GLOBAL DIVERSITY AND DISTRIBUTION OF WHITEFLY-TRANSMITTED GEMINIVIRUSES OF COTTON

J. K. Brown
Department of Plant Sciences
University of Arizona
Tucson, AZ

Abstract

Geminivirus diseases of cotton are on the rise, worldwide, yet few have been studied in adequate detail to permit the implementation of rational approaches to disease control. The rising costs of managing the whitefly vector, coupled to substantial losses caused by geminivirus-incited diseases now hinder cotton production by requiring inputs that are beyond economic feasibility. The requirement for geminivirus disease resistance in numerous cotton cultivars and multiple, diverse geographic cotton production areas of the world presents a new and unique challenge. To meet this need, baseline information concerning the identity, the distribution, and the relevant characteristics of cotton-infecting geminiviruses and virus strains, thereof, are now required. This study addresses this problem by attaining and applying molecular sequence analysis to key regions of the genomes of cotton-infecting geminivirus collected from cotton growing regions of the world. Specifically, we are examining the sequence similarities of the conserved the coat protein or AV1 gene, and the similarities and particular features associated with diagnostic nucleotides found in the LIR/CR that are involved in regulating essential aspects of the disease cycle. This effort represents the first cataloging and mapping of geminivirus identity and distribution, and the first investigation of the breadth of geminiviral relationships, or the 'diversity' of geminiviruses of cotton, worldwide. It seeks to understand relationships between cotton-infecting geminiviruses and of these viruses and other well-characterized or 'reference' geminiviruses from diverse crop and weed species. This data base of molecular and biotic information will serve as the cornerstone for the rational selection of virus species and strains toward developing cotton cultivars with resistance customized to protect against disease caused by geminiviruses relevant to the production area. This approach will also permit the first precise evaluation of the breadth of disease resistance in a cultivar by permitting challenge-inoculation with narrowly and broadly divergent virus genotypes, thereby providing both a predictive capacity for sustainability of disease resistance and a safeguard to achieve long term protection against indigenous and introduced, exotic geminiviruses of cotton.

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 1:155-161 (1998)
National Cotton Council, Memphis TN

Introduction

Whitefly-transmitted geminiviruses affecting crop plants were first recognized as a unique family of plant viruses in the mid 1970's when virions of bean, and then tomato and cassava-infecting geminiviruses were visualized by electron microscopy (Brown, 1994; Brown and Bird, 1992). These viruses were further shown to be novel because they were, and still are, the only plant viruses known to contain single-stranded circular DNA genomes, and consequently, have been recognized as a separate family of plant viruses, the Geminiviridae. The viruses placed within subgroup III of this family are transmitted by a single species of whitefly, *Bemisia tabaci* (Gennadius), and one other putative species within *Bemisia* that has been tentatively considered a member of a possible species complex. Viruses in the other two subgroups within the Geminiviridae are transmitted only by leafhoppers or planthoppers. The unique particle morphology and circular, single-stranded DNA genome make this group of viruses distinct from all others known to date. Since their initial recognition, the whitefly-transmitted geminiviruses have become recognized as virulent, emergent pathogens in a multitude of dicotyledonous crops, and are of particular importance in cotton and vegetables. Diseases incited by whitefly-transmitted geminviruses cause reduced growth and hence, significant reductions in yield and quality. Infection at seedling or subsequent, early developmental stages typically results in little to no production in most widely grown vegetable and cotton cultivars. To date, there are few geminivirus-resistant crop varieties available, though much effort is in progress to control geminivirus diseases through development of resistant varieties that are suitable to specific locations and market needs.

Until quite recently, whitefly-transmitted (WFT) geminiviruses were known primarily to infect non-cultivated plants in dry subtropical and tropical climates, with the few exceptions being diseases of cassava, tobacco, and tomato described primarily in the Old World. From the 1970's to the present, whitefly-transmitted viruses were suspected to be involved in increasingly more prevalent diseases of crops, including cotton and vegetables that were grown in large monoculture systems in both the Old and New Worlds. The increasingly wider distribution and higher population levels of whitefly vector populations on a global scale can be directly implicated in the development of epidemics and incidental infections in crops throughout the world. This is thought to be due to the increased use of pesticides to control primary pests and to the cultivation of higher yielding cotton and specialty horticultural varieties, while at the same time, the emphasis on incorporating disease and pest resistance, and the impending development of insecticide resistance in pest and non-pest insects were not considered of timely importance.

