitates oligomerization.

2.6.9 Geminivirus satellite rep and betaC1 proteins

Monopartite begomoviruses contain DNA A, similar to bipartite begomovirus but are often associated with one or more satellite DNA (sub-viral) components, about 1.4 kb in size. Betasatellite, most common amongst the satellite DNA, functions as virulence determinant in monopartite virus. It encodes for one ORF (betaC1), encodes multifunctional protein, sometime associated with silencing mechanisms. Replication of DNA betasatellite is dependent upon replication of helper cognate DNA A molecule. Protein betaC1 interacts with Sucrose nonfermenting – 1 relating kinase (SNF-1 Kinase) and S- adenosyl homocysteine hydrolase, suppresses RNA silencing and thereby, boosts DNA accumulation of their helper virus and symptom severity (Shen Q et. al., 2012; Yang X. et. al., 2012).

Alphasatellites (formerly DNA1) are circular, single stranded DNA molecules, about half a size of DNA A, are often found to be associated with begomovirus/betasatellite complexes. Alphasatellites encode their own replication associated protein, similar to nanovirus, hence are dispensable for their own replication (Idris AM et. al., 2011; Nawaz-UL-Rehman MS. et. al., 2010). Though, they are proficient for own replication but require helper begomovirus (DNA A/β satellite complex) for movement and transmission (Briddon RW and Stanley J, 2006). Both of these satellites encompass signature nonanucleotide sequence of begomovirus (TAATATTAC) and replicate through rolling circle mode of replication.
2.7 Geographic Differentiation

Phylogenetic studies show that begomoviruses are clustered into two distinct groups, the Old World (OW) viruses (eastern hemisphere, Europe, Africa, Asia) and the New World (NW) viruses (western hemisphere, the Americas) (Padidam M et al., 1999; Rybicki EP, 1994). OW Begomovirus have a number of distinguishable genomic characteristics that distinguish OW and NW viruses. All NW begomoviruses are always bipartite. They have lost Pre CP gene and are largely dependent upon associated DNA – B molecule for replication and movement. OW begomovirus can be both bipartite and monopartite, where some monopartite begomoviruses are also associated with satellite DNA molecule. In addition, all OW begomoviruses carry AV2 ORF in DNA – A which is not present in NW begomoviruses (Rybicki EP, 1994; Stanley J et al., 2005). Rybicki EP (1994) proposed that most NW viruses may have evolved more recently after the continental separation of the Americas from Gondwana approximately 130 million years ago. Whiteflies movement and carrier of diseased plants are major transport mode from Asia to the Americas and may have transmitted viruses that were the ancestors of NW viruses. Under different agroclimatic conditions, the exported OW begomoviruses evolved independently till date to NW begomoviruses with strict adherence to cognate DNA B molecule and loss of gene function of Pre CP ORF.

2.8 Vector Transmission

Being obligate parasites, viruses depend entirely on the host for their propagation. Cellulose wall and intact cuticle of plant cells restrict virus entry, directly, into the plant cell. However, seed transmission, vegetative propagation or mechanical inoculation through wounding can led to direct entry of viruses into the plant cell. Yet in nature, most plant viruses prefer to transmit by specific vectors from one host to another. Such virus – vector relationship is specific to some extent and involves virus – ligand interaction. Viruses display their specificity through coat protein, helper component or transmission factor (Whitefield AE et al. 2015).
Geminiviruses are strictly vector transmitted viruses. No seed transmission is reported to date. Whitefly *Bemisia tabaci* (Gennadius) (Hemiptera Aleyrodidae), is cosmopolitan arthropod insect pest which transmits begomoviruses in circulative non-propagative persistent manner. The location of begomovirus in *B. tabaci* is filter chamber cells and the accessory salivary gland. Begomovirus enters into the host plant when ovuliferous white fly sucks the phloem sap of plant and transfers virion particles. Viral particles enter the cells through vascular tissue where they are uncoated and entered into nucleus with the help of CP. This cycle repeats when whitefly sucks phloem sap and sits on healthy plant. Viral particles are ingested by stylets along with phloem sap of infected plants, move through alimentary canal to midgut. From midgut, they enter into the haemolymph and finally salivary gland for transmission to healthy plant during next feeding cycle. (Brault V *et. al.*, 2010). Begomoviral CP interacts with insect GroEL and HSP 16. These interactions might stabilize and shelter virion particles into haemolymph of the insect vector (Hanley-Bowdoin L *et. al.*, 2013).

