Global Diversity and Distribution of Cotton-Infecting Geminiviruses: An Essential Requisite to Developing Sustainable Disease Resistance

Judith K. Brown, University of Arizona, Tucson, USA

Introduction

Until quite recently, whitefly-transmitted (WFT) geminiviruses were restricted primarily to weed hosts in the subtropics and tropics, with rare exceptions. For example, in Africa, cassava mosaic virus (ACMV) has been problematic in cassava since the turn of the century, primarily because the virus has been

maintained in cropping systems by the practice of vegetatively propagating infected cassava from slips. In Sudan, since the 1950s, leaf curl of cotton was reported as an annual disease, though the magnitude of yield loss varied from year to year. Tomato-infecting geminiviruses were documented in India and the Middle East at about the same time and were periodically yield limiting. In the Western Hemisphere, geminivirus incited diseases were of little importance until the 1960s when disease outbreaks occurred in bean and tomato crops in Brazil, the Caribbean Basin, Mexico, and Venezuela. However, by the 1980s, geminiviruses became recognized as emergent virus pathogens on a global scale, and are of widespread importance in cotton-vegetable agroecoystems.

Bemisia tabaci (Genn.), known as the cotton or sweet potato whitefly, is the sole whitefly vector of geminiviruses. Until the 1960s, *B. tabaci* was recognized as an infrequent pest and virus vector, and hence, was considered of little consequence to crop production. The increased importance of geminivirus diseases and their whitefly vector have been exacerbated by two recent events: the nearly global introduction and establishment of an exotic whitefly vector type of *B. tabaci*, the B biotype and the widespread and simultaneous development of insecticide resistance in local whitefly vector populations throughout the subtropics/tropics. High levels of vector populations in these agroecosystems, many only recently expanded with the aid of irrigation projects, has resulted in the mobilization of many new geminiviruses from weed to crop species, and subsequently, routine and damaging diseases caused by geminviruses in cropping systems for the first time. Among the crops most damaged by geminivirus diseases are beans, cotton (*G. arboreum, G. barbadense, G. hirsutum*), all cucurbits, pepper, tomato, and watermelon.

This global upsurge in *B. tabaci* in agroecosystems has greatly facilitated higher rates of virus transmission, subsequently leading to increased disease incidence and a wider biogeographic distribution of geminiviruses. Consequently, the growing trend is toward increased baseline levels of virus inoculum in agroecosystems leading to higher disease incidences and unprecedented yield losses due to infection by essentially unknown or poorly studied plant viruses. Therefore, most problematic to controlling these new virus pathogens is a general lack of knowledge about their identity, distribution, and host range, and about the biogeographic diversity and distribution of whitefly vector populations that can vary considerably with respect to host range, vectoring capacity, reproductive abilities, and frequency of resistance to particular insecticides. Finally, the unavailability of disease resistant crop cultivars does not promise feasible solutions to combat economic losses in the near future.

Background

History of Geminviruses in Cotton

During the 1950-1980s, decreased cotton production in Central America and the Caribbean Basin was attributed in part to whitefly-transmitted geminviruses (J. Bird, per comm.). In addition, following the establishment of the B biotype whitefly, enormous losses were felt due to feeding damage and honeydew contamination of lint, and subsequently to appearance of many new geminivirus diseases. Indeed, cotton production in Puerto Rico, the South Coast of Guatemala, and Dominican Republic is now either non-existent or greatly reduced, and in Nicaragua, although production continues, geminiviruses are a growing constraint (Brown et al., 1991; Kraemer, 1966; P. K. Anderson, pers. comm.). Several geminivirus diseases of cotton have been described in Brazil and possibly, in Paraguay (Costa, 1976), but these viruses are as yet uncharacterized. In the US (Arizona and California) and northern Mexico (Sonora and Sinaloa), cotton leaf crumple disease occurred sporadically since the 1950s when it was first described, however, typically late season infections have precluded substantial crop losses in most years. Cotton-infecting viruses associated with whitefly infestations have also been reported in Texas since the 1950s, and although preliminary data suggest they are distinct from leaf crumple, these viruses are also poorly studied.

The most worrisome virus diseases of cotton occur primarily in the Eastern Hemisphere. Particularly serious losses have been documented in India, Pakistan, and in Sudan, Egypt, and several other African countries. In this part of the world, geminivirus diseases in cotton began to escalate in the 1980s, increasing steadily in distribution since then. Whitefly-transmitted geminivirus diseases of cotton now occur annually in these areas and widespread epidemics have become a common occurrence (Brown, 1992; Idris, 1990; Mansoor et al., 1993; Varma, pers. comm.). For example, in Pakistan in 1993/94, about two million hectares of cotton were infected by cotton leaf curl virus and production was reduced to 1.4 million tons from an all time high of 2.2 million tons in 1991/92 (ICAC estimates). In 1994, leaf curl, or a related geminivirus, first threatened cotton production in the Indian Punjab, while at the same time, there were increased reports of cotton leaf crumple virus in Mexico and the SW US. Although geminiviruses were likely present or are 'indigenous' to these areas, the contemporary diseases in cotton appear to be more widely distributed and to cause substantially more damage than their previous counterparts, perhaps because varieties grown for contemporary markets lack disease tolerance or resistance that may have been a characteristic of varieties grown in the past. Nonetheless, the effects of these diseases are compounded by rising vector populations and the likely possibility that new geminiviruses may yet emergence. Increased disease pressures and whitefly populations have made necessary heavy applications of insecticides to reduce whitefly populations, pointing to the urgent need for virus disease resistant cotton varieties.