Members of the *Bemisia tabaci* (Genn.) complex, among which are the tobacco, cotton, or sweet potato whitefly, and

B. argentifolli , or the B biotype of *B. tabaci* , are the only known whitefly vectors of subgroup III geminiviruses (Brown, 1996). *B. tabaci* is readily capable of establishing to extreme population levels, particularly in crops grown under irrigated, arid conditions in both field and greenhouse systems. In addition, this whitefly has the potential to colonize a wide range of dicotyledonous species, among which are primarily vegetable and fiber species of great importance to worldwide agricultural production efforts. Recent studies indicate that there are numerous apparently isolated populations of *B. tabaci* that vary somewhat in their capacity to develop high population densities and cause direct feeding damage, in the extent of their host ranges, and in the efficacy with which they can transmit geminiviruses (Bedford et al., 1994; Brown and Bird, 1995; Brown et al., 1995a,b,c).

Unprecedented population explosions of whitefly vector populations are now problematic on a regular basis in cotton and vegetable agroecosystems in mild climate production zones, worldwide. This phenomena is related to the expansion and duration of monoculture production of vegetables and cotton, both made possible by greater irrigation capacities, to the cultivation of high yielding varieties that harbor little if any resistance to geminiviruses, and to the direct and indirect affects of pesticide usage. The B biotype is considered an anomaly, in that it is one of the recognized *B. tabaci* populations that has been well-characterized, and is known to produce extremely high population densities in many crop species, and causes due to feeding damage and honeydew contamination. In addition, this whitefly has been shown to transmit numerous geminiviruses of New and Old World origin (Bedford et al., 1994; Brown et al., 1995a,b,c). Since the establishment of the B biotype in cotton-vegetable agroecosystems in the sunbelt states of the US and throughout Latin America which occurred between 1986-1996, it is considered a primary factor in the emergence of geminiviruses. There is a direct correlation between hosts that the B biotype favors for colonization and the host species infected by emerging viruses in cotton. Also of importance are geminiviruses in cabbage and cole crops, eggplant, melons, okra, pepper, squash, tomato, come of which likely have the potential to infect cotton as well. Indeed, cotton and many vegetables once readily cultivated for a profit in the drier valleys of the Dominican Republic, Guatemala, Jamaica, Puerto Rico, Nicaragua, and most recently, in Brazil and (possibly) Paraguay face production constraints because of this new vector and the impact of indigenous *B. tabaci* which are also important vectors of geminiviruses as well as pests in cotton (Brown et al., 1991; Kraemer, 1966; R. Caballero, pers. comm.).

Cotton leaf crumple disease of cotton was first reported in the irrigated desert southwest in the 1950's (Brown, 1992). This disease was later shown to be caused by a whitefly-transmitted geminivirus, CLCV is typically not economically limiting to production in the region because in most years, although disease is widespread, infection occurs late in the season and does minimal damage (Brown and Nelson, 1984; 1987; Brown et al., 1987; Butler et al., 1986). From the 1950 to the 1980's, declining cotton production in Central America and the Caribbean Basin was attributed, in part, to whitefly-transmitted geminviruses (J. Bird, per comm.). In addition, the establishment of an exotic and highly polyphagous whitefly vector in New World locations beginning in the mid-1980's, referred to as the B biotype, resulted in enormous agricultural losses (Brown, 1994; 1996; Brown and Bird, 1992; Brown et al., 1995a, b,c).

In the Eastern Hemisphere, cotton production has experienced increasing threats from diseases caused by locally occurring geminiviruses that are transmitted by the indigenous *B. tabaci* populations (Brown, 1992; 1998). To date, the B biotype has not been reported as a pest or vector in the Eastern Hemisphere, except as a vector of a geminivirus of watermelon in Yemen (Brown et al., 1995a; D. Walkey and P. Jones, personal communication). Among the additional countries with virus problems associated with *B. tabaci* populations are Cameroon, India, Malawi, Mali, South Africa, Sudan (Idris, 1990), and Pakistan (J. K. Brown, 1998, APS Cotton Compendium, in press) though for the most part, these viruses are generally poorly studied. In Pakistan, at least one whitefly-transmitted geminivirus was described from cotton and it is attributed as a primary causal agent in the 1993-1995 epidemic (Mansoor et al., 1993). It is likely that many distinct virus species infect cotton on a worldwide basis, though little work has been done in this important area until recently. Consequently, there is much to be learned about the identity, distribution, and potential threats these emerging viruses pose to cotton production.