### 2.9 Begomovirus infection and movement in host

The insect pests transfers virion particles are insect pests are phloem-feeders, hence load virus particles to the phloem tissue of the plant. Phloem tissues contain actively dividing cells; hence provide accessory machinery for viral replication. Further, they are connected directly to symplast of the whole plant via plasmodesmata. Probably, for this reason geminivirus possesses phloem specific gene expression which is at low level in other tissues (Morra MR & Petty IT, 2000). Occasionally, viruses found in non-phloem tissue, mostly mesophyll cells, often with RNA viruses. RNA silencing is also another mechanism by which plants restrict virus particles to phloem cells and restrict their entry to mesophyll cells.

Upon entry virus particle uncoats and nascent ssDNA enters the host nucleus. Inside the nucleus, they replicate and new DNA molecules are transported into cytosol by NSP. MP then transfers virion particles to cell periphery and then from one cell to another cell. Then distribution occurs from one plant to another plant with the help of movement proteins. In case of monopartite begomovirus, involved CP and Pre CP functionally replace NSP and MP. It has also been reported that βC1 protein encoded by betasatellite can be an alternative for movement of DNA B
molecule (Saeed M et. al., 2007). In some cases, βC1 protein alone can play role in viral particle movement (Briddon RW et. al., 2001; Cui X. et. al., 2005).

2.10 Geminivirus Replication

Meristems are the region in the plant cells (at the shoot apex and the root apex) which continuously divide, proliferate and thus give growth to specific region in which they are located. In addition to different biochemical pathways, plant cells are totipotent in nature by which any plant cell can undergo dedifferentiation and then redifferentiate into specific lineage. Viruses are totally depending upon cellular factors and other replication machineries for their multiplication, movement and disease incidence. In this analogy, it seems that meristems are most prone to virus infection due to excessive availability of replication machineries but is not true. Meristems divide at very higher rate and viruses cannot keep pace with such higher divisional rates hence meristems are considered essentially as “Virus Free” zone. Intercellular movement of viruses’ into host cell is achieved by vascular system of plants. Movement through cellular junctions, plasmodesmata, is too slow to keep pace with actively dividing cells. Meristems lack vascular system is also another reason that they are virus free zones. Also, the silencing mechanisms has higher activity in these regions, do not allow virus replication (Baulcombe DC., 1999, Ratcliff F et. al. 1997). When virus infects the cell, the cell is already differentiated i.e. in the G0 phase of cell cycle, where replication machinery is not available or available at minimal rate which is not sufficient for virus replication. Replication factors are almost exclusively present in proliferating cells, at S phase of cell cycle. Hence, it is speculated that geminivirus induce cellular cell cycle change from G0/G1 to S phase for making the cell a good shelter for its own replication and other downstream replication processes.

Replication dogma of geminiviridae follows a path of ssDNA to dsDNA (A Replicative form) to ssDNA via rolling circle mode of replication. Similar to RCR employed by various coliphages (Ø x 174) and various bacterial plasmids. The characteristic feature of RCR mode is that it requires nicking and closing activity for the initiation and completion of the process.
The process of replication completes in a two-phase, where in the first phase, these virion sense strand act as template for complementary sense strand, to form ds replicative form (RF) of DNA. The replicative machineries for such conversion of ssDNA to dsDNA should consist of host encoded DNA polymerase along with other auxiliary factors, necessary for replication. In the second phase, this RF act as the template for both transcription – translation and generation of new virion sense strands.