The rising costs of managing the whitefly vector, coupled to losses due to geminivirus diseases now hinder cotton produc-

tion by the demand for inputs beyond economic feasibility. The clear need for virus-resistance in high yielding, high quality cotton varieties that are tailored to disease prone areas presents a new challenge. If this challenge is not met, there is no certainty that cotton production at present day levels will be possible, nor is there a guarantee that market demands will be achieved in the absence of crop sustainability. To meet the growing demand for geminivirus resistant cotton varieties and to achieve their development, baseline information is required concerning the identity, the distribution, and the relevant characteristics of the most threatening geminivirus pathogens of cotton. Only then, can germplasm be developed with sustainable disease resistance against the specific geminivirus pathogens in a region. Also essential is sound knowledge about the range of virus strains against which cultivars are effectively protected, and about the distribution of commonly occurring geminviruses, given that tolerant or resistant cultivars will be highly desirable in other cotton growing regions where similar and possibly distinct strains of geminivirus pathogens and whitefly vector complexes occur.

The Impact of *B. tabaci* Vector - Geminivirus Complexes

Indigenous viruses such as CLCV in cotton have caused sporadic disease in the southwestern US for over thirty years. Epidemics have typically been associated with mild winters that yield early season whitefly populations and the cultivation of ratooned cotton infected with CLCV during the previous season (Dickson, 1954; Allen et al., 1960; Brown et al., 1983; Erwin, 1959; Russell, 1981). In epidemic years, cotton yields are reduced by 50-80%, particularly in those fields near ratooned cotton which serve as virus and whitefly reservoirs. Although the precise reasons for periodic epidemics of CLCV have not been ascertained, years of high disease incidence are associated with mild winters that promote early season increases in vector populations. Until recently, disease resistant varieties have not been considered worth the economic investment and programs to eradicate ratooning practices have been implemented.

During the late 1970s early 1980s, sporadic geminivirus epidemics in cotton and vegetables were generally regarded as anomalies that would not establish as persistent diseases of annual importance. However, the growing inability to reduce whitefly vector populations in cotton-vegetables agroecosystems, partially due to insecticide resistance, rapidly resulted in increased baseline virus inoculum levels in the US and Mexico. The establishment of the exotic Old World B biotype whitefly in SW cotton in 1988-90 resulted in such dramatic direct feeding damage and honeydew-contaminated lint that losses due to CLCV went largely unnoticed. Today, there is growing awareness of the persistence of this disease and incidence has risen over time. At about the same time, similar upsurges in local (or endemic) whitefly vector populations led to serious disease situations in many world locations, and cotton leaf curl virus emerged as an economic threat to production in Pakistan. In addition to the rising impact of locally occurring virus diseases, it is also feared that new geminiviruses will either emerge or be introduced from neighboring locales due to the widespread cultivation of high-yielding, disease susceptible varieties. These predictions are driven by observed patterns of increasing virus disease incidences in the US and the Caribbean Basin following the establishment of the highly prolific B biotype whitefly vector. Indeed, such trends have been born out by recent reports of geminiviruses in cotton in Guatemala (Brown et al., 1993), India (Varma, pers. comm.), Nicaragua (R. Caballero, pers. comm.), Pakistan, Paraguay, The Dominican Republic (Brown et al., 1991), and the US-Texas, (Brown, unpublished).

The global increase in whitefly pressures and local, or 'indigenous' WFT geminiviral pathogens (Brown, 1990; Brown and Bird, 1992), suggests that epidemics caused by WFT geminiviruses will continue to rise. Clearly, WFT geminivirus epidemics have already become routine in certain vegetable-cotton agroecosystems, and although information is incomplete, these viruses are likely harbored in a variety of plant genera or species that may be related or unrelated to cotton. As a result, substantial efforts have been underway to achieve stringent control of the whitefly vector, combining new insecticide chemistries with biological control agents. As yet, there has been little effort toward reducing the impact of geminivirus diseases through the development of host plant resistance. Indeed, the documentation of many new and emerging WFT geminiviruses in cotton has launched an effort to critically examine the identity, distribution, and biogeographic and genetic variation among cotton-infecting geminiviruses, rapidly becoming of importance on a worldwide basis.

Geminiviruses of Cotton

In general, WFT geminiviruses have been poorly studied, despite their recent emergence as important pathogens. Factors that have hindered their characterization include the requirement to rear the whitefly vector for experimental virus transmission in the laboratory, their limitation to phloem tissues which makes virus particle 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. The overall knowledge about virus diseases of cotton has been reviewed recently (Brown, 1992; 1997); and it is quite clear that whitefly-transmitted geminiviruses are among the most limiting and poorly studied pathogens of cotton on a global basis.

Biological characteristics such as virus host range, disease symptomatology, and virus-vector relationships remain ill-defined

for most geminiviruses of cotton. Among the best studied is the CLCV from Arizona, which has a narrow host range within the Malvaceae and Leguminosae (Brown and Nelson, 1984). The discovery that common bean was an alternate host of CLCV facilitated visualization of geminivirus particles by transmission electron microscopy, resulting in the first confirmation that the CLCV was morphologically like other geminiviruses. This unique virion morphology is now accepted as a sole trait of the Geminiviridae, together with that of a single-stranded, circular DNA genome.

