UNIVERSIDAD DE MÁLAGA

Identification of new Sub-Saharan African begomoviruses and *Bemisia tabaci* species boundaries

Happyness Gabriel Mollel

Tesis Doctoral

Málaga

2019

AUTOR: Happyness Gabriel Mollel

http://orcid.org/0000-0003-4745-8628

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

Departamento de Biología Celular, Genética y Fisiología
Facultad de Ciencias
Programa de Doctorado de Biotecnología Avanzada

Identification of new Sub-Saharan African begomoviruses and *Bemisia tabaci* species boundaries

Memoria de
TESIS DOCTORAL
presentada por la Licenciada en Biotecnología

Happyness Gabriel Mollel

para optar al grado de
Doctora por la Universidad de Málaga

Directores
Dr. Jesús Navas Castillo
Dr. Elvira Fiallo Olivé
Dr. John Colvin

Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora"
(IHSM-UMA-CSIC)

Málaga, octubre 2019

Jesús Navas Castillo, Doctor en Ciencias Biológicas, Investigador Científico del Consejo Superior de Investigaciones Científicas y Profesor Asociado del Departamento de Biología Celular, Genética y Fisiología de la Universidad de Málaga,

INFORMA que la licenciada **Happyness Gabriel Mollel** ha realizado en el Laboratorio de Virología del Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM-UMA-CSIC), bajo su **dirección** y **tutoría**, el trabajo que con el título "Identification of new Sub-Saharan African begomoviruses and *Bemisia tabaci* species boundaries**"** constituye su Tesis Doctoral para aspirar al grado de Doctor en Biología y

AUTORIZA su lectura en la Universidad de Málaga, haciendo constar que las publicaciones que avalan esta tesis no han sido utilizadas en tesis defendidas con anterioridad.

Y para que así conste y tenga los efectos que correspondan en cumplimiento de la legislación vigente, firma el presente informe en Algarrobo-Costa (Málaga) el 1 de octubre de 2019.

Elvira Fiallo Olivé, Doctora en Ciencias Biológicas, Investigadora Contratada del Consejo Superior de Investigaciones Científicas,

INFORMA que la licenciada **Happyness Gabriel Mollel** ha realizado en el Laboratorio de Virología del Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM-UMA-CSIC), bajo su **co-dirección**, el trabajo que con el título "Identification of new Sub-Saharan African begomoviruses and *Bemisia tabaci* species boundaries**"** constituye su Tesis Doctoral para aspirar al grado de Doctor en Biología y

AUTORIZA su lectura en la Universidad de Málaga, haciendo constar que las publicaciones que avalan esta tesis no han sido utilizadas en tesis defendidas con anterioridad.

Y para que así conste y tenga los efectos que correspondan en cumplimiento de la legislación vigente, firma el presente informe en Algarrobo-Costa (Málaga) el 1 de octubre de 2019.

Natural Resources Institute

John Colvin, PhD in Entomology and Professor of Entomology and Plant-Virus Epidemiology at the Natural Resources Institute (NRI), University of Greenwich, United Kingdom,

DECLARES that Ms. **Happyness Gabriel Mollel** has carried out her doctoral research under his **co-supervision**, with the title "Identification of new Sub-Saharan African begomoviruses and *Bemisia tabaci* species boundaries" which constitutes her Doctoral Thesis to attain the degree of Doctor in Biology and

AUTHORIZE its defence at the University of Malaga, stating that the publications supporting this thesis have not been used in theses previously defended.

In recognition whereof, I sign this report in Chatham Maritime (Kent, UK) on October 4, 2019.

John Colvin.
4TH October 2019.

This work was carried out in the Plant Virology Laboratory at Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" (IHSM-UMA-CSIC) (Algarrobo-Costa, Málaga, Spain) and stays at the Natural Research Institute (NRI), University of Greenwich (Chatham Maritime, United Kingdom). It was funded by the Bill & Melinda Gates Foundation (Head Grant Agreement OPP1058938) through a sub-contract (B0436x12) from the NRI and by the Ministerio de Economía y Competitividad (MINECO, Spain) (project AGL2013-48913-C2-1-R co-financed by the European Regional Development Fund).

Part of the results of this Doctoral Thesis have been published in JCR journals:

Mollel HG, Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017) A novel East African monopartite begomovirus-betasatellite complex that infects *Vernonia amygdalina*. *Archives of Virology* 162: 1079-1082 **[CHAPTER 1]**

Mollel HG, Sseruwagi P, Ndunguru J, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017) Desmodium mottle virus, the first legumovirus (genus *Begomovirus*) from East Africa. *Archives of Virology* 162: 1799-1803 **[CHAPTER 2]**

Mollel HG**,** Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2019) African basil (*Ocimum gratissimum*) is a reservoir of divergent begomoviruses in East Africa. *Plant Disease* (in press) DOI: 10.1094/PDIS-08-19-1675-RE **[CHAPTER 3]**

And presented at the scientific conferences:

- XIV Congreso Nacional de Virología. Cádiz, 2017.
- XIX Congreso Nacional de la Sociedad Española de Fitopatología. Toledo, 2018.

DEDICATION

This thesis is dedicated to my dear husband Juma William Yabeja and lovely son William Juma Yabeja whose their affection, love, encouragement and prayers made me able to get such success and honor, and above all the Almighty God who has led me this far. Every challenging work needs self-effort as well as guidance of elders. I also dedicate this work to my sisters and brothers, and my parents who always loved me unconditionally and taught me to work hard for things I aspire to achieve in life since I was little girl.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my university supervisor, Dr. Jesús Navas-Castillo (IHSM-UMA-CSIC, Spain) and co-supervisors: Dr. Elvira Fiallo-Olivé (IHSM-UMA-CSIC, Spain) and Professor John Colvin (NRI, University of Greenwich, UK), for giving me the opportunity to undertake this study under their supervision, for their constructive criticism and encouragement throughout my study.

Special thanks to the Ministerio de Economía y Competitividad (MINECO, Spain) and European Regional Development Fund (AGL2013-48913-C2-1-R), and Bill & Melinda Gates Foundation through a subcontract (B0436x12) from the Natural Resources Institute, University of Greenwich, UK, for funding my tuition fees, research and welfare.

I would like to express my sincere appreciations to the team at IHSM-UMA-CSIC especially Ms. Remedios Tovar, Ms. Elisa Navas-Hermosilla, Mr. Miguel López and Mr. Juan Cárdenas García, and team at NRI especially Dr. Sharon van Brunschot, Mr. Stephen Young, Dr. Simon Springate, Dr. Soňa Vyskočilová, Dr. Rekha RA Swamy, Ms. Rebecca Grimsley, Ms. Jackie Atim and Ms. Annet Namuddu for their assistance during my study. I would also like to thank my fellow students and other scientists at IHSM-UMA-CSIC for their support and encouragement during my study.

Finally, I would like to thank my parents who nurtured me and sacrificed much of their scarce resources for my education. Their support has enabled me to acquire this higher level of education.

TABLE OF CONTENTS

ABSTRACT

ABSTRACT

Begomoviruses (genus *Begomovirus*, family *Geminiviridae*) are an extremely successful group of ssDNA viruses that have a very wide host range, infecting dicotyledonous plants and causing devastating diseases to important crops, particularly in tropical and sub-tropical regions. Begomoviruses also infect numerous weeds and other wild plants that may act as reservoirs for plant viruses. In 2015, a survey was conducted in Uganda aimed at collecting and characterizing viruses on symptomatic cultivated and non-cultivated plants, suggesting begomovirus infections. Complete begomovirus genomes and DNA satellites were amplified by rolling circle amplification, cloned and sequenced. Based on the sequence analysis, this study reports five novel begomovirus species and a new betasatellite: i) vernonia crinkle virus (VeCrV) in association with vernonia crinkle betasatellite (VeCrB) infecting *Vernonia amygdalina* (family Compositae), ii) desmodium mottle virus (DesMoV) infecting *Desmodium* sp. [Fabaceae (Leguminosae)] and iii) ocimum yellow vein virus (OcYVV), ocimum mosaic virus (OcMV) and ocimum golden mosaic virus (OcGMV) infecting *Ocimum gratissimum* (Lamiaceae). Phylogenetic analysis and genome organization of the DNA-As of these novel begomoviruses showed that they belong to the Old Word phylogenetic group, with the exception of DesMoV that belonged to the legumovirus group. DesMoV and OcGMV are closely related to begomoviruses infecting crops, such as soybean in Nigeria and tomato in Uganda, respectively. However, the remaining begomoviruses are mostly related to begomoviruses infecting wild plants in Africa and Asia. These results support previous studies that indicate that non-cultivated plants are melting pots for recombination and emergence of novel begomoviral and DNA satellite genomes that could be a threat to economically important crops.

This study also investigated the mating interactions between whitefly populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) colonizing cassava (*Manihot esculenta*) in sub-Saharan Africa (SSA) and morning glory (*Ipomoea indica*) in Spain. *B. tabaci* is a group of more than 40 cryptic species that include some of the most devastating insect pests and plant-virus vectors of crops worldwide. Reciprocal crosses were conducted to examine mating interactions within and between putative *B. tabaci* species (based on the phylogenetic species delimitation criterion of 3.5% mtCOI sequence divergence): SSA2 populations from Spain and SSA1-SG1, SSA1-SG3, SSA2 and SSA3 from sub-Saharan Africa. Hybrid females of first filial generation (F_1) were confirmed molecularly using the nuclear gene *pre-messenger RNA processing factor 8* (*prpf8*) and restriction fragment length polymorphism analysis. The reciprocal crosses between SSA2 from Spain and SSA2, SSA3 and SSA1-SG3 from sub-Saharan Africa produced hybrid females in both directions of the crosses in F_1. However, no hybrid females were produced in a reciprocal cross between SSA2 from Spain and SSA1-SG1 from sub-Saharan Africa. The digestion of amplified *prpf8* nuclear gene of F_1 hybrid females exhibited a mixed restriction pattern derived from both whitefly populations from Spain and

ABSTRACT

sub-Saharan Africa, hence verifying the progeny were true hybrids. In addition, the production of both female and male in the second filial generation (F_2) among siblings of reciprocal crosses of SSA2 from Spain, and SSA2 and SSA3 from sub-Saharan Africa confirmed the fertility and viability of F_1 hybrid females. These findings confirm that whitefly populations SSA2 from Spain interbreed with SSA2 and SSA3 from sub-Saharan Africa, and produce fertile offspring, hence suggesting that these populations belong to the same biological species, even though the partial mtCOI sequence divergence between SSA2 and SSA3 exceeded 3.5%.

RESUMEN EXTENDIDO

RESUMEN EXTENDIDO

El género *Begomovirus* (familia *Geminiviridae*) es el mayor género de virus de plantas y de toda la Virosfera con respecto al número de especies que comprende. Actualmente, contiene 409 especies reconocidas por el Comité Internacional de Taxonomía de Virus (ICTV). Los begomovirus constituyen un grupo extremadamente exitoso de virus emergentes que infectan una amplia gama de plantas dicotiledóneas que incluyen importantes cultivos hortícolas y de fibra, principalmente en las regiones tropicales y subtropicales. Las enfermedades causadas por begomovirus se caracterizan por una serie de síntomas que incluyen mosaicos, rizado, deformación y moteado foliar así como amarilleo de venas. Algunos ejemplos de begomovirus económicamente importantes incluyen el complejo del rizado amarillo del tomate extendido por todo el mundo, los virus del mosaico de la yuca presentes en el África subsahariana, los virus del rizado foliar del algodón sobre todo en Pakistán e India y un número elevado de begomovirus que infectan el tomate en América del Sur, América Central y el Caribe. Los begomovirus se transmiten de forma circulativa por las moscas blancas pertenecientes al complejo *Bemisia tabaci* (Hemiptera: Aleyrodidae) a una gran variedad de plantas cultivadas y no cultivadas. Debido a las fuertes pérdidas de alimentos y producción comercial de cultivos importantes como la yuca, el tomate, las legumbres de grano y el algodón, los begomovirus representan una amenaza importante para la seguridad y sostenibilidad alimentaria mundial. Durante las últimas dos décadas, las epidemias de begomovirus emergentes han causado enormes pérdidas y amenazado la producción de cultivos, particularmente en las regiones tropicales y subtropicales. Se ha estimado que las pérdidas económicas anuales debidas a infecciones por begomovirus en todo el mundo son de 1.3-2.3 miles de millones de dólares en el cultivo de yuca en África, 300 millones de dólares en las leguminosas de grano en India, 5 mil millones de dólares en el algodón en Pakistan y 140 millones de dólares en el tomate en Florida.

Los genomas de los begomovirus están compuestos por una o dos moléculas de ADN circular monocatenario de 2.5-2.8 kb que se encapsidan en partículas cuasi-icosaédricas de morfología gemelar. Según su organización genómica, los begomovirus pueden ser monopartitos o bipartitos. Los genomas de los begomovirus bipartitos consisten en dos componentes, denominados ADN-A y ADN-B, cada uno de 2.5-2.8 kb, mientras que los begomovirus monopartitos tienen un componente que se asemeja al ADN-A de los begomovirus bipartitos. Los componentes del genoma de los begomovirus poseen una región intergénica no codificante de 280-350 nucleótidos que contiene una caja TATA y secuencias conservadas conocidas como iterones que son específicas del sitio de unión de la proteína asociada a la replicación. El análisis filogenético de las secuencias nucleotídidicas del genoma de begomovirus monopartitos y del componente ADN-A de los bipartitos, junto con su organización genómica, ha permitido clasificar a las especies del género en cuatro linajes: begomovirus del Viejo

Mundo (Old World, OW), begomovirus del Nuevo Mundo (New World, NW), sweepovirus y legumovirus, los dos últimos con un rango de huéspedes limitado.

Los ADN satélites son agentes subvirales que carecen de genes que codifiquen funciones para su replicación, por lo que dependen de virus ayudantes para la multiplicación en la célula huésped infectada. Los ADN satélites asociados a begomovirus no tienen similitud de secuencia con los virus ayudantes, aparte de la presencia de una estructura en horquilla que se requiere para la acción de la proteína asociada a la replicación. Los ADN satélites asociados con los begomovirus se clasifican en dos familias: *Alphasatellitidae*, una familia que comprende las subfamilias *Geminialphasatellitinae* y *Nanoalphasatellitinae* con cuatro y siete géneros, respectivamente y la familia recientemente aceptada *Tolecusatellitidae* que contiene dos géneros, *Betasatellite* y *Deltasatellite*.

Los betasatélites, anteriormente denominados DNA-β, son moléculas circulares de ADN de simple cadena (ADNsc) de ~1350 nt, aproximadamente la mitad de los componentes del genoma de los begomovirus que no comparten una identidad de secuencia significativa con sus virus auxiliares a excepción de la horquilla que actúa como origen de replicación. Los betasatélites dependen estrictamente de los begomovirus auxiliares para su replicación, encapsidación y el movimiento dentro y entre plantas, aparentemente por transencapsidación en la cápsida del virus auxiliar. Poseen una región rica en adenina, una secuencia (~ 120 nts) que está altamente conservada entre todos los betasatellites y un único gen que codifica para la proteína βC1. Todas las funciones de los betasatélites están relacionadas con la presencia de la proteína βC1, incluyendo la supresión del silenciamiento génico transcripcional y postranscripcional, movimiento del virus en la planta, regulación de los niveles de ADN viral e inducción o incremento de síntomas de la enfermedad en las plantas infectadas.

Los alfasatélites, anteriormente denominados DNA 1, son moléculas circulares de ADNsc de aproximadamente la mitad del tamaño (~ 1400 nts) de los componentes del genoma de los begomovirus. No son ADN "satélites" en sentido estricto, ya que codifican su propia proteína asociada con la replicación. Poseen una región rica en adenina y la mayoría de ellos contienen una estructura en horquilla con un nonanucleótido común a los nanovirus. Aunque los alfasatélites se asocian principalmente con begomovirus monopartitos del Viejo Mundo y frecuentemente en asociación con complejos begomovirus-betasatélite, también se han encontrado en el Nuevo Mundo en asociación con begomovirus bipartitos. Algunos alfasatélites disminuyen la sintomatología del virus auxiliar mientras que en otros casos la aumentan.

Los deltasatélites son moléculas circulares de ADNsc de ~ 700 nt, aproximadamente un cuarto del tamaño de los componentes genómicos de begomovirus, sin una similitud de secuencia

significativa con los virus auxiliares. Están asociados con begomovirus monopartitos del Viejo Mundo, begomovirus bipartitos del Nuevo Mundo y sweepovirus. A diferencia de los betasatélites y alfasatélites, los deltasatélites no codifican ninguna proteína. Necesitan del begomovirus auxiliar para la replicación y el movimiento en las plantas y la transmisión por *B. tabaci*. Contienen una secuencia con similitud con la región conservada de los satélites presente en los betasatélites, una región rica en adenina, una estructura en horquilla similar a la de los virus auxiliares y una horquilla secundaria adicional. La presencia de deltasatélites en la planta huésped puede causar una reducción de la acumulación de los begomovirus auxiliares y/o la gravedad de los síntomas.

Las moscas blancas son hemípteros de la familia *Aleyrodidae*, compuesta por más de 160 géneros y más de 1500 especies. Dos moscas blancas, *Bemisia tab*aci (sensu lato) y *Trialeurodes vaporariorum*, causan considerables pérdidas económicas para la agricultura en todo el mundo. Se cree que *B. tabaci* se originó en África, desde donde se extendió al norte de África, la cuenca del Mediterráneo, Asia, Australia y las Américas, siendo actualmente el grupo de especies de mosca blanca más ampliamente distribuido y económicamente más importante del mundo. *B. tabaci* ha incrementado su importancia como plaga en las últimas décadas y es un vector de virus de plantas pertenecientes a alrededor de 430 especies pertenecientes a los géneros *Begomovirus* (familia *Geminiviridae*), *Carlavirus* (familia *Betaflexiviridae*), *Crinivirus* (familia *Closteroviridae*), *Ipomovirus* (familia *Potyviridae*), *Polerovirus* (familia *Luteoviridae*) y Torradovirus (familia *Secoviridae*). *B. tabaci* causa daños directos a las plantas infestadas a través de la extracción de grandes cantidades de savia del floema, lo que puede resultar en una reducción de rendimiento superior al 50%. Además, la alimentación directa de las moscas blancas puede provocar la maduración irregular de los frutos, clorosis de las hojas, blanqueo del tallo y la inducción de trastornos fitotóxicos que resultan en plateado en algunos genotipos de algunas especies de plantas. Estos efectos pueden ser muy graves cuando grandes poblaciones de *B. tabaci* colonizan una planta. El daño indirecto se debe a la excreción de melaza en la superficie de las hojas y frutos que sirve como medio para el crecimiento del hongo de la negrilla que oscurece las hojas y los frutos, impidiendo el proceso de fotosíntesis de la planta y reduciendo la calidad de los frutos. Sin embargo, el daño indirecto más importante causado por *B. tabaci* se debe a su papel como vector de cientos de virus vegetales que infectan plantas cultivadas y no cultivadas. Basándose en la comparación de secuencias nucleotídicas parciales del gen de la citocromo oxidasa 1 mitocondrial (mtCO1), se ha propuesto que *B. tabaci* es un complejo de 11 grupos genéticos que integran a más de 40 especies morfológicamente indistinguibles, separadas por un mínimo de divergencia nucleotídica del 3.5%. Estas especies crípticas son genéticamente diversas y difieren en características biológicas como

especificidad de la transmisión de virus, inducción de trastornos fitotóxicos, invasividad, resistencia a insecticidas y rango de plantas huésped. Se han llevado a cabo diversos estudios reproductivos -cruzamientos recíprocos- para comprender el nivel de relación entre poblaciones de *B. tabaci*, morfológicamente indistinguibles. La compatibilidad reproductiva entre las supuestas especies crípticas de *B. tabaci* se ha analizado en base a la fecundidad y la fertilidad obtenidas mediante cruzamientos intra- e inter-poblacionales.

En base a estos antecedentes, el objetivo general de esta tesis fue generar información sobre los begomovirus que infectan plantas silvestres en el este de África e investigar los límites biológicos de supuestas especies (basadas en la divergencia del gen mtCO1) de *B. tabaci* de África subsahariana (SSA). Esto se logró abordando los objetivos específicos que se estructuran en los siguientes capítulos de la tesis.

- Un nuevo complejo de begomovirus monopartito-betasatélite del este de África que infecta a *Vernonia amygdalina*.
- Desmodium mottle virus, el primer legumovirus (género *Begomovirus*) del este de África.
- La albahaca africana (*Ocimum gratissimum*) es un reservorio de begomovirus divergentes en Uganda.
- Interacciones de apareamiento entre supuestas especies crípticas de *Bemisia tabaci* de la cuenca del Mediterráneo y África subsahariana.

En 2015, se llevó a cabo un muestreo en Uganda con el objetivo de recolectar y caracterizar los begomovirus presentes en plantas sintomáticas, principalmente no cultivadas. Se amplificaron genomas completos de begomovirus y DNA satélites asociados mediante amplificación por círculo rodante (RCA), se clonaron y secuenciaron. En base a los análisis de secuencia, en este estudio se han caracterizado cinco nuevas especies de begomovirus y un nuevo betasatélite: i) vernonia crinkle virus (VeCrV) y vernonia crinkle betasatellite (VeCrB) infectando a *Vernonia amygdalina* (familia Compositae), ii) desmodium mottle virus (DesMoV) infectando a *Desmodium* sp. (Fabaceae) y iii) ocimum yellow vein virus (OcYVV), ocimum mosaic virus (OcMV) y ocimum golden mosaic virus (OcYVV) infectando a *Ocimum gratissimum* (Lamiaceae). El análisis filogenético y la organización genómica del ADN-A de estos nuevos begomovirus mostraron que pertenecen al grupo filogenético Old Word, con la excepción de DesMoV que pertene al grupo de los legumovirus. DesMoV y OcGMV están estrechamente relacionados con begomovirus que infectan cultivos, como soja en Nigeria y tomate en Uganda, respectivamente. Sin embargo, los begomovirus restantes están relacionados principalmente con begomovirus que infectan plantas silvestres en África y Asia. Estos resultados apoyan estudios previos que indican que las plantas no cultivadas se comportan como

auténticos crisoles donde pueden tener lugar fenómenos de recombinación y aparición de nuevos genomas de begomovirus y ADN satélites que pueden ser una amenaza para cultivos económicamente importantes.

La utilización de la técnica Amplificación por Circulo Rodante (RCA), que empleala ADN polimerasa del bacteriófago φ29, análisis mediante polimorfismo de longitud de fragmentos de restricción (RFLP), clonación y secuenciación, permitió descubrir que diez muestras foliares recogidas de otras tantas plantas sintomáticas no cultivadas en Uganda estaban infectadas por begomovirus y dos de ellas también estaban infectadas por un betasatélite. Todos los begomovirus caracterizados y el betasatélite representan aislados de especies nuevas para la ciencia. Muchos otros estudios han utilizado RCA en combinación con RFLP con éxito en el diagnóstico y caracterización de virus con genomas de ADN circulares monocatenarios que incluyen begomovirus y ADN satélites. Las muestras sintomáticas se obtuvieron como parte de un estudio en el sureste de Uganda llevado a cabo a lo largo de las carreteras principales donde se detectaron y tomaron muestras de plantas silvestres que mostraban síntomas sospechosos de una infección por begomovirus. Además, se recolectaron muestras de algunas plantas que crecían en huertos familiares.

Los virus y el ADN satélite identificados en este estudio en plantas no cultivadas pertenecen al género *Begomovirus* (familia *Geminiviridae*) y al género *Betasatellite* (familia *Tolecusatellitidae*), respectivamente. De acuerdo con el umbral de demarcación de especies de begomovirus del 91% para el ADN-A de begomovirus bipartitos o el genoma completo de begomovirus monopartitos, se descubrieron cinco especies distintas de begomovirus. Estos begomovirus incluyen: vernonia crinkle virus (VeCrV) y su betasatélite asociado vernonia crinkle betasatellite (VeCrB) aislados de *Vernonia amygdalina* (Compositae); desmodium mottle virus (DesMoV) identificado en la mala hierba *Desmodium* sp. (Fabaceae) y ocimum yellow vein virus (OcYVV), ocimum mosaic virus (OcMV) y ocimum golden mosaic virus (OcGMV) que infectaban *Ocimum gratissimum* (Lamiaceae). Con la excepción de DesMoV, los nuevos begomovirus se agrupan en diferentes linajes dentro del principal grupo filogenético de begomovirus, los begomovirus del Viejo Mundo. DesMoV se agrupó dentro del grupo filogenético de los legumovirus, siendo el primer legumovirus descrito en el este de África. Los análisis filogenéticos también mostraron que algunos de los nuevos begomovirus están solo lejanamente emparentados con los begomovirus previamente caracterizados. Con respecto a VeCrB, de acuerdo con el umbral de demarcación de especies de betasatélites del 91%, representa una nueva especie. Esta es la primera caracterización molecular completa de begomovirus que infectan plantas silvestres en Uganda y la primera identificación de un betasatélite en el mismo país. En conjunto, estos hallazgos indican que parece existir una gran diversidad de especies de begomovirus que

infectan plantas silvestres en Uganda y probablemente en el resto del este de África. Los begomovirus que infectan plantas no cultivadas apenas han recibido atención por parte de los investigadores en esta región de África, pero debe tenerse en cuenta que son una amenaza potencial para cultivos importantes. En relación con esto, debe enfatizarse que, en base al análisis de identidad de secuencia, algunos de los nuevos begomovirus están relacionados con begomovirus que infectan cultivos en África. Por ejemplo, DesMoV y OcGMV están relacionados con soybean mild mottle virus que infecta soja en Nigeria y tomato leaf curl Uganda virus que infecta tomate en Uganda, respectivamente. Además, West Africa asystasia virus1, descrito inicialmente infectando la planta silvestre *Asystasia* sp., se ha encontrado posteriormente infectando plantas comerciales de yuca en Camerún. En otras partes del mundo se ha demostrado que los begomovirus que infectan las malas hierbas pueden infectar los cultivos próximos de forma natural. Por ejemplo, sida micrantha mosaic virus y macroptilium yellow spot virus fueron aislados de plantas de judía en Brasil, ageratum enation virus de plantas de tomate en India y euphorbia mosaic virus de plantas de tabaco en Cuba y plantas de pimiento en México. Por otra parte, algunos estudios han demostrado la transmisión de begomovirus por el insecto vector, *B. tabaci*, de huéspedes silvestres a especies cultivadas. Por lo tanto, estos resultados refuerzan la importancia de considerar las plantas no cultivadas al desarrollar estrategias de control sostenible para las enfermedades virales en una región determinada.

En otras partes del mundo también se han aislado diferentes begomovirus de los géneros de plantas analizados en este trabajo. *Desmodium* sp. se ha encontrado infectado por rhynchosia yellow mosaic Yucatán virus y desmodium leaf distortion virus en México y macroptilium yellow spot virus en Brasil. *Vernonia cinerea* se ha encontrado infectada por vernonia yellow vein Fujian virus y betasatélites y alfasatélites asociados en China y vernonia yellow vein virus y un betasatélite asociado en India. En el género *Ocimum*, se han encontrado infecciones de tomato leaf curl virus y un betasatélite asociado en *Ocimum sanctum* en India y de chilli leaf curl virus, tomato yellow leaf curl virus y tomato leaf curl betasatellite en *Ocimum basilicum* en Omán.

En el presente estudio se han detectado eventos de recombinación tanto en begomovirus como en los betasatélites. Los eventos de recombinación detectados en VeCrV y VeCrB involucran a begomovirus que infectan pimiento, tomate, *Siegesbeckia glabrescens* y *Vernonia cinerea* en Asia y betasatéites que infectan yuca, tabaco y algodón en África. Los begomovirus involucrados en los eventos de recombinación del ADN-A de OcYVV, OcMV y OcGMV infectan la yuca, el tomate y la soja en Asia, el tomate en las islas del Océano Índico y la yuca, el tomate y *Desmodium* sp. en África. Los begomovirus recombinantes naturales han estado directamente involucrados en la aparición de nuevas enfermedades y epidemias en los cultivos en muchos países. Los resultados obtenidos en este trabajo respaldan la idea de estudios previos de que las plantas silvetres son nichos ecológicos donde

pueden tener lugar fenómenos de recombinación y aparición de nuevas especies de virus y satélites. Por otra parte, la detección de tres especies distintas de begomovirus en *Ocimum gratissimum* sugiere que esta planta puede actuar como reservorio en el que pueden surgir virus recombinantes. En África occidental se han descrito otras plantas silvestres, como *Asystasia gangetica*, también con infecciones mixtas de distintas especies de begomovirus (asystasia begomovirus 1 y asystasia begomovirus 2).