Subsequent to whitefly transmission, viral nucleic acid makes its way to enter to the host cell nucleus. With the help of cellular factors, viral nucleic acid is converted to transcriptionally active double stranded replicative form. This dsDNA is transcriptionally active that transcribes and translates various virus proteins, necessary for its assembly and movement. One of such protein is Replication initiator protein (Rep), which is required to initiate and extend rolling circle mode of replication, to generate pool of ssDNA for the viral spread in the next stage (Desbiez C et. al., 1995; Heyraud-Nitschke F, 1995; Laufs J, 1995). This is the first stage in viral life cycle, where some of the viral encoded proteins also participate in replication process. Rep is oligomeric, multifunctional, viral encoded protein, possesses various replication associated activities, such as helicase, topoisomerase, site specific binding, nicking and ligase. To initiate RCR, Rep recognizes highly conserved stem loop structure located in the CR. Geminivirus-signature nonanucleotide sequence (TAATATT ↓AC) is located in the loop (Heyraud-Nitschke F, 1993a, b). In order to initiate rolling circle replication, Rep nicks the DNA between 7th and 8th nucleotide (arrow) in the nonamer sequence. Such nicking will cleave phosphodiester linkage between adjacent nucleotide, releases free 5’ phosphate and 3’ OH groups. Rep binds to 5’ phosphate, while 3’ OH terminus serve as a primer. As the replication proceed, parental plus strand is being replaced from the intact minus strand template to complete one circle. Such completion of replication is recognized by multifunctional rep protein associated at 5’ phosphate which nicks, releases and ligates to release unit length viral sense strand (Stenger DC et. al., 1991; Hanley-Bowdoin L et. al., 1999).

Jupin I et. al. (1995) show that different isolates of TYLCV exhibit strict specificity in the interaction between the Rep protein and IR and is required for their replication in vivo. First 116 amino acids and left part of IR is specificity determinant and act in cis – trans manner. Such conversion of viral sense strand to complementary (-) sense strand, so as to produce ds
replicative form, requires identification and activation of (-) strand origin of replication region by replication factors. In case of mastervirus, a small stretch of DNA is associated to (+) strand and is the region recognized by replication machineries to initiate replication (Erdmann JB et. al., 2010). IR spans cis acting signals from 2581 to 221 nucleotides and defines (-) strand origin in begomovirus (Bisaro DM, 1996; Saunders K et. al., 1991; Gutierrez C, 2000; Preiss W and Jeske H, 2003)

After the replication, the covalently closed circular DNA in its relaxed form needs to be converted to supercoiled structures by a mechanism, similar as catalyzed by DNA gyrase as in prokaryotes. Such supercoiling essentially occurs by recruitment of host histones or histone like proteins during replication process to form geminiviral minichromosome (Pilartz M and Jeske H, 1992) Micrococcal nuclease digestion produces DNA fragments that also supply a proof of formation of minichromosome during replication.

The sequence recognized by the Rep is most important, since even a point mutation in the sequence or alteration/disruption in structure will be not recognized by Rep. In addition to stem loop nonamer, upstream sequence elements are also required for replication assembly and may facilitate Rep binding in a sequence specific manner. In TGMV, direct repeat sequence of 13 bp occupies position in TATA box and controls the expression of Rep and transcription start site. This 13 bp sequence is constituted by direct repeat sequence of 5 bp (underlined) with 3 bp spacer in between (5'-GGTAGTAAGGTAG-3') Rep binding at the loop requires dimerization. It has been proposed that rep will act as a dimer or oligomer due to presence of only one tyrosine molecule, which is required for cleavage, at active site of rep (Marsin S and Forterre P, 1999; Orozco BM et. al., 1997; 1998). Analysis of Rep binding sites in begomovirus revealed a consensus sequence, GG-AGTAC/TC/TGG-AG. Electrophoretic mobility shift assay followed by in vitro Dnase I footprinting assays with purified rep demonstrated two footprints covering the region of sequence GCAATTGGTCTCTCAA and TGAATCGGTGCTCTGGGG. The repeated motifs in the sequence (Bold and Italics) are involved in rep binding.