The increase in *B. tabaci* population levels in agroecosystems has clearly facilitated higher rates of geminivirus transmission in cultivated and weed hosts, thereby potentially increasing the number and distribution of sources of virus inocula that survive during and between cropping seasons. The end result is an increase in the baseline levels of geminivirus inoculum in agroecosystems and economically limiting constraints on production due to increased costs of controlling whitefly populations and accompanying yield losses in heavily diseased fields. Most problematic in devising measures to control these new diseases is the paucity of knowledge about the biotic and genetic characteristics of the most problematic viruses, and their molecular epidemiology. For example, most viruses remain unidentified and in general have never been studied. As a result, information about their distribution, host range, and virus-vector relationships are not known.

Information is also lacking about the biological and genetic variation within and between whitefly vector populations, and about the distribution of those that are particularly troublesome. Indeed, the inability to distinguish between

different vector populations, despite the recognition that unique biotic characteristics are associated with particular populations that affect host range and virus spread, has been a profound problem for pest management programs. Though recent studies point to considerably variability with respect to whitefly vector host range, vectoring capacity, reproductive abilities, and resistance to insecticides, information relevant to control by targeting particular populations is generally unavailable for most whitefly vector populations. Given that *B. tabaci* populations exhibit different biotic characters but cannot be readily distinguished using morphological traits, identification based upon relevant molecular variation is necessary.

Discussion

In general, WFT geminiviruses of cotton are poorly studied, despite their recent emergence as important virus pathogens. Factors that have hindered their characterization in the past were the need to rear the whitefly vector for experimental virus transmission, the limitation of geminiviruses to phloem tissues making traditional virus purification difficult, an abundance of secondary products and polysaccharides in cotton leaves that interfere with virus isolation, and their characteristically narrow host ranges that limit the potential to discover alternative, less recalcitrant host species for experimental studies. A review of the literature on virus diseases of cotton (Brown, 1992; 1998) confirms that whitefly-transmitted geminiviruses are the most economically limiting of virus pathogens of cotton, yet are among the most poorly studied of cotton pathogens on a global basis.

Recent advances in the application of molecular biological methods to the characterization of geminiviruses have facilitated the investigation of cotton infecting geminiviruses. New approaches involve the application of polymerase chain reaction (PCR) and universal subgroup III that amplify the middle or core region of the coat protein gene (Wyatt and Brown, 1996), or geminivirus-specific primers to direct the amplification of key regions of the virus genome (Idris and Brown, 1998). Partial genome sequences of relevant genes or genomic regions are useful for establishing virus identity and relationships between viruses in the absence of a complete genomic sequence, the latter, a time-consuming and arduous task. Used in conjunction with biotic information about the isolate, and when compared for multiple viruses, coat gene and LIR/CR sequences provide important clues about relationships between whitefly-transmitted geminiviruses, as based upon the specific targeted or marker sites within the viral genome.

The geminivirus coat protein gene (AV1) and large viral intergenic/common region (LIR/CR) sequences have been shown to be useful for establishing the relative identity of geminiviruses (Brown et al., 1998, in preparation). The coat protein gene is useful because it contains sequences that are highly conserved and regions that are variable to such a degree that phylogenetic inferences can be correlated to biotic and geographic facts, making these sequences 'informative' (Padidam et al., 1995). The coat contains sequences that are conserved to function in the formation of the characteristic 'geminate' coat protein that encapsidates the ssDNA genome, plays a role in virus movement in the plant (Pooma et al., 1996), and is required for vector-mediated transmission (see review and refs in Brown, 1996).

The large intergenic region is considered to be an informative sequence of the geminivirus genome because it contains viral regulatory sequences essential for the disease cycle (Eagle and Hanley-Bowdoin, 1997) and for potentially interacting with other geminiviruses when they occur in a mixture in the same host, a possible means by which additional genotypes can emerge, possibly with distinct biological properties (Arguello-Astorga et al., 1994). Specifically, these sequences are important for predicting the likelihood of cross-replication, or pseudo-recombination, between compatible, and therefore, closely related isolates (in question), and is conserved at key sites to perform essential functions in the disease cycle. Regions involved are not entirely clear, but several have been implicated in functions key to virus disease cycle. Among these are modules of direct or inverse repeated or 'iterated' sequences (sequences, directionality, and the specific number of repeated iterons) (Arguello-Astorga et al., 1994) thought to be involved in binding of rep protein/host factor complexes during viral replication, a guanine-rich or G-box postulated to be involved in binding or orientation of the viral replication (Rep) protein and which inadvertently also regulates transcription of the Rep gene (AC1), and the nanonucleotide sequence TAATATTAC conserved in all geminiviruses, that contains the origin of replication (Lazarowitz et al., 1992; Eagle & Hanley-Bowdoin, 1997). The nick site or the site at which the circular DNA is linearized for replication is between the A and C of the nanonucleotide (Heyraud et al. 1995).