Se han encontrado betasatelites en asociación con begomovirus de plantas cultivadas y no cultivadas en numerosos países. VeCrB es el cuarto betasatélite que se detecta en plantas silvestres en África. Anteriormente, se habían identificado cotton leaf curl Gezira betasatellite, ageratum leaf currl Cameroon betasatellite y tomato leaf curl Togo betasatellite. La detección de un cuarto betasatélite pone de manifiesto una creciente complejidad de estos agentes subvirales en África.

Un objetivo de esta tesis fue determinar los límites de especies en el grupo genético de África subsahariana de *B. tabaci* a través de estudios de compatibilidad reproductiva. Las moscas blancas, incluida *B. tabaci*, son insectos haplodiploides, con progenie masculina haploide producida a partir de huevos no fertilizados y progenie hembra diploide a partir de huevos fertilizados. Para determinar si dos poblaciones de mosca blanca se comportan como especies biológicas distintas, se deben realizar estudios de compatibilidad reproductiva que revelen la proporción de progenie femenina obtenida de cruces recíprocos intra e interpoblacionales. En este trabajo se investigaron las interacciones de apareamiento entre poblaciones de mosca blanca de *B. tabaci* que infestan yuca (*Manihot esculenta*) en África subsahariana (SSA) e *Ipomoea indica* en España. Concretamente, se utilizó una población de África subsahariana 2 de España (SSA2-Sp) y las poblaciones subsaharianas: SSA1-SG1-Ug (Uganda), SSA1-SG3-Tz (Tanzania), SSA2-Ng (Nigeria), SSA2-Ug (Uganda) y SSA3-Ng (Nigeria). Se realizaron cruces recíprocos para examinar las interacciones de apareamiento dentro y entre supuestas especies del complejo *B. tabaci* (según el criterio de delimitación de especies basado en una divergencia de la secuencia del gen mtCOI mayor del 3.5%). Las hembras híbridas de primera generación filial (F_1) se confirmaron molecularmente usando el gen *pre-messenger RNA processing factor 8* (*prpf8*) y análisis de polimorfismo de longitud de fragmentos de restricción. Los cruces recíprocos entre SSA2 de España y SSA2, SSA3 y SSA1-SG3 de África subsahariana produjeron hembras híbridas en ambas direcciones en la F_1. Sin embargo, no se produjeron hembras híbridas en un cruce recíproco entre SSA2 de España y SSA1-SG1 de África subsahariana. La digestión del gen *prpf8* amplificado de las hembras híbridas F_1 exhibió un patrón de restricción mixto derivado tanto de las poblaciones de moscas blancas de España como del África subsahariana, lo que confirmó la naturaleza híbrida de esta progenie. Además, la producción de hembras y machos en la segunda generación filial (F_2) entre hermanos de cruces recíprocos de SSA2 de España y SSA2 y SSA3 de

África subsahariana confirmaron la fertilidad y la viabilidad de las hembras híbridas F_1. Estos hallazgos confirman que las poblaciones de mosca blanca SSA2 de España se cruzan con SSA2 y SSA3 del África subsahariana y producen progenie fértil, lo que sugiere que estas poblaciones pertenecen a la misma especie biológica, a pesar de que la divergencia de secuencia mtCOI entre SSA2 y SSA3 excede del 3.5%. El emparejamiento físico de machos y hembras, que es un comportamiento de apareamiento característico en *B. tabaci* se observó en todos los cruces interpoblacionales, incluidos los que se establecieron entre SSA2-Sp y SSA1-SG1-Ug, aunque en este caso no se obtuvo descendencia femenina. El resultado negativo de este cruzamiento es similar al obtenido recientemente al cruzar SSA2-Ug con SSA1-SG1-Ug. Además, estudios realizados en poblaciones naturales han demostrado que no existe flujo génico entre poblaciones de SSA1 de África oriental y central y de SSA2 de Kenia y Camerún. La poca progenie femenina encontrada al cruzar las poblaciones SSA1-SG3-Tz y SS2-Sp probablemente podría deberse a la cópula forzada entre los individuos de ambas poblaciones debido a las condiciones de confinamiento establecidas en el laboratorio, indicando asimismo un alto nivel de incompatibilidad reproductiva. Poca progenie femenina, menos del 30%, también se ha observado en cruzamientos entre SSA1-SG3 (TzCas-Mtw) y SSA2 (UgCas-Nam) y MEAM1 y otras especies, lo que se ha atribuido al uso de hembras ya apareadas. Sin embargo, las condiciones experimentales utilizadas en este estudio permite que solo las moscas blancas adultas vírgenes participen en los cruzamientos recíprocos, evitando así resultados falsos positivos. Además, la progenie femenina producida en los cruzamientos realizados se confirmó molecularmente utilizando un marcador nuclear que demostró la verdadera naturaleza de los híbridos y verificó que no hubo contaminación de distintos genotipos durante el experimento. La producción de progenie femenina en dos generaciones, F_1 y F_2, observada entre SSA2-Sp y SSA2-Ug, SSA2-Ng y SSA3-Ng fue una buena indicación de un apareamiento exitoso entre ellas y por lo tanto mostró que estas poblaciones pertenecen a la misma especie biológica. Trabajos recientes también han mostrado la producción de híbridos al cruzar poblaciones de SSA1-SG1-Ug y SSA1-SG2-Ug en condiciones controladas de laboratorio. Por otra parte, en condiciones naturales, se ha detectado la presencia de flujo génico entre poblaciones de SSA2 y SSA4.

De acuerdo con la propuesta de delimitación de especies en el complejo *B. tabaci* basada en el 3,5% de divergencia de la secuencia parcial del gen mtCOI, se ha sugerido que las poblaciones deberían considerarse como especies en función de este umbral. Considerando los resultados obtenidos en los cruzamientos llevados a cabo en este trabajo, el mencionado umbral del 3.5% fue un buen indicador de la delimitación de especies, con la excepción de un cruzamiento entre SSA2-Sp y SSA3-Ng que, a pesar de divergir en un 5.78%, eran reproductivamente compatibles. Este no es el primer caso de compatibilidad observado entre poblaciones que divergen en > 3.5% en el gen mtCOI.

Así, existen evidencias de compatibilidad de apareamiento casi completa en una dirección y compatibilidad parcial en la otra dirección entre Asia II-3 y Asia II-9, que divergen entre sí en un 4,57%. Además, se ha demostrado que MED y MED-ASL son especies diferentes a pesar de haberse considerado previamente como una sola especie en base al umbral del 3.5%. Existen sin embargo numerosos estudios que respaldan el umbral del 3,5% para la diferenciación de especies biológicas, involucrando por ejemplo los genotipos Q1 y Q2 pertenecientes a la especie MED, MEAM1 y MED, SSA1-SG1-Ug y SSA1-SG2-Ug, Asia I y Asia II-7 y Asia II-6 y Asia II-1. En conclusión, los resultados del presente trabajo se suman al creciente número de ejemplos que muestran que la divergencia de secuencia de 3.5% del gen mtCOI todavía puede usarse como una regla general para los límites de las especies del complejo *B. tabaci*, a la vez que pone de manifiesto la necesidad de la confirmación experimental mediante cruzamientos recíprocos.

En resumen, los resultados de este trabajo indicaron que las plantas no cultivadas son huéspedes comunes de nuevos begomovirus en Uganda, en algunos casos asociados a ADN satélites. Además, como han mostrado trabajos previos, algunas especies de plantas silvestres son vulnerables a las infecciones por diferentes begomovirus lo que puede facilitar la aparición de nuevas especies o cepas por recombinación. Este puede dar lugar a variantes más virulentas y destructivas que sus virus parentales. Además, la identificación del primer betasatélite en el este de África evidencia una complejidad adicional de estos patosistemas. Los resultados obtenidos con un número limitado de muestras provenientes de Uganda sugieren que muy probablemente haya un número muy elevado de begomovirus por descubrir en África, sobre todo en malas hierbas y otras plantas silvestres que rodean los campos de los agricultores y amenaza sus cultivos.

En este estudio no se demostró si los begomovirus caracterizados podrían infectar y causar enfermedades en plantas cultivadas, con la excepción de algunas plantas de albahaca africana. Sin embargo, hay ejemplos en la literatura que demuestran que virus inicialmente encontrados en plantas silvestres han saltado a distintos agrosistemas. Por ejemplo, la yuca, un cultivo que se originó en América Latina, se infectó con varios begomovirus tras su introducción en África, lo que sugiere que estos virus se transmitieron a partir de plantas nativas de la región. Los daños potenciales que los begomovirus que infectan malas hierbas y otras plantas silvestres podrían causar a los cultivos podrían ser más graves que los causados en sus huéspedes naturales si, por ejemplo, se vuelven más virulentos y/o son más fácilmente transmisibles por sus vectores, las especies del complejo *B. tabaci*. La diversidad de begomovirus y ADN satélites asociados que infectan plantas no cultivadas en África debiera estudiarse en mayor profundidad, para así comprender su papel real en la epidemiología de las enfermedades que puedan causar en los cultivos, ya sea actuando como fuentes primarias de inóculo o

como una origen continuo de nuevos virus, lo que podría alterar por ejemplo las estrategias de control basadas en la utilización de variedades resistentes o tolerantes. Además, es prioritario realizar estudios de transmisión en condiciones controladas para determinar si los begomovirus caracterizados se pueden transmitir a cultivos importantes en el continente.

En cuanto a los experimentos de cruzamientos recíprocos puede concluirse, teniendo en cuenta los datos de flujo genético obtenidos por otros grupos y el concepto de especie biológica, que las poblaciones españolas de SSA2, que colonizan primariamente plantas de *I. indica*, pertenecen a la misma especie que las poblaciones del África subsahariana de SSA2 de Uganda y Nigeria, a la que también pertenecerían las poblaciones de SSA3 de Nigeria, todas ellas colonizadoras de plantas de yuca. Así se demuestra que estas poblaciones, aunque alopátricas, han conservado la capacidad de cruzarse y producir descendencia fértil al menos en condiciones experimentales. Aunque SSA2-Sp se cruzó con las poblaciones africanas de SSA2 y SSA3 y prudujo descendencia viable, se requiere investigación adicional para comprender si podría colonizar las plantas de yuca y si es capaz de transmitir las principales enfermedades virales de este cultivo u otros begomovirus de importancia económica. También sería interesante estudiar el rango de huésped SSA2-Sp y otras poblaciones de SSA para determinar la importancia potencial que podrían suponer para cultivos importantes.

La complejidad de las enfermedades emergentes causadas por begomovirus y sus ADN satélites asociados, junto con la explosión de las poblaciones de mosca blanca del complejo *B. tabaci*, requiere un esfuerzo investigador por parte de virólogos, entomólogos y mejoradores para desarrollar medidas de control sostenibles y así minimizar el daño que causan a cultivos de enorme importancia económica a escala global.

INTRODUCTION

The genus *Begomovirus*

The genus *Begomovirus* (family *Geminiviridae*) is the largest genus of plant viruses and in the entire Virosphere with respect to the number of species that it comprises. Currently, it contains 409 species recognized by the International Committee on Taxonomy of Viruses (ICTV) (https://talk.ictvonline.org/taxonomy/). Begomoviruses constitute an extremely successful group of emerging viruses infecting a wide host range of dicotyledonous plants including important vegetables and fiber crops, mainly in tropical and sub-tropical regions (Navas-Castillo *et al*., 2011). Begomoviruses exhibit various disease symptoms including leaf mosaics, curling, deformation, mottling, enations and crinkling, and vein yellowing and clearing (Rojas *et al*., 2018).

Examples of economically important begomoviruses include: tomato yellow leaf curl virus complex on tomato in different regions (Rojas *et al*., 2018), cassava mosaic viruses on cassava in sub-Saharan Africa (Legg *et al*., 2015), cotton leaf curl viruses on cotton in Pakistan and India (Briddon *et al*., 2001; Mansoor *et al*., 2003), and other important begomoviruses infecting tomato in South America (Gilbertson *et al*., 2015; Macedo *et al*., 2018; Márquez-Martín *et al*., 2011) (**Figure 1**).

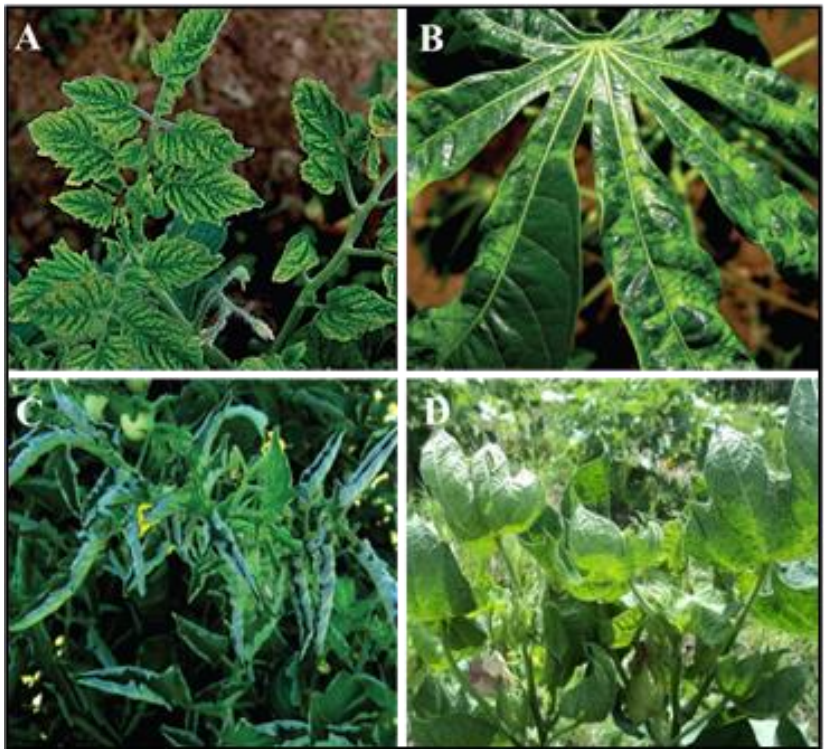

Figure 1. Symptoms of begomovirus infection. (**A**) *Tomato yellow leaf curl virus* on tomato, (**B**) Mixed infection of *African cassava mosaic virus* and *East African cassava mosaic virus* on cassava, (**C**) *Tomato mottle leaf curl virus* on tomato **D**) *Cotton leaf curl virus* on cotton. Figures A-C are reproduced from Rojas *et al.* (2018) and D from Farooq *et al.* (2014).

Begomoviruses are transmitted in a circulative manner by whiteflies of the *Bemisia tabaci* group of species (Hemiptera: Aleyrodidae) to a large variety of cultivated and non-cultivated plants (Brown and Czosnek, 2002; Navas-Castillo *et al*., 2011). Due to heavy losses of food and cash crops, such as cassava, tomato, grain legumes and cotton, begomoviruses represent an important threat to global food security and sustainability. During the last two decades, epidemics of re-emerging and newly emerging begomoviruses have caused huge crop losses and threatened crop production, particularly in the tropics and sub-tropics. The annual economic losses due to begomovirus infections worldwide have been estimated to be US $1.3-2.3 billion in Africa for cassava (Thresh and Cooter, 2005), US $300 million for grain legumes in India (Varma and Malathi, 2003), US $5 billion for cotton in Pakistan (Briddon and Markham, 2000) and US $ 140 million for tomato in Florida (Moffat, 1999).

Genome organization and replication

Begomovirus genomes are composed of one or two circular single stranded (ss) DNA molecules of 2.5-2.8 kb that are encapsidated in twinned quasi-icosahedral particles. Based on their genome organization, begomoviruses can be monopartite or bipartite. Bipartite begomoviruses consist of two components, referred to as DNA-A and DNA-B, each of 2.5-2.8 kb, while monopartite begomoviruses have one component resembling the DNA-A of bipartite begomoviruses (**Figure 2**) (Brown *et al*., 2015; Zerbini *et al*., 2017). Begomovirus genome components possess a non-coding intergenic region (IR) of 280-350 nucleotides (nt) that contains a TATA box and conserved sequences known as iterons that are specific to the binding site of the replication-associated protein (Rep) (Argüello-Astorga *et al*., 1994; Argüello-Astorga and Ruiz-Medrano, 2001; Fontes *et al*., 1994). Most begomoviruses include two iterated sequences located upstream the TATA box and one inverted repeat. However, some begomoviruses have the iteron downstream of the TATA box, which is usually shorter, imperfect and maintaining a core of at least five nucleotides (Argüello-Astorga *et al*., 1994). Cognate DNA-A and DNA-B components of bipartite begomoviruses share a common region (CR) of approximately 200 nts, located within the IR, that contains cis-element for replication and control of gene expression (Argüello-Astorga *et al*., 1994; Argüello-Astorga and Ruiz-Medrano, 2001; Eagle *et al*., 1994). The CR exhibits a high degree of sequence identity of both genome components of bipartite begomoviruses (Brown *et al*., 2012). The virion-strand origin of DNA replication consists of a predicted hairpin structure containing a conserved nonanucleotide, TAATATTAC sequence in the loop and repeated upstream motifs (Stanley, 1995; Stanley *et al*., 2005).

Begomoviruses replicate through an intermediate double stranded DNA molecule in the nuclei of infected host plant cells and depend upon host DNA replication machinery (Bisaro *et al*.,

1996; Jeske, 2007). Their genomes are among the smallest known virus genomes. Begomoviruses utilize a bidirectional mode of transcription and overlapping genes for efficient coding of proteins (Rojas *et al*., 2005). The DNA-A component of bipartite begomoviruses can replicate autonomously and produce virions but requires the DNA-B component for movement.

Begomovirus ORFs and proteins codified

Monopartite begomovirus genomes and DNA-A of bipartite begomoviruses contain five or six open reading frames (ORFs) while DNA-B of bipartite begomoviruses contain two ORFs. The proteins coded by these ORFs and their functions are summarized below.

i) ORF (A)V1 codes for coat protein (CP). CP is used for encapsidation, insect transmission and movement in plants (Hanley-Bowdoin *et al*., 2000; Padidam *et al*., 1996; Sharma and Ikegami, 2009). It also interferes with nicking of DNA thus limiting the viral DNA copy number during rolling circle replication (RCR) (Rojas *et al*., 2005; Yadava *et al*., 2010).

ii) ORF (A)V2 codes for pre-coat protein (pre-CP). This is a pathogenicity determinant (Padidam *et al*., 1996) and it is involved in virus movement in plants (Sharma *et al*., 2011). In addition, it acts as a suppressor of RNA silencing (Yadava *et al*., 2010; Zrachya *et al*., 2007). ORF V2 is not present in New World begomoviruses.

iii) ORF (A)C1 codes for the replication associated protein (Rep). Rep is an RCR-initiator protein that recognizes iterons and nicks within the nonanucleotide sequence to initiate viral DNA replication (Hanley-Bowdoin *et al*., 2004; Nash *et al*., 2011). It also executes ATPase and helicase activities and stimulates the DNA synthesis by controlling the cell cycle through interaction with the plant retinoblastoma-related protein factor (pRBR) (Choudhury *et al*., 2006; Pant *et al*., 2001).

iv) ORF (A)C2 codes for a transcription activator protein (TrAp). It activates the late virion-sense transcription for bipartite begomoviruses, serves as a suppressor of RNA silencing in bipartite begomoviruses (Trinks *et al*., 2005; Yang *et al*., 2007). It also overcomes virus induced hypersensitive cell death (Hussain *et al*., 2007; Mubin *et al*., 2010).

v) ORF (A)C3 codes for a replication enhancer protein (REn). REn establishes conducive environment for virus replication (Pedersen and Hanley-Bowdoin, 1994) and stimulates viral DNA replication (Pasumarthy *et al*., 2011).

iv) ORF (A)C4 codes for the (A)C4 protein. The function of (A)C4 still remains unclear but some viruses used it as pathogenicity determinant and suppressor of post-transcriptional gene silencing (PTGS) of siRNAs (Gopal *et al*., 2007; Pandey *et al*., 2009; Vanitharani *et al*., 2004).

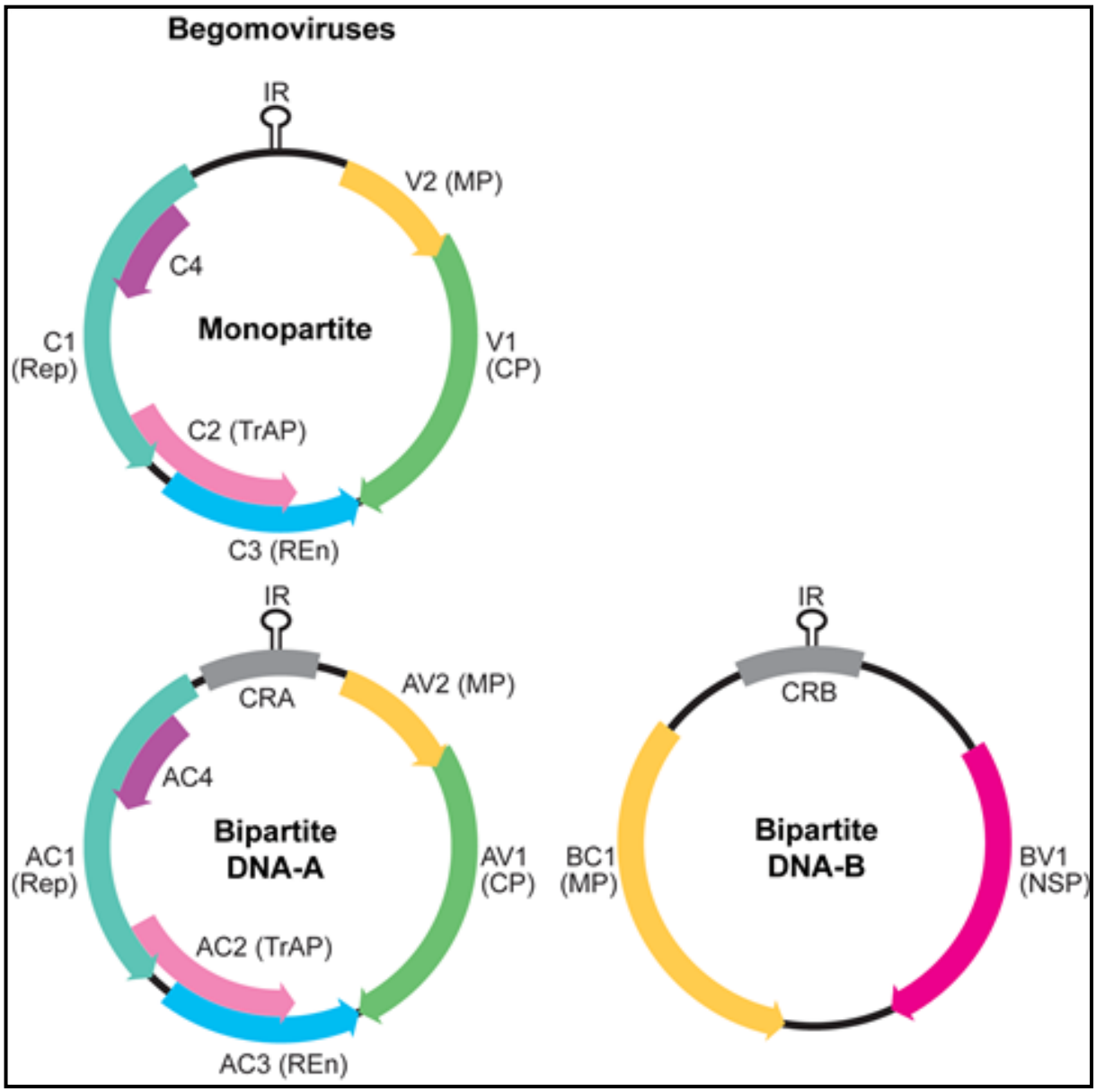

Figure 2. Genomic organization of monopartite and bipartite begomoviruses. ORFs are denoted as either being contained on the virion-sense (V) or complementary-sense (C) strand. The "common region" (CRA and CRB) represent intergenic sequences that are shared between the two genomic components of bipartite viruses. The position of the stem-loop containing the conserved TAATATTAC sequence, located within the intergenic regions (IR, top; CRA and CRB, bottom) is shown. Coat protein, CP; Pre-coat protein, Pre-CP; Replication-associated protein, Rep; Transcriptional activator protein, TrAP; Replication enhancer, REn; Movement protein, MP; Nuclear shuttle protein, NSP (Figure reproduced from Brown *et al.*, 2012).

The DNA-B component contains two ORFs in opposite orientation coding for proteins involved in inter- and intra-cellular movement of the virus within the host plant.

i) ORF BV1 codes for the nuclear shuttle protein (NSP). It transports viral DNA from the nucleus to the cytoplasm (Sanderfoot *et al*., 1996) and act as a pathogenicity determinant (Hussain *et al*., 2005).

ii) ORF BC1 codes for a movement protein (MP). It coordinates viral DNA movement across plasmodesmatal boundaries (Sanderfoot and Lazarowitz, 1996) and it is also responsible for viral pathogenic properties (Jeffrey *et al*., 1996).

Phylogeny of begomoviruses

At present, based on the phylogenetic analysis of complete nucleotide sequences of DNA-A and genome organization, whitefly transmitted-begomoviruses can be classified into four lineages; Old World (OW) begomoviruses, New World (NW) begomoviruses, sweepoviruses and legumoviruses (Briddon *et al*., 2010; Ilyas *et al*., 2009; Trenado *et al*., 2011) (**Figure 3**).

Old World and New World begomoviruses

OW and NW begomoviruses are originated from the Old World (Africa, Asia and Europe) and New World (The Americas), respectively. OW begomoviruses can have monopartite or bipartite genomes whereas most of the NW begomoviruses have bipartite genomes. A few exceptions of monopartite begomoviruses natives to the NW exist, including tomato leaf deformation virus, reported from Peru and Ecuador (Márquez-Martín *et al*., 2011; Melgarejo *et al*., 2013; Sánchez-Campos *et al*., 2013). Monopartite begomovirus genomes and DNA-A components of OW begomoviruses contain ORF V2, which is absent in NW begomovirus genomes.

Legumoviruses and sweepoviruses

The genomes of most legumoviruses are bipartite, although a DNA-B component has not been identified for cowpea golden mosaic virus (CPGMV), Dolichos yellow mosaic virus (DoYMV) and soybean mild mottle virus (SbMMoV) (Alabi *et al*., 2010; Maruthi *et al*., 2006). Phylogenetically, they segregate basal to the main branch of OW begomoviruses (Briddon *et al*., 2010; Fauquet *et al*., 2008). Sweepoviruses are monopartite begomoviruses infecting sweet potato (*Ipomoea batatas*) and other members of the family *Convolvulaceae* (Fauquet and Stanley, 2003; Ilyas *et al*., 2009; Lozano *et al*., 2009; Trenado *et al*., 2011). Sweepovirus group clusters separately from the rest of the begomovirus species and appears to belong to a branch distinct from the Old and New World groups (Briddon *et al*., 2010; Trenado *et al*., 2011).