Recent advances in molecular cloning have facilitated a preliminary study of viral gene sequences, required to achieve virus identification. In this study, the conserved viral coat protein gene of CLCV from Arizona was compared to that of an isolate causing leaf distortion of cotton in Texas, and two isolates from Guatemala that cause leaf curl or yellow mosaic symptoms. Results indicated that at least three distinct viruses were involved, and that there may be several strains or close relatives of CLCV in the Americas. However, much remains to be learned about the similarities and differences among these New World isolates. The only other cotton virus for which genetic level information is available is an isolate of CLCuV from Pakistan. A geminivirus has been cloned from infected cotton and the genome has been partially sequenced. But, in the absence of additional virus sequences needed for comparison, it was not yet possible to predict the relationship of the Pakistan leaf curl virus to others from other world areas.

Knowledge about the genetic variation and relevant biological characteristics of WFT geminiviruses are now imperative to permit identification and to discover the distribution and degree of importance that geminiviruses pose to cotton production. A panel of well-characterized viruses are needed as sources of virus diversity against which germplasm may be screened. These same viruses, once cloned, can be sequenced in their entirety to learn more about their relationships to one another (i.e. genotype variability), and can also serve as sources of virus genes that can be engineered to produce disease resistant transgenic cotton using virus-derived resistance. Detailed comparison of select geminiviruses at the level of individual genes or key sequences involved in regulating the virus disease cycle (capsid protein, replicase, regulatory regions, movement proteins) will lend insights toward virus diversity and the global distribution of viruses and related strains utilizing molecular based information that is now accessible for the first time. Knowing who and where the most serious viruses are is clearly the first step toward developing sustainable disease resistance in cotton, irrespective of traditional plant breeding or genetically engineered plant approaches.

The Whitefly Vector

Members of the whitefly *B. tabaci* species complex are the only known vectors of subgroup III (whitefly-transmitted subgroup) geminiviruses, worldwide (Bedford et al., 1994). This whitefly has a potentially broad host range among (Brown et al., 1995) with certain populations expressing preferences for cotton, vegetables, and ornamentals. Geminiviruses are transmitted in a persistent, circulative manner by their whitefly vector, meaning that once virus is taken up from a host plant, transmission can occur within several hours, and continuously for the life of the vector. This type of virus-vector relationship makes persistently transmitted viruses as the most difficult to study, as well as to manage in the field. Possible future strategies for disease management also involve interference with this highly specific virus- vector relationship, however, at present, little is understood about the mechanisms involved in this process. A greater understanding of the mechanisms governing whitefly-mediated geminivirus transmission will lead to possibilities for interfering with those processes, feasibly through expression of anti-transmission factors in transgenic plants. Such factors, when expressed as transgenes in transgenic plants, are envisioned to 'neutralize' transmission by mimicking the virus and thereby, binding to sites in the whitefly vector that are essential to whitefly-mediated transmission. Saturation of essential sites with 'modified virus' will preclude virus binding, and hence, interfere with transmission.

Studies devoted to investigating the importance of host races, strains, or biotypes within the *B. tabaci* species complex have provided important insights necessary for managing whitefly and disease problems. These efforts will continue to shed new light on our understanding of the morphologically identical members of the *B. tabaci* complex are surprisingly, 'cryptic' in that they may differ entirely in such biological characteristics as host range, vector capabilities, fecundity, and insecticide resistance. Certain widespread and troublesome populations of the *B. tabaci* complex, for example, the B biotype, can be identified and tracked for the first time, using molecular gene markers, or biotype-specific DNA sequences that serve as an important fingerprint (Brown et al, in preparation). Sound information is available about the biological characteristics of this population, and can be obtained for others. Indeed, had it been understood that although *B. tabaci* is morphologically the same regardless of its biogeographic niche, while at the same time genetically heterogeneous and having likewise distinctive biological traits, the introduction of the B biotype could have possibly been prevented or at least recognized much sooner. The concept of whitefly biotypes or strains along with similar studies of virus strains and quasi-species also requires rigorous investigation at the biological and molecular levels in order to explain presently puzzling variation observed in whitefly vectoring capacities with certain geminivirus-vector-host complexes. Presently, however, very little is know about biology and genetic diversity for most whitefly biotypes or geminivirus strains or about the nature of the specific interactions that lead to either high, moderate, or low frequency transmission events.

Whether there are vector biotypes with a capacity to more efficiently transmit certain geminiviruses to and from cotton is not known. Also relevant to vector characteristics that affect transmission of geminiviruses is knowledge of unique behavioral characteristics that may be peculiar to a particular whitefly race or biotype, for example, long distance dispersal behavior and extremely broad host range (exhibited by the B biotype, for example) coupled to vectoring capacity and resistance to insecticides, that must be taken into account in implementing successful biological control programs that rely on release and/or enhancement of whitefly parasitoids, fungal or viral pathogens or predators to reduce whitefly populations.