Alignment of DNA sequences and analysis of sequences based upon calcuated distance, by parsimony analysis to predict phylogenetic relationships, or analyis using other sophisticated algorithms to evaluate degrees of divergence and evolutionary histories that are based upon complete or partial sequences permits the construction of a 'tree' or dendrogram that depicts those relationships in a lineage, or tree. These lineages can be used to infer evolutionary relationships that provide an indication of the number of nucleotide changes or differences between different taxa in question. These trees provide important clues about the geographic origin of particular geminiviruses, and about biological and genotypic characteristics that are necessary to establish virus distribution maps and to explain complex disease epidemiologies.

PCR primers were designed previously to amplify viral chromosome fragments by taking advantage of sequences that flank the LIR/CR which contains sequences conserved or identical in all subgroup III viruses (Idris and Brown, 1998). Obtaining the same LIR sequences from an A and B viral component indicates the virus has a bipartite genome organization and that these sequences are likely from cognate components, or of the same virus genome. In contrast, the presence of a single component indicates a monopartite organization. Geminiviruses in the Eastern Hemisphere typically contain one or two chromosomes, whereas viruses originating in the Western Hemisphere thus far all contain two chromosomes. These PCR diagnostics, then, provide important clues to geminivirus identity and geographic origin, as well as insights into virus relationships.

Phylogenetic analysis of geminiviruses based upon the coat protein or LIR/CR sequences position the most closely related viruses within the same 'cluster' or group, whereas, those that are not as closely related are placed on different branches with their closest sequence relatives. Inclusion of the leafhopper/ planthopper relatives within the Geminiviridae in the analysis reveals clear separation of whitefly-transmitted viruses from subgroup I and II viruses. Within the whitefly subgroup, the viruses are clearly further separated by geography of origin (Eastern or Western Hemisphere), and at times by a further sub-geographic separation (Brown, 1996; Brown and Wyatt, 1995; 1996; Padidam et al., 1995).

Using this approach, the introduction of a geminivirus from one geographic world region into another can be readily detected. In line with predictive approach for the rational selection of relevant isolates for resistance screening or as sources of viral genes toward pathogen-derived resistance, we are examining the prospects of applying phylogenetic relationships to development of disease resistant germplasm with well-defined breadths of virus resistance. In this approach, it is possible to make predictions about the genotype and numbers of closely or distantly related geminiviruses in germplasm from breeding programs and in plants protected by virus-derived resistance. Though the approaches differ in possible resistance mechanisms, both hope to achieve resistance to as many viruses and strains as possible to accomodate for extant virus pathogens and those that may emerge from extant relatives in the future. We hypothesize that germplasm with resistance to a geminivirus from the same or an adjacent geographic region may also afford protection against other closely related viruses, and provide less or little protection against more divergent viruses from distant geographic regions, the viruses having evolved either under more or less the same conditions and/or possibly from a biogeographically related common ancestor. Consequently, germplasm or transgenic plants expressing the gene of a particular geminivirus genotype will be most likely be effective against most closely related viruses as opposed to those evolving in biogeographic isolation. Clearly, this hypothesis cannot claim *a priori* knowledge about any particular mechanism operating in a resistant genotype, or whether there is a traceable evolutionary basis for a mechanism that is necessarily congruent with geminivirus pathogen evolution. None the less, the broad theoretical and practical utilities of this diagnostic and predictive tool can not be underestimated.