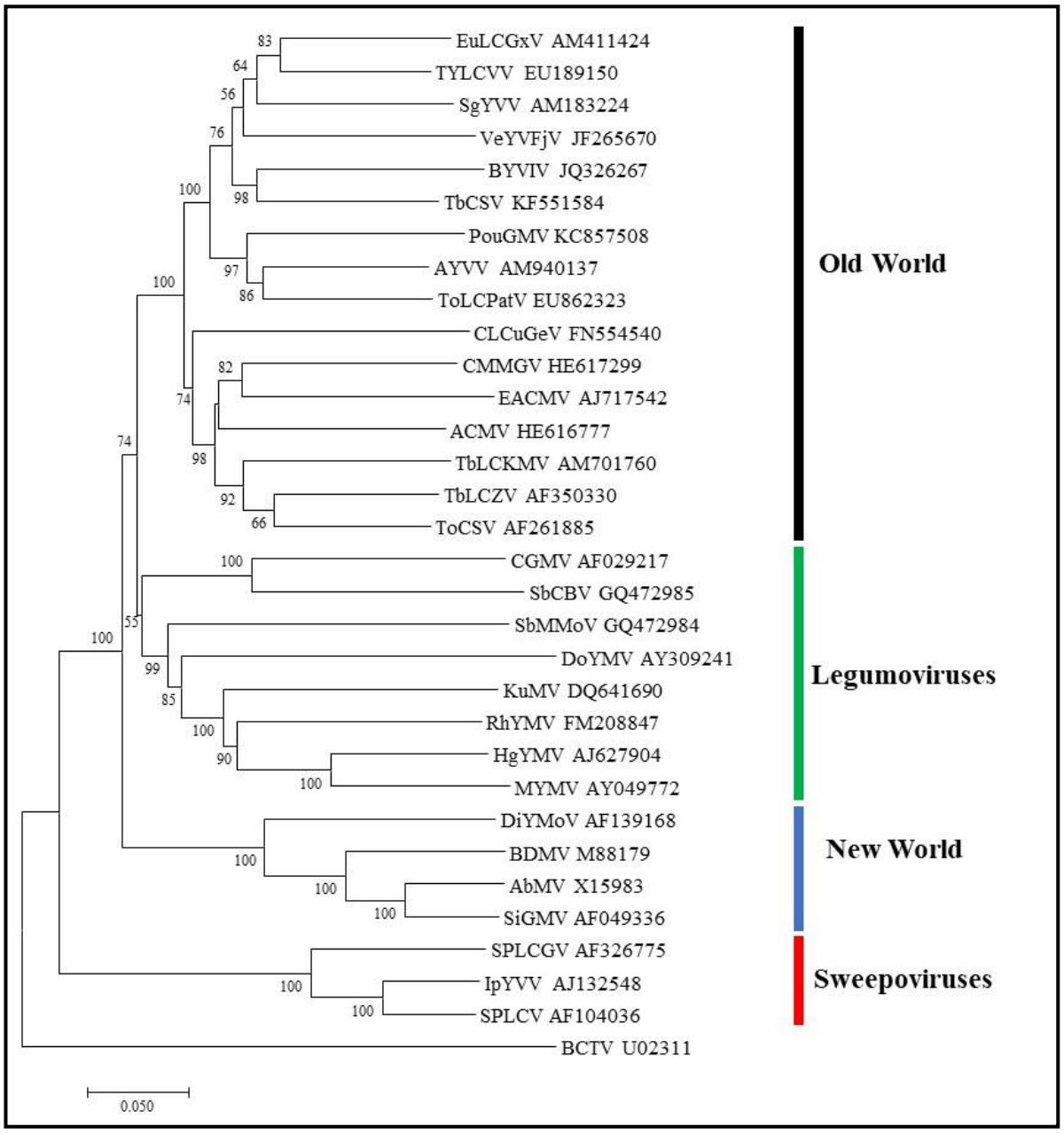

Figure 3. Begomovirus groups phylogeny based on the complete nucleotide sequences of DNA-A components retrieved from the GenBank. Sequences were aligned using the ClustalW algorithm and the tree was constructed using Neighbor-joining algorithm implemented by MEGA7. The tree was rooted by using beet curly top virus (BCTV, genus *Curtovirus*) as an outgroup. Bootstrap values (1,000 replicates) are given at the branch nodes. Only bootstrap values higher than 50% are indicated.

DNA satellites associated with begomoviruses

DNA satellites are subviral agents which lack genes encoding functions for replication, so rely on their helper virus for multiplication in the infected host cell (Mayo *et al*., 2005). DNA satellites have no similarities to begomoviruses in genome sequence apart from the presence of a stem-loop structure that is required for the Rep protein (Zhou, 2013). DNA satellites associated with begomoviruses are classified into two families; *Alphasatellitidae*, a family that comprises two subfamilies- *Geminialphasatellitinae* and *Nanoalphasatellitinae* with four and seven genera respectively (Briddon *et al*., 2018), and the recently accepted family *Tolecusatellitidae* which contains two genera, *Betasatellite* and *Deltasatellite* (https://talk.ictvonline.org/ICTV/proposals/2016.021a-kP.A.v2.Tolecusatellitidae).

DNA satellites are mostly found to be associated with majority of the monopartite begomoviruses (Briddon and Stanley, 2006; Mansoor *et al*., 1999; Fiallo-Olivé *et al*., 2013; Lozano *et al*., 2009). This type of begomovirus-complexes frequently prevails in the Old Word including Asia and Africa. However, recently alphasatellites and deltasatellites have been found associated with bipartite begomoviruses from the New World (Fiallo-Olivé *et al*., 2012, 2016; Paprotka *et al*., 2010; Romay *et al*., 2010). The genome organization of begomoviruses and associated DNA satellites (betasatellites, alphasatellites and deltasatellites) are represented in Figure 4.

Betasatellites

Betasatellites, formerly DNA-β, are circular, ssDNA molecules ~1350 nt, about half of the begomovirus genome components (Saunder *et al*., 2000). They share no significant sequence identity with their helper viruses other than a potential stem-loop structure containing the ubiquitous nonanucleotide TAATATTAC sequence that serves as the origin of virion-strand DNA replication in begomoviruses (Briddon *et al*., 2003). They completely rely on helper begomoviruses for replication, encapsidation and movement within and between plant apparently by trans-encapsidation in the helper virus CP (Mayo *et al*., 2005; Nawaz-ul-Rehman and Fauquet, 2009; Saunder *et al*., 2000). Characteristically, betasatellites have a sequence rich in adenine (A-rich region), a sequence (~120 nts) that is highly conserved among all betasatellites (satellite conserved region, SCR) and an ORF, βC1, located in the complementary sense strand which encodes a protein of ~118 amino acids (Briddon *et al*., 2003). Betasatellites lack the iterons of their helper begomoviruses and Rep protein is believed to instead recognize iteron-like sequences residing between the A-rich and SCR sequences to initiate RCR (Saunders *et al*., 2008).

All functions of betasatellites reported so far are related to the presence of βC1 protein, including suppression of transcriptional and posttranscriptional gene silencing (Cui *et al*., 2005; Zhou, 2013), virus movement in plant (Saeed *et al*., 2007), up-regulation of viral DNA levels (Briddon *et al*., 2001; Saunders *et al*., 2000), interaction with a variety of host factors and CP of helper viruses (Eini *et al*., 2009; Kumar *et al*., 2006) and induction of disease symptoms in infected plants (Cheng *et al*., 2011).

Alphasatellites

Alphasatellites, formerly DNA 1, are circular ssDNA molecules of about half size (~1400 nts) of the begomovirus genome components. They are not true satellites molecules, because they are capable of self-replicating, so are designed as satellite-like molecules (Mansoor *et al*., 1999; Mayo *et al*., 2005). They contain a single gene in the virion-sense orientation that encodes a Rep protein and an A-rich region. Majority of alphasatellites contain a predicted hairpin structure with a loop containing a nonanucleotide TAGTATTAC sequence that is common to nanoviruses (Briddon *et al*., 2004; Mansoor *et al*., 1999), however, an alphasatellite with a distinct nonanucleotide sequence (TAATATTAC) that is common to geminiviruses has been reported in Kenya (Briddon *et al*., 2004).

Although alphasatellites are associated mostly with OW monopartite begomoviruses, and frequently in association with begomovirus-betasatellites complexes, they have also been characterized in the NW in association with bipartite begomoviruses (Mar *et al*., 2017; Paprotka *et al*., 2010; Romay *et al*., 2010). Some research has shown that alphasatellites cause a decrease in viral DNA levels in the plant (Briddon *et al*., 2004). Nevertheless, an unusual alphasatellite associated with tomato yellow leaf curl virus (TYLCV) has been described to ameliorate begomovirus symptoms significantly and reduce betasatellites DNA accumulation in plants (Idris *et al*., 2011).

Deltasatellites

Deltasatellites are circular ssDNA molecules of ~700 nt, about one quarter the size of a begomovirus genomic component with no significant sequence similarity with the helper viruses. They have been found to be associated with monopartite OW begomoviruses (Dry *et al*., 1997), bipartite NW begomoviruses (Fiallo-Olivé *et al*., 2012) and sweepoviruses (Lozano *et al*., 2016). Unlike betasatellites and alphasatellites, deltasatellites do not encode for any protein. They need the helper begomovirus for replication and movement in plants and transmission by *B. tabaci* (Fiallo-Olivé *et al*., 2016; Hassan *et al*., 2016). They contain a sequence with similarity to the SCR of betasatellites, an A-rich sequence, a predicted stem-loop structure containing the nonanucleotide TAATATTAC, and an additional (secondary) predicted stem-loop (Fiallo-Olivé *et al*., 2012; Lozano *et al*., 2016). The

presence of deltasatellites in the host plant can cause a reduction of begomovirus accumulation and/or symptom severity (Fiallo-Olivé *et al*., 2016; Hassan *et al*., 2016).

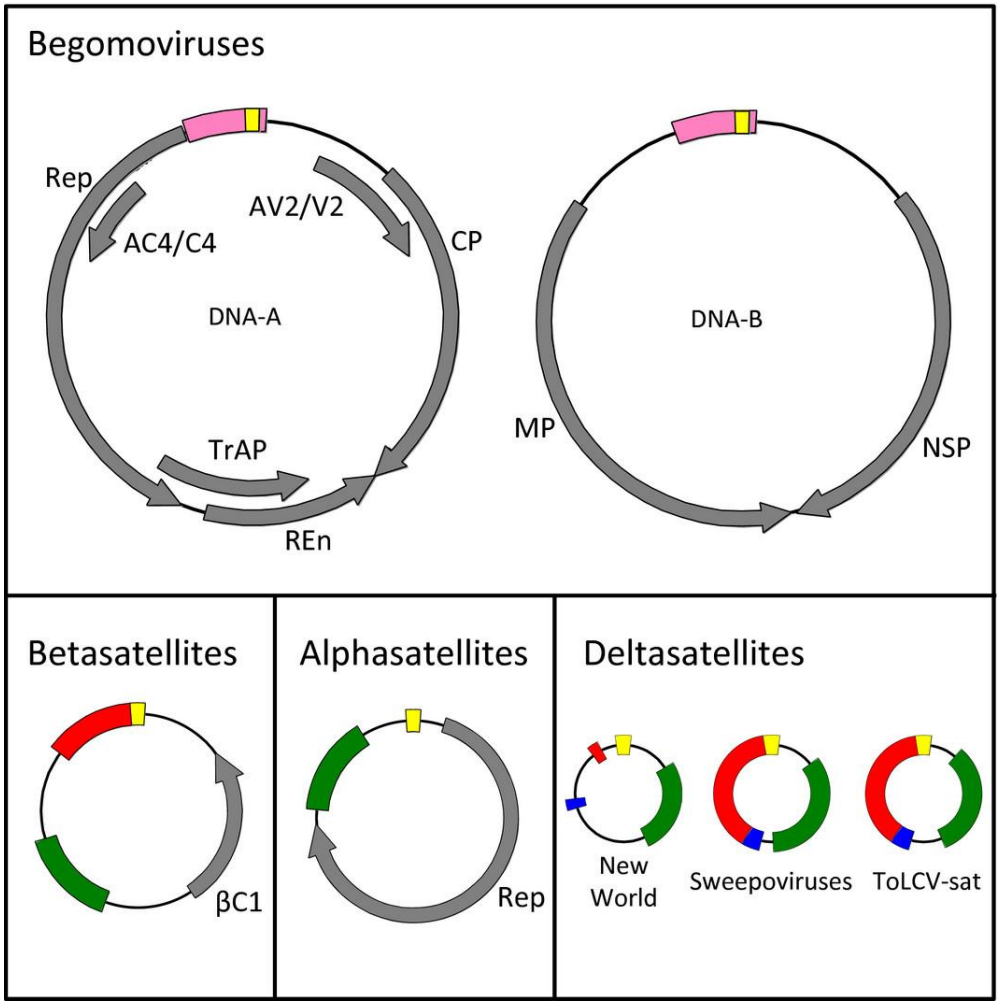

Figure 4. Schematic representation of begomovirus genome and associated DNA satellites (betasatellites, alphasatellites and deltasatellites). Deltasatellites include molecules associated with New World bipartite begomoviruses (New World), Convolvulaceae-infecting begomoviruses (sweepoviruses), and tomato leaf curl virus (ToLCV-sat). Begomovirus genomes include the conserved stem-loop (yellow box) which is located in the common region (pink box) and encoded proteins (grey arrows): AV2/V2, coat protein (CP), replication-associated protein (Rep), AC4/C4, transcriptional activator protein (TrAP), replication enhancer (REn), movement protein (MP) and nuclear shuttle protein (NSP). Monopartite begomoviruses lack a DNA-B component and AV2 is absent in New World begomoviruses. Main genome features for betasatellites, alphasatellites and deltasatellites are shown: conserved stem-loop (yellow box), satellite conserved region (red box), A-rich region (green box) and secondary stem-loop (blue box). Proteins encoded by betasatellites (βC1) and alphasatellites (Rep) are represented by grey arrows (Figure reproduced from Fiallo-Olivé *et al*., 2016).

Taxonomy and nomenclature of begomoviruses and betasatellites

The classification of viruses into species, strains and variants is established by the ICTV Study Groups. Due to the small size of begomovirus genomes and common use of rolling circle amplification (RCA) (Inoue-Nagata *et al*., 2004), there is an increasing number of viral genomic sequences deposited in public databases. The *Geminiviridae* Study Group of the ICTV has proposed an updated set of demarcation criteria for classification and naming begomovirus species (Brown *et al*., 2015). It has been established that a new DNA-A or monopartite begomovirus genome sequence with <91% pairwise identity to any other published, would belong to a new begomovirus species whereas an isolate with sequence identity ≥91% with an isolate belonging to an established species would be classified as a member of that species. Similarly, it has been established a threshold of <94% for begomovirus strains. New virus species should follow a standard nomenclature structure which usually includes the infected host plant, the symptoms observed in the field and the word "virus" (Brown *et al*., 2015; Zerbini *et al*., 2017).

According to the recently proposed species demarcation threshold for the genus *Betasatellite* (family *Tolecusatellitidae*), any new sequence of a betasatellite <91% identical to any published betasatellite sequence, it would belong to a novel betasatellite species (https://talk.ictvonline.org/files/proposals/taxonomy_proposals_plant1/m/plant02/6357). The naming of betasatellite species and isolates follows the agreement set down for the helper begomoviruses (Briddon *et al*., 2008).

Factors that govern genetic diversity of begomoviruses

The evolution of begomoviruses and their genetic diversity are mostly driven by mutation, recombination and pseudo-recombination which contribute significantly to the appearance of new viral variants, increasing their potential of adaptation to different hosts and environmental conditions.

Mutation

Mutation is a change in the nucleotide sequence of a short region of a genome. It occurs during several different biological processes like replication slippage and can also be induced by UV-light and chemical treatments. Incorporation of a non-complementary nucleotide during duplication of DNA or RNA gives rise to point mutations, which alters the genetic information. The mutation rate of RNA viruses was thought to be higher than DNA viruses (Domingo and Holland, 1997), however it was noticed that the mutation rate of ssDNA begomoviruses was similar to that of RNA viruses (Duffy and Holmes, 2008, 2009; Fondong and Chen, 2011; Ge *et al*., 2007). It has been shown that

accumulation of point mutations contributes to the diversification of natural begomovirus populations (Lima *et al*., 2017; Van Der Walt *et al*., 2008).

Recombination

Recombination is a process in which exchange of segments occurs between two strands of DNA or RNA during replication. The genetic diversity of begomoviruses and their evolution is being driven by interspecific homologous recombination. Many studies have reported that an important number of begomovirus epidemics are directly linked to recombinant viruses, for example, cassava mosaic disease (CMD) in Africa (Bull *et al*., 2006; Pita *et al*., 2001), tomato yellow leaf curl disease (TYLCD) in Mediterranean basin (García-Andrés *et al*., 2007a; Navas-Castillo *et al*., 2000), and cotton leaf curl disease (CLCuD) in the Indo-Pakistan sub-continent (Mansoor *et al*., 2003; Zhou *et al*., 1998). Also, recombination plays an important role in the recurrent speciation of viruses causing TYLCD (Fiallo-Olivé *et al*., 2019a).

Pseudo-recombination (reassortment)

Pseudo-recombination is another source of genetic variability, which occur by the exchange (reassortment) of genomic components between bipartite begomoviruses, belonging or not to the same species (Briddon *et al*., 2010; Hou and Gilbertson, 1996; John *et al*., 2008; Rojas *et al*., 2005; Silva *et al*., 2014).

Begomoviruses infecting non-cultivated crops

Non-cultivated plants, including weeds and other wild plants, are widely distributed throughout the world. There are about 250,000 species of plants worldwide (Heywood, 1993), and approximately 3% of them behave as weeds (Holm *et al*., 1979). Benefits associated with weeds in agriculture include soil stability, source of organic matter, human and animal consumption, insect's repellent and herbal medicine (Hillocks, 1998; Stepp, 2004; Stepp and Moerman, 2001). However, a negative impact of weeds in agriculture includes reducing crop yield and quality as well as serving as alternative hosts for crop diseases or provide shelter to insect to overwinter (Hillocks, 1998; Oerke, 2006). Weed plants are infected by many plant viruses, acting as primary inoculum sources, serving as alternative host for the existing viruses and allowing transmission to the economic important crops within the region.

Numerous species of non-cultivated plants, especially of the families Euphorbiaceae, Fabaceae, Malvaceae and Solanaceae, are known hosts of begomoviruses (Morales and Anderson, 2001; Rojas *et al*., 2018). These weed/wild hosts can serve as reservoirs for infection of nearby crops

(Barbosa *et al*., 2009; Bedford *et al*., 1998; García-Andrés *et al*., 2006), as overwintering area (García-Andrés *et al*., 2006) and as 'mixing vessels' for interspecific coinfection and recombination (García-Andrés *et al*., 2006; Silva *et al*., 2012). Several studies have suggested that the diversity of begomovirus populations in non-cultivated hosts is higher than that observed in crop-infecting begomoviruses (Silva *et al*., 2012; Wyant *et al*., 2011).

Although weeds are reservoir host of existing begomoviruses and source of new viruses worldwide, a limited number of studies have been conducted to determine their effects to important crops in the African continent, especially in East Africa. Thus, most of the studies in this area have been focused on characterization of begomoviruses infecting crops (Bull *et al*., 2006; Gibson *et al*., 1996; Lefeuvre *et al*., 2007; Leke *et al*., 2011; Ndunguru *et al*., 2005; Tiendrébéogo *et al*., 2010; Wasswa *et al*., 2011; Zhou *et al*., 2008; Zinga *et al*., 2013).

The whitefly *Bemisia tabaci*

Whiteflies are hemipterans in the family *Aleyrodidae*, composed of more than 160 genera and more than 1,500 species (Martin and Mound, 2007). Two whiteflies, *Bemisia tabaci* (sensu lato) and *Trialeurodes vaporariorum*, cause considerable economic losses to agriculture worldwide (Brown, 1994; Byrne *et al*., 1990; Martin 1987). *B. tabaci* is believed to have originated in Africa (De Barro *et al*., 2011) and it is the most widely distributed and economically important group of whitefly species worldwide (De Barro *et al*., 2011; Dinsdale *et al*., 2010). *B. tabaci* has gained increased significance as a pest in the last decades (Brown, 1994) and it is a vector of around 430 plant virus species which include members of the genera *Begomovirus* (*Geminiviridae*), *Carlavirus* (*Betaflexiviridae*), *Crinivirus* (*Closteroviridae*), *Ipomovirus* (*Potyviridae*), *Polerovirus* (*Luteoviridae*) and *Torradovirus* (*Secoviridae*) (Ghosh *et al*., 2019; Jones, 2003; Navas-Castillo *et al*., 2011; Polston *et al*., 2014).

B. tabaci causes direct damage to plants through extraction of large quantities of phloem sap, which can result in greater than 50% yield reductions. In addition, direct feeding of the whiteflies can result in irregular ripening of fruits, chlorosis of leaves, stem blanching and induction of phytotoxic disorders resulting in silvering in few genotypes of some plant species (Byrne *et al*., 1990; Costa and Brown, 1991). The effects are severe when large populations of *B. tabaci* colonize a plant. The indirect damage is due to excretion of honey dew onto the surface of the leaves and fruits that serves as a media for sooty mould growth which darkens leaves and fruits and hence prevents the photosynthesis process in the plant and reduces the quality of fruits and fibre (Byrne and Bellows, 1991). Moreover, the most important indirect damage caused by *B. tabaci* is due to its role as a vector of hundreds of plant viruses infecting cultivated and non-cultivated plants.

Life history

B. tabaci adults feed and lay eggs on the undersides of leaves (**Figure 5A**). The number of eggs per female vary depending on the host plant and environmental conditions (Byrne and Bellows, 1991). The eggs are ovoid and have a pedicel which is a primary channel through which water is absorbed from the plant (Buckner *et al*., 2002), and a guide for spermatozoa during fertilization (Byrne and Bellows, 1991). The eggs hatch at the apical end after completion of development to release crawlers (first instar) (Byrne and Bellows, 1991). The first instar nymphs have functional legs, and antennae with two and three apparent segments, respectively (Gill, 1990). The first instar nymphs move quickly in their search for a feeding site and once successful, they remain sessile until the adult emerges. The second and third instars nymphs differ in size; they are oval and appear to only have one segment in their legs and antennae (Gill, 1990; Legg, 1994). The fourth instar or 'pupa' is shield shaped, broadly elliptical with two red eye spots at the anterior end visible beneath the translucent integument (Byrne and Bellows, 1991).

Adult eclosion of *B. tabaci* is affected by temperature and photoperiod, and mostly occurs during the first hours of the photophase. The body length and wing expanse of adults differ in size after emergence from the pupa case (**Figure 5B**). On average, females measure is 0.91 mm and 2.13 mm in body length and wing expanse, while males are 0.85 mm and 1.81 mm in body length and wing expanse, respectively (Gill, 1990). Newly emerged adults are sexually immature, and copulation takes place 1 to 8 h after emergence in summer, or up to three days later during spring and fall (Byrne and Bellows, 1991). *B. tabaci* reproduces parthenogenetically with the mated females producing both haploid males and diploid females, whilst the unmated females produce only haploid male offspring (Byrne and Bellows, 1991). The ratio of female to male usually differs greatly under field conditions and depends on host species, temperature and time of the year (Sharaf *et al*., 1985). The average longevity of the adults differs with females having a longer life span (35 days) than the males (20 days) (Butler *et al*., 1983).

The life cycle of *B. tabaci* from egg to adult (**Figure 6**) differs significantly depending on the host plant it feeds on (Coudriet *et al*., 1985). On average, twelve generations are attained annually under field conditions (Butler *et al*., 1983; Fishpool and Burban, 1994; Legg, 1994). Generally, higher temperatures result in faster development times compared to lower temperatures (Gerling *et al*., 1986).

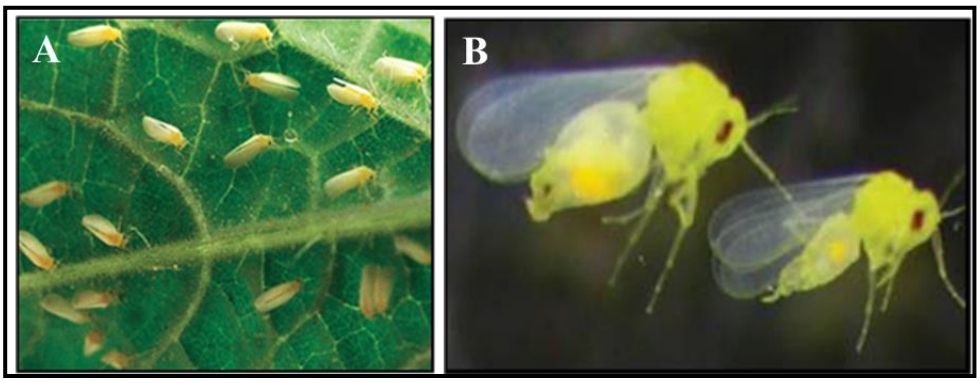

Figure 5. *Bemisia tabaci* infestation of adults underneath a leaf (**A**) (Figure reproduced from Rojas *et al*., 2018) and female (left) and male (right) *B. tabaci* adults (**B**) (Figure reproduced from Mugerwa, 2018).

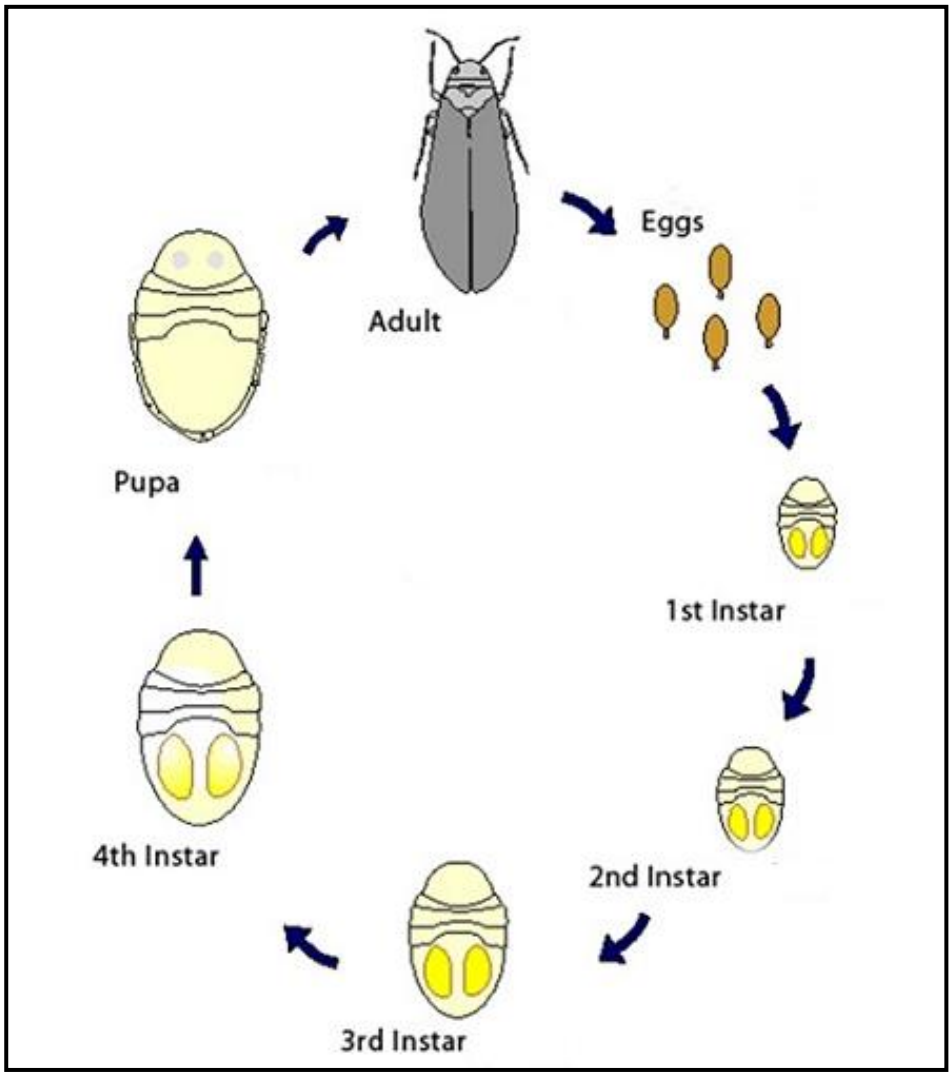

Figure 6. Life cycle of *B. tabaci* (Figure adapted from https://biocontrol.ucr.edu/bemisia.html)

Relative humidity and host plant are among the several factors which affect *B. tabaci* survival from egg to adult (Gerling *et al.*, 1986; Xu *et al.*, 2011; Zang *et al.*, 2006). High mortality rates can occur for whitefly immature development stages; eggs and early nymph instar stages have been reported to occur at low (≤20%) and high (≥80%) relative humidity during temperatures fluctuation, 22-30°C. However, when the second to fourth instars are established, they are more resistant to the effects of low and high humidity (Gerling *et al.*, 1986). *B. tabaci* survival rates vary on different host plants. For instance, an Asia II-2 population was reported to have a significantly higher survival rate on cotton as compared to squash, however, they did not survive on cabbage, kidney beans and tobacco (Zang *et al.*, 2006). In contrast, an Asia II-1 population survived on all plants used in the experiment with the highest survival rates on kidney beans and squash. In the same study, a MEAM1 population had a significantly higher survival rate on cabbage and cotton as compared to kidney beans, squash, and tobacco (Zang *et al.*, 2006).

Host range, colonization and population dynamics

In the late 1990s, *B. tabaci* was considered by some to be a single, polyphagous whitefly species colonizing over 500 plant species from 74 families (Brown *et al*., 1995b). Since then, however, the weight of evidence shows that *B. tabaci* is a group of more than 40 cryptic biological species (see section 2.3 below). The existence of monophagous *B. tabaci* populations (Brown *et al*., 1995b; Burban *et al*., 1992), is consistent with there being multiple, cryptic, biological species within what is still known as *B. tabaci*.