Multifunctional roles of Rep may be determined by following processes of replication. First, Rep is the only protein that recruits at origin and initiates RCR by nicking. Such process is analogous to helicase activity of prokaryotic or eukaryotic DNA replication. Rep may directly act as
helicase or assist host cellular helicase for such nicking. Rep binding at specific site causes distortion and results into localized extrusion that results in hairpin structure. Third, Rep recruits host cellular replication machineries for initiation of replication (Orozco BM et. al., 1997). Rep recruitment require specific binding site. The other rep molecule recognizes the presence of already bound another rep molecule or rep associated unit length sequence, as a substrate for cleavage.

![Diagram of Rep binding sites](image)

**Figure 2.9:** Modular organization of Begomovirus-CR. Common Region of DNA A shows various replication and upstream transcription regulatory elements. Rep binding iteron sites are on both sides to stem-loop structure. Stem-loop structure contains signature nonanucleotide sequence, which is essential as start site of replication. Rep binds and nicks between 7th and 8th nucleotide position (arrow). Transcription start site and regulatory element (AG motif) are upstream of stem-loop structure.

In the late replication cycle, the virion sense ssDNA produced by RCR are sequestered by NSP/CP and transfer to cytoplasm from nucleus, where they are directed to plasma membrane by MP and transfer to neighboring healthy cells.

**2.11 Transcription**

As replication proceeds, newly formed ssDNA has three fates. First, it forms a complex with NSP complex and gets transported to cytoplasm and packed with coat protein. Secondly, pool of ssDNA serves as a template for next cycle of rolling circle replication. Third, part of ssDNA also serves as a template for RNA polymerase for transcription. Transcription in geminivirus is
bidirectional and regulated temporally, where early transcript genes are located on complementary strand and late transcript genes are on viral sense strand (Shimada-Beltran H and Rivera-Bustamante RF, 2007; Thommes PA and Buck KW, 1994). Transcript analysis of Geminiviruses reveals that they are poly cistronic. Promotor elements of geminivirus structured as eukaryotic transcription system such as TATA box, CAAT box, Polyadenylation signals and splicing sites. LIR/IR/CR of geminiviruses houses replication start site as well as transcription promoter site. The viral mRNAs have been well studied in a number of begomoviruses: ACMV (Townsend R et. al., 1985), ABMV (Frischmuth S et. al., 1991), TLCV (Mullineaux PM et. al., 1993), MYMV (Shivaprasad PV et. al., 2005). Proteins are expressed from subgenomic RNAs (sgRNAs) with leaky scanning for the expression of some of the overlapping proteins.

2.12 Geminivirus and plant cell cycle modulation

The characteristic feature of multicellular eukaryotes is the “Division of Labor” to the each group of specific cells. Most of these cells perform their assigned work after they have differentiated from the ancestor pool. Such cell, because they are already differentiated into specific lineage, are devoid of replication machinery and are transcriptionally inactive, as of other dividing cells. Such is true for plant cell as well, where the proliferation activity is confined to the limited region, called meristems (shoot and root apex), developing leaves and the cambium.

Surprisingly, geminivirus infection is excluded to this highly proliferating region, and is limited to cells, which are already differentiated and possess very low levels of replication machineries. This contradiction of requirement of replication machineries by virus particle and unavailability of the same by host cells leads researchers to view for geminiviral interaction with plant cell cycle (Hanley – Bowdoin L et. al., 1999; Gutierrez C., 2000).

In animal cells, passage of cells from G1 to S phase transition is tightly regulated by a family of proteins called cyclins - cyclin dependent kinase (CDK) – CDK Inhibitor (CDKI). Retinoblastoma (RB) protein modulates the activity of E2F – DP family of transcription factors (Bartek J, 2001; Reed SI, 1997). CDK-Cyclin complex phosphorylates RB, results in the release of E2f-DP bound to RB leads to expression of genes required for G1/S transition (Figure 2.10,
Figure 2.10: Geminivirus infection and cell cycle modulation. In resting G0 phase, cell is replicatively inactive. Retinoblastoma – like protein (RBR) binds to E2F transcription factor, hence not allowing E2F to bind to its binding site. As a result, transcription is halted. Upon, geminivirus infection, geminiviral encoded Rep or RepA protein interact with RBR, so as to release E2F. Released E2F binds to E2F binding site, downstream, resulting into active transcription and entering of cell from G1 to replicatively active S phase.
Nevins JR, 2001; Harbour JW; 2000). Recent findings suggest that, rep protein of mastrevirus, could interact with RB through LXCXE motif, and modify cell cycle to enter into S phase (Dahiya A et. al., 2000). In begomovirus, such LXCXE motif is absent but the region that interacts with RB protein is mapped to (region containing 4 helix) rep oligomerization site (Arguello-Astorga G et. al., 2004; Kittlemann K et. al., 2009).