Here, the results of a preliminary analysis of LIR/CR marker sequences by parsimony (PAUP) (Fig 1) and mean calculated pairwise distances expressed as percent similarities are shown, along with iterations identified in the common region of select cotton isolates (Table 1; Table 2). Also shown are the results of parsimony analysis of the coat protein gene (Fig 2.) and corresponding mean % similarities calculated from the pairwise mean distance matrix (Table 3) for select geminivirus isolates from cotton. Distance and phylogenetic analyses of the coat protein gene of these isolates indicate that they have an distinct origin in either an Old or a New World site. For example, cotton leaf curl virus-Pak 1 (CLCuV-Pak1) is without a doubt distinct from all other geminiviruses found thus far in either locale, and is clearly an Old World virus species, quite distinct from the well-studied cotton leaf crumple virus of the southwestern US and Mexico desert. Among the New World isolates examined here, cotton leaf crumple virus (CLCV), and possibly, strains thereof, was documented as the sole geminivirus species in Arizona (CLCVAZCgr94 &95), California (CLCVCalif 94&95), and in Caborca and the Mexicali Valley of Sonora, Mexico (MexCaborcaCYM, MexicaliVLCr 17&18), all three sites in which the disease has been previously documented. In addition, a close relative or strain of CLCV was documented for the first time in Guatemala (Guat57cot94). Also, at least one and possibly two apparently distinct and as yet undescribed geminiviruses of cotton were found in Texas (CotV9Tx61, Txcotgrmos, TxcotLCr, TxMontAl and TxcotLCr & CotV7Tx) and another in Guatemala (Guat2cot). Studies will now be required to substantiate these preliminary findings, including obtaining infectious clones and complete DNA sequences with which to examine the capacity for *trans* activation of replication and movement functions between components of distinct isolates. Additional geminivirus isolates from the Dominican Republic (1992), Brazil (1997-98), Sudan (1997), India (1998), and other cotton growing regions of the world are under presently investigation using this approach. The goal is to provide a comprehensive data set that will permit an increasingly greater global perspective on the distribution of cotton-infecting geminiviruses.

The initial goal of this effort is to determine the identity and map the geographic distribution of the most prevalent geminiviruses of cotton on a global basis. Second, viral sequence data will be used to establish the phylogenetic relationships between cotton-infecting geminiviruses in relation to well-studied geminiviruses, using the analogous sequences as reference sequences having a biotic and

geographic basis. Third, particular regulatory sequences will be evaluated to predict if interactions may be possible between viruses if present in mixed infections, a result that also lends insights to relationships at the virus strain versus species level. The long range goal is to catalog and map whitefly-transmitted geminiviruses of cotton on a global basis, and a establish a rational means for the selection of relevant viruses and/or strains toward developing customized geminivirus disease resistant cotton varieties. The most important virus species will then be selected according to the criteria of association with substantial disease losses and a widespread distribution in cotton. These viruses will be subjected to molecular cloning to obtain a full length infectious virus clones and their complete DNA sequences, thus achieving for the first time, the isolation of these viruses in 'pure culture', reproduction of disease symptoms after inoculating cotton with the infectious clones, and the end result being, characterized geminiviruses of cotton for disease resistance efforts.

Summary

Detailed comparison of select geminiviruses at the level of individual genes or key sequences involved in establishing and completing the virus disease cycle (capsid protein, replicase, regulatory regions, movement proteins) will lend insights into virus diversity and the global distribution of viruses and related strains. This project is accomplishing these goals for the first time by obtaining and utilizing molecular based information obtained directly from cotton infecting geminivirus genomes in conjunction with important biotic characteristics. Knowledge of the identity and geographic distribution of the most threatening geminiviruses in cotton and weed hosts is the first step toward developing sustainable disease resistance in cotton through traditional plant breeding or genetically engineered plant approaches. Selection of relevant geminiviruses and obtaining infectious clones of these viruses will provide a panel of genetically and biologically diverse isolates against which germplasm may be screened to identify resistant genotypes.

These viruses, once cloned, can be sequenced in their entirety to learn more about their relationships to one another (i.e. genotype variability), and will then also serve as sources of geminivirus genes that can be engineered and expressed in relevant cultivars to produce disease resistant transgenic cotton. Virus-derived resistance approaches to disease control rely upon the expression of an inactive or mutated, cloned viral gene in a transgenic plant, the gene having been obtained from the target virus, and hence conferring protection of the territory by being present first. This approach can be thought of as immunization of plants whereby, a mutated form of the protein is engineered and used to 'transform' the plants genome to include the virus gene. When a transgenic plant makes the mutant viral 'gene product' or protein, the presence of the protein infers with the function of that particular virus gene when the virus in inoculated to the plant by the vector whitefly. The result is a plant protected from virus infection. Finally, evaluating the breadth of resistance against our panel of well-characterized virus genotypes provides an excellent opportunity to target customized virus resistance to those cotton growing areas in which the particularly virulent and widespread geminiviruses or strains occur.

References

Arguello-Astorga, G. R., R. G. Guevara-Gonzalez, L. R. Herrera-Estrella, and R. F. Rivera-Bustamante. 1994. Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. Virology 203: 90-100.