The population dynamics of *B. tabaci* species are affected by climatic conditions such as relative humidity, temperature, rain and wind, and host plant availability (Avidov, 1956; Byrne and Bellows, 1991). Furthermore, natural enemies such as parasitoids and predators, and insecticides also affect survival (Bellows and Arakawa, 1988; Fishpool *et al*., 1995; Horowitz *et al*., 1984; Horowitz, 1986; Naranjo and Ellsworth, 2005).

Systematics

Genetic complexity of *B. tabaci* was first documented in the 1950's when morphologically indistinguishable populations were reported to vary biologically and ecologically (Bird, 1957). As these named species were morphologically indistinguishable and their nymph morphology was plastic (varying according to the host plant species they developed on), they were synonymized into a single species called *Bemisia tabaci* (Russell, 1957) (Table 1).

Table 1 Whitefly species synonymized as *B. tabaci* (Russel, 1957).

No	Species	Described by	Year	Country
1	*Aleurodes tabaci*	Gennadius	1889	Greece
2	*Aleurodes inconspicua*	Quaintance	1900	Florida, US
3	*Bemisia incospicua*	Quaintance and Baker	1914	Florida, US
4	*Bemisia emiliae*	Corbett	1926	Sri Lanka
5	*Bemisia signata*	Bondar	1928	Brazil
6	*Bemisia costa-limai*	Bondar	1928	Brazil
7	*Bemisia bahiana*	Bondar	1928	Brazil
8	*Bemisia gossypiperda*	Misra and Lamba	1929	India, Pakistan
9	*Bemisia achyranthes*	Singh	1931	India
10	*Bemisia hibisci*	Takahashi	1933	Taiwan
11	*Bemisia longispina*	Priesner and Hosny	1934	Egypt
12	*Bemisia gossypiperda var. mosaicivectura*	Ghesquiere	1934	Congo
13	*Bemisia goldingi*	Corbett	1935	Nigeria
14	*Bemisia nigeriensis*	Corbett	1935	Nigeria
15	*Bemisia rhodesiaensis*	Corbett	1936	Zimbabwe
16	*Bemisia tabaci*	(Gennadius) Takahashi	1936	Mariana Islands
17	*Bemisia manihotis*	Frappa	1938	Madagascar
18	*Bemisia vayssierei*	Frappa	1939	Madagascar
19	*Bemisia (Neobemisia) hibisci*	(Takahashi) Visnya	1941	Taiwan
20	*Bemisia (Neobemisia) rhodesieaensis*	(Corbett) Visnya	1941	Zimbabwe
21	*Bemisia lonicerae*	Takahashi	1957	Japan
22	*Bemisia minima*	Danzig	1964	Georgia
23	*Bemisia miniscula*	Danzig	1964	Georgia

This synonymization, based on classical taxonomy, caused problems and generated controversy for entomologists, due to the distinct biological traits exhibited by different populations (Burban *et al*., 1992; Costa and Russell, 1975; Legg *et al*., 1994). A new naming system developed, therefore, for the morphologically indistinguishable *B. tabaci* populations showing different biological traits. These were known as biotypes or host races (Burban *et al*., 1992; Brown *et al*., 1995a). For example, in the southern U.S, a *B. tabaci* population capable of inducing silverleaf symptoms in squash (*Cucurbita* sp.) was named the "B biotype", due to the production of a distinct electromorph profile, using esterase typing, from that of the indigenous American *B. tabaci* population, named the "A biotype" (Costa and Brown, 1990; Costa *et al*., 1993a). Several other biotypes or host races including the 'E' on *Asystasia* sp. in Benin and 'J' polyphagous in Nigeria (Bedford *et al*., 1994), 'N' *Jatropha gossypifolia* (Bird and Maramorosch, 1978), 'cassava' and 'okra' (Burban *et al*., 1992), 'Sida' and a polyphagous non-cassava biotype in Brazil (Bird and Maramorosch, 1978; Costa and Russell, 1975), S on *Ipomoea indica* and Q biotype in Spain (Banks *et al*., 1999; Guirao *et al*., 1997) have been reported. This led to the development of a concept that *B. tabaci* composed of a series of biotypes (Bedford *et al*., 1994; Costa and Brown, 1991). To date, over 35 'biotypes' have been named (Banks and Markham 2000; Liu *et al*., 2012; Perring, 2001) based on esterase profiles (Brown *et al*., 1995b) and biological assays such as mating compatibility and virus transmission ability (Bedford *et al*., 1994; Costa *et al*., 1993a; Maruthi *et al*., 2004, 2001).

Among the most commonly studied molecular techniques in whitefly systematics includes: protein polymorphism involving isozyme variation in esterases (Brown *et al*., 1995b; Wool *et al*., 1989), DNA-based molecular techniques, like random amplified polymorphic DNA (RAPD) (Shankarappa *et al*., 2007), PCR fingerprinting (De Barro and Driver, 1997; Gawel and Bartlett, 1993; Guirao *et al*., 1997), amplified fragment length polymorphism (AFLP) markers (Cervera *et al*., 2000), and mitochondrial DNA marker genes; mitochondrial cytochrome oxidase-I (mtCOI) (Brown, 2000; Frohlich *et al*., 1999; Simon *et al*., 1994), ribosomal RNAs, 16S rDNA (Prokaryotes) (Clark *et al*., 1992; Frohlich *et al*., 1999) and 18S rDNA (eukaryotes) (Campbell *et al*., 1993, 1994), and ribosomal nuclear marker of the internal transcribed spacer I (ITSI) region sequences (De Barro *et al*., 2000).

The mtCOI marker has been used extensively to study the genetic variability and evolutionary relationships among populations of *B. tabaci* from different geographical locations and host-plant species (Frohlich *et al*., 1999). In 1997, two distinct *B. tabaci* populations colonizing cassava in Uganda were identified using mtCOI and considered to be 'local' Ug1 or 'invader' Ug2, which were subsequently described as sub-Saharan Africa 1 and sub-Saharan Africa 2, respectively (Legg *et al*., 2002). In a later study, wide distribution of sub-Saharan Africa 1 was detected using the mtCOI marker in the cassava growing areas in Kenya, Uganda and Tanzania, nevertheless, sub-Saharan

Africa 2 was not identified from that study (Mugerwa *et al*., 2012). In addition, the mtCOI marker has been used to detect five distinct geographic populations of cassava *B. tabaci*, including a distinct southern African clade (Berry *et al*., 2004). Furthermore, the presence of B (=MEAM1), Q (=MED) and sub-Saharan Africa 5 whiteflies collected on cassava and non-cassava host in South Africa was identified using mtCOI (Esterhuizen *et al*., 2013).

Based on phylogenetic analysis and pairwise comparisons of mtCOI gene distance between populations of *B. tabaci* worldwide, Dinsdale *et al*. (2010) and Hu *et al*. (2011) provided a framework to suggest that *B. tabaci* is a cryptic species complex containing 11 higher genetic groups and at least 24 morphologically indistinguishable species. Based on field surveys conducted in India (Chowda-Reddy *et al*., 2012) and other countries, more putative species have been added to the list. While phylogenetic species themselves have limited biological significance, because their criteria and degrees of difference are intrinsically subjective (Avise 2000), the presence of a good delineation of phylogenetic species within a cryptic species complex provides a realistic structure (De Barro *et al*., 2011). Members of a cryptic species complex are indistinguishable morphologically and so ultimate differentiation between them comes from evidence of reproductive isolation associated with nuclear marker.

The *B. tabaci* species complex is globally distributed and the putative species are named based on their geographic locations and 3.5% pairwise genetic divergence identified by Dinsdale *et al*. (2010) as a putative boundary separating different species among the *B. tabaci*. These include Mediterranean; Middle East-Asia Minor 1; Middle East-Asia Minor 2; Indian Ocean; Asia I; Australia/Indonesia; Australia; China; China 2; Asia II 1; Asia II 2; Asia II 3; Asia II 4; Asia II 5; Asia II 6; Asia II 7; Asia II 8; Italy; sub-Saharan Africa 1 (SSA1); sub-Saharan Africa 2 (SSA2); sub-Saharan Africa 3 (SSA3); sub-Saharan Africa 4 (SSA4); sub-Saharan Africa 5 (SSA5); sub-Saharan Africa 6 (SSA6); sub-Saharan Africa (SSA7); sub-Saharan Africa 8 (SSA8); sub-Saharan Africa 9 (SSA9); sub-Saharan Africa 10 (SSA10); sub-Saharan Africa 11 (SSA11); sub-Sahara Africa 12 (SSA12); sub-Sahara Africa 13 (SSA13); New World; Uganda; Asia II 9; Asia II 10; Asia III; and China 3; and Asia I-India and New World 2 (Boykin *et al*., 2012; Boykin and De Barro, 2014; Lee *et al*., 2013; Mugerwa *et al*., 2018).

The 11 high-level genetic groups described by De Barro *et al.* (2011) still exhibit a substantial amount of mtCO1 diversity and delimitation cannot entirely be based on this partial mtDNA gene sequence (**Figure 7**). With nuclear markers coupled with crossing experiments between genetic groups, proper species boundaries will be identified (Hsieh *et al.*, 2014; Liu *et al.*, 2012). The genetic groups generated by the mtCO1 marker were supported by three single copy nuclear genes, namely *nuclear shaker cognate w* (*shaw*), *RNA polymerase II* (*RNApyII*) and *pre-mRNA processing factor 8*

(*prpf8*) (Hsieh *et al.*, 2014). Although the phylogenetic trees generated by partial mtCO1 and concatenated nuclear sequences showed different relationships among genetic groups, their species statuses were strongly supported (Hsieh *et al.*, 2014).

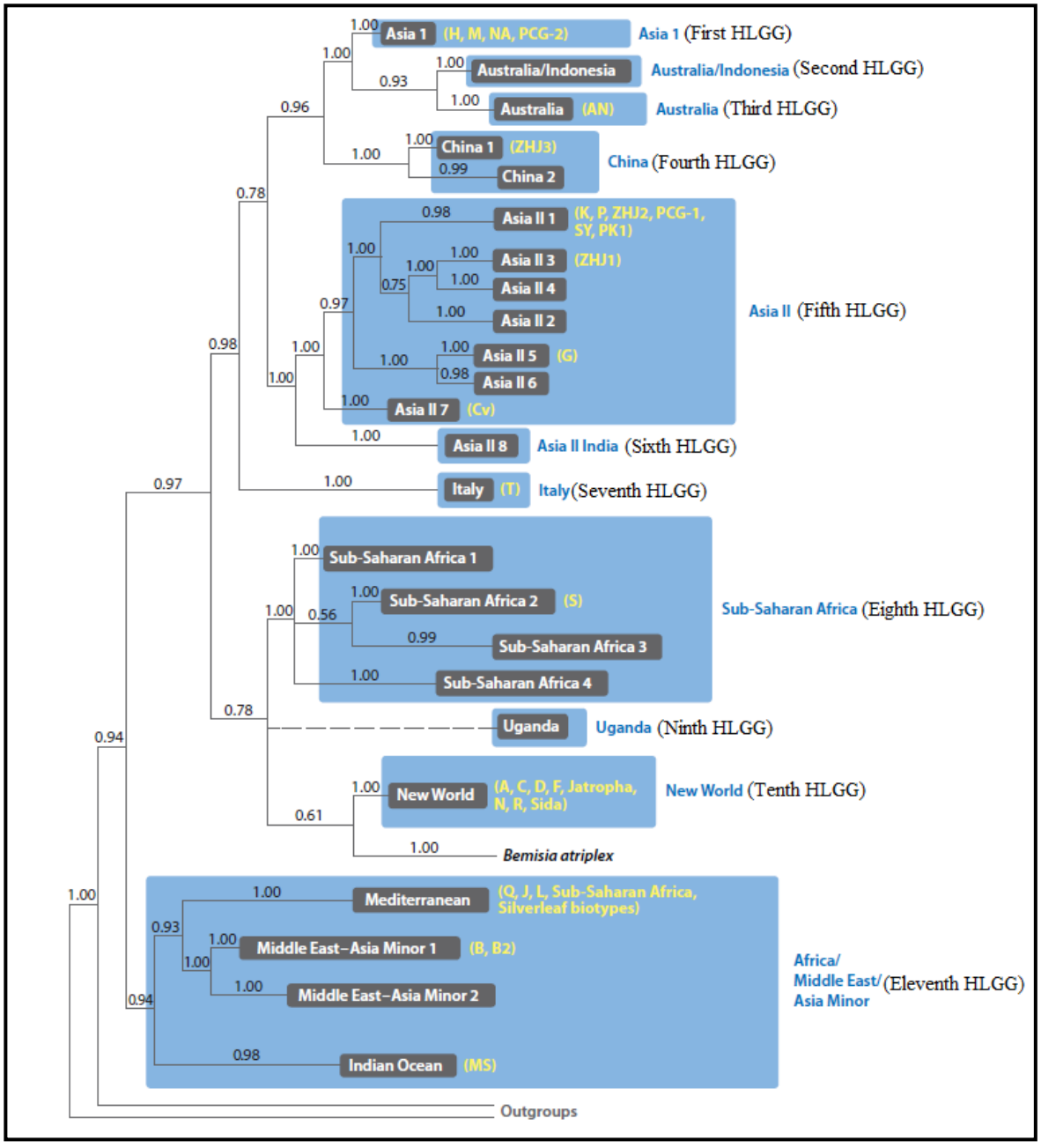

Figure 7. The *Bemisia tabaci* species group phylogeny generated by Bayesian analysis using partial mitochondrial cytochrome oxidase 1 gene sequences. Evolutionary relationships of the 11 high-level genetic groups-HLGG (blue boxes) and 24 low-level genetic groups-LLGG (black boxes) as described in De Barro *et al.* (2011).

Mating studies

Several mating studies have been conducted in order to understand the level of relatedness among the morphologically indistinguishable *B. tabaci* populations. Reproductive compatibility between putative cryptic species of *B. tabaci* has been reported based on realized fecundity and fertility between intra- and inter-species crosses (Xu *et al.*, 2010). Many crossing experiments have been conducted between MEAM1 and non-MEAM1 *B. tabaci* species (Bedford *et al.*, 1994; Costa *et al.*, 1993a; Liu *et al.*, 2007; Luan *et al.*, 2008; Perring *et al.*, 1993). Perring *et al*. (1993) reported no female progeny was produced on the reciprocal crosses between the indigenous New World species (formerly A biotype) and the newly introduced MEAM1 species. This was indicative of complete reproductive isolation between New World species and MEAM1 species. However, when the same crosses were performed by Costa *et al*. (1993a), very few females were produced, and hence concluded that the few observed F_1 female progeny was indicative of reproductive incompatibility between these species. The electrophoretic analyses of non-specific esterases were carried out on the F_1 female progeny and these all had profiles identical to the female parent (Costa *et al*., 1993a). If the F_1 females had been the result of successful crossing between A and B biotypes, however, the expected result would have been a combination of the two esterase profiles. These data and the conclusion of Costa *et al*. (1993a), therefore, was that the A and B biotypes were reproductively incompatible, even though they were considered to be the same species, at that time. Other researchers, nevertheless, have argued that the observed female progeny produced in that study (Costa *et al.*, 1993a) was due to the use of already mated female parents in inter-population crosses (Luan *et al*., 2008). Subsequently studies have supported the reproductive incompatibility observations (Bedford *et al*., 1994) reported by Perring *et al*. (1993). Reproductive incompatibility between MEAM1 and New World species was confirmed on mating behaviour experiment when no copulation took place (Perring and Symmes, 2006).

Crossing experiments carried out between MEAM1 and Australian indigenous species resulted into a high proportion of male progeny (≥70%) just as showed in unmated females for these species. Few female offspring (≤30%) produced in these crossing was attributed to successful fertilization and a high degree of incompatibility between these species (De Barro and Hart, 2000). Reciprocal crosses between MEAM1 with Asia II-1 (ZHJ2) and Asia II-2 (ZHJ1), and Asia II-2♂ x Asia II-1♀ resulted in no female progeny being produced, demonstrating incompatibility between these species. Although, few female hybrids, ≤2%, were produced in a one direction cross between Asia II-1♂ and Asia II-2♀, no hybrids were detected from field-collected samples (Xu *et al*., 2010).

Other studies between different members of the *B. tabaci* species complex have been conducted to determine reproductive compatibility. Qin *et al*. (2016), reported reproductive incompatibility and partial reproductive compatibility between different *B. tabaci* species which

included: Asian genetic groups, China 1, MEAM1 and MED. Complete reproductive isolation in reciprocal crossing experiments between the MED, Asia II-3, and Asia II-1 species was also demonstrated when no female progeny was produced (Wang *et al*., 2010). The reproductive incompatibility recorded in these studies above has partly been associated with the observed differences in the courtship and mating behaviours of these species (De Barro and Hart, 2000; Liu *et al*., 2007; Perring and Symmes, 2006; Qin *et al*., 2016). In addition, reproductive isolation may result from a pre-mating barrier, a post-mating pre-zygotic barrier, or a post-zygotic barrier (Merrell, 1981) for populations that are geographically isolated.

Apart from other genetic groups of *B. tabaci*, mating studies among sub-Saharan African (SSA) whitefly populations colonizing cassava has also been conducted. No evidence of mating incompatibility was reported between reciprocal crosses of six CMD pandemic and four non-CMD pandemic cassava *B. tabaci* populations from East Africa (Maruthi *et al*., 2001). Furthermore, gene flow between whitefly populations colonising cassava and okra populations has been reported by Omondi *et al*. (2005). In another study focusing on reciprocal crossings between allopatric and sympatric cassava *B. tabaci* populations from Africa and India, mating incompatibility was observed between allopatric cassava *B. tabaci* populations from Africa and India (Maruthi *et al*., 2004). Similar results were obtained between sub-Saharan Africa population 'ASL' and MED whiteflies from the Mediterranean Basin (Vyskočilová, *et al*., 2018). A summary of the crossing experiments available in the literature is presented in Table 2.

Table 2. Summary of the mating data among 17 putative *Bemisia tabaci* species from the published crossing studies

Species	Males																	References
(Females)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
1. Asia I	C	N							N	N	N	N	N					Bedford *et al*., 1994; Qin *et al*., 2016
2. Asia II 1	N	C	N			N	N				N		N					Bedford *et al*., 1994; Luan et al., 2008; Qin *et al*., 2016; Xu *et al*., 2010
3. Asia II 2		P	C								N							Liu et al., 2007; Xu *et al*., 2010; Zang and Liu, 2007
4. Asia II 3				C				P										Qin *et al*., 2016
5. Asia II 5					C									N	N			Maruthi *et al*., 2004
6. Asia II 6	N		N			C				N		N						Qin *et al*., 2016
7. Asia II 7	N						C											Qin *et al*., 2016
8. Asia II 9	N			P				C				N						Qin *et al*., 2016
9. Australia									C					N				De Barro and Hart, 2000; Liu *et a*l., 2007
10. China 1	N					N				C								Qin *et al*., 2016
11. New World	N	N									C		N					Bedford *et al*., 1994; Costa *et al*., 1993a; Perring *et al*., 1993
12. Italy												C	N					Demichelis *et al*., 2005
13. MED	N					N			N				C	P	P			Liu *et al*., 2007, Xu *et al*., 2010, Zang and Liu, 2007
14. MEAM 1	N	N				N			N		N		P	C				Bedford *et al*., 1994; Byrne *et al*., 1995; Costa *et al*., 1993a; De Barro and Hart, 2000; Liu *et al*., 2007; Luan *et al*., 2008; Maruthi *et al*., 2004; Ronda *et a*l., 1999; Perring *et al*., 1993; Xu *et al*., 2010; Zang and Liu, 2007
15. SSA1					N								P		C	P		Maruthi *et al*., 2001; Omondi *et al*., 2005
16. SSA2					N										P	C		Maruthi *et al*, 2001; Maruthi *et al*., 2004
17. Uganda									N								C	Maruthi *et al*., 2004

In the table, 'C' Complete compatibility; 'N' No reproductive compatibility; 'P' Low number of F_1 females.

Recently, however, Mugerwa (2018) conducted crossing studies within SSA1 sub-groups, and among SSA1 sub-groups and SSA2 population. SSA2 has not only been reported in sub-Saharan Africa, but it has also been found in the Mediterranean basin including Spain colonising the ornamental plant morning glory (*Ipomoea indica*) (Banks *et al*., 1999; De la Rúa *et al*., 2006; Hadjistylli *et al*., 2015; Laarif *et al*., 2015).

SSA2 populations from Spain are related, based on mtCOI partial sequence, to whitefly populations colonizing cassava in sub-Saharan Africa and phylogenetically, are grouped in a single clade with SSA2 associated with severe epidemics infecting cassava collected in Uganda (Legg *et al*., 2002). Apart from the genetic data available for SSA2 populations from Spain, it is not known whether or not they interbreed with SSA2 or other genetic groups of *B. tabaci* colonizing cassava in sub-Saharan Africa.

Management and control of begomoviruses and their vector, *B. tabaci*

In order to combat and minimize the effects of begomoviruses and *B. tabaci* to the crop productivity, sustainable and durable control strategies are needed. Begomoviral diseases can be controlled in three phases: i) before the growing season in which resistant cultivar or virus- and vector-free planting materials, time of planting and field placement are selected, ii) during the planting season whereby vector populations are managed by insecticide, rouging virus-infected plants and biological control using predators, parasitoids and fungi, and iii) after the growing season whereby the viral inoculum sources should be immediately removed and destroyed (Rojas *et al*., 2018).

In addition, regular surveillance for disease symptoms at early stages of planting is of important to confine the spread of the disease to nearby fields (Snehi *et al*., 2015). Moreover, removal of commonly growing weeds surrounding farmers' field is very crucial, because they have been identified as key inoculum sources of viral infections (Barbosa *et al*., 2009; Bedford *et al*., 1998; García-Andrés *et al*., 2006). Alternatively, begomoviruses diseases could be controlled based on integrated pest and disease management (IPDM) program using a combination of a variety of management methods that are cost effective and cause the least damage to the environment (Moriones *et al*., 2011).

At present, the main approach employed to manage *B. tabaci* populations is the use of insecticides. This is problematic, however, due to, the widespread resistance that *B. tabaci* has developed to most of the insecticides in use (Castle *et al*., 2010; Horowitz *et al*., 2007; Palumbo *et al*., 2001). Therefore, other methods such as cultural, mechanical and biological control could be a better approach to manage *B. tabaci* that damages crop and transmits viral diseases (Horowitz *et al*., 2011).

The imposition of a crop-free period of several weeks, use of plant barriers and UV-absorbing polyethylene films have been reported to decrease whitefly populations (Hilje *et al*., 2001; Antignus *et al*., 1996). The use of yellow sticky traps has been a usefully approach in monitoring whiteflies populations (Gerling and Horowitz 1984). Furthermore, the use of natural enemies such as parasitoids and predators for biological control is considered a very effective way for controlling whitefly infestations (De Barro and Coombs, 2009; Gerling *et al*., 2001; Gould *et al*., 2008; Otim *et al*., 2008).

Additionally, the breeding of resistance against begomoviruses and the whitefly vector have been suggested to be the most promising and efficient approach for managing the transmission of begomoviruses and whitefly infestations (Horowitz *et al*., 2011). Hence, the use of virus- and vector-resistant cultivars could reduce virus disease spread and loss of crops due to whitefly damages respectively (García-Andrés *et al*., 2009; Rojas *et al*., 2018).

OBJECTIVES

OBJECTIVES

Begomoviruses and their whitefly-vector species, within the *Bemisia tabaci* complex, have become major problems globally, due to the huge economic losses that they inflict on important agricultural crops. Begomoviruses also infect non-cultivated plants that may act as reservoirs of viral diversity, thus posing a serious threat to cultivated plants. In addition, the current species delimitation methods for species within the *B. tabaci* complex, and its limitations, highlights the need for further research into biological species identification and associated traits, such as differential virus-transmission abilities. Basic information about begomoviruses infecting non-cultivated plants and the identification of species boundaries in the *B. tabaci* complex are also of vital practical importance in the selection and implementation of the best control measures, in order to reduce the economic damage caused by begomovirus-whitefly complexes.

The over-arching objective of this study, therefore, was to generate information about begomoviruses infecting wild plants in East Africa and investigate the biological species boundaries between the sub-Saharan Africa (SSA) mtCO1 phylogenetic groups of *B. tabaci*. This was achieved by addressing the following specific objectives, structured into the following chapters of the thesis.

i. A novel East African monopartite begomovirus-betasatellite complex that infects *Vernonia amygdalina*.
ii. Desmodium mottle virus, the first legumovirus (genus *Begomovirus*) from East Africa.
iii. African basil (*Ocimum gratissimum*) is a reservoir of divergent begomoviruses in Uganda.
iv. Mating interactions between *Bemisia tabaci* putative cryptic species from the Mediterranean basin and sub-Saharan Africa.

CHAPTER 1

A NOVEL EAST AFRICAN MONOPARTITE BEGOMOVIRUS-BETASATELLITE COMPLEX THAT INFECTS *VERNONIA AMYGDALINA*

Chapter published as
Mollel HG, Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017) A novel East African monopartite begomovirus-betasatellite complex that infects *Vernonia amygdalina*. *Archives of Virology* 162: 1079-1082

CHAPTER 1: A NOVEL EAST AFRICAN MONOPARTITE BEGOMOVIRUS-BETASATELLITE COMPLEX THAT INFECTS *VERNONIA AMYGDALINA*

1.1 Introduction

The genus *Begomovirus* is the largest of the seven genera in the plant virus family *Geminiviridae* (Brown *et al*., 2012; Varsani *et al*., 2014). Begomoviruses are transmitted by the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) to a large variety of cultivated and wild plant species (Navas-Castillo *et al*., 2011). Begomoviruses have a circular, single stranded DNA genome, monopartite or bipartite, encapsidated in twinned icosahedral particles. Bipartite begomoviruses have two genome components, referred to as DNA-A and DNA-B, of similar size (2.5-2.8 kb), while monopartite begomoviruses have only one component, which is similar to DNA-A of bipartite begomoviruses.

The DNA-A virion-sense strand encodes coat (CP) and pre-coat (pre-CP) proteins, the latter of which is present only in Old World (OW) begomoviruses. The DNA-A complementary-sense strand encodes the replication-associated protein (Rep), a transcriptional activator protein (TrAP), a replication enhancer protein (REn) and C4 protein. DNA-B encodes a nuclear shuttle protein (NSP) on the virion-sense strand and a movement protein (MP) on the complementary-sense strand. There are more than 300 accepted begomovirus species according to the recently updated demarcation criteria for the genus, which consider a DNA-A pairwise identity of 91% as the species threshold (Brown *et al*., 2015).

Recombination is a phenomenon that is crucial for speciation and evolution in the family *Geminiviridae* and contributes to the richness in species of the genus *Begomovirus*. This stresses the importance of recombination studies when analysing new begomoviruses.

Several types of DNA satellites have been described to be associated with begomoviruses: betasatellites (Briddon *et al*., 2003), alphasatellites (Briddon *et al*., 2004) and deltasatellites (Lozano *et al*., 2016). Betasatellites are circular, single stranded DNA molecules about half the size of the begomovirus genome components that have been described to be associated with OW monopartite begomoviruses and are essential for induction of typical disease symptoms (Briddon *et al*., 2003). Betasatellite genomes contain an open reading frame in the complementary-sense strand encoding the βC1 protein, an A-rich region, a conserved stem-loop and a satellite conserved region.

Vernonia amygdalina Delile (family Compositae), known as bitter leaf, is a wild shrub that grows in tropical Africa and used in traditional medicine to treat malaria (Masaba, 2000). In this chapter, the molecular characterization of a new monopartite begomovirus and isolated betasatellite from *V. amygdalina* plants from Uganda are described.

1.2 Materials and methods

Samples from *V. amygdalina* plants showing crinkled leaves were collected in March 2015 from two locations in Uganda (sample UG7 from Naama [00°24.691' N, 31°59.927' E] and UG9 from Kawungera [00°27.761' N, 31°39.171' E]) (**Figure 8**). Morphological identification of the plant samples was confirmed molecularly by DNA barcoding using chloroplast rbcL and matK genes (Hollingsworth *et al*., 2009). Total DNA was extracted from leaf tissue using a modified CTAB method (Permingeat *et al*., 1998) and used as a template for rolling-circle amplification (RCA) using φ29 DNA polymerase (TempliPhi kit, GE Healthcare), and polymerase chain reaction (PCR) for the detection of DNA-B component using degenerate primers (Rojas *et al*., 1993). Amplified RCA products were digested with a set of restriction enzymes (*Bam*HI, *Eco*RI, *Hind*III, *Nco*I, *Nhe*I and *Sal*I). Fragments of RCA products digested with *Eco*RI (~2.8 kb) and *Nco*I (~1.3 kb) were cloned into pBlueScript II SK (+) (Stratagene) and pGEM-T Easy Vector (Promega), respectively. Recombinant plasmid DNAs were introduced into *Escherichia coli* DH5α by electroporation, and selected clones were sequenced at Macrogen Inc. (Seoul, South Korea).