Geminivirus Detection and Identification toward Disease Control

Very little is known about the identity or the distribution of the most geminivirus pathogens of cotton. Toward this end, the Arizona (JKB) laboratory has developed a polymerase chain reaction (PCR)-based method that permits highly sensitive detection of a gene fragment present in all whitefly-transmitted geminiviruses. Amplification of this diagnostic fragment (550 bp) and its visualization by agarose gel electrophoresis is the first available assay that confirms the absence or presence of geminivirus infection in plants. Obtaining the specific nucleotide sequence of the virus gene fragment readily permits a comparison of this viral sequence with sequences from well-characterized geminiviruses, providing an invaluable identification tool. Using this approach, we are archiving coat protein gene sequences of geminiviruses by compiling a database that can be accessed on the World Wide Web. After matching an input sequence to the most closely related sequence in the data base, relevant information about that virus and other close or more distant relatives can be located in data base linkages. This interactive component will permit the first comparative virus identification of geminiviruses of cotton. Information about biological and genetic diversity amongst these viruses can also be applied to study or track viruses over broader geographic areas and crop species. This is necessary because there is scant information concerning crops and/or weeds that serve as virus reservoirs to bridge the disease cycle. Further, there is no capacity for testing germplasm in resistance trials in the field or greenhouse with well-defined virus genotypes selected from disease prone areas. Reliable laboratory-based detection and identification methods are the cornerstones of developing disease resistance.

The following is a summary of the research undertaken in the US laboratory during 1996-97 as supported by the Common Fund for Commodities in a collaborative effort between scientists at John Innes Centre and collaborators at NIBGE, Faisalabad, Pakistan under the direction of Dr. Kauser Malik, Director of NIBGE. Supplemental funding for portions of this work has been provided by Cotton Incorporated, Raleigh, NC, USA, and through a donation by Delta and Pine Land Company, Scott, MS, USA.

DNA Sequence Database and Phylogenetic Inferences for Whitefly-Transmitted Geminiviruses of Cotton: Towards Disease Resistance

Project Activities and Progress to August, 1997

Objective 1

Subgroup III Geminivirus Isolate Collection

Shown in Table 1 are the geminivirus isolates of cotton currently obtained for archiving and for molecular evaluation in the Arizona laboratory. We have thus far, obtained isolates from the Western Hemisphere: US (Arizona, California, Texas), Mexico (Sonora), Guatemala, Puerto Rico, and The Dominican Republic. In 1997 we expect to obtain isolates from Brazil, Nicaragua, Paraguay, and from other cotton growing regions where viruses diseases prevail. The Eastern Hemisphere isolates we have obtained include viruses from India, Pakistan, and Sudan. We continue efforts to obtain additional isolates through the assistance of colleagues, cotton breeders, commercial companies, and travel to specific locations when opportunities arise.

Rationale

The continual acquisition of geminivirus isolates from affected areas is absolutely essential to achieve representative Old and New World viruses for the geminivirus collection archive, working collection, and for the WWW:// cotton geminivirus database (GEMINI DETECTive). The virus collection and database will be used for virus identification by comparison of input virus sequences and to classify viruses based upon phylogenetic comparison to yield a differential panel of virus genotypes for disease resistance efforts. Defining comparative genetic variability and relationships of geminiviruses and implementing sequence data to track virus distribution will permit the first global map of geminivirus genotype distribution in cotton. This tool is a key component of an optimal strategy toward resistance that will ultimately permit varieties to be tailored to resist specific viruses that cause diseases in the particular locale. This approach relies upon a capacity to select specific geminivirus isolates representative of those particular geographic areas for challenging germplasm in traditional breeding programs and as sources of virus genes in transgenic approaches to disease resistance. This initial goal involves the development of effective methods to determine the range and nature of virus genotypes that infect cotton germplasm. This objective has direct application toward developing virus-specific and/or broad spectrum resistance in cotton intended for specific or multiple sites where particular geminiviruses are

known to be yield-limiting pathogens of cotton.

Objective 2

PCR of Geminivirus Coat Protein Fragment for Detection and DNA Sequences for Virus Identification

Development of PCR Primers and Universal Detection of Whitefly-Transmitted Geminiviruses in Cotton

We have developed the first method that permits universal amplification of a geminivirus coat protein gene fragment 550 bp in size from all geminivirus-infected plants (Wyatt and Brown, 1996). Diagnostic size viral gene fragments have been cloned and the DNA sequence has been determined using automated technology. Geminivirus sequences are compared with those of previously studied viruses previously obtained, and entered into the Arizona GEMINI-DETECTive data base on the world wide web. Comparison of input sequences with those of well-characterized viruses provides a rapid, accurate, user-friendly approach for identifying whitefly-transmitted geminiviruses, and for establishing the discovery of a previously undescribed virus. New sequences are added to the database and the virus is further studied in biological and molecular terms. Using this approach, we are presently tracking and mapping the distribution of cotton- infecting geminiviruses, and will ultimately be able to point to the most abundant and widely distributed viruses that can be selected for disease resistance efforts. Consistent sampling over time will allow for the study of multiple samples per each major production zone and hence, accurate information about predominant virus genotypes. This approach will also facilitate the rapid discovery of new or emerging viruses.