Bedford, I.D., P.G. Markham, J.K. Brown and R.C. Rosell. 1994. Geminivirus trnsmission and biological characterization of whitefly (*Bemisia tabaci*) biotypes from different world regions. Ann. appl. Biol. 125: 311-325.

Brown J. K. and J. Bird. 1995. Variability within the *Bemisia tabaci* species complex and its relationship to new epidemics caused by geminiviruses. CEIBA (Zamorano) 36, 73-80.

Brown, J. K. 1994. Current Status of *Bemisia tabaci* Genn. as a pest and vector in world agroecosystems. FAO Plant Prot. Bull. 42: 3-32.

Brown, J. K. 1996. Chapter 5 in: Molecular Biology and Epidemiology of Subgroup III, Geminiviridae. Plant-Microbe Interactions Review Series, G. Stacey and N. Keen, eds (Chapman and Hall) pp 125-195.

Brown, J. K. 1988. Plant Virus and MLO-like Diseases of Cotton in: The Cotton Compendium. American Phytopathological Society, St. Paul, Mn., T. Kirkpatrick and C.S. Rothrock (Eds) (in press).

Brown, J. K., D. Frohlich, and R. Rosell, 1995a. The Sweetpotato/Silverleaf Whiteflies: Biotypes of *Bemisia tabaci* Genn., or a Species Complex? Ann. Rev. Entomology 40: 511-534.

Brown, J. K., J. Bird, D. R. Frohlich, R. C. Rosell, I. D. Bedford, and P. G. Markham. 1995b. The relevance of variability within the *Bemisia tabaci* species complex to epidemics caused by Subgroup III geminiviruses. Pages 77-92 in: *Bemisia '95: Taxonomy, Biology, Damage, Control and Management* (D. Gerling and R. T. Mayer, eds). Intercept Publications, Wimborne, UK.

Brown, J. K., S. Coats, I . D. Bedford, P. G. Markham, J. Bird, and D. R. Frohlich. 1995c. Characterization and distribution of esterase electromorphs in the whitefly,

Bemisia tabaci (Genn.) (Homoptera:Aleyrodidae). Biochemical Genetics: 33: 205-214.

Brown, J.K. 1992. Virus Diseases of Cotton, pages 275-330 in: Cotton Diseases. R. J. Hillocks, ed. Commonwealth Agricultural Bureaux International, Oxon, United Kingdom. 415 pp.

Brown, J.K. and J. Bird. 1992. Whitefly-transmitted geminiviruses in the Americas and the Caribbean Basin: past and present. Plant Dis. 76:220-225.

Brown, J.K. and Nelson, M.R. 1984. Geminate particles associated with cotton leaf crumple disease in Arizona. Phytopathology 74:987-990.

Brown, J.K. and Nelson, M.R. 1987. Host range and vector relationships of cotton leaf crumple virus. Plant Dis. 71:522-524.

Brown, J.K., Mihail, J.D., and Nelson, M.R. 1987. The effects of cotton leaf crumple virus on cotton inoculated at different growth stages. Plant Dis. 71:699-703.

Butler, G.D., Jr., Brown, J.K. and Henneberry, T.J. 1986. Effect of cotton seedling infection by cotton leaf crumple virus on subsequent growth and yield. J. Econ. Entomol. 79:208-211.

Eagle, P. A. and L. Hanley-Bowdoin. 1997. *cis* Elements that contribute to geminivirus transcriptional regulation and the efficiency of DNA replication. Journal of Virology 71: 6947-6955.

Heyraud, F., S. Schumacher, J. Laufs, S. Schaefer, J.Schell, and B. Gronenborn. 1995. Determination of the origin cleavage and joining domain of geminivirus rep proteins. 1995. Nucleic Acids Res. 23: 910-916.

Idris, A. M. 1990. Cotton leaf curl virus disease in Sudan. Med. Fac. Lanbow, Rijksunir, Gennt, 55, (2a), 1990.

Idris, A. M. and J. K. Brown. 1998. Sinaloa tomato leaf curl virus: biological and molecular evidence for a new subgroup III geminivirus. Phytopathology (accepted for publication).

Lazarowitz, S. G., L. C. Wu, S. G.Rogers, and J. S. Elmer. 1992. Sequence-specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. 1992. Plant Cell 4:799-809.

Mansoor, S., Bedford, I. D., Pinner, M. S., Stanley, J., and Markham, P. G. 1993. A whitefly-transmitted geminivirus associated with cotton leaf curl disease in Pakistan. Pakistan J. of Botany 25: 105-107.