Initial sequence similarity comparison was performed using the BLAST program (http://www.ncbi.nih.gov/). Sequence alignments were performed using MUSCLE (Edga *et al*., 2004), pairwise identity scores were calculated using SDT (Sequence demarcation tool) (Muhire *et al*., 2014), and MEGA 7 was used for phylogenetic analysis (Kumar *et al*., 2015). To detect putative recombinant fragments, a search for potential parental begomoviruses and betasatellites in the GenBank database was conducted using SWeBLAST (Fourment *et al*., 2008) with a window size of 200 and a step size of 200. The sequences with the highest SWeBLAST scores were selected for alignment using MUSCLE (Edga *et al*., 2004) and subsequent recombination analysis using the RDP4 package with default settings (Martin *et al*., 2015).

1.3 Results and discussion

Both samples, UG7 and UG9 generated similar restriction patterns, which suggested the presence of a putative monopartite begomovirus (~2.8 kb) and a putative DNA satellite (~1.3 kb) in each sample. None of the samples tested for DNA-B yielded amplification products, confirming the presence of a monopartite begomovirus. Cloned ~2.8 kb DNA from samples UG7 (2791 nt, KX831132) and UG9 (2791 nt, KX831133) had a genome organization typical of OW monopartite begomoviruses, with CP and pre-CP proteins encoded in the virion-sense strand, and Rep, TrAP, Ren and C4 proteins encoded in the complementary-sense strand. Begomoviruses from samples UG7 and UG9 showed the highest nucleotide sequence identity (73.1% and 73.2%, respectively) to an isolate of the monopartite

begomovirus tomato leaf curl Vietnam virus (EU189150). DNA of ~1.3 kb cloned from the same samples (UG7, 1365 nt, KX831134; UG9, 1364 nt, KX831135) showed the typical genome organization of betasatellites (A-rich region, stem-loop, satellite conserved region and βC1 gene). Betasatellites from UG7 and UG9 showed the highest nucleotide sequence identity (67.1% and 68.2%, respectively) to vernonia yellow vein Fujian betasatellite (JF733779) found in *Cyanthillium cinereum* (L.) H.Rob. (syn. *Vernonia cinerea*) in China.

The DNA-A-like genomes and betasatellites isolated from samples UG7 and UG9 were 99.4% and 97% identical, respectively. According to the begomovirus species demarcation threshold of 91% (Brown *et al*., 2015), the monopartite begomovirus reported here should be considered to belong to a new species. The name vernonia crinkle virus (VeCrV) is proposed for the novel begomovirus and that the two isolates be designated [Uganda-Naama UG7-2015] and [Uganda-Kawungera UG9-2015]. According to the recently proposed betasatellite species demarcation threshold of 91% (https://talk.ictvonline.org/files/proposals/taxonomy_proposals _plant1/ m/ plant02 /6357), the betasatellite found in the same samples would represent a novel betasatellite, for which the name vernonia crinkle betasatellite (VeCrB) is proposed and that the two isolates be designated [Uganda-Naama UG7-2015] and [Uganda-Kawungera UG9-2015].

Recombination is commonly detected in begomovirus and betasatellite genomes (Fiallo-Olivé *et al*., 2010a; Lefeuvre *et al*., 2007; Mubin *et al*., 2009). Recombination analysis showed the presence of recombinant fragments in both VeCrV and the associated VeCrB (Table 3). Interestingly, recombination events detected in VeCrV involve genomes from Asia and Africa.

Phylogenetic analysis showed that VeCrV isolates clustered with two OW begomoviruses from Africa, tobacco leaf curl Zimbabwe virus (AF350330) and tobacco leaf curl Comoros virus (AM701760) (**Figure 9A**). In contrast, both VeCrB isolates grouped with betasatellites from Asia (vernonia yellow vein Fujian betasatellite [JF733779] and vernonia yellow vein betasatellite [FN435836]) (**Figure 9B**). As additional begomovirus and betasatellite sequences from East Africa are discovered, this will enable more precise phylogenetic relationships of this new begomovirus and associated betasatellite to be determined.

This is the first report of a begomovirus-betasatellite complex infecting plants of the genus *Vernonia* in Africa and the first identification of a betasatellite in Uganda.

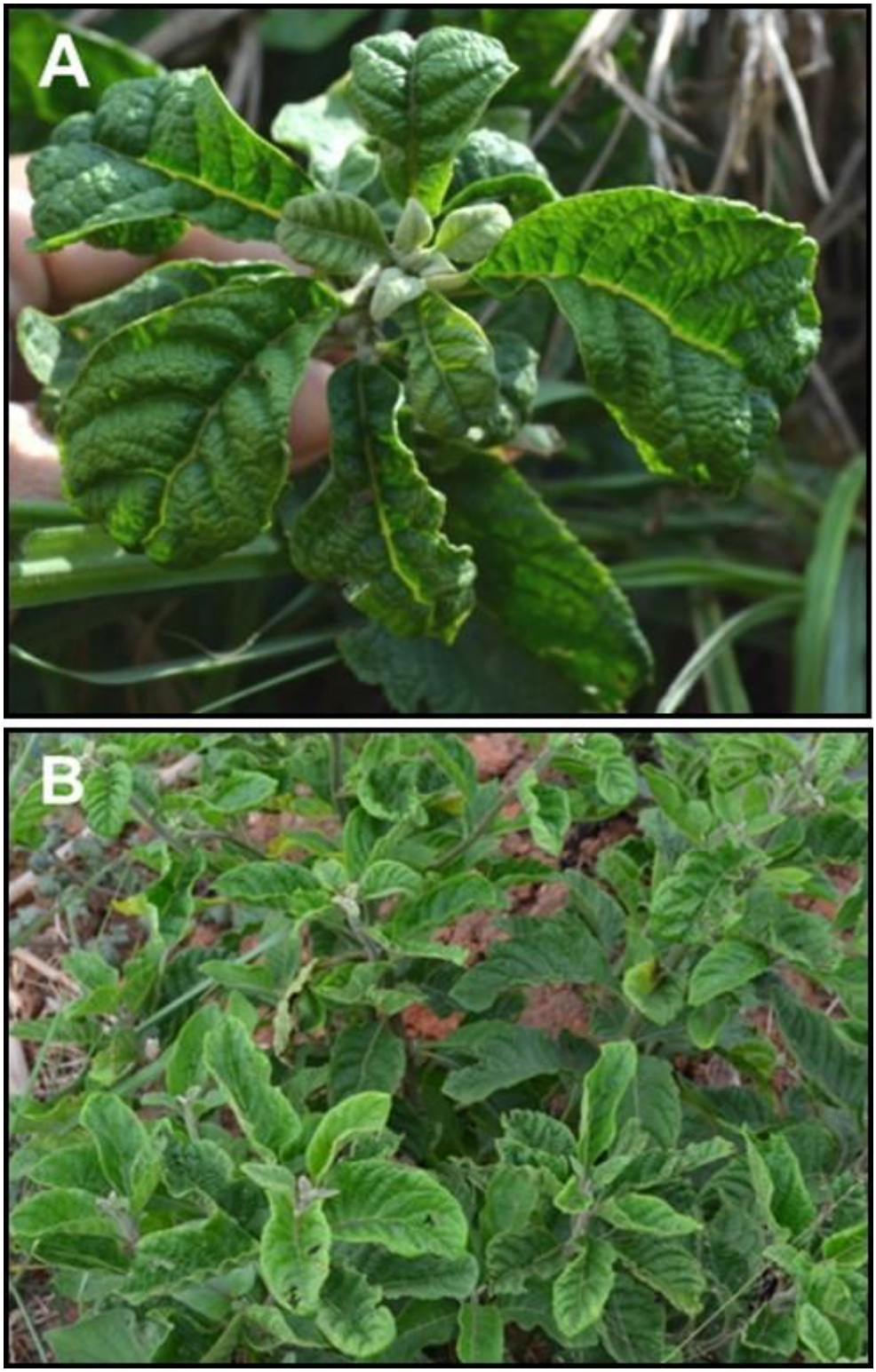

Figure 8. *Vernonia amygdalina* plants analyzed in this work showing crinkle symptoms on leaves (**A**, sample UG7; **B**, sample UG9).

Table 3. Recombination events within the DNA sequences of VeCrV and VeCrB detected by at least five methods included in the RDP package.

Recombinant	Recombination breakpoints	Parent-like sequences		*p* value	Methods that detected recombination*
		Major	Minor		
VeCrV	305-471	ChiLCINV (KT948070)	TbLCTHV (KT322140)	$7.032x10^{-25}$	G, B, M, C, S
VeCrV	564-857	TYLCMLV (LM651403)	TbLCZV (AM701756)	$1.300x10^{-13}$	R, G, B, M, C, S, 3S
VeCrV	2128-2361	EACMV (KJ888092)	TbLCTHV (KT322140)	$2.577x10^{-18}$	R, G, B, M, C, S, 3S
VeCrV	2464-2563	CoYMV (KT454834)	ToLCLKV (AF274349)	$4.226x10^{-07}$	R, G, B, C, S
VeCrB	1022-1060	SgYVGdB (AM238695)	VeYVB (FN435836)	$4.659x10^{-05}$	R, G, B, M, C

*The method with the lower *p* value obtained for each region is underlined.

R, RDP; G, GENCONV; B, BootScan; M, MaxChi; C, Chimaera; S, SiScan; 3S, 3Seq.

ChiLCINV, chilli leaf curl India virus; TbLCTHV, tobacco leaf curl Thailand virus; TYLCMLV, tomato yellow leaf curl Mali virus; TbLCZV, tobacco leaf curl Zimbabwe virus; EACMV, East African cassava mosaic virus; CoYMV, cotton yellow mosaic virus; ToLCLKV, tomato leaf curl Sri Lanka virus; SgYVGdB, Siegesbeckia yellow vein Guangxi betasatellite; VeYVB, Vernonia yellow vein betasatellite.

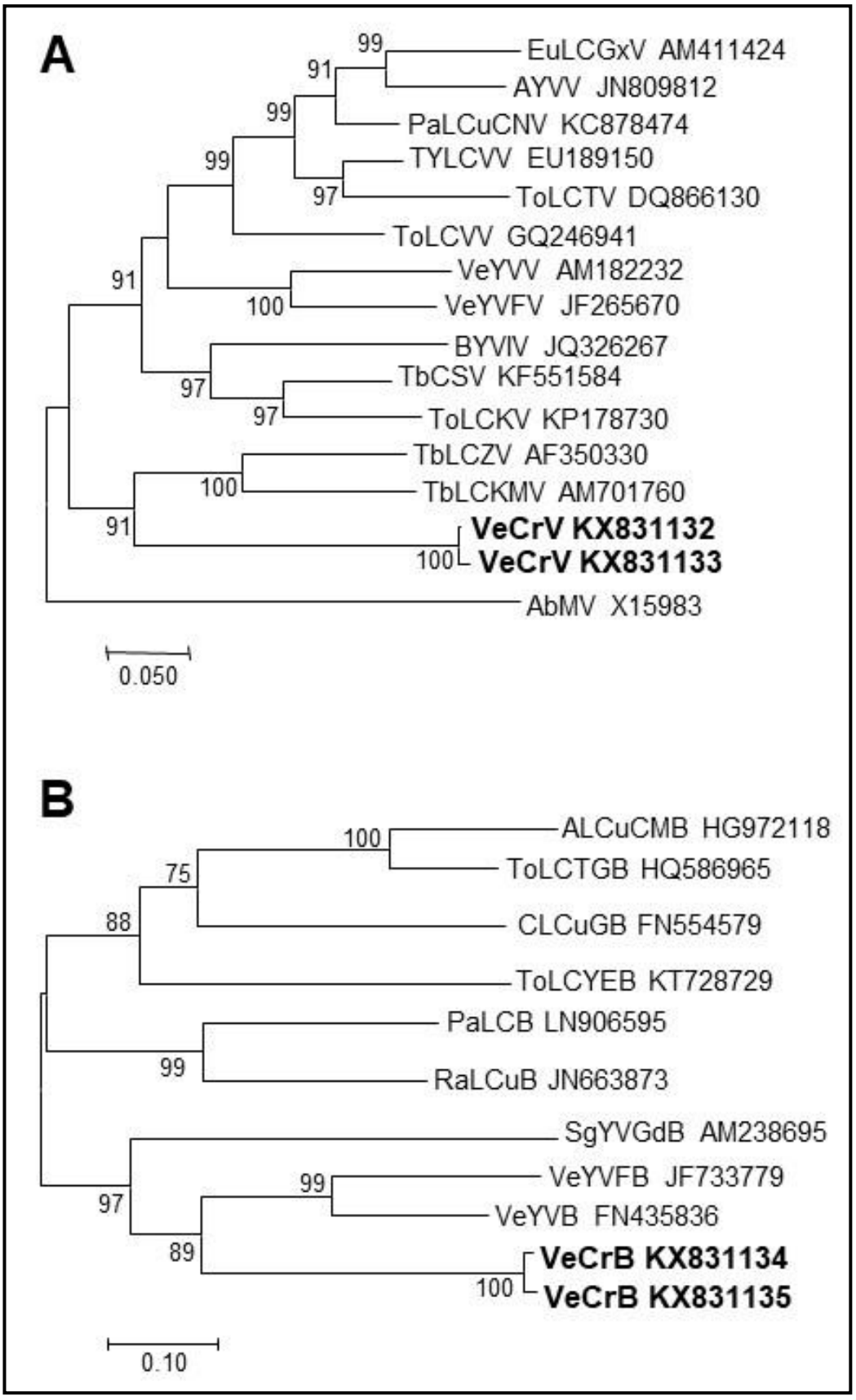

Figure 9. Phylogenetic tree showing the relationships between vernonia crinkle virus and other monopartite Old World begomoviruses (**A**) and vernonia crinkle betasatellite and other betasatellites (**B**). The trees were constructed by the maximum-likelihood method using the MEGA 7 program. Only bootstrap values [75% (1000 replicates) are shown. Viruses used to generate the phylogenetic tree (**A**) are as follows: AYVV, ageratum yellow vein virus; BYVIV, bhendi yellow vein India virus; EuLCGxV, euphorbia leaf curl Guangxi virus; PaLCuCNV, papaya leaf curl China virus; TbCSV, tobacco curly shoot virus; TbLCKMV, tobacco leaf curl Comoros virus; ToLCKV, tomato leaf curl Karnataka virus; TbLCZV, tobacco leaf curl Zimbabwe virus; ToLCTV, tomato leaf curl Taiwan virus; ToLCVV, tomato leaf curl Vietnam virus; TYLCVV, tomato yellow leaf curl Vietnam virus; VeYVFV, vernonia yellow vein Fujian virus; VeYVV, vernonia yellow vein virus. The DNA-A sequence of abutilon mosaic virus (AbMV), a New World begomovirus, was used as an outgroup. Betasatellite sequences used for the phylogenetic tree are as follows: ALCuCMB, ageratum leaf curl Cameroon betasatellite; CLCuGB, cotton leaf curl Gezira betasatellite; PaLCuB, papaya leaf curl betasatellite; RaLCuB, radish leaf curl betasatellite; SgYVGdB, siegesbeckia yellow vein Guangxi betasatellite; ToLCTGB, tomato leaf curl Togo betasatellite; ToLCYEB, tomato leaf curl Yemen betasatellite; VeYVFB, vernonia yellow vein Fujian betasatellite; VeYVB, vernonia yellow vein betasatellite. The bar below each tree indicates nucleotide substitutions per site.

CHAPTER 2

DESMODIUM MOTTLE VIRUS, THE FIRST LEGUMOVIRUS (GENUS *BEGOMOVIRUS*) FROM EAST AFRICA

Chapter published as
Mollel HG, Sseruwagi P, Ndunguru J, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017) Desmodium mottle virus, the first legumovirus (genus *Begomovirus*) from East Africa. *Archives of Virology* 162: 1799-1803

CHAPTER 2: DESMODIUM MOTTLE VIRUS, THE FIRST LEGUMOVIRUS (GENUS *BEGOMOVIRUS*) FROM EAST AFRICA

2.1 Introduction

The family *Geminiviridae* comprises seven genera, differentiated based on genome organization, nucleotide sequence identity and biological properties: *Begomovirus, Mastrevirus, Eragrovirus, Curtovirus, Turncurtovirus, Topocuvirus* and *Becurtovirus* (Brown *et al*., 2012; Varsani *et al*., 2014). The genus *Begomovirus* is the largest in the family, with 322 accepted species (Brown *et al*., 2015; https://talk.ictvonline.org/ictv_wikis/geminiviridae/m/files_gemini/5120). Begomoviruses are transmitted by the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) and frequently cause important plant diseases around the world (Navas-Castillo *et al*., 2011). Bipartite begomoviruses possess two genome components (DNA-A and DNA-B), which are essential for virus infectivity and the size of each component ranges between 2.5 and 2.8 kb. DNA-A and DNA-B share ~200 nt in the common region (CR), located within the intergenic region, that contains cis elements for replication and control of gene expression. The CR exhibits a high degree of sequence identity between both genome components of bipartite begomoviruses (Brown *et al*., 2012). Based on the phylogenetic analysis of complete nucleotide sequences of DNA-A, begomoviruses can be classified into four lineages, Old World (OW), New World (NW), sweepoviruses and legumoviruses (Briddon *et al*., 2010).

Legumoviruses, or legume-infecting begomoviruses from the OW, are amongst the most atypical begomoviruses (Ilyas *et al*., 2010). They are distinct from the numerous legume-infecting begomoviruses that occur in the Americas and in phylogenetic analyses they group in a cluster different from those of OW and NW begomoviruses (Briddon *et al*., 2010; Fauquet *et al*., 2008). The difference between legumoviruses and typical OW begomoviruses could have arisen due to genetic isolation involving either a host-range barrier or lack of movement of whitefly vectors between legumes and non-leguminous plants, thereby preventing genetic exchange between both groups of viruses (Qazi *et al*., 2007). The genomes of most legumoviruses are bipartite, although a DNA-B component has not been identified for cowpea golden mosaic virus (CPGMV), Dolichos yellow mosaic virus (DoYMV) and soybean mild mottle virus (SbMMoV) (Alabi *et al*., 2010; Maruthi *et al*., 2006). Little attention has been paid to legumoviruses infecting wild plants. Scarce examples include DoYMV infecting *Lablab purpureus* (sin. *Dolichos lablab*) (Maruthi *et al*., 2006) and horsegram yellow mosaic virus (HgYMV) infecting *Macrotyloma uniflorum* (Barnabas *et al*., 2010) in India, kudzu mosaic virus (KuMV) infecting *Pueraria montana* in Vietnam (Ha *et al*., 2008), soybean chlorotic blotch virus (SbCBV) infecting *Centrosema pubescens* in Nigeria (Alabi *et al*., 2010) and Rhynchosia yellow mosaic virus (RhYMV) infecting *Rhynchosia minima* in Pakistan (Ilyas *et al*., 2009).

CHAPTER 2

The genus *Desmodium* is composed of about 370 accepted species native to tropical East Asia, Africa and America. Some *Desmodium* species are considered as weeds, although others containing potent secondary metabolites that are used in agriculture in push-pull technology (Cook *et al*., 2007). This chapter describes the first legumovirus from *Desmodium* species from East Africa.

2.2 Materials and methods

In this study, leaf samples of two *Desmodium* sp. (family Fabaceae) plants showing mottle symptoms (**Figure 10**) were collected in Kikonge, southwestern Uganda, in March 2015 (00°22.641' N; 32°11.252' E [sample UG4], 00°22.640' N; 32°11.252' E [sample UG5]). Morphological identification of the plant samples at the genus level was confirmed molecularly by DNA barcoding using chloroplast rbcL and matK genes (Hollingsworth *et al*., 2009).

Total nucleic acids were extracted from leaf samples using a modified CTAB method (Permingeat *et al.*, 1998). To test for the presence of begomoviruses in the samples, putatively causing the observed symptoms, nucleic acids were used as a template for rolling circle amplification (RCA) using φ29 DNA polymerase (TempliPhi kit, GE Healthcare) and amplified RCA products were digested with a set of restriction enzymes (*Cla*I, *Bam*HI, *Eco*RI, *Hind*III, *Nco*I, *Nhe*I and *Sal*I) (Inoue-Nagata *et al*., 2004).

Putative full length begomoviral genomic components (~2.8 kb) were cloned from each sample (*Eco*RI and *Hind*III for UG4 and *Cla*I and *Eco*RI for UG5) into pBlueScript II SK (+) (Stratagene) and selected clones (one per sample and restriction enzyme) were sequenced at Macrogen Inc. (Seoul, South Korea). Sequences were assembled with SeqMan software included in the DNASTAR package (DNASTAR Inc.). Open reading frames (ORFs) were identified using Open Reading Frame Finder (NCBI) and confirmed using the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) on their deduced amino acid sequences.

Initial sequence identity comparison was performed using the BLAST program, sequences were aligned with MUSCLE (Edga *et al*., 2004) and pairwise identity scores were calculated using SDT (Sequence demarcation tool) (Muhire *et al*., 2014). A phylogenetic analysis using maximum likelihood (ML) was used after selecting the best-fit model of nucleotide substitution based on corrected Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC) as implemented in MEGA 6 (Tamura *et al*., 2013). Recombination analysis was performed using RDP4 (Martin *et al*., 2015) after alignment, with MUSCLE, of the sequences selected with SWeBLAST (with a window size of 200 and a step size of 200) (Fourment *et a*l., 2008). SWeBLAST avoids the significant problem of deciding which sequences to compare, thus allowing identification of putative

parents of recombinant sequences (four sequences for DNA-A and nine sequences for DNA-B). Only recombination events detected with at least five methods with p-values lower than 10^{-2} were considered.

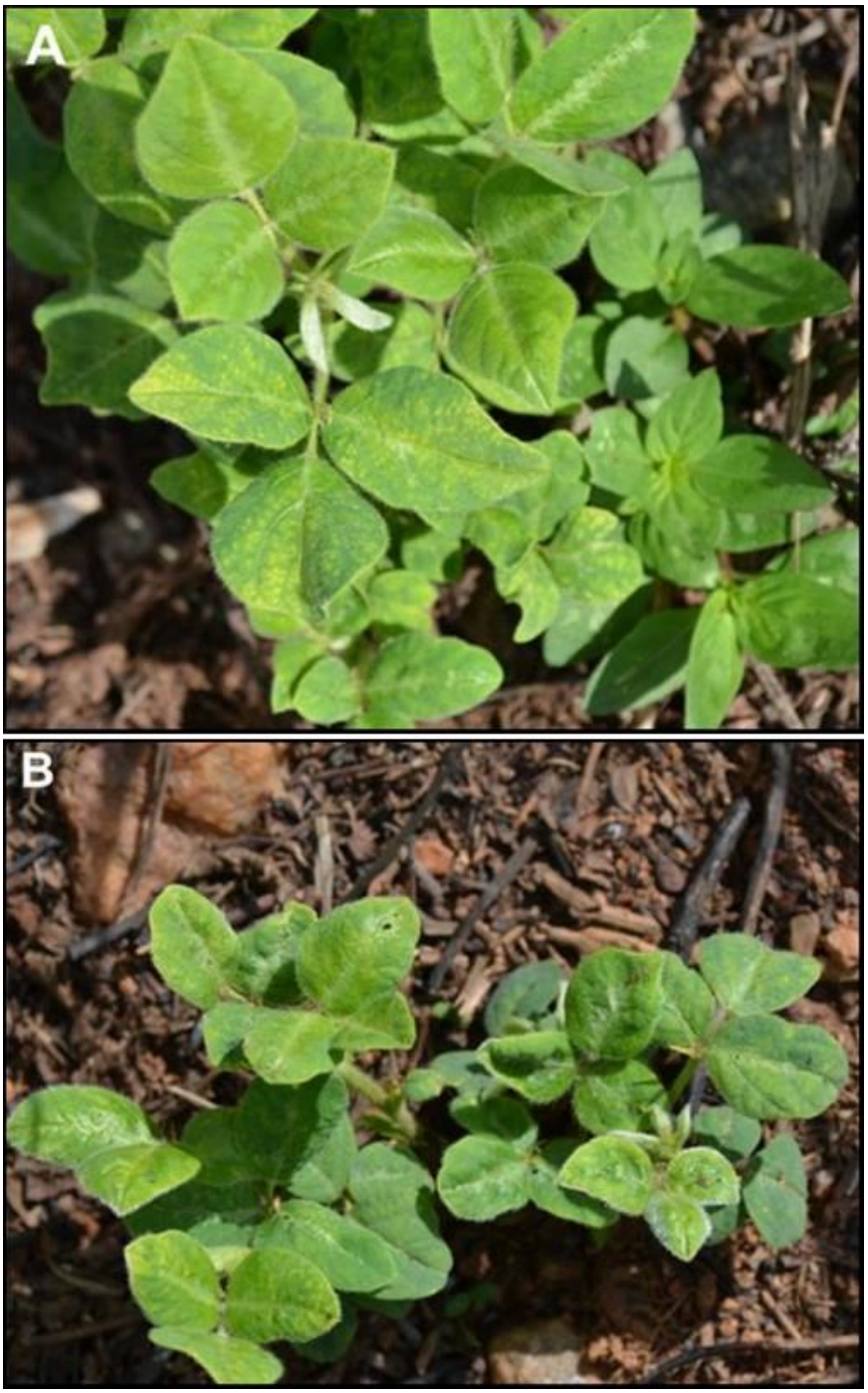

Figure 10. *Desmodium* sp. plants analyzed in this work showing mottle symptoms on leaves. (**A**) sample UG4, (**B**) sample UG5.

2.3 Results and discussion

The cloned genome components from each sample were shown to correspond to begomoviral DNA-A and DNA-B components. DNA-A component from sample UG4 (2767 nt, *Eco*RI fragment, KY294724) and UG5 (2767 nt, *Cla*I fragment, KY294725) showed the highest nucleotide sequence identity (79.9% and 80%, respectively) to SbMMoV (GQ472984), a legumovirus found in soybean (*Glycine max*) in Nigeria (Alabi *et al*., 2010). The DNA-B component from sample UG4 (2715 nt, *Hind*III fragment, KY294726) exhibited the highest nucleotide sequence identity (65.4%) with an isolate of African cassava mosaic virus (ACMV) (KJ887741) from Madagascar (De Bruyn *et al*., 2016), while UG5 DNA-B (2713 nt, *EcoR*I fragment, KY294727) exhibited the highest nucleotide sequence identity (66.4%) with another isolate of ACMV (HE616778) from Burkina Faso (Tiendrébéogo *et al*., 2012).

Pairwise nucleotide identities between DNA-A and DNA-B from samples UG4 and UG5 were 100% and 99.7%, respectively, confirming that the virus identified from both samples belonged to the same begomovirus species. Also, the virus showed a typical genome organization of Old World bipartite begomoviruses. In accordance, therefore, with current taxonomic guidelines for the genus *Begomovirus* (a new DNA-A sequence with less than 91% pairwise identity to any other published begomovirus DNA-A sequence will belong to a new begomovirus species) (Brown *et al*., 2015), the isolates described here ([Uganda-Kikonge UG4-2015] and [Uganda-Kikonge UG5-2015]) represent a novel species for which the name Desmodium mottle virus (DesMoV) is proposed.

DNA-A and DNA-B from both samples showed a CR of 179 nt (DNA-A) and 156 nt (DNA-B) with sequence identities of 90.2% (sample UG4) and 90.8% (sample UG5). The difference in length of the CR is due to a deletion in DNA-B. Both components from each sample showed three copies of iterons (AATCGGGGGT) (one is inverted and the most proximal to TATA box is imperfect), indicating that DNA-A and DNA-B isolated from each sample constitute a cognate pair.

A phylogenetic tree based on alignment of the DNA-A sequences obtained here with those of selected begomoviruses (including one sequence from each legumovirus species) showed that they grouped in a cluster with the legumoviruses (**Figure 11A**). However, DNA-Bs grouped with ACMV, a typical OW begomovirus (**Figure 11B**). Similar phylogenetic relationships have been described previously for SbCBV, the only bipartite legumovirus identified in Africa until now (Alabi *et al*., 2010). This is an example of the distinct evolutionary history undergone by the DNA-A and DNA-B genome components, as shown previously for other begomoviruses (Briddon *et al*., 2010; Bull *et al*., 2006; Fiallo-Olivé *et al*., 2010a; Pita *et al*., 2001; Qazi *et al*., 2007). No recombination event was detected in any of the genome components of DesMoV.