Laboratory Analysis of Virus Samples Analyzed to 1996 by Universal Polymerase Chain Reaction and Future Work

All samples listed in Table 1 have been confirmed to contain geminivirus DNA, based upon the diagnostic subgroup III-specific PCR assay described above. From studies thus far, we are discovering many newly emerging geminiviruses in cotton, worldwide, and that most have not been previously studied. Because this is the only global effort of its kind, samples from all cotton growing locations in which suspect geminiviruses are present are needed for analysis. To accomplish this effort, we have solicited the help of many colleagues and hope in the upcoming seasons, to receive additional material. The UA laboratory will also engage in local and regional collecting trips to accomplish this goal, and we will continue to rely on our collaborators in Pakistan, India, the United Kingdom (Rothamsted Experiment Station), and elsewhere throughout the world to assist in obtaining additional samples. Samples will be received in the Arizona laboratory from any colleague in any world location wishing to provide material for this study under our recently awared USDA PPQ permit to import plant leaf material for this effort. Sampling both cotton and nearby weed species are desirable. The UA laboratory has the responsibility to obtain and assemble coat protein gene and viral common region sequences for all cotton infecting geminiviruses into the data base.

Objective 3

Biolistic Inoculation of Cotton with Geminivirus DNA Extracts and Full Length Infectious Clones of Cotton-Infecting Geminiviruses: A Working Collection

We previously developed a simple biolistic method for inoculating seedling cotton with DNA extracts from virus-infected material, and have shown this possible for the virus isolates from Arizona, California, Texas, and Guatemala (see Table 1). This method will be used to inoculate cotton seedlings with infectious clones when available from our laboratory and other laboratories working on this project. While perfecting this method to permit inoculation of cotton seedlings with full length infectious virus clones, we have found greater success when viral genomic clone inserts are excised from the cloning vector prior to inoculation. We have been successful with this method in demonstrating infectivity of infectious clones or DNA extracts containing viral DNA. This will be the preferred method for preliminary evaluation of infectivity of cloned virus genomes, and can also be applied to screening of elite germplasm in late stages of resistance efforts (it is too expensive and time-consuming to use in breeding programs where hundreds to thousands of plants must be screened). Select virus clones could feasibly be engineered into an *Agrobacterium* binary vector to facilitate inoculation of large numbers of plants.

Objective 4

PCR of Target Regions for Phylogenetic Comparisons PCR Amplification, Cloning, and Sequencing of the Core Region of the Coat Protein Gene of Cotton-Infecting Geminiviruses

All samples except those collected in 1996 have been cloned and the DNA sequences obtained for the core coat protein gene. Phylogenetic trees reconstructed from these sequence data indicate that viruses are likely endemic in the locations where they have been collected, and no evidence of introductions of exotic strains or viruses have been detected. Also, it is clear that cotton leaf crumple virus from Arizona is of New World origin (recent), whereas, the Pakistan cotton leaf curl virus is of Eastern Hemisphere origin. All DNA core coat protein sequences obtained thus far, have or will soon be placed into the data base. Ultimately, a file will be compiled for each well characterized isolate that will include the core coat protein gene sequence, geographic information, host range data, photos of symptoms in cotton and relevant hosts, and any other useful information that will facilitate virus identification and disease management in the short term. By November 1997, the GEMINI-DETECTive data base will be accessible on the World Wide Web via the IPM Network, hosted by the North Carolina State University National Science Foundation Center for Integrated Pest Management, under the direction of Dr. Ron Stinner,

Table 1. PCR Positive Whitefly-transmitted Geminivirus Isolates Under Further Study in the Arizona Laboratory

Isolate Designation	Symptom in Cotton	Geographic Source	Source Plant	Year
Sudan okra (cot leaf curl)	minor vein thickening	Sudan	okra	1995
txcot94*	mosaic, stunting	TX, USA	cotton	1994
clcvaz1	leaf crumple	AZ, USA	cotton	1982
cottex92*	foliar stunting	TX, USA	cotton	1992
hibis93*	N/A	AZ, USA	hibiscus	1993
txsida93	N/A	TX, USA	Malvastrum	1993
cot1guat*94	leaf crumple	Guatemala	cotton	1994
cotguat94ver*	mosaic	Guatemala	cotton	1994
cot2guat94*	leaf crumple	Guatemala	cotton	1994
clcvaz2	leaf crumple	AZ, USA	cotton	1993
cotguat92ym*	yellow mosaic	Guatemala	cotton	1992
cotdr*	mosaic	Dom Rep	cotton	1992
mx94okra*	yellow mosaic	Mexico	okra	1995
guatabut94	N/A	Guatemala	Abutilon (weed)	1994
abmvroth	N/A	West Indies	Abutilon (ornamental)	1800s
hibis94*	N/A	AZ, USA	hibiscus	1994
cotguatlc94*	leaf curl	Guatemala	cotton	1994
cotegypt95	leaf curl	Egypt	cotton	1995
cottex96	mosaic	TX, USA	cotton	1996
kenaftex96	leaf curl	TX, USA	kenaf	1996
cotmex96	leaf crumple	Sonora, Mexico	cotton	1996
cotmex96	yellow mottle	Sonora, Mexico	cotton	1996
cotmex96	leaf crumple	Sonora, Mexico	cotton	1996
cotindia96	leaf curl	Punjab, India	cotton	1996
cotpak1	leaf curl	Faisalabad, Pakistan	cotton	1996
cotpak2	leaf curl/mild	Faisalabad, Pakistan	cotton	1996
cotaz	leaf crumple	AZ, USA	cotton	1996
sidatex	mosaic	TX, USA	*Sida glabra*	1996
cotsud1	small vein	Sudan	cotton	1996
cotsud2	big vein	Sudan	cotton	1996
cotcabMx	CLCV	Mexico	cotton	1997
cotRMVPR	Cotton veinal yellows	Puerto Rico	*R. minima*	1997

* Denotes new disease report.