This strategy of modulation of host cell cycle adapted by geminivirus is similar to that of animal oncoviruses (Merits A et. al., 2000; Horser CL et. al., 2001; Gutierrez C et. al., 2004). Geminivirus protein dependent interaction with RB releases bound transcription factors, would initiate expression of genes require for S phase transition, thus creating cellular machineries available for replication.

2.13 Present knowledge and relevant bibliography relating to viruses identified in this study.

2.13.1 Mungbean yellow mosaic India virus (MYMIV)

History

Mungbean yellow mosaic virus “Indian” strain was first observed in late fifties by Nariani TK (1960). It produces typical mosaic symptoms on leaves of infected plants and naturally transmitted by white fly (Nene YL, 1973). Biological properties of MYMV were reviewed by Honda Y and Ikegami M (1986). These legume infecting viruses are not sap-transmissible and infect selected leguminous plants. Upon infection, they produce biologically indistinguishable symptoms, making specific identification of the viruses difficult (Pant V et. al., 2001). Based on viral epitope profile, Swanson MM et. al. (1992) distinguishes “legumoviruses” into two broad classes: one comprised of dolichos yellow mosaic virus and second includes yellow mosaic viruses that infect most of grain legumes. In addition to India, the virus is widely prevalent in the Indian subcontinent, Sri Lanka, Bangladesh and Pakistan (Honda Y, 1986). An epidemic of yellow mosaic disease of mungbean was also identified in Thailand in the 1980s (Honda Y et. al., 1983). MYMIV infection is confined to Northern India, Pakistan, Nepal, Bangladesh and Indonesia. While, MYMV is mostly confined to Thailand, Vietnam and Eastern Ghats and Deccan plateau of India (Islam MN et. al., 2012; Tsai WS et. al., 2013). MYMV is important
economically as it infects five major leguminous species, blackgram, mungbean, French bean, pigeonpea and soybean, causing yield loss of about $300 million annually (Varma A et. al., 1992).

Properties

Geminate viral particles measure size of 18 X 30 nm (Honda Y et. al., 1983, Honda Y, 1986). Morinaga T et. al. (1983) identified two identical sized ss circular DNA molecule of different nucleotide sequence encompassed into geminate particle of MYMV. Serological relationships of MYMV were reported with ACMV, BGMV and TYLCV (Honda Y et. al., 1983). Thermal Inactivation Point (TIP) of MYMV is 45-50°C, Dilution End Point (DEP) is $10^{-2}$ to $10^{-3}$ and Longetivity in vitro of Virus (LIV) is 1 – 2d (Honda Y et. al., 1983).

Host Range

Natural infection of MYMV has been reported onto seven species into three genera of Fabaceae: Dolichos biflorus (Williams FJ et. al., 1968), Glycine max (Ahmad M and Harwood RF, 1973), Phaseolus aconitifolius (Ahmad M and Harwood RF, 1973), P. aureus (Nariani TK, 1960), P. mungo (Vanitha Rani et. al., 1996), P. vulgaris (Singh RN, 1979), P. lunatus (Shahid MS et. al., 2012), Macrotyloma uniflorum, cajanus cajan (Biswas KK and Varma A, 2000).