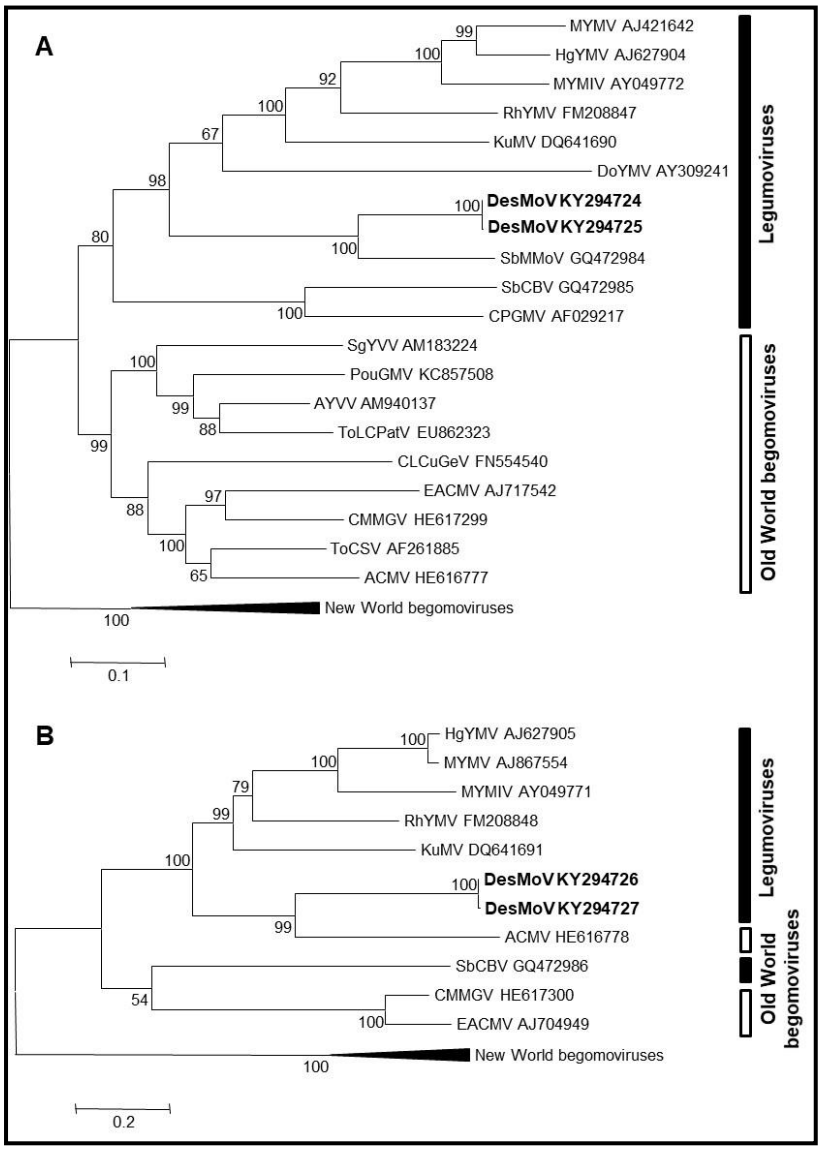

Figure 11. Phylogenetic trees illustrating the relationship of isolates of Desmodium mottle virus (DesMoV) DNA-A (**A**) and DNA-B (**B**) to other begomoviruses. The trees were constructed by the maximum-likelihood method (1000 replicates) with the MEGA 6 program using the best fit model, TN93+G+I for DNA-A and HKY+G+I for DNA-B. ACMV, African cassava mosaic virus; AYVV, Ageratum yellow vein virus; CLCuGeV, cotton leaf curl Gezira virus; CMMGV, cassava mosaic Madagascar virus; CPGMV, cowpea golden mosaic virus; DoYMV, Dolichos yellow mosaic virus; EACMV, East African cassava mosaic virus; HgYMV, horsegram yellow mosaic virus; KuMV, kudzu mosaic virus; MYMIV, mungbean yellow mosaic India virus; MYMV, mungbean yellow mosaic virus; PouGMV, Pouzolzia golden mosaic virus; RhYMV, Rhynchosia yellow mosaic virus; SbCBV, soybean chlorotic blotch virus; SbMMoV, soybean mild mottle virus; SgYVV, Siegesbeckia yellow vein virus; ToCSV, tomato curly stunt virus; ToLCPatV, tomato leaf curl Patna virus. A set of New World begomoviruses was used as the outgroup. The bar below each tree indicates nucleotide substitutions per site.

DesMoV is the first legumovirus to be described from East Africa to date and because it is phylogenetically closely related to begomoviruses that infect soybean or cowpea in West Africa, it may represent a potential threat to these crops. Additional research work to investigate its host-plant range and whitefly transmission characteristics, therefore, should be initiated to assess the threat to crops posed by this newly discovered begomovirus.

CHAPTER 3

AFRICAN BASIL (*OCIMUM GRATISSIMUM*) IS A RESERVOIR OF DIVERGENT BEGOMOVIRUSES IN UGANDA

Chapter published as
Mollel HG, Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2019) African basil (*Ocimum gratissimum*) is a reservoir of divergent begomoviruses in East Africa. *Plant Disease* (in press) DOI: 10.1094/PDIS-08-19-1675-RE

CHAPTER 3: AFRICAN BASIL (*OCIMUM GRATISSIMUM*) IS A RESERVOIR OF DIVERGENT BEGOMOVIRUSES IN UGANDA

3.1 Introduction

Begomoviruses (genus *Begomovirus*, family *Geminiviridae*) are plant viruses with circular single-stranded DNA genomes encapsidated in twin (geminate) quasi-icosahedral virions. A high number of begomoviruses cause severe damage to many vegetable and fiber crops worldwide. They are transmitted by whiteflies (Hemiptera: Aleyrodidae) of the *Bemisia tabaci* complex (Navas-Castillo *et al.*, 2011). Begomovirus genomes are monopartite or bipartite (Zerbini *et al.*, 2017). The genomes of bipartite begomoviruses consist of DNA-A and DNA-B components, each of 2.5-2.6 kb. DNA-A encodes the coat protein (CP), a putative movement protein (MP) (present only in Old World begomoviruses), replication associated protein (Rep), a transcriptional activator (TrAP), a replication enhancer (REn) and the C4 protein. DNA-B encodes a nuclear shuttling protein (NSP) and movement protein (MP). Both components share a common region of approximately 200 nt within an intergenic region (IR) that includes the replication origin. Rep initiates viral DNA replication by binding to iterative sequences (iterons) within the IR and introducing a nick into the conserved TAATATT↓AC sequence. The genomes of monopartite begomoviruses resemble the DNA-A component of bipartite begomoviruses. In some cases, begomoviruses are associated with DNA satellites, namely betasatellites, alphasatellites and deltasatellites (Lozano *et al.*, 2016; Zhou, 2013).

In East Africa, begomoviruses causing cassava mosaic disease have been extensively studied (reviewed by Rey and Vanderschuren 2017). However, little attention has been devoted to viruses infecting minor crops, weeds and other wild plants. In Uganda, for example, isolates of only two begomovirus species have been fully characterized infecting wild plants: Vernonia crinkle virus (associated with a betasatellite) from *Vernonia amygdalina* (Compositae) and Desmodium mottle virus (the first legumovirus from East Africa) from *Desmodium* sp. (Fabaceae) (Mollel *et al*., 2017a, b).

Ocimum spp., generically named as basils, are aromatic plants belonging to the family Lamiaceae (Labiatae). Members of the genus include annual herbs, suffrutices and shrubs native to the tropical and warm temperate regions, with the greatest number of species in Africa (Paton 1992; Paton *et al*., 1999). *O. gratissimum*, African basil, is widespread in the Paleotropics, from India to West Africa, and naturalized in tropical America. In Africa, where *O. gratissimum* forms a variable polymorphic complex, it is a common culinary herb also used in traditional medicine against a range of illnesses such as coughs, fever, ear infection and abdominal pain (Kokwaro, 2009). Also, it is of major potential economic importance due to the production of essential oil that has a large antimicrobial spectrum (Lawrence, 1992). The objective of this study was to search and characterize

the putative begomoviruses infecting wild and cultivated *O. gratissimum* plants in Uganda. The full-length genome of DNA-A and DNA-B was cloned and sequenced from six field-collected plants expressing typical begomovirus symptoms. The genome and phylogenetic analyses showed the presence of three novel begomoviruses, two of them highly divergent. To our knowledge there have been no previous reports of viruses infecting *O. gratissimum* nor infecting *Ocimum* spp. in Africa.

3.2 Materials and methods

3.2.1 Sample collection and plant identification

As part of a survey for begomoviruses in Uganda, leaf samples from six *O. gratissimum* plants were collected from different locations in the Central and Western regions of the country in March 2015 (**Figure 12**; Table 4). The survey was conducted along main roads, where wild-plant specimens were collected and in some backyard gardens, where cultivated plants were sampled. All plants exhibited symptoms suspicious of begomovirus infection (**Figure 13**). Morphological identification of the plant species was confirmed molecularly by DNA barcoding using chloroplast rbcL (ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) and matK (maturase K) genes (Hollingsworth *et al*., 2009).

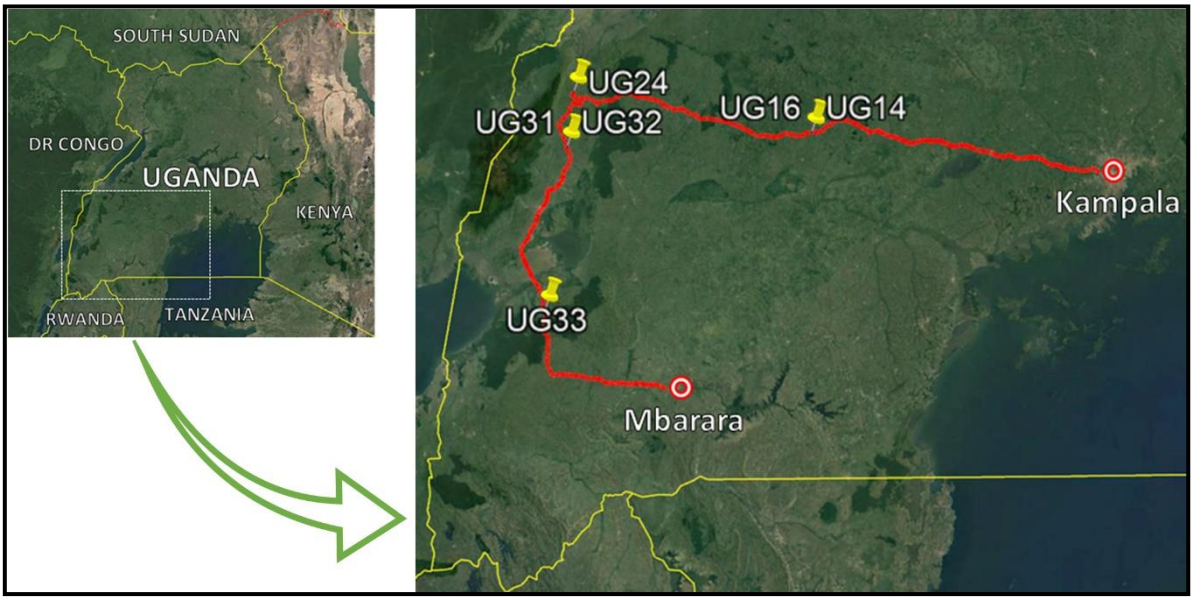

Figure 12. Map showing the survey route in southwestern Uganda and the locations of the six *Ocimum gratissimum* plants sampled and analyzed in this work.

Table 4. Samples of *Ocimum gratissimum* collected in southwestern Uganda.

Sample code	GPS coordinates	Field type	Symptoms
UG14	0°29.817'N 31°16.003'E	backyard garden (cultivated)	yellow vein and mosaic
UG16	0°29.828'N 31°15.993'E	backyard garden (cultivated)	yellow vein and mosaic
UG24	0°41.002'N 30°14.063'E	roadside (wild)	mosaic
UG31	0°26.706'N 30°12.203'E	roadside (wild)	golden mosaic
UG32	0°26.726'N 30°12.206'E	roadside (wild)	yellow vein and mosaic
UG33	0°15.669'S 30°06.138'E	roadside (wild)	golden mosaic

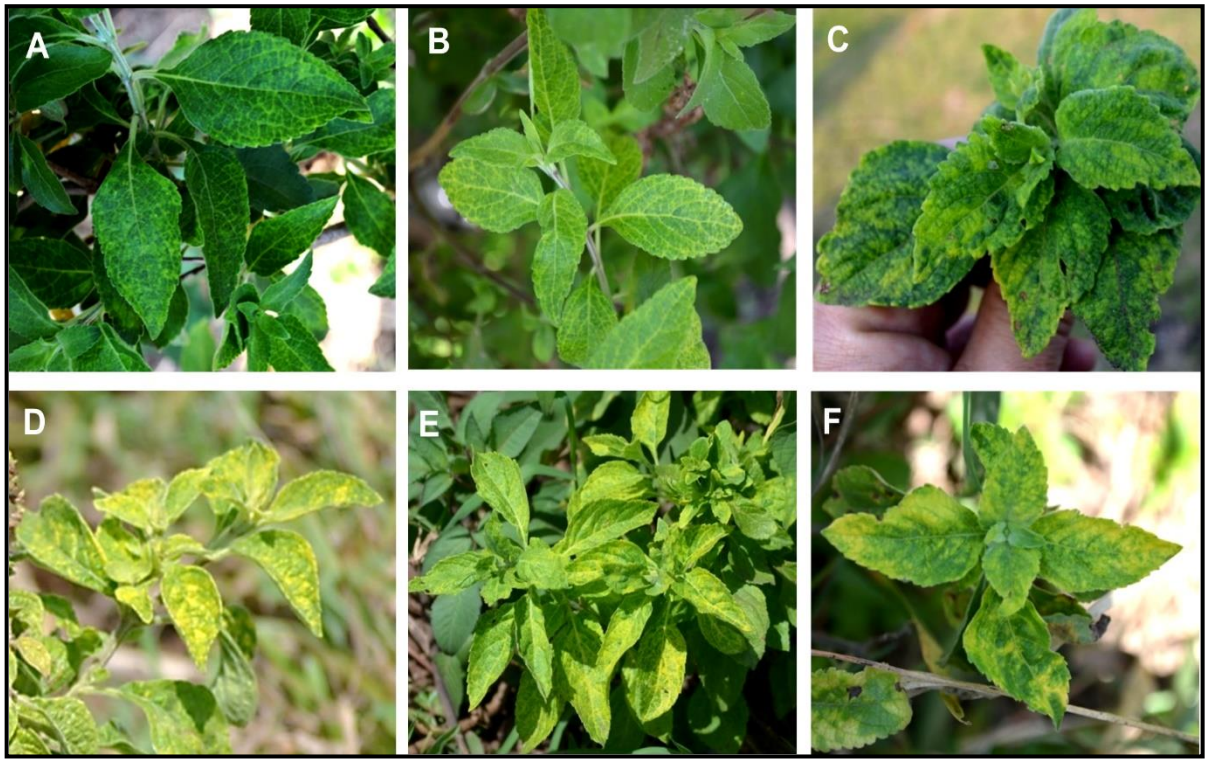

Figure 13. Photographs of the collected *Ocimum gratissimum* plants, showing typical begomovirus symptoms: UG14 (**A**), UG16 (**B**), UG24 (**C**), UG31 (**D**), UG32 (**E**) and UG33 (**F**).

3.2.2 Virus component cloning

Total DNA was extracted from leaf tissue using a modified CTAB (Cetyl Trimethyl Ammonium Bromide) method (Permingeat *et al.*, 1998) and used as a template for rolling-circle amplification (RCA) using φ29 DNA polymerase (illustra TempliPhi 100 Amplification Kit, GE Healthcare, Buckinghamshire, UK**)**. Amplified RCA products were digested with *Hpa*II, a 4-nt restriction enzyme commonly used for preliminary restriction fragment length polymorphism (RFLP) analysis of RCA products obtained from samples infected by ssDNA viruses, and restriction products analyzed on a 1% agarose gel. Then, RCA products were digested with a set of 6-nt restriction enzymes (*Bam*HI, *Eco*RI, *Hind*III, *Nco*I, *Nhe*I and *Sal*I). Selected fragments (~2.7 kb) of RCA products putatively corresponding to full length begomoviral components were cloned into pBlueScript II SK (+) (Stratagene, La Jolla, CA) or closed pGEM-T Easy Vector (Promega, Madison, WI). Recombinant plasmid DNAs were transformed into *Escherichia coli* DH5α by electroporation and selected clones were sequenced at Macrogen Inc. (Seoul, South Korea).

3.2.3 Sequence analysis

Initial nucleotide sequence similarity comparison was performed using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences were aligned with MUSCLE (Multiple Sequence Comparison by Log-Expectation) (Edgar, 2004) and pairwise identity scores were calculated using SDT (Sequence demarcation tool) (Muhire *et al*., 2014). *In silico* digestion analysis of cloned begomovirus sequences was performed with Restriction Analyzer (http://www.molbiotools.com/restrictionanalyzer.html).

3.2.4 Phylogenetic analysis

MEGA 7 (Molecular Evolutionary Genetics Analysis) was used for phylogenetic analysis using Maximum likelihood (ML) method (Kumar *et al*., 2015). The best-fit model of nucleotide substitution was selected based on corrected Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC) as implemented in MEGA 7 (Kumar *et al*., 2015).

3.2.5 Recombination detection analysis

Detection of potential recombinant sequences, identification of likely parental sequences, and localization of recombination breakpoints was carried out with RDP4 (recombination detection program) (Martin *et al*., 2015). The analysis was performed with default settings for the different

detection methods and a Bonferroni-corrected *P*-value cut-off of 0.05. Only recombination events detected with five or more methods with p-values lower than 10^{-2} were considered. Sequences used for recombination analysis were selected using SWeBLAST (Sliding Window Web-based BLAST) with a window size of 200 and a step size of 200 (Fourment *et al*., 2008) from an alignment generated with MUSCLE (Edgar, 2004).

3.3 Results

3.3.1 Genome characterization of novel begomovirus species

Digestion with *Hpa*II of RCA-amplified products from the six *O. gratissimum* leaf samples yielded restriction patterns supporting the suspected begomovirus infections, based on the symptomatology observed in the field (**Figure 14A**). Symptoms included yellow vein and/or (golden) mosaic. Sequencing of cloned DNA fragments following digestion with 6-nt enzymes confirmed the presence of DNA-A and DNA-B genomic components in the six samples (Table 5). The genome organization of the six begomoviruses was that typical of Old World bipartite begomoviruses, with six and two proteins encoded by DNA-A and DNA-B, respectively (GenBank accession numbers MN313658 to MN313669).

Initial nucleotide sequence similarity comparison for DNA-A and DNA-B performed using BLASTN and/or BLASTX showed similarities with other Old World begomoviruses. Pairwise nucleotide identity analysis showed that most DNA-A and DNA-B from the six samples showed the highest nucleotide identities with begomoviruses from Asia, i.e. China, Pakistan and India (Table 5). DNA-A of samples UG14, UG16, UG24 and UG32 showed the highest identity (72.4%, 72.6%, 71.8% and 72.3%, respectively) with the monopartite begomovirus Clerodendrum golden mosaic Jiangsu virus (ClGMJsV), isolated from the weed *Clerodendrum cyrtophyllum* (Lamiaceae) in China (Li and Zhou 2010). DNA-B from the same samples showed the highest identities (64.5%, 64.6%, 62.8% and 63.8%, respectively) with isolates of tomato leaf curl New Delhi virus (ToLCNDV) from pumpkin collected in Pakistan and potato in India. DNA-A of samples UG31 and UG33 showed the highest identities (82.7% and 83.3%, respectively) with the monopartite begomovirus tomato leaf curl Uganda virus (ToLCUV), isolated from tomato in Uganda (Shih *et al*., 2006a), whereas DNA-B from the same samples showed the highest identities (62.2% and 63.4%, respectively) with tomato leaf curl Gujarat virus isolated from tomato in India.

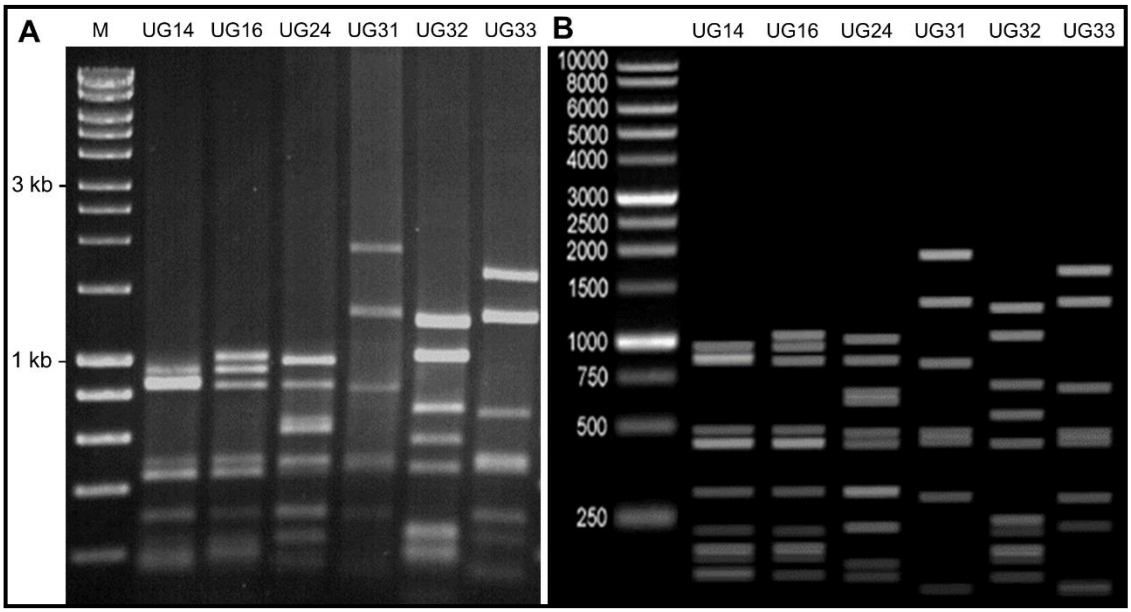

Figure 14. Restriction fragment length polymorphism (RFLP) analysis. **A,** RFLP performed by digestion of rolling circle amplification (RCA) products obtained from *O. gratissimum* plant DNA extracts with the restriction enzyme *Hpa*II revealed on a 1% agarose gel. M: molecular weight marker (HyperLadder 1kb, Bioline). **B**, *In silico* RFLP analysis of the DNA-A plus DNA-B sequences obtained from each sample was performed with Restriction Analyzer (http://www.molbiotools.com/restrictionanalyzer.html).

Table 5. Nucleotide percentage identity of the full-length DNA-A and DNA-B components of the begomoviruses isolated from *Ocimum gratissimum* plants, with the most closely related begomovirus genomes available in GenBank.

***Ocimum gratissimum* begomoviruses**						**Begomoviruses with the highest nucleotide identity**			
Virus[a]	**Sample**	**DNA component**	**GenBank acc. no.**	**Enzyme used for cloning**	**Size (nt)**	**Virus**	**Country**	**GenBank acc. no.**	**% identity**
OcYVV	UG14	DNA-A	MN313667	*Eco*RI	2804	ClGMJsV	China	FN396966	72.4
		DNA-B	MN313666	*Eco*RI	2752	ToLCNDV	Pakistan	KT948073	64.5
	UG16	DNA-A	MN313665	*Sal*I	2804	ClGMJsV	China	FN396966	72.6
		DNA-B	MN313664	*Eco*RI	2753	ToLCNDV	Pakistan	KT948073	64.6
	UG32	DNA-A	MN313663	*Eco*RI	2805	ClGMJsV	China	FN396966	72.3
		DNA-B	MN313662	*Eco*RI	2755	ToLCNDV	Pakistan	KT948073	63.8
OcMV	UG24	DNA-A	MN313669	*Nco*I	2794	ClGMJsV	China	FN396966	71.8
		DNA-B	MN313668	*Nco*I	2760	ToLCNDV	India	KC874496	62.8
OcGMV	UG31	DNA-A	MN313661	*Nco*I	2752	ToLCUV	Uganda	DQ127170	82.7
		DNA-B	MN313660	*Nco*I	2678	ToLCGUV	India	KP235538	62.2
	UG33	DNA-A	MN313659	*Eco*RI	2750	ToLCUV	Uganda	DQ127170	83.3
		DNA-B	MN313658	*Eco*RI	2678	ToLCGUV	India	KP235538	63.4

[a] Complete virus names are shown in **Figure 17** legend.

Based on their sequence identities, the DNA-As of the begomoviruses described in this work are clustered in three groups (**Figure 15**). In accordance with the begomovirus species demarcation criterion (<91% nucleotide identity for DNA-A) (Brown *et al.*, 2015), the begomoviruses described here should be assigned to three new species. The following names are proposed for them: Ocimum yellow vein virus (OcYVV, isolates [Uganda-UG14-2015], [Uganda-UG16-2015] and [Uganda-UG32-2015] for samples UG14, UG16 and UG32, respectively), Ocimum mosaic virus (OcMV, isolate [Uganda-UG24-2015] for sample UG24) and Ocimum golden mosaic virus (OcGMV, isolates [Uganda-UG31-2015] and [Uganda-UG33-2015] for samples UG31 and UG33, respectively).

A

	UG14	UG16	UG32	UG24	UG31	UG33
UG14	100					
UG16	99.4	100				
UG32	97.3	97.1	100			
UG24	90	90	90.1	100		
UG31	70.1	70.3	70	69.2	100	
UG33	70.1	70.2	70.5	69.6	96.3	100

B

	UG14	UG16	UG32	UG24	UG31	UG33
UG14	100					
UG16	99.3	100				
UG32	97.2	97	100			
UG24	83.1	83.1	82.7	100		
UG31	81	81.2	81.1	77.3	100	
UG33	81.8	81.7	81.1	77.8	93.4	100

Figure 15. Identities between the begomovirus isolates characterized in this work. Identities for DNA-A (**A**) and DNA-B (**B**) were obtained using SDT (Muhire *et al.*, 2014). Each group, which corresponds to a single begomovirus species (<91% DNA-A identity between species), is highlighted with a different color.

DNA-A and DNA-B genomic components isolated from each sample showed common regions with identities of 83.9% (UG14), 83.3% (UG16), 91.4% (UG24), 94.6% (UG31), 89.5% (UG32) and 93.3% (UG33). Furthermore, iterons were essentially identical in sequence and position in both components from the same sample, demonstrating that both DNA-A and DNA-B from each sample constitute a cognate pair (**Figure 16**).

In silico digestion analysis of cloned begomovirus sequences produced a RCA-RFLP pattern from each sample (**Figure 14B**) identical to those experimentally obtained (**Figure 14A**). This analysis unequivocally confirms that the only begomovirus present in an amplifiable amount in each sample is the one cloned and sequenced in this work.

Figure 16. Alignment of the DNA-A and DNA-B common region of Ocimum yellow vein virus from samples UG14, UG16 and UG32, Ocimum mosaic virus from sample UG24 and Ocimum golden mosaic virus from samples UG31 and UG33. The TATA box is boxed. Arrows indicate the position and orientation of the iterons.

3.3.2 Phylogenetic and recombination analysis of OcYVV, OcMV and OcGMV

Phylogenetic analysis of DNA-A and DNA-B of the begomoviruses described in this work showed that they are related to Old World begomoviruses from Africa and Asia (**Figure 17**). OcGMV (from samples UG31 and UG33) DNA-A grouped with tomato leaf curl Uganda virus as sister species. OcYVV (from samples UG14, UG16 and UG32) and OcMV (from sample UG24) DNA-A were highly divergent and grouped in a single and differentiated cluster supported by a high bootstrap value (**Figure 17A**). However, DNA-B of all begomoviruses described in this work cluster together

(**Figure 17B**). RDP analysis showed a recombinant origin for all DNA-A molecules sequenced in this work (Table 6). Begomoviruses involved in recombination events of DNA-A infect cassava, tomato and soybean in Asia, tomato in the Indian Ocean Islands, and cassava, tomato and *Desmodium* sp. in Africa (Lefeuvre *et al*., 2007; Mollel *et al*., 2017b; Shih *et al*., 2006b). Minor parents of the three recombination events detected in both isolates of OcGMV are unknown, but in one of them (nt 2322/2320 - 2626/2641) the most closely related sequence was, surprisingly, from a New World begomovirus, Sida micrantha mosaic virus (FN436005; Paprotka *et al*., 2010). With respect to the DNA-Bs, only the one of OcMV was a recombinant. The minor parent of the recombination event detected in OcMV DNA-B, interestingly, is an isolate of OcYVV and the major parent, although unknown, would be related to OcGMV.