Center Director. As of August, 1997, the first phase of the data base has been accomplished, and the site contains relevant disease information and viral DNA sequences for ten geminivirus isolates that infect cotton and other Malvaceous plants.

Efforts are also underway to clone and sequence PCR products of the large viral intergenic region (LIR), viral regulatory sequences that are nearly identical on interacting A and B viral chromosomes, or are found in a characteristic location on the chromosome of viruses with a single chromosome. PCR can be used to amplify from single or multiple chromosomes by taking advantage of certain sequences flanking the LIR and are conserved or nearly identical in all subgroup III viruses. Obtaining two LIR sequences that are nearly identical that are associated with A and B viral chromosomes from the same plant indicates the virus has two genomic components (bipartite), as do viruses in the Eastern and Western Hemisphere, such as cotton leaf crumple (Brown et al., unpublished). Thus far, several cotton infecting viruses having a single chromosome (monopartite) have been found in the Eastern Hemisphere, and cotton leaf curl appears to be a likely candidate for a single chromosome virus. Inoculating plants with clones that have the same or compatible LIR sequence results in development of disease symptoms, permitting virologists to demonstrate that the virus clones are not only infectious when present at the same time (both are needed to achieve infection of bipartite viruses), but also establishes that the pathogen has been isolated in entirety, and can now be accurately named and subsequently characterized. These compatible clones can then be reliably used to artificially inoculate plants without the whitefly to mediate transmission. Clearly, this approach permits studies of virus host range, symptom phenotype, and the capacity to experimentally inoculate germplasm for resistance screening, or challenge inoculation of transgenic plants engineered for resistance.

This sequence is also an important indication of the genetic nature of a geminivirus to be targeted for transgenic plant mediated virus-derived resistance. 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 a type of 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. Because single and double chromosome-containing geminiviruses have distinct sets of genes with different functions, it is essential to know what type of chromosome organization applies. Hence, it is possible to target a viral gene shared in common between single and double chromosome viruses, or one that is found only in one type of virus organization. The approach described here to examine the coat gene and LIR sequences of geminiviruses, without having to clone and obtain the entire genome first, will permit us to rapidly determine if the pathogen in question has a single or double chromosome, further permitting diagnostic molecular differentiation among uncharacterized geminiviruses.

Marker Sequence-based Predictions of Geminivirus Relationships: An Essential Tool for Resistance Objectives

The DNA sequences of the geminivirus coat protein gene and LIR are useful to establish geminivirus identification. These sequences when compared among multiple viruses can provide important clues about the evolutionary relationships of whitefly-transmitted geminiviruses. Evolutionary histories of viruses can be viewed as 'trees' using aligned virus DNA sequences. Such trees can be used to predict virus relationships based upon sequence similarities and differences, and provide information about the geographic origin of the virus and about its biology. Trees generated with geminivirus coat protein or LIR sequences place most closely related viruses in the same cluster on a branch, while those that are not closely related are placed on a different branch with their closest sequence relative. These two regions of the virus gnome have been shown useful in making predictions about virus identification without cloning and sequencing the entire virus chromosome. Either of these two sequences yield trees that show relationships of viruses by separating viruses (sequences) into clusters or most closely related groupings. The large cluster of whitefly transmitted viruses are separated from two large clusters of related viruses transmitted by leafhopper vectors. Once placed in the whitefly subgroup, viruses are clearly separated by geography of origin (Eastern or Western Hemisphere), and at times by a further sub-geographic separation (Brown, 1996 (abstr); Brown and Wyatt, 1995 (abstr); 1996 (abstr); Padidam et al., 1995). Thus, the introduction of a geminivirus from one geographic world region can be readily detected using this approach. In addition, germplasm that is protected against an Eastern Hemisphere virus could be tested for its ability to protect against infection by other closely geographically related viruses, as well as the more distantly related relatives from another region. Clearly, the broad theoretical and practical utilities of this predictive tool should not be underestimated.

In the large intergenic region are found regulatory sequences that are postulated as important in predicting the likelihood of cross-replication or pseudo-recombination between compatible, and putatively, genetically similar or, the most closely related viruses. Proposed for use with these predictions are families of iterated sequences found in the intergenic region (sequences, directionality, and the specific number of repeated iterons) (Arguello-Astorga et al., 1994), and these viral sequences thought to be involved in binding of rep protein/host factor complexes during viral replication. We are currently obtaining these marker sequences to facilitate predictive inferences of virus genomic subclusters, i.e. clusters containing evolutionarily related viruses of cotton. These data will assist in determining if viral genomic groupings inferred by DNA sequences of viral marker gnome regions can be corroborated with responses of genetically diverse cotton germplasm when infected with our library of prototype isolates. Representative viruses will be selected from those listed in Table 1. We continue to obtain key viral DNA sequences for all isolates with which to infer relationships and further catalog identity and distribution of the viruses.