2.13.2 Horsegram Yellow Mosaic Virus (HgYMV)

HgYMV was first reported by Williams FJ et. al. (1968) in India. HgYMV was found as causal agent of yellow mosaic disease (YMD). Incidence of disease ranged from 60-100% in summer and early rainy season. YMD is characterized by yellow mosaic patches on leaves, reduced leaf size and dwarfism in severely affected plants (Muniyappa V et. al., 1987). Occurrence of HgYMV was found limited to Southern India (Borah BK and Dasgupta I, 2012; Varma A and Malathi VG, 2003)

Properties

Geminate viral particles measure size of 15 -18 X 30 nm (Muniyappa V and Veeresh GK, 1984) isolated from horsegram infected with HgYMV.
Host Range


2.13.3 Rhynchosia yellow mosaic virus (RhYMV)

Rhynchosia mosaic disease was first investigated by Bird J and Sanchez J (1971). The symptoms include presence of bright yellow mosaic patches intermingled with green leaves. Bird J et. al. (1975) identified that RhYMV is transmitted by white fly in a persistent, circulative manner.

Properties

RhYMV induces typical geminiviral nuclear inclusions in phloem cells of *Rhynchosia minima* (Christie RG and Bird J, 1984; Christie RG et. al., 1986). No other cytological studies were reported for RhYMV.

Host range

RhYMV has been reported in natural infection of *Cajanus cajan*, *Glycine max*, *P. lathyroides*, *Macroptilium lathyroides*, *R. minima*, *Nicotiana tabaccum*, *Abelmoschus esculentus*, *Gossipium hirsutum* and *Ipomea quinquefolia* (Bird J and Sanchez J, 1971).

2.14 Begomovirus infection on French bean

Number of begomoviruses namely, bean golden mosaic virus (BGMV), bean yellow golden mosaic virus (BGYMV), bean dwarf mosaic virus (BDMV), bean yellow dwarf virus (BYDV), Bean calico mosaic virus (BCaMV) are found to infect *P. vulgaris* naturally (Costa AS, 1965; Galvez GE and Morales FJ, 1989). Field spread of these viruses was more efficiently carried out by adult female white flies that male, in persistent manner. Further, epidemics of disease is more severe, if suitable white fly reproductive host plants such as tomato, egg plant and tobacco are in
the adjacent area of crop field. Macroscopic symptom studies revealed that BGMV and BGYMV produces undifferentiated visible symptoms, yet, molecular data suggest that BGMV and BGYMV are related but distinct begomovirus species (Morales FJ, 2003). Recently much progress has been made for finding diversity and natural occurrence of begomovirus infection on French bean. This include, bean yellow chlorosis virus (BYCV) and bean white chlorosis mosaic virus (BWCMV; Fiallo-Olive E et. al., 2013), yellow mosaic virus (YMV; Jyothi V et. al., 2013), Tomato yellow leaf curl virus (FbLCV; Kamaal N et. al., 2013), French bean leaf curl virus (TYLCCNV; Dong JH et. al., 2006), Mungbean yellow mosaic India virus (MYMIV, Shahid MS et. al., 2012; Singh RN, 1979), Horsegram yellow mosaic virus (HgYMV, Monger WA et.al., 2010) and Tobacco curly shoot virus (TbCSV; Venkataravanappa V et. al., 2012). The common symptoms of infection include partial or complete brilliant golden and/or yellow mosaic on leaves, stunted growth, less number of pods with reduced pod size.

2.15 “Omics” technologies and Plant-Virus interactions

Although, begomoviruses possess small genomes, they encode number of proteins through host cellular machineries to multiply as well as command host cellular metabolism. Such prevalence of viral infection may reduce the production or complete loss, depending upon the severity of infection and spread. Begomoviral life cycle depends upon the differential expression of genes at specific time point of the host life cycle. Systemic viral infection, maintenance and spread rely on how host protein/s modulate/s and participate/s upon viral infection. Viral incidence onto host plant efficiently changes the proteome expression profile, studies of which, may throw a light on resistant strategies owned by host plant to combat viral incidence. The understanding of functional aspects of viral infection will result in formulating better universal resistant strategies against these viruses.