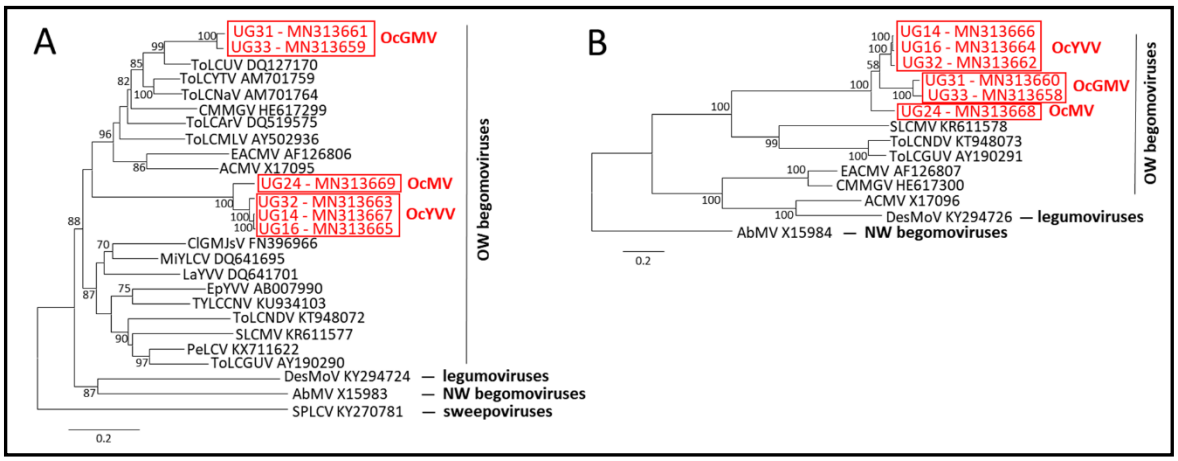

Figure 17. Phylogenetic trees illustrating the relationship of Ocimum yellow vein virus (OcYVV), Ocimum mosaic virus (OcMV) and Ocimum golden mosaic virus (OcGMV) DNA-A (**A**) and DNA-B (**B**) to other Old World begomoviruses: ACMV, African cassava mosaic virus; CMMGV, cassava mosaic Madagascar virus; ClGMJsV, Clerodendrum golden mosaic Jiangsu virus; EACMV, East African cassava mosaic virus; EpYVV, Eupatorium yellow vein virus; LaYVV, Lindernia anagallis yellow vein virus; MiYLCV, Mimosa yellow leaf curl virus; PeLCV, Pedilanthus leaf curl virus; SLCMV, Sri Lankan cassava mosaic virus; ToLCArV, tomato leaf curl Arusha virus; ToLCGUV, tomato leaf curl Gujarat virus; ToLCMLV, tomato leaf curl Mali virus; ToLCYTV, Tomato leaf curl Mayotte virus; ToLCNaV, tomato leaf curl Namakely virus; ToLCNDV, tomato leaf curl New Delhi virus; ToLCUV, tomato leaf curl Uganda virus; TYLCCNV, tomato yellow leaf curl China virus. The legumovirus Desmodium mottle virus (DesMoV); the sweepovirus sweet potato leaf curl virus (SPLCV) and the New World begomovirus Abutilon mosaic virus (AbMV) were included as outgroups. The trees were constructed by the maximum-likelihood method (1000 replicates) using the MEGA 7 program with the best-fit model, TN93+G+I for DNA-A and HKY+G+I for DNA-B. The bar below each tree indicates the number of nucleotide substitutions per site. Only bootstrap values higher than 50% are shown.

Table 6. Recombination events within the DNA-A sequence of *O. gratissium* begomoviruses detected by at least five methods included in the RDP4 package.

Recombinant genomes (GenBank acc. no.)	Recombination breakpoints	Parent-like sequences[a] (GenBank acc. no)		*p* value	Methods that detected recombination[b]
		Major	Minor		
OcYVV DNA-A (MN313667)	1183-1378	SLCMV (KR611577)	TYLCCNV (KU934103)	$3.289x10^{-04}$	R, B, M, C, S, 3S
	2131-2279	ToLCYTV (AM701759)	DesMoV (KY294724)	$1.830x10^{-06}$	R, B, M, C, 3S
OcYVV DNA-A (MN313665)	1183-1378	SLCMV (KR611577)	TYLCCNV (KU934103)	$3.289x10^{-04}$	R, B, M, C, S, 3S
	2131-2279	ToLCYTV (AM701759)	DesMoV (KY294724)	$1.830x10^{-06}$	R, B, M, C, 3S
OcYVV DNA-A (MN313663)	986-1410	SLCMV (KR611577)	TYLCCNV (KU934103)	$3.289x10^{-04}$	R, B, M, C, S, 3S
	2120-2278	ToLCYTV (AM701759)	DesMoV (KY294724)	$1.830x10^{-06}$	R, B, M, C, 3S
OcMV DNA-A (MN313669)	989-1369	SLCMV (KR611577)	TYLCCNV (KU934103)	$3.289x10^{-04}$	R, B, M, C, S, 3S
	2132-2272	ToLCYTV (AM701759)	DesMoV (KY294724)	$1.830x10^{-06}$	R, B, M, C, 3S
OcMV DNA-B (MN313668)	2608-582	Unknown	OcYVV (MN313666)	$1.403x10^{-25}$	R, G, B, M, C, S, 3S
OcGMV DNA-A (MN313661)	361-481	EACMV (HG530114)	Unknown	$1.185x10^{-04}$	R, G, B, M, C, 3S
	1971-2109	ToLCArV (DQ519575)	Unknown	$2.143x10^{-03}$	R, G, B, M, C
	2322-2626	ToLCNaV (AM701764)	Unknown	$4.931x10^{-11}$	R, G, B, M, C, S, 3S
OcGMV DNA-A (MN313659)	359-488	EACMV (HG530114)	Unknown	$1.185x10^{-04}$	R, G, B, M, C, 3S
	1969-2107	ToLCArV (DQ519575)	Unknown	$2.143x10^{-03}$	R, G, B, M, C
	2320-2641	ToLCNaV (AM701764)	Unknown	$4.931x10^{-11}$	R, G, B, M, C, S, 3S

[a] Complete virus names are shown in **Figure 17** legend.

[b] The method with the lowest p-value obtained for each region is underlined. R, RDP; G, GENCONV; B, BootScan; M, MaxChi; C, Chimaera; S, SiScan; 3S, 3Seq.

3.4 Discussion

Begomoviruses are a major constraint to production of economically important crops around the world, including sub-Saharan Africa, where a great effort has been devoted to characterizing viruses infecting major crops, exemplified by cassava (Rey and Vanderschuren 2017). Nevertheless, begomoviruses infecting minor crops, weeds and other wild plants are poorly studied in this continent. However, understanding the viral diversity in minor crops and wild plants is important as it has epidemiological implications. Thus, these plants can act as reservoirs of viruses infecting important crops (García-Arenal and Zerbini 2019; Navas-Castillo *et al*., 2011) and even as niches for the development of new strains or species by recombination (García-Andrés *et al*., 2006).

In the present study, symptomatic cultivated and wild *O. gratissimum* plants from Uganda were found to be commonly infected with begomoviruses, being the first report of a viral infection in this species. This is also the first report of a virus infecting *Ocimum* spp. in Africa. Outside Africa, basil (*O. basilicum*) has been reported to be infected by alfalfa mosaic virus (*Alfamovirus*) and Impatiens necrotic spot virus (*Orthotospovirus*) in United States (Poojari and Naidu 2013; Wintermantel and Natwick 2012) and pepino mosaic virus (*Potexvirus*) in Italy (Davino *et al*., 2009). On the other hand, holy basil [*O. tenuiflorum* (syn. *O. sanctum*)] has been reported to be infected by cucumber mosaic virus (*Cucumovirus*) in India (Khan *et al*., 2011). With regard to begomovirus infections, *O. basilicum* has been found infected by chili leaf curl virus and tomato yellow leaf curl virus in Oman (Ammara *et al*., 2015) and *O. tenuiflorum* by tomato leaf curl Gujarat virus in India (Gaur *et al*., 2012; Nehra *et al*., 2019).

The three novel begomovirus species, reported here for the first time, belonged to the Old World phylogenetic group of begomoviruses, but the DNA-A of two of them were highly divergent. In contrast, all the DNA-B components were closely related to one another and grouped in a single cluster. This is an example that demonstrates that the two genome components, DNA-A and DNA-B, have distinct molecular evolutionary histories. This phenomenon reflects component exchange between different viruses, which has been shown to be much more frequent in Old World begomoviruses than in New World begomoviruses (Briddon *et al*., 2010). This has been shown previously, for example, for the New World begomoviruses Rhynchosia rugose golden mosaic virus (Fiallo-Olivé *et al*., 2010a) and pepper golden mosaic virus (Brown *et al*., 2005) and the Old World begomoviruses Desmodium mottle virus (Mollel *et al*., 2017b) and those belonging to the cassava mosaic virus complex (De Bruyn *et al*., 2016). In the case of *O. gratissimum*, it can be hypothesized that the DNA-B components of the three begomoviruses have evolved to adapt to this host. It would be interesting to test if they are also well adapted to other species of the genus *Ocimum*.

Begomoviruses have been mostly described from annual crops but, particularly in the last few years, the availability of RCA and next-generation sequencing techniques have revealed a higher diversity in this genus of plant viruses, including those present in wild plants, many of them perennial. The finding of novel begomovirus species in *O. gratissimum* is a clear example of the hidden viral diversity present in understudied hosts and geographic areas. Our survey was carried out in southwestern Uganda following a transect along the main local roads with some stops to visit backyard gardens. Only six *O. gratissimum* plants were observed with symptoms but all of them were infected by begomoviruses belonging to three new species. Africa is largely underrepresented in relation to other tropical regions as Latin America or the Indian subcontinent, with respect to plant-virus identification. Increased efforts, therefore, should be implemented to survey new areas and wild host-plants, which will surely be rewarded by the discovery of novel viruses.

The results of phylogenetic and recombination analyses obtained in this work emphasize the interconnection between cultivated and wild plants in relation to the exchange of viruses between them (Cooper and Jones 2006; García-Arenal and Zerbini 2019). There are a number of examples of begomoviruses initially identified in wild plants and later in crops as is the case of Euphorbia mosaic virus in the New World, which was first identified in pioneering research in the 1950´s, infecting several species of the genus *Euphorbia* (Costa and Bennett 1950; Bird *et al.*, 1975; Debrot and Centeno 1986) and later found causing emergent diseases in sweet pepper and tobacco crops (Gregorio-Jorge *et al.*, 2010; Fiallo-Olivé *et al.*, 2010b). A similar example from the Old World is Ageratum yellow vein virus, firstly, isolated from the common weed *Ageratum conyzoides* (Tan *et al*., 1995) and later from soybean (Samretwanich *et al*., 2001) and tomato (Andou *et al*., 2010).

Additional work is needed to determine the geographical distribution of the begomoviruses described in this work across Africa and their potential harmfulness to crops including other cultivated species of the genus *Ocimum* and economically important crops. OcGMV, one of the begomovirus species characterized here, is closely related with a begomovirus isolated from tomato also in Uganda (Table 5, **Figure 17**), so OcGMV could potentially infect this crop too. Also, it is important to consider that *O. gratissimum*, being a perennial plant, could be a reservoir acting as a melting pot for accumulation of different begomoviruses prone to recombine and driving the generation of new recombinant viruses, as has been shown for other wild plants (García-Andrés *et al*., 2006).

Although the begomoviruses discovered in *O. gratissimum* are the suspect causal agent of the symptoms shown by the hosts, Koch's postulates still need to be proven, as well as the possibility that these viruses infect certain crops. This is important, because their abilities to infect crops may result in future plant-virus epidemics in East Africa.

On the other hand, it would be interesting to determine which species of the *B. tabaci* complex present in East Africa are responsible for the transmission of the begomoviruses characterized in this work. Low populations of *B. tabaci* were found infesting some of the *O. gratissimum* plants sampled in this work but they were not genetically characterized. In a study carried out in 2003/04 to identify the host-plant distribution of *B. tabaci* in Uganda, Sseruwagi *et al*. (2005) reported the occurrence of the genotype 'Uganda 3' [currently named Sub-Saharan Africa 6 species (Mugerwa *et al*., 2018)] only on wild *O. gratissimum* plants. However, SSA6 has been found recently to occur on a wider host range in Uganda, including several wild plant species (Mugerwa *et al*., 2018).

The importance of *O. gratissimum* as host and reservoir of begomoviruses and whiteflies is highlighted by the results shown here and this knowledge also provides new insights into the epidemiological situation of plant viruses in Africa. These findings have also the potential to contribute to the design of more effective and specific control strategies against the panoply of viruses that have been and are being described in this continent, which otherwise constitute probably only the tip of the iceberg of virus threats to economically and socially important crops.

CHAPTER 4

MATING INTERACTIONS AMONG WHITEFLY POPULATIONS OF *BEMISIA TABACI* BELONGING TO THE SUB-SAHARAN HIGH-LEVEL GROUP FROM SPAIN AND SUB-SAHARAN AFRICA

CHAPTER 4: MATING INTERACTIONS AMONG WHITEFLY POPULATIONS OF *BEMISIA TABACI* BELONGING TO THE SUB-SAHARAN HIGH-LEVEL GROUP FROM SPAIN AND SUB-SAHARAN AFRICA

4.1 Introduction

The whitefly *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) cryptic species complex is widely distributed throughout the tropical and sub-tropical regions of the world where it has become a major crop pest (Brown *et al.*, 1995b; Martin *et al.*, 2000). *B. tabaci* is a phloem-feeding insect that causes damage on a wide range of crops including cassava, tomato, cucubirts, beans, cotton and sweet potato in different parts of the world. *B. tabaci* causes direct damage to the plants through feeding on the host plant by both nymphs and adults and indirect damage by excreting honeydew on the surfaces of the leavess (Byrne and Bellows, 1991). The honeydew acts as a substrate for the growth of sooty mould, which interferes with photosynthesis and reduces the quantity and quality of the product (Gerling and Mayer, 1996). In addition, *B. tabaci* is the main vector of begomoviruses that are the most devastating plant viruses of important vegetable and fiber crops in tropical and sub-tropical regions (Navas-Castillo *et al.*, 2011; Rojas *et al.*, 2018).

B. tabaci originated in sub-Saharan Africa and spread to northern Africa, the Mediterranean basin, Asia, Australia, the Neotropics and southern North America (Boykin *et al.*, 2007; De Barro, 2012; Mugerwa *et al.*, 2018). Based on the comparison of the partial *mitochondrial cytochrome oxidase 1* (mtCO1) gene, *B. tabaci* has been proposed to be a complex of 11 high-level genetic groups (De Barro *et al.*, 2011). These genetic groups are composed of more than 40 morphologically indistinguishable species, which are separated by a minimum of 3.5% mtCOI nucleotide divergence (Dinsdale *et al.*, 2010; De Barro and Ahmed, 2011; Boykin and De Barro, 2014). They are genetically diverse and differ greatly in their biological characteristics (Liu *et al.*, 2009; Firdaus *et al.*, 2013; Alemandri *et al.*, 2015) including: specificity of begomovirus transmission (Polston *et al.*, 2014; Maruthi *et al.*, 2002), inducement of phytotoxic disorders (Brown *et al.*, 1995a; Sseruwagi *et al.*, 2005), invasiveness (Brown *et al.*, 1995a; Liu *et al.*, 2007; Wang *et al.*, 2012), resistance to insecticides (Horowitz and Ishaaya, 2014), and host-plant range (Zang *et al.*, 2006; Xu *et al.*, 2011).

The relatedness of *B. tabaci* species complex has been mainly studied using mtCOI gene, however proper species boundaries of *B. tabaci* has been proposed to be achieved with nuclear markers coupled with mating data from crossing experiment between different genetic groups (Liu *et al.*, 2012; Hsieh *et al.*, 2014). *B. tabaci* is an arrhenotokous insect that produces females and males from fertilized and unfertilized eggs, respectively (Byrne and Devonshire, 1996), thus facilitating reproductive compatibility studies. Mating studies within the *B. tabaci* species complex have been

conducted with genetic groups from different parts of the world, revealing reproductive incompatibility based on fecundity and fertility data (De Barro and Hart, 2000; De Barro *et al*., 2005; Maruthi *et al*., 2004; Perring *et al*., 1993; Perring and Symmes, 2006; Qin *et al*., 2016; Liu *et al*., 2007; Sun *et al*., 2011; Vyskočilová *et al*., 2018; Xu *et al*., 2010). The Sub-Saharan Africa (SSA) high-level group is among the major genetic groups of the *B. tabaci* species complex. Five putative species were initially described within this group, SSA1 to SSA5, mostly infecting cassava (*Manihot esculenta*) (Berry *et al*., 2004; Esterhuizen *et al*., 2013; Sseruwagi *et al*., 2006; Mugerwa *et al*., 2012) and associated to transmission of begomoviruses causing cassava mosaic disease (CMD) (Chant, 1958; Maruthi *et al*., 2002) and ipomoviruses causing cassava brown streak disease (CBSD) (Maruthi *et al*., 2005b). Eight more SSA species have been discovered SSA6 to SSA13 (Mugerwa *et al*., 2018; Sseruwagi *et al*., 2005). SSA1 and SSA2 are the most abundant whitefly populations colonizing cassava in sub-Saharan Africa, and they were associated with the rapid spread of the CMD pandemic, particularly in East Africa (Legg, 1999). SSA2 has not only been reported in sub-Saharan Africa, but it has also been found in the Mediterranean basin including Spain colonising the ornamental plant morning glory (*Ipomoea indica*) (Banks *et al*., 1999; De La Rúa *et al*., 2006; Hadjistylli *et al*., 2015; Laarif *et al*., 2015). Besides to the available mtCOI sequences of Mediterranean and sub-Saharan SSA2 populations, no information on the gene flow between the two groups of populations is available. Therefore, to understand the reproductive interactions between them, and with other SSA populations, reciprocal crossing experiments were conducted and the fertility of the female hybrids of first filial (F_1) generation was analyzed.

4.2 Materials and Methods

4.2.1 Insect rearing

B. tabaci colonies were reared on aubergine plants (*Solanum melongena* 'Black Beauty') in an insectary with controlled conditions at NRI, UK. Plants were grown from seeds in a whitefly-free room at 28 ± 2°C, 50 - 60% relative humidity (RH) and a 14:10 light:dark (L:D) photoperiod. Whitefly colonies were maintained in nylon cages of 44 × 44 × 44 cm (BugDorm, US) at 28 ± 2°C, 60% RH and 14:10 L:D photoperiod. Colonies from sub-Saharan Africa were established from field populations collected from cassava in Uganda [in 2013 (SSA1-SG1) and 2016 (SSA2)], Tanzania [in 2013 (SSA1-SG3)] and Nigeria [in 2016 (SSA2 and SSA3)]. Colonies from Spain were established from pure colonies previously established at IHSM, Spain, from a field population collected from *Ipomoea indica*. The purity of core whitefly colonies was checked by sequencing the partial mtCOI

gene amplified using primer pairs C1-J-2195 and TL2-N-3014 (Frohlich *et al*., 1999). The nucleotide divergence of the partial mtCOI between the populations analysed in this work as shown in Table 7.

Table 7. Pairwise comparison of the partial mitochondrial cytochrome oxidase 1 nucleotide sequence expressed as percentage of nucleotide divergence between populations of *Bemisia tabaci* used in the crossing experiments. The grey areas show the values < 3.5%, proposed as a threshold for species in the *B. tabaci* complex.

Bemisia tabaci populations	**1**	**2**	**3**	**4**	**5**	**6**
1. SSA1-SG1-Ug	0					
2. SSA1-SG3-Tz	1.22	0				
3. SSA2-Ug	8.07	7.76	0			
4. SSA2-Ng	7.61	7.61	0.76	0		
5. SSA2-Sp	7.31	7.31	1.37	1.22	0	
6. SSA3-Ng	7.00	7.00	5.94	6.09	5.78	0

4.2.2 F_1 and F_2 reciprocal crossing experiments

Reproductive compatibility among whitefly populations SSA2-Spain (SSA2-Sp) and SSA1-SG1-Uganda (SSA1-SG1-Ug), SSA1-SG3-Tanzania (SSA1-SG3-Tz), SSA2-Uganda (SSA2-Ug), SSA2-Nigeria (SSA2-Ng) and SSA3-Nigeria (SSA3-Ng) from sub-Saharan Africa was determined by reciprocal crossing experiments. The crossing study involved single and group mating experiments. All experiments were conducted in the insectary and conditions mentioned above.

All crosses were carried out using newly emerged virgin adults isolated by excision of a small leaf piece containing a single 4th instar nymph, individually placed into glass tubes and left at room temperature in the insectary for adult emergence. Emerged adults were sexed using a binocular stereomicroscope (Nikon, Japan; 10x magnification) prior to mating experiments. For single mating experiments, controlled crosses were carried out with one female (♀) and three males (♂) from the same population, whilst for the reciprocal crosses, female and males originating from different populations were used. The four adults were released onto aubergine plants in three to four true-leaf stages, rooted in soil and enclosed in Lock-and-Lock pots (LLs) whitefly-proof cages (Wang *et al*., 2011). The LLs had side openings and an opening at the top covered by 160 µm nylon mesh for air circulation. At least 24 h before introducing the adults, all leaves, except a fully expanded one, were removed to facilitate contact between the mating partners. Similar experiment was conducted as group mating experiment in which five females and ten males were used. Survival of the parental adults was

monitored periodically. Parents were collected after seven days and stored in 90% ethanol at -20°C. All emerged adults were collected, sexed and stored for subsequent molecular analysis to verify the genetic make-up of F_1 females.

4.2.3 Molecular verification of F_1 hybrid females

The parentage of female progeny was studied using the nuclear marker *pre-mRNA processing factor 8* gene (*prpf8*) (Gershoni *et al.*, 2009). Partial *prpf8* gene sequence was amplified by polymerase chain reaction (PCR) with the primer set, prp8F (5'-GCCTTGGGAGGTGTTGAAG-3') and prp8R (5'-GGCTTGCATCCAGGGTACC-3') (Hsieh *et al*., 2014).

Genomic DNA was extracted from hybrid females stored in 90% ethanol using the modified Chelex extraction method for whiteflies (White *et al*., 2009). At least four individuals from each reciprocal cross were used to confirm the identities of the female progeny. Whiteflies were homogenised individually using zirconium oxide beads (Web Scientific) and a Bullet Blender Gold (Next Advance)) in 75 µl of 10% Chelex 100 Resin solution (Bio-Rad) containing 6 µl Proteinase K (10µg/µl). The mixture was centrifuged at 13,500 rpm for 1 min and incubated at 37 °C for 1 hour, and subsequently at 96 °C for 8 min. The supernatant was used as a template DNA for amplification by PCR. PCR assay was conducted using 2 µl of template DNA in a total reaction volume of 30 µl consisting of 5x MyTaq buffer HS, 1 unit of MyTaq DNA polymerase (Bioline), 0.2 µM of each primer and ultrapure water. PCR conditions were: 1 cycle of 5 min at 94 °C; 35 cycles of 1 min at 94 °C, 1 min at 56 °C, and 2 min at 72 °C; and a final cycle of 10 min at 72 °C. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) of amplified *prpf8* gene for both reciprocal and control crosses was performed using the restriction enzyme *Apo*I. The digested products were separated on 3% agarose gel and bands were visualized by GelRed (Biotium) staining and viewed with a UV light source.

4.2.4 Fertility and viability test and statistical analysis

The fertility and viability of F_1 hybrids were tested in a separate set of reciprocal crossing experiments for populations that produced both female and male offspring. Self-crossing of F_1 siblings from both directions of the crosses were collected and released onto young aubergine plants with a single leaf. Then the procedures were followed as described in section 4.2.2 above.

Statistical analysis was performed using R (R Core Team, 2013). Counts of offspring generated in crossing experiments were analysed by a generalised linear model with negative binomial error distribution and a log link using the MASS library (Venables and Ripley, 2002). For

the proportion of female progeny, a generalised linear model with quasibinomial error distribution and logit link of the proportional data was used. Multiple comparisons of offspring count, and female proportions in the mating experiments were performed by Tukey's test (Tukey, 1949) using the multcomp package (Hothorn *et al*., 2008) and significant differences were demonstrated by compact letter display. A separate analysis was also carried out for comparison of proportions of females produced in F_1 and F_2 generations.

4.3 Results

4.3.1 F_1 reciprocal crossing experiments

All the reciprocal crosses among SSA2-Sp and SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng produced both female and male progeny for both single and group mating experiments (Tables 8 and 9). However, no female progeny was produced in F_1 from a reciprocal cross between SSA2-Sp x SSA1-SG1-Ug, whilst the control crosses produced 33-79% females (Table 8).

Table 8. Results (means and standard errors) from reciprocal crosses of single mating experiments among the Spanish population, SSA2-Sp and populations from sub-Saharan Africa: SSA1-SG1-Ug, SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng of the *Bemisia tabaci* putative species.

Crosses (1♀ × 3♂)	**Replicates**	**Mean no. of progeny**	**Mean no. of female**	**Mean % females**
Control				
♀SSA1-SG1-Ug × ♂SSA1-SG1-Ug	7	11.71 ± 4.03^{ab}	9.29 ± 4.80	79.00 ± 11.00^{d}
♀SSA2-Sp × ♂SSA2-Sp	9	66.0 ± 19.17^{b}	21.78 ± 9.76	33.00 ± 5.00^{abc}
♀SSA1-SG3-Tz × ♂SSA1-SG3-Tz	7	46.0 ± 15.21^{b}	18.29 ± 9.32	40.00 ± 7.00^{ab}
♀SSA2-Ng × ♂SSA2-Ng	7	30.43 ± 10.14^{ab}	9.29 ± 4.80	31.00 ± 8.00^{abc}
♀SSA2-Ug × ♂SSA2-Ug	6	23.83 ± 8.62^{ab}	10.33 ± 5.75	43.00 ± 10.00^{abcd}
♀SSA3-Ng × ♂SSA3-Ng	9	33.00 ± 9.68^{ab}	23.11 ± 10.35	70.00 ± 7.00^{bcd}
Reciprocal				
♀SSA1-SG1-Ug × ♂SSA2-Sp	6	26.83 ± 9.68^{ab}	0.00 ± 0.00	0.00 ± 0.00
♀SSA2-Sp × ♂SSA1-SG1-Ug	6	5.00 ± 1.98^{a}	0.00 ± 0.00	0.00 ± 0.00
♀SSA1-SG3-Tz × ♂SSA2-Sp	6	101.33 ± 35.91^{b}	15.67 ± 8.64	15.00 ± 4.00^{ab}
♀SSA2-Sp × ♂SSA1SG3-Tz	8	37.13 ± 11.52^{ab}	4.38 ± 2.18	12.00 ± 5.00^{ab}
♀SSA2-Ng × ♂SSA2-Sp	11	54.55 ± 14.36^{b}	18.55 ±7.54	34.00 ± 5.00^{abc}
♀SSA2-Sp × ♂SSA2-Ng	10	25.00 ± 7.00^{ab}	6.60 ± 2.89	26.00 ± 7.00^{abc}
♀SSA2-Ug × ♂SSA2-Sp	10	45.50 ± 12.59^{ab}	19.30 ± 8.22	42.00 ± 6.00^{abcd}
♀SSA2-Sp × ♂SSA2-Ug	5	115.40 ± 44.77^{b}	45.40 ± 27.12	39.00 ± 5.00^{abcd}
♀SSA3-Ng × ♂SSA2-Sp	9	45.22 ± 13.19^{ab}	15.22 ± 6.86	34.00 ± 6.00^{abc}
♀SSA2-Sp × ♂SSA3-Ng	12	26.75 ± 6.83^{ab}	20.25 ± 7.87	76.00 ± 6.00^{cd}

Different superscript letters indicate statistically significant differences ($P < 0.05$) between crosses (Tukey's test).

Table 9. Results (means and standard errors) from reciprocal crosses of group mating experiments among the Spanish population, SSA2-Sp and populations from sub-Saharan Africa: SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng of the *Bemisia tabaci* putative species.