Padidam, M., R. N. Beachy, and C. F. Fauquet. 1995. Classification and identification of geminiviruses using sequence comparisons. J. Gen. Virol. 76: 249-263.

Pooma, W. W.K. Gillette, J. L. Jeffrey, and I. T. D. Petty. 1996. Host and viral factors determine the dispensibility of coat protein for bipartite geminivirus systemic movement. Virology 218: 264-268.

Swofford, D. 1993. *PAUP*: Phylogenetic Analysis Using Parsimony, version 3.1.1. Washington D.C.: Smithsonian Institution.

Wyatt, S. D., and J. K. Brown. 1996. Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. Phytopathology 86: 1288-1293.

Table 1. Whitefly-transmitted geminivirus isolates that were PCR-positive for presence of subgroup III gemniviruses and the * denotes those used in this study

Isolate Designation	Symptom in Cotton	Geographic Source	Plant	Year
abmvroth		West Indies	*Abutilon*	1850?
azhibis93		AZ, USA	hibiscus	1993
clcvaz	leaf crumple	AZ, USA	cotton	1982
clcvazcgr94*	leaf crumple	AZ, USA	cotton	1994
clcvazcgr95*	leaf crumple	AZ, USA	cotton	1995
mparvbmac97		AZ, USA	*M. parviflora*	1997
DrcotMosaic		Dom Rep	cotton	1992
egyptcot95	leaf curl	Egypt	cotton	1995
guat2cot94*	leaf crumple	Guatemala	cotton	1994
guat57cot94*	leaf crumple	Guatemala	cotton	1994
guatabut94		Guatemala	*Abutilon*	1994
guatcot294*	mosaic	Guatemala	cotton	1994
guatcot92ym	yellow mosaic	Guatemala	cotton	1992
guatmalv weed22*		Guatemala	Malv weed	1994
guatsidaspp*		Guatemala	Sida spp.	1994
hibis94		AZ, USA	hibiscus	1994
indiacot96	leaf curl	Punjab, India	cotton	1996
mexcabcym96*	yellow mottle	Sonora, Mex	cotton	1996
mexcablcr96*	leaf crumple	Sonora, Mex	cotton	1996
mexicalivlcr96*	leaf crumple	Sonora, Mex	cotton	1996
mx94okra		Mexico	okra	1995
pakclcuv*	leaf curl	Faisal, Pak	cotton	1993
pakcot	leaf curl/mild	Pakistan	cotton	1996
prrmv	veinal yellows	Puerto Rico	*R. minima*	1997
sudan okra	leaf curl /suspect	Sudan	okra	1995
sudan okra	leaf curl /suspect	Sudan	okra	1997
sudcot	leaf curl	Sudan	cotton	1996
texcot92	Stunting	TX, USA	cotton	1992
texcot96	Mosaic	TX, USA	cotton	1996
texkenaf96		TX, USA	kenaf	1996
tx28coty grmos94*	Yellow, green moss	TX, USA	cotton	1994
txcot29LCr94*	Mosaic, stunting	TX, USA	cotton	1994
txcotLCr96*	leaf curl	TX, USA	cotton	1996
txcot796*	leaf curl, mosaic	TX, USA	cotton	1994
txmontal23cot*	Mosaic, leaf cupping	TX, USA	cotton	1996
txsida		TX, USA	*Sida glabra*	1996
txsida93		TX, USA	*Malvastrum*	1993

Table 2. Range of mean pairwise distance comparisons of large intergenic/common region sequences for selected geminivirus isolates from cotton.

Isolate Iterations # Taxa in LIR/CR	Greatest% Similar to Reference Gv	Least % Similar to Reference Gv
CLCVAz N=3	54.1-54.5 TTMoV	42.3-42.6 TLCrV 43.6 ACMV
CLCVCab N=2	54.1-54.5 ToMoV 54.9 BGMV(PR&Jam)	43.2 TLCrV 43.6 ACMV
CLCVMexV N=5	53.4 TTMoV 54.3 ToLCVCh	43.0-43.2 TLCrV 43.7 ACMV
Texcot-1 (Leaf Curl) n=2	54.1 TTMoV 54.8% BGMV(PR&Jam)	41.6 TLCrV, ACMV
Texcot-2 Mosaic/stunting n=4	56.6% AbMV (WI)	40.8% ACMV

Table 3. Range of mean pairwise distance comparisons of coat protein gene (*AV1*) sequences for selected geminivirus isolates from cotton.