Over past few year, advances in omics technologies led researchers to gain inside into molecular events happening at plant-microbe interactions. Recently, functional genomics strategies, including proteomics, transcriptomics and metabolomics, have contributed much to our understanding of gene and protein function and expression profiles. Using these ‘omics’ technologies and next generation sequencing approaches, defense-related genes and proteins expressed during phytopathogen infections have been identified and enormous literature is
documented for such interactions. Yadav RK and Chattopadhyay D (2014) studied transcriptomic profile of soybean gene expression under early infection by MYMIV to understand early stage local responses and differences in gene expression in resistant and susceptible soybean cultivars. All together 444 transcripts categorized in 11 functional classes showed significant changes. Genes involved in hypersensitive response, programmed cell death and resistance response pathways are upregulated and expression of genes for photosynthesis and sugar transport are reduced (Yadav RK and Chattopadhyay D, 2014). Recently, Kushwaha N et. al. (2015) showed that an interactome network of host genes are differentially expressed upon ChLCV and identified range of proteins involved in key cellular processes including transcription, replication, photosynthesis, and defense. However, it has been well known that transcriptomic expression profile not necessarily mirror proteomic profile of the cell. Posttranscriptional, Posttranslational modifications, epigenetic regulations, alternate splicing and spatial expression are much evident layers of gene regulation in eukaryotes. Such plethora of regulatory mechanisms do not guarantee that changes in gene or mRNA expression will be accounted by parallel change in protein expression, which are key players of cellular activities.

Further, keeping in view, much diversity of pathogen and host plants, each plant-pathogen interaction is unique and hence defense response encountered by host plant is unique too. Moreover, these responses are largely affected by agroecoclimatic factors such as temperature, humidity, nutrition, variety and age of host plants. Hence, complete picture of molecular plant-pathogen interactions is still in nascent state. Progresses in mass spectrometry techniques and real time update of protein identification databases made proteomics and functional genomics a primary choice of researchers to study plant-pathogen interactions. Through, proteomics approach large numbers of candidate genes expression and their functional profile have been identified during phytopathogen interactions. Proteomics approach has been used to study many begomoviral-plant interactions. Through gateway cloning technologies and proteomics approach, role of AC2 ORF of ToCMoV in establishing disease has been identified. AC2 ORF of begomovirus was shown to modulate plant defense processes by increasing the proteome profile of several oxidative stress-related and pathogenesis-related proteins, photosynthesis and energy production systems related proteins (Carmo LST et. al., 2013). Kundu S et. al. (2013) employed biochemical techniques coupled with proteomic analyses to study compatible and incompatible
plant-virus interactions between *Vigna mungo* and MYMIV. The results achieved shows early accumulation of the defense and stress related proteins are the key factors in inducing resistance. The changes in primary carbohydrate metabolism harmonize stress response in incompatible interactions. Further, Kundu S. *et al.* (2011) reported the role of salicylic acid (SA) in inducing resistance to MYMIV infection in *Vigna mungo* has by proteomics. They have identified 29 differently expressed host proteins involved in stress responses upon SA treatment, Proteins are functionally annotated to metabolism, photosynthesis, transport and signal transduction. Yellowing symptoms observed on susceptible variety of *Vigna mungo* is characteristic symptom of MYMIV infection, and is due to degradation of photosynthetic apparatus. Increase abundance of photosynthetic genes is a marking signature of restoration of photosynthesis to provide more metabolites required for repartition of resources towards defense.

Proteomics approach was also employed to study plant-virus interactions other than begomovirus. Some of this study includes, tobacco mosaic tobamovirus infecting Chilli (*Lee BJ et. al.*, 2006), Plum Pox Potyvirus infecting peach (*Diaz-Vivancos P et. al.*, 2006), Pepper mild mottle virus infecting *N. benthamiana* (*Rahoutei J*, 1999; 2000), Rice infection by rice yellow mottle sobemovirus (*Delalande F et. al.*, 2005).

Successful systemic infection of viruses to the plant requires efficient replication of virus inside host cells. Subsequently, it requires that new particles should move through plasmodesmata for cell-cell movement and later, to the vascular phloem system to circulate systemically to the plant. Proteomics approaches to study plant-virus interactions only elucidate participation of host proteins into viral replication cycle. Involvement of plant proteins and signaling molecules in viral movement is difficult to study and yet to be demonstrated due to many reasons. Viral movement in plants is tissue specific. Tissue type, such as parenchyma (cell – cell movement) and phloem cells (long distance movement), are difficult to isolate hence proteomics study involving virus movement is impeded.