Objective 5

Cloning and Sequencing of Cotton Leaf Crumple Geminivirus and other New World Geminiviruses

The most prevalent virus in the US is cotton leaf crumple, but several new viruses including several from Texas and Sonora, Mexico appear to be of new importance. These viruses will be cloned and sequences of the entire genomes obtained beginning with the highest priority virus, CLCV. Likewise, the Pakistan laboratory has done extensive work toward cloning cotton geminivirus isolates from their locale. Molecular clones of the most prevalent cotton-infecting geminiviruses will be obtained as components of the working virus collection. As stated, infectious virus clones are required for disease resistance efforts, and to carry out additional investigations of virus-vector biology and the molecular epidemiology of priority viruses.

To this end, we have cloned and nearly completely sequenced the cotton leaf crumple virus from Arizona. In 1997, we will rectify the several areas of sequence discrepancy in the cloned virus. Because symptoms incited by the cloned CLCV are milder than expected, we will re-clone the virus to obtain cloned A and B components of a more typically virulent virus. We have also obtained partial clones of two geminivirus isolates (Brown, 1992; 1994) from the south coast of Guatemala, and will continue to investigate the molecular characteristics of these two isolates. Presently, one isolate appears to be closely related to cotton leaf crumple, while the other appears to be a distinct virus. A Texas-USA isolate from cotton will be the third priority isolate for the New World viruses, while isolates from India and Africa are Eastern Hemisphere priorities. To date, it has

been impossible to obtain quality extracts for such work with Indian or African (Sudan, Mali, South Africa) isolates, however, efforts are underway to circumvent this problem.

Related Scientific Cooperation and Travel for the Project

Pakistan

To obtain important collections virus isolates associated with the recently reported leaf curl outbreak in the Indian and Pakistan Punjab regions, the US team traveled to Pakistan and India in November 1996. In Pakistan, we visited with the National Institute of Biotechnology and Genetic Engineering (NIBGE), met with research scientists and the director of the institute, Dr. Kauser Malik. During our visit, we also obtained virus samples, reviewed research progress in the US and Pakistan laboratories, and exchanged information on the status of cotton leaf curl in Southern Asia. This travel was supported primarily by the Common Fund for Commodities-Cotton and in part by funds from USDA/OICD/FAS for scientific exchange.

India

Important contacts were made at Punjab Agricultural University, Ludhiana, Punjab, India with Dr. A.S Khehra, Vice Chancellor, and Dr. K.S. Aulakh, Director of Research, and with Dr. L.S. Randhawa, Principal Cotton Breeder. Dr. Randhawa is interested in collaboration on cotton virus/student exchange/supply of DNA extracts/possible joint publications. Dr. Gupta, Dept. of Biotechnology is interested in identification of geminiviruses of vegetable crops and possible alternate weed hosts. He is interested in detection techniques and co-operative efforts on virus characterisation. Dr. Dhanju, Vegetable Virologist studying under Dr. Gupta is excellent at field identification of disease symptoms, and he is undertaking variety trials assaying for virus resistance, based upon symptom development, only. Our hosts generously provided transportation, housing, and essential assistance in making collections of research materials (cotton, okra, and malvaceous weeds near cotton fields though to harbor geminiviruses).

CLCuV-infected cotton was first noticed in 1994 close to the Indo-Pakistan border. In 1995, 1000 ha were infected in the Punjab, which increased to 5000 ha in 1996 with a concomitant infection of 10,000 ha in Rajasthan (state SW of Punjab). Symptoms include upward and downward leaf curling, vein thickening and enations. Although the farmers are not yet worried, the cotton breeders are becoming concerned. PAU cotton breeder, Dr. Randhawa, has identified two dominant CLCuV resistance genes in *G. hirsutum* and will release cotton variety LHH144 (reputed to be resistant), containing at least one resistance gene, next year. In the future, PAU, as the main provider of commercial seed (free of charge) to the Punjabi farmers, will only release resistant varieties. As an interim measure, PAU have recommended the growing of *G. arboreum* (known to be resistant to CLCuV) instead of *G. hirsutum* in a 5 km wide area along the Indo-Pak border in order to establish a CLCuV-free buffer zone. If, in the future, CLCuV develops into a major threat to Indian cotton production, farmers have plans to change over to sugar cane, or rice in areas where the water table is rising. These farmers are therefore less reliant upon cotton for economic viability than their Punjabi counterparts in Pakistan.

Workers in PAU's cotton breeding department have identified approx. 10-20 weed species as CLCuV virus and/or whitefly hosts (including *Sida* spp. and *Abutilon* spp.) following vector transmissions from weeds to cotton and back to the weeds. Details are not clear, however, with no confirmation of whitefly biotype nor accurate virus identification. Local citrus orchards provide favourable for establishing whitefly populations, therefore PAU tests all cotton breeding lines in these conditions. Also, transmission tests are undertaken in the greenhouse with the use of clip cages on young cotton leaves.

The main breeding sites are at Ludhiana (main campus) and Faridkot (sub-station), where the initial crosses are made, along with varietal and hybrid trials (F_1 and F_2 generations). Progeny is then sent out to testing centres at Bathinda (not visited), Muktsar and Abohar. At Muktsar, segregating populations are grown, and hybrid seed production undertaken. Alternate blocks of *G. hirsutum* and *G. arboreum* are planted in order to reduce the potential for cross-pollination between hybrids of the same species (NB. *G. hirsutum* and *G. arboreum* are sexually-incompatible). Pollination is carried out by honey bees. At Abohar, testing of intra-*hirsutum* hybrids and production of varietal seed is done, along with agronomic trials. All promising varieties/hybrids are tested at each of these sites prior to release in an assessment of their characteristics under different climatic conditions. Once a variety/hybrid is deemed suitable for release, the seed is sent to selected larger farmers for use in large-scale demonstration plots. Smaller farmers are invited to see the improved material, and hence the seed is distributed from PAU throughout the state.