Isolate/Range # Taxa	W/I Range of Each Taxon % Similarity	Greatest and Least % Similar with Reference Gv
CLCVAz n=4	.001-.012	98.9-98.8
CLCVCalif n=4	.011-.030	97-98.9
CLCVCab n=4	.001-.003	99.7-99.9
CLCVMexV n=2	.003	99.7
Guat2 n=2	.008	99.2
Guat57 n=1	-	-
Texcot-1 (Leaf Curl) n=2	.009	99.1
Texcot-2 Mosaic n=4	.000-.012	0-98.8

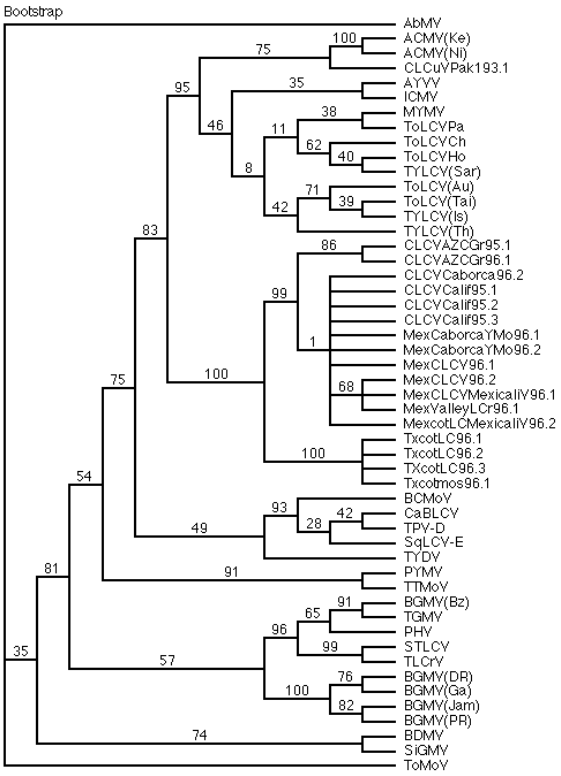

Figure 1. Single most parsimonious tree (unrooted) for the large intergenic region (LIR) relationships of 24 uncharterized virus isolates and reference taxa of the Geminiviridae. Maize streak virus (subgroup I) and beet curly top virus (subgroup II) are included as outgroups. Bootstrap sampling conditions 10 replicates for each of 100 iterations and the heuristic search option with TBR branch swapping. Bootstrap percentages are indicated on the nodes.

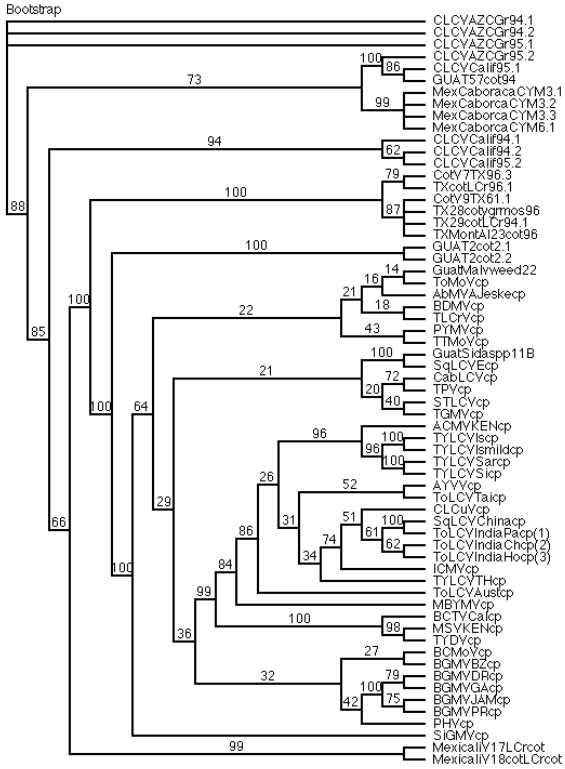

Figure 2. Single most parsimonious tree (unrooted) for the coat protein gene (AV1) relationships of 24 uncharterized virus isolates and reference taxa of the Geminiviridae. Maize streak virus (subgroup I) and beet curly top virus (subgroup II) are included as outgroups. Bootstrap sampling conditions 10 replicates for each of 100 iterations and the heuristic search option with TBR branch swapping. Bootstrap percentages are indicated on the nodes