### 2.16 Resistance strategies employed against geminivirus infection

Geminiviruses are the major threat to plants and affect the production drastically. Controlling the vector that transmits the virus was the only available resistance strategy employed until now
(Legg JP and Thresh JM, 2000). But in recent years, increased knowledge of molecular biology of geminivirus and its interaction with host cell cycle proteins opens up new doors for new and more promising resistance strategies. Such novel mechanisms, although, only experimentally tested, can be readily applied to the virus infected zones in the field. The resistance strategy mainly employs the principle of reduction of viral products at any given time of viral life cycle so as to produce none or defected virus particle. There are mainly four such strategies employed for resistance (Vanderschuren H. et. al., 2007).

First is the DNA interference. It has been known that some of the subviral DNA molecules associated with geminivirus genome negatively regulate DNA accumulation in the host cells through an unknown mechanism (Briddon rw and Stanley J, 2006, Cui X et. al., 2005). The potential inhibition mechanism can become a potential member of a resistant strategy.

Silencing at the transcription and post transcription level is an effective mechanism to abolish viral infection (Tenllado F et. al., 2004). Transcriptional gene silencing (TGS) and Post transcriptional gene silencing (PTGS) is a mechanism adopted by eukaryotic cell to combat with viruses and transposons. Expression of ds RNA, homologues to viral coding region may easily evoke silencing cascade leading to cleavage of viral transcript. Another way of silencing at the RNA level is the epigenetic modification of viral promoters. Increased methylation of the transgene promoter triggers silencing (Jones L et. al., 1999 and Al-Kaff NS et. al., 2000). Seemanpillai M et. al. (2003) also observed the same phenomenon in TLCV by methylation at TLCV derived promoter. Despite the effective nature of TGS, it can’t control promoter regulation completely and basal level of expression is always there. PTGS could be an alternative option where target RNA are degraded and is no more available for translation (Kumagai MH et. al., 1995; Ruiz MT et. al., 1998). Thus, TGS or PTGS together could provoke necessary silencing cascade to down regulate viral spread effectively.

Expression of viral proteins is another strategy for resistance. Earlier studies have proved that expression of TMV Coat Protein (CP) reduces accumulation of TMV particles (Abel PP et. al., 1986, Beachy RN, 1999). Such strategy is largely employed for resistance against RNA virus but not found equally effective against DNA virus (Frischmuth S and Stanley J, 1998). Such mechanisms are known as pathogen derived resistances (PDRs).
Geminiviral life cycle depends upon the differential expression of genes at specific time point of the life cycle. Interruption in this sequence leads to incomplete viral formation. Expression of antibodies against common epitopes of the proteins, like rep, is another alluring idea for resistance.

A ribozyme is an RNA molecule that is capable of catalysing cleavage of target RNAs, similar to the action of protein enzymes. Ribozymes are considered as a potential mechanism for down regulation of the expression of the gene during the course of development. This strategy was used to target mRNA of Rep of MYMIV (Chilakamarthi U et. al., 2007; Mishra SK et. al., 2014). The suitable change in the recognition site which covers the conserved catalytic site is able to target and cleave any RNA.

Lopez – Ochoa L et. al. (2006) reported that peptide which binds to rep protein and consequently interferes with viral infection and spread. Peptide aptamers are “designer” proteins that interact and inhibit the function of specific protein of interest. Such aptamer based antiviral strategy was more explored in case of animal viruses like hepatitis b virus (Butz K et. al., 2001), rabies virus (Real E et. al., 2004) and human papilloma virus (Butz K et. al., 2000) etc.

Increased threat of geminivirus to the crops greatly demands an efficient and nearly universal strategy which is practically applicable. Recent advances into the mechanism of disease transmission have provided new direction for designing such strategies via genetic engineering. One should keep in mind, during employing this strategy, about its effect, reversal and potential effect on the qualitative and quantitative production of plant products.

*****