The summer monsoon (kharif) season is also a time during which farmers grow other vegetable crops, including okra, tomato, pepper (*Capsicum* spp.) and eggplant (brinjal). Severe geminivirus infections are reported to have occurred in okra, tomato and pepper in the preceding years. By the time of our visit, most tomato and okra had been harvested but plots of pepper breeding lines were severely affected by a geminivirus-like disease. The symptoms included leaf curl, yellowing and shortening of internodes. We also observed geminivirus-like symptoms in eggplant which had not been previously noted. Whiteflies were also evident in eggplant plots. Other crop species with suspected geminivirus infections were also sampled. Weed hosts were also plentiful, and several were collected from Ludhiana and Abohar. The most commonly infected weed was *Ageratum* with a yellow vein symptom. Whiteflies were abundant on stands of this plant. At other times, congress grass (a

dicot) is reputed to be frequently infected with a leaf- and stem-distorting symptom. There is a report that the B biotype is present in Pakistan and possibly across the border into the Indian Punjab as well.

Brazil

J. K. Brown traveled to Brazil in late November and visited a cotton growing region near Campinas to establish contacts through which to acquire virus samples from infected cotton in the spring when the diseases are anticipated. During my (JKB) visit to Brazil in November 1996, I traveled to the cotton growing area to see newly planted cotton, visit with the retired virologist, Dr. A. S. Costa to learn more about the whitefly-transmitted geminiviruses in cotton in Brazil and surrounding locales, and made arrangements to receive relevant materials when available from assistants in the research station in Campinas. According to Dr. A.S. Costa (Emeritus) in the Virology Department, at least three distinct geminiviruses infect cotton in the region, annually. One is found toward the end of the season in conjunction with the foliar reddening disease, currently of unknown etiology, but postulated to be a luteovirus. This trip was funded in part by local concerns, and partially by the USDA OICD/FAS in support of international cooperative exchange of scientists and information, and toward the geminivirus database project.

Nicaragua

A Ph.D. graduate student from the laboratory, Rafael Caballero, visited his homeland in 1997 and traveled to the cotton growing area to collect geminiviruses samples for the project. The UA laboratory has a USDA PPQ Federal Quarantine permit to transport plant material into the laboratory from Nicaragua for extraction of viral DNA.

Sudan

A Ph.D. student in the laboratory, A. M. Idris, spent six months in Sudan during 1996 on a Rockefeller Foundation Fellowship. During this time, he obtained DNA extracts of several cotton infecting geminivirus isolates, and these isolates are currently under study in the AZ laboratory.

Benefits of Travel to The Project

Many rewards were gained from personal visits by project scientists to discuss and gain new insights into geminivirus disease problems around the world, and these will clearly benefit the Common Fund project. During past and most recent field visits to Pakistan, India, Brazil, Mexico, Puerto Rico, The Dominican Republic (1990-92), and within the US (Arizona, California, Texas), our laboratory made contact with several key colleagues that aided us in obtaining virus infected materials toward the stated goals. For example, several faculty members at the University of Sao Paulo, at NIBGE in Faisalabad, and in the Punjab Agricultural University indicated they would assist us in future collections, and that we could depend upon them for laboratory facilities, and for extraction of cotton and virus-infected weed samples. In turn, we have provided these laboratories with protocols for DNA isolation from cotton and weed samples and will maintain contact during the next several years to stay abreast of current and new problems and assist them in their work, as we are able. These efforts will help all involved to remain in productive and cooperative contact, thereby facilitating early communications of timely disease phenomena in the field.

References

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

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

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 Epidemiolgy of Subgroup III, Geminiviridae. Plant-Microbe Interactions Review Series, G. Stacey and N. Keen, eds (Chapman and Hall) pp 125-195.

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

Brown J.K. and Bird, J. 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., Bird, J., Frohlich, D.R., Rosell, R.C., Bedford, I.D. and Markham, P.G. 1995. 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., Coats, S., Bedford, I.D., Markham, P.G., Bird, J. and Frohlich, D.R. 1995. Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera:Aleyrodidae). *Biochemical Genetics*, 33: 205-214.

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., 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.

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

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., Nelson, M.R. and Lambe, R.C. 1986. Cotton leaf crumple virus transmitted from naturally infected bean from Mexico. *Plant Dis*, 70:981.

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.

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

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.

Wilson, F.D. and Brown, J.K. 1989. Inheritance of resistance to cotton leaf crumple virus. Page 22 In: J.M. Brown (ed.). *Proc. of Beltwide Cotton Prod. Res. Conf.*, National Cotton Council of America, Memphis, TN, USA.

Wilson, F.D. and Brown, J.K. 1991. Inheritance of resistance to cotton leaf crumple virus in cotton. *J. of Heredity,* 82:508-509.

Wilson, F.D., Brown, J.K. and Butler, Jr., G.D. 1990. Reaction of cotton cultivars and lines to cotton leaf crumple virus. *J. of the Az.-Nev. Acad. Sci*, 23:7-10.

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