Crosses (5♀ × 10♂)	**Replicates**	**Mean no. of progeny**	**Mean no. of female**	**Mean % females**
Control				
♀SSA1-SG3-Tz × ♂SSA1-SG3-Tz	4	88.75 ± 19.24^{abc}	34.00 ± 15.56	38.00 ± 9.00^{b}
♀SSA2-Sp × ♂SSA2-Sp	4	174.50 ± 37.2^{c}	56.00 ± 25.46	32.00 ± 6.00^{ab}
♀SSA2-Ng × ♂SSA2-Ng	3	40.67 ± 10.53^{a}	13.33 ± 7.24	33.00 ± 15.00^{ab}
♀SSA2-Ug × ♂SSA2-Ug	3	43.00 ± 11.10^{a}	12.67 ± 6.89	29.00 ± 14.00^{ab}
♀SSA3-Ng × ♂SSA3-Ng	4	67.00 ± 14.66^{abc}	41.50 ± 18.94	62.00 ± 11.00^{b}
Reciprocal				
♀SSA1-SG3-Tz × ♂SSA2-Sp	8	122.62 ± 18.64^{bc}	3.00 ± 1.13	2.00 ± 2.00^{a}
♀SSA2-Sp × ♂SSA1-SG3-Tz	6	133.00 ± 23.30^{bc}	3.67 ± 1.56	3.00 ± 2.00^{a}
♀SSA2-Ng × ♂SSA2-Sp	5	185.20 ± 35.34^{c}	71.40 ± 28.96	39.00 ± 6.00^{b}
♀SSA2-Sp × ♂SSA2-Ng	4	37.75 ± 8.51^{a}	13.25 ± 6.23	35.00 ± 14.00^{ab}
♀SSA2-Ug × ♂SSA2-Sp	7	144.29 ± 23.37^{bc}	61.14 ± 20.99	42.00 ± 6.00^{b}
♀SSA2-Sp × ♂SSA2-Ug	5	98.40 ± 19.02^{abc}	45.60 ± 18.59	46.00 ± 8.00^{b}
♀SSA3-Ng × ♂SSA2-Sp	5	152.60 ± 29.21^{c}	57.80 ± 23.49	38.00 ± 6.00^{b}
♀SSA2-Sp × ♂SSA3-Ng	6	43.83 ± 7.99^{ab}	31.50 ± 11.79	72.00 ± 10.00^{b}

Different superscript letters indicate statistically significant differences ($P < 0.05$) between crosses (Tukey's test)

Mean counts of F_1 adults in these successful crosses in single mating experiments, as well as female proportions were not significantly different from the controls (Table 8; **Figure 18A**). In group mating experiments, the mean counts of F_1 adults in the successful crosses differed significantly from controls except for a cross between SSA2-Sp and SSA1-SG3-Tz (Table 9), and with regards to female proportions no significant difference was observed from the controls except in a cross between SSA2-Sp and SSA1-SG3-Tz (**Figure 18B**). In general, the proportion of females in controls and F_1 reciprocal crosses for both single and group mating ranged from 2 to 40% for SSA2-Sp x SSA1-SG3-Tz, 26 to 39% for SSA2-Sp x SSA2-Ng), 29 to 46% for SSA2-Sp x SSA2-Ug and 32 to 76% for SSA2-Sp x SSA3-Ng.

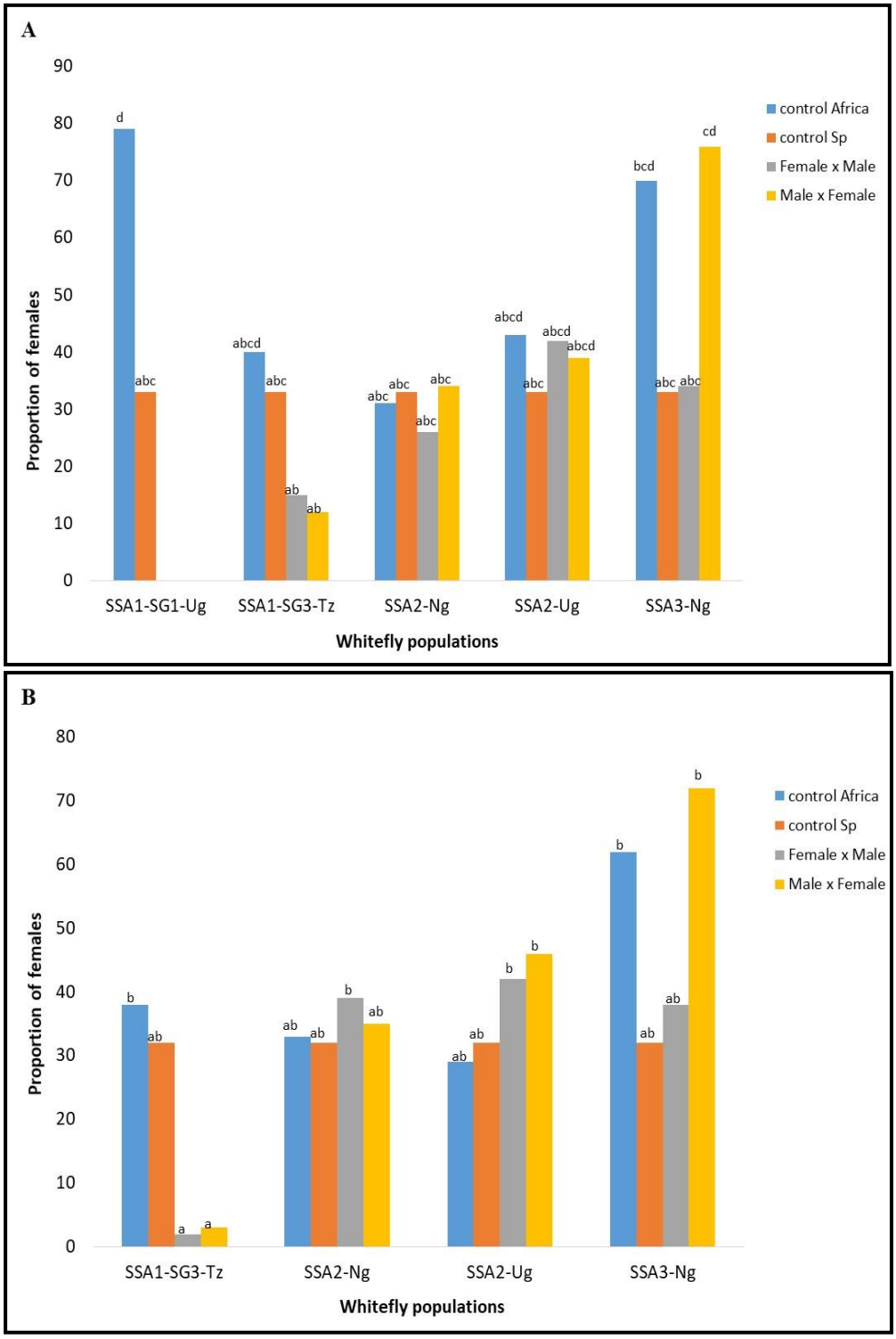

Figure 18. Proportion of females produced in single crossing experiments (**A**) and group crossing experiments (**B**) among whitefly populations of *B. tabaci* species. The compact letter display codes show which bars are significantly different at the $p<0.05$ level. Bars with the same letter are not significantly different.

4.3.2 F_1 hybrids verification by nuclear marker

The nuclear marker, *prpf8* gene was successfully amplified from all controls (Spanish and African populations) and reciprocal crosses: SSA2-Sp x SSA1-SG3-Tz, SSA2-Sp x SSA2-Ng, SSA2-Sp x SSA2-Ug and SSA2-Sp x SSA3-Ng that produced hybrids female. The digestion of amplified *prpf8* gene using restriction enzyme *Apo*I yielded different restriction patterns that distinguished the Spanish population from sub-Saharan Africa populations. The Spanish population generated fragments of about 450 bp, 321 bp and 195 bp, whereas sub-Saharan Africa populations produced fragments of approximately 503 bp, 450 bp and 195 bp **(Figure 19).**

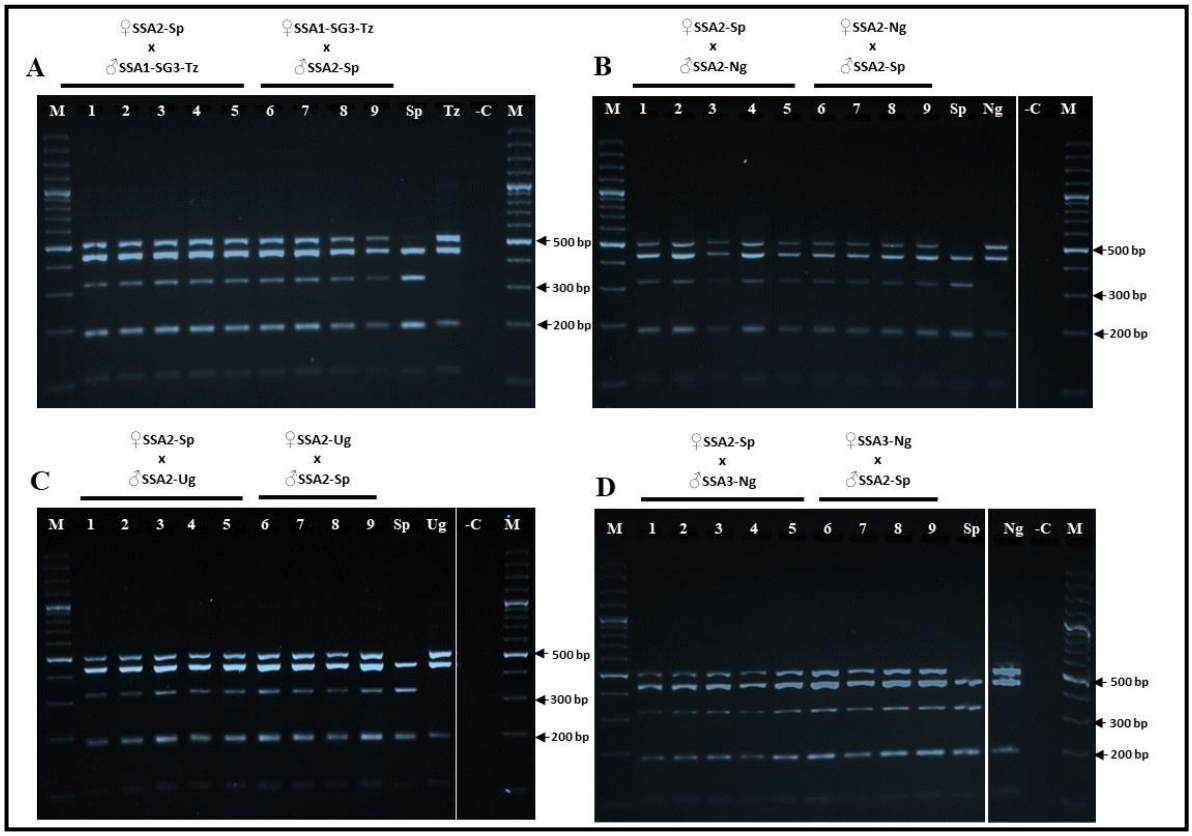

Figure 19. RFLP patterns of amplified *prpf8* nuclear gene of hybrid females digested with *Apo*I enzyme. The patterns of reciprocal crosses of populations of *B. tabaci* from Spain and sub-Saharan Africa were compared with their control crosses. Lane 1-9; hybrids female for SSA2-Sp x SSA1-SG3-Tz cross (**A**), SSA2-Sp x SSA2-Ng cross (**B**), SSA2-Sp x SSA2-Ug cross (**C**), and SSA2-Sp x SSA3-Ng cross (**D**). Sp: Control for SSA2-Sp, and Tz: Control for SSA1-SG3-Tz, Ng: Control for SSA2-Ng, Ug: Control for SSA2-Ug and Ng: Control for SSA3-Ng respectively. -C: Negative control (PCR reaction mix without template DNA) and M: DNA marker, 100 bp plus.

4.3.3 Fertility and viability test

Both male and female offspring were produced in six combinations of F_1 hybrids in both directions of the crosses in all replicates of F_2 generation. These successful crosses included SSA2-Sp x SSA2-Ug, SSA2-Sp x SSA2-Ng and SSA2-Sp x SSA3-Ng. However, SSA2-Sp in combination with SSA1-SG3-Tz produced hybrid females only in one direction of the cross derived from female SSA2-Sp and male SSA1-SG3-Tz (Table 10).

There were significant differences in the proportions of female progeny between crosses ($P<0.05$), but not between F_1 and F_2 generations (Table 11; **Figure 20**). Also, the interaction was not significant, as there was no evidence from our data that crosses had different outcomes in F_1 and F_2.

Table 10. Results (means and standard errors) of F_1 and F_2 generations among the Spanish population, SSA2-Sp and populations from sub-Saharan Africa: SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng of the *Bemisia tabaci* putative species.

Crosses	Replicates	Mean no. of progeny	Mean no. of female	Mean % females
F_1 generation				
♀SSA1-SG3-Tz × ♂SSA2-Sp	8	122.63 ± 23.58	3.00 ± 1.16	2.00 ± 2.00^{a}
♀SSA2-Sp × ♂SSA1-SG3-Tz	6	133.00 ± 29.50	3.67 ± 1.60	3.00 ± 2.00^{a}
♀SSA2-Ng × ♂SSA2-Sp	5	185.20 ± 44.84	71.40 ± 30.03	39.00 ± 5.00^{bc}
♀SSA2-Sp × ♂SSA2-Ng	4	37.75 ± 10.58	13.25 ± 6.44	35.00 ± 15.00^{bc}
♀SSA2-Ug × ♂SSA2-Sp	7	144.29 ± 29.60	61.14 ± 21.76	42.00 ± 5.00^{bc}
♀SSA2-Sp × ♂SSA2-Ug	5	98.40 ± 24.02	45.60 ± 19.26	46.00 ± 7.00^{bc}
♀SSA3-Ng × ♂SSA2-Sp	5	152.60 ± 37.02	57.80 ± 24.35	38.00 ± 6.00^{bc}
♀SSA2-Sp × ♂SSA3-Ng	6	43.83 ± 9.97	31.50 ± 12.21	72.00 ± 9.00^{c}
F_2 generation				
♀SSA2-Sp × ♂SSA1-SG3-Tz	3	89.00 ± 28.09	25.33 ± 13.95	28.00 ± 9.00^{abc}
♀SSA2-Ng × ♂SSA2-Sp	7	64.71 ± 13.47	13.86 ± 5.08	21.00± 6.00^{ab}
♀SSA2-Sp × ♂SSA2-Ng	3	46.00 ± 14.77	23.67 ± 13.05	51.00 ± 14.00^{bc}
♀SSA2-Ug × ♂SSA2-Sp	6	178.50 ± 39.46	69.67 ± 26.75	39.00 ± 5.00^{bc}
♀SSA2-Sp × ♂SSA2-Ug	3	66.00 ± 20.97	29.00 ± 15.93	44.00 ± 11.00^{bc}
♀SSA3-Ng × ♂SSA2-Sp	3	94.67 ± 29.85	41.33 ± 22.57	44.00 ± 9.00^{bc}
♀SSA2-Sp × ♂SSA3-Ng	3	89.00 ± 28.09	25.33 ± 13.95	28.00 ± 9.00^{abc}

Different superscript letters indicate statistically significant differences ($P < 0.05$) between crosses (Tukey's test)

Table 11. Analysis of deviance for the proportion of female progeny from reciprocal crossing experiment for F_1 and F_2 generations.

	Df	Deviance Resid.	Df	Resid. Dev	F	Pr(>F)
NULL			73	1969.86		
cross	7	1195.82	66	774.04	16.4157	8.596e-12 ***
gen	1	0.76	65	773.28	0.0730	0.787941
cross:gen	6	229.40	59	543.88	3.6739	0.003625 **

Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

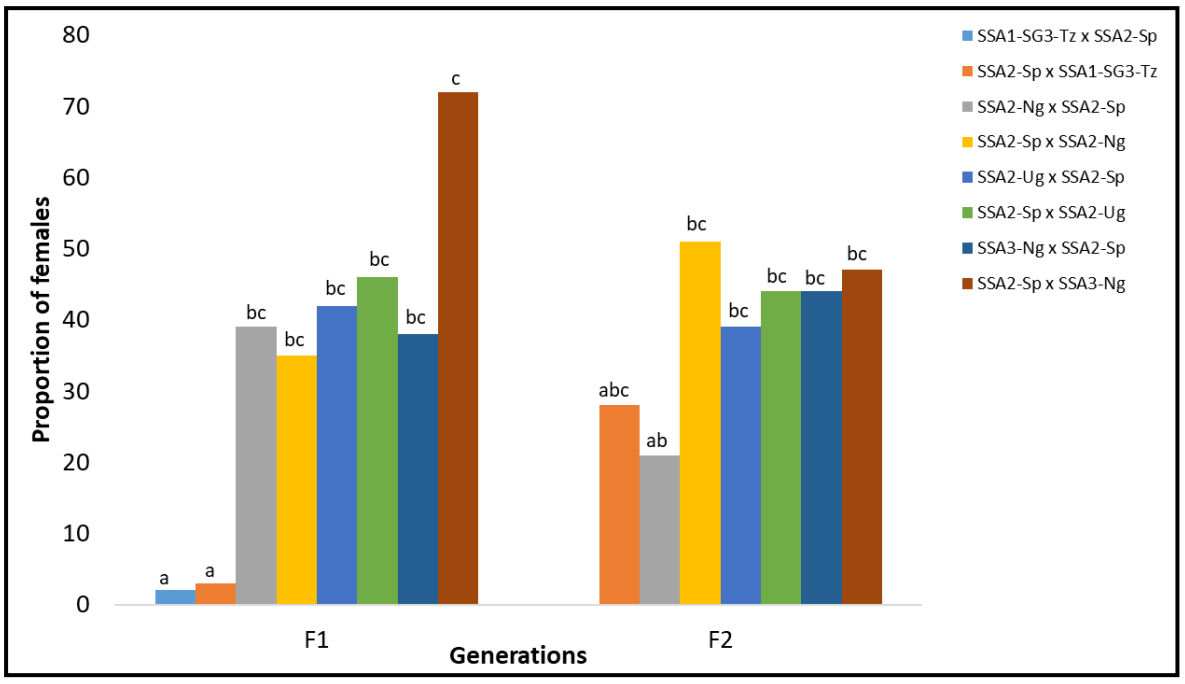

Figure 20. The proportion of hybrid females from reciprocal crosses among SSA2-Sp and SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng for two generations (F_1 and F_2). The compact letter display codes show which bars are significantly different at the $p<0.05$ level. Bars with the same letter are not significantly different.

4.4 Discussion

B. tabaci is a cryptic species complex comprising members that are indistinguishable morphologically and therefore, ultimate differentiation between them comes from evidence of reproductive isolation (Liu *et al*., 2012). Proper species boundaries of *B. tabaci* has been proposed earlier to be achieved with nuclear markers coupled with mating data from crossing experiment between genetic groups (Liu *et al.*, 2012; Hsieh *et al.*, 2014). Based on the laboratory reciprocal crossing experiments performed and *prpf8* nuclear marker, this study found mating interactions between whitefly populations, SSA2-Sp colonizing *I. indica* in Spain and SSA1-SG1-Ug, SSA1-SG3-Tz, SSA2-Ug, SSA2- and SSA3-Ng colonizing cassava in sub-Saharan Africa. The reciprocal crossing experiment carried out in a laboratory-controlled conditions confirmed complete mating compatibility between SSA2-Sp and SSA2-Ug, SSA2-Ng and SSA3-Ng, and partial compatibility between SSA2-Sp and SSA1-SG3-Tz in both directions of the crosses. However, complete mating incompatibility was also verified between SSA2-Sp and SSA1-SG1-Ug. Furthermore, self-crossing experiments of F_1 siblings confirmed the fertility and viability of SSA2-Sp x SSA2-Ug, SSA2-Sp x SSA2- Ng and SSA2-Sp x SSA3-Ng hybrids in F_2 generation.

The production of both female and male progeny in the F_1 generation among SSA2-Sp and SSA2-Ug and other genetic groups, SSA2- and SSA3-Ng indicated that they are capable of interbreeding and producing offspring. The close relationship observed between SSA2-Sp and SSA from sub-Saharan Africa in this study is not surprising because their genetic relationship was previously suggested based on the species delimitation of 3.5 % mtCOI nucleotide divergence as established by Dinsdale *et al*. (2010). Furthermore, all six type of self-crossing of the F_1 hybrids of SSA2-Sp and SSA2-Ug, SSA2-Ng and SSA3-Ng produced both female and male progeny in F_2 generation in both directions of the crosses regardless of which population was the source of maternal and paternal background in the F_1 generation, hence supporting that they belong to the same species of *B. tabaci*. This is not the first report for allopatric populations to produce female progeny, for instance, Vyskočilová *et al*. (2018) reported the gene flow between Q1 from Spain and Q2 from Israel. However, reproductive isolation has been reported between allopatric populations colonizing cassava in India and sub-Saharan Africa (Maruthi *et al*., 2004). This study, therefore, provided further supportive evidence by laboratory crossing experiments in which the progeny counts, and sex ratios could be observed and compared, and viability could be tested.

Mating compatibility within whitefly populations of SSA genetic group of *B. tabaci* that colonize cassava in sub-Saharan Africa has also been generated in the laboratory (Maruthi *et al*., 2001; Mugerwa, 2018). Moreover, gene flow between whitefly populations from cassava and

whiteflies from okra plants has been reported in Ghana (Omondi *et al.*, 2005). Besides this study of the SSA high-level genetic group, mating studies from other genetic groups performed under laboratory conditions demonstrated reproductive incompatibility among whitefly populations that occur in the same or distinct geographical region (Qin *et al.*, 2016; Vyskočilová *et al*., 2018).

Due to the presence of unique fragments of 321 bp and 503 bp for Spanish and African populations, respectively, they could be distinguished from one another and from the hybrid females using PCR-RFLP of the nuclear marker *prpf8* gene. Analysis of amplicons of hybrid females revealed heterozygosity by producing mixed restriction patterns in F_1 hybrid females of both populations, compared to the homozygous ones from control individuals, confirming their hybrid status. Another nuclear marker of a different gene, *Glutamate Carrier 1* **(***GC1*) has also been used to verify the presence of hybrids of Spain Q1 and Israel Q2 by using PCR-RFLP, and sequencing the amplicons of nuclear gene of hybrid females (Vyskočilová *et al*., 2018).

When using the biological species concept definition, these experiments reveal that the SSA2-Sp population belongs to the same species as the SSA2-Ug, SSA2- and SSA3-Ng whiteflies colonizing cassava in sub-Saharan Africa. These populations are differentiated in allopatry but retained the ability to interbreed in the laboratory conditions and perhaps it may happen in the field under natural conditions. The existence of hybrids between cassava whiteflies SSA2 and SSA4 collected from the field has been detected by Chen *et al*. (2019) and Wosula *et al.* (2017). In addition, Brown *et al*. (2004) reported the hybrids between the SSA1 (local) and SSA2 (invader) *B. tabaci* associated with severe mosaic disease of cassava in East Africa. Moreover, hybrids of different genetic groups of *B. tabaci* species complex from field collected samples have also been reported. For instance, based on the population genetic studies, Terraz *et al.* (2014) detected hybrids between Q1 and Q2, both belonging to MED species. Similar result was obtained between MED and MEAM1 (Tahiri et *al*., 2013; McKenzie *et al*., 2012). Furthermore, hybrids under natural conditions have also been detected for MEAM1 species and indigenous species from Australia (Gunning *et al*., 1997) and the Indian Ocean (Dellate *et al*., 2006).

The mating incompatibility between SSA2-Sp and SSA1-SG1-Ug and SSA1-SG3-Tz observed in the laboratory crossing study was due to failure of production of female progeny and low proportions of female progeny, respectively. The incompatibility between SSA2-Sp and SSA1-SG3-Tz was further studied when a mixture of the two populations were allowed to mate and very few hybrid females were detected. The low proportions of female progeny produced among SSA2-Sp and SSA1-SG3-Tz could be due to the two populations being kept in captivity under favourable laboratory conditions (Sun *et al*., 2010). The incompatibility observed in this study could mean that SSA2-Sp is a distinct species from SSA1-SG1-Ug and SSA1-SG3-Tz populations of *B. tabaci*.

Reproductive incompatibility has also been observed within the SSA genetic group among allopatric populations, SSA1-SG1-Ug and SSA1-SG3-Tz, SSA2-Ug and SSA1-SG3-Tz of *B. tabaci,* and sympatric cassava whiteflies, SSA2-Ug and SSA1-SG1-Ug and SSA1-SG2-Ug (Mugerwa, 2018). Additionally, other studies among different genetic groups within *B. tabaci* species complex showed complete reproductive isolation. For instance, Elbaz *et al*. (2010) and Sun *et al*. (2011) reported reproductive isolation between B (MEAM1) and Q (MED) biotypes of *B. tabaci*. Moreover, reciprocal crossing studies among sub-Saharan "MED-ASL" populations and MED (Q1 from Spain and Q2 from Israel) showed mating incompatibility between them (Vyskočilová *et al.*, 2018).

In conclusion, based on the available genetic data coupled with mating data from reciprocal crossing experiments and the biological species concept, SSA2-Sp belongs to the same species as SSA2-Ug, SSA2-Ng and SSA3-Ng populations that colonize cassava in sub-Saharan Africa. According to this study, it shows that these two populations are isolated geographically but retained the ability to interbreed and produce fertile offspring when occurring together. Furthermore, the results also demonstrated that SSA2-Sp is a distinct species from SSA1-SG1-Ug and SSA1-SG3-Tz populations colonizing cassava. The results from this study support only partially the species threshold of 3.5% mtCOI sequence divergence, because SSA2-Sp and SSA3-Ng, that diverged from each other by 5.78% (Table 7), were reproductively compatible.

Although SSA2-Sp mated with SSA2-Ug, SSA2-Ng and SSA3-Ng colonising cassava in sub-Saharan Africa and produce fertile progeny, further research is required in order to understand how these population differ biologically, eg whether SSA2-Sp can develop on cassava crops and if it is able to transmit the major cassava viral diseases, and other begomoviruses of economic importance. It would be also important to study the host range of SSA2-Sp in other African crops and non-cultivated plants that could act as host reservoirs for overwintering of the whiteflies.

GENERAL DISCUSSION, CONCLUDING REMARKS AND PROSPECTS

In the last two decades, descriptions of begomovirus infections in crops have increased greatly in tropical and sub-tropical regions of the world (Navas-Castillo *et al*., 2011). This has been caused in part by the increase in the populations of the vectors, the whiteflies of the *B. tabaci* complex. The efficient dissemination and high polyphagy of *B. tabaci* have enabled the transmission of begomoviruses from wild to cultivated plants, thus driving the emergence of serious diseases (Fiallo-Olivé *et al*., 2019b; Gilbertson *et al*., 2015). The significant role that weeds and other wild plants have played as sources of novel begomoviruses and reservoir hosts of begomoviruses for important crops worldwide has been reported (García-Arenal and Zerbini 2019; Navas-Castillo *et al*., 2011).

This study aimed at: i) characterizing new begomoviruses infecting non-cultivated plants in East Africa, specifically Uganda, and ii) identify the boundaries between species of the *B. tabaci* complex belonging to the Sub-Saharan high-level genetic group.

i. Frequent occurrence of new begomoviruses infecting wild plants in Uganda

With the aid of rolling circle amplification (RCA) using bacteriophage φ29 DNA polymerase, restriction fragment length polymorphism (RFLP) and sequencing, ten leaf samples collected from symptomatic non-cultivated plants in Uganda were found to be infected by begomoviruses, and two of them were also infected by a betasatellite. All characterized begomoviruses and the betasatellite represent new species. Many other studies have used RCA in combination with RFLP successfully, in diagnosing and characterizing viruses with single-stranded circular DNA genomes including begomoviruses and DNA satellites (Jeske, 2018).

The symptomatic samples were obtained as part of a survey for begomoviruses in Uganda, conducted along main roads where wild-plant specimens showing symptoms suspicious of begomovirus infection were spotted and sampled. In addition, some of the plant samples were collected in backyard gardens.

The viruses and the DNA satellite identified in this study from non-cultivated plants belong to the genus *Begomovirus* (family *Geminiviridae*) and genus *Betasatellite* (family *Tolecusatellitidae*), respectively. In accordance with begomovirus species demarcation threshold of 91% for DNA-A of bipartite begomoviruses or complete genome of monopartite begomoviruses (Brown *et al*., 2015), five distinct begomovirus species were discovered (**Figure 21**). These begomoviruses include: vernonia crinkle virus (VeCrV) and its associated vernonia crinkle betasatellite (VeCrB) isolated from *Vernonia amygdalina* (Compositae); desmodium mottle virus (DesMoV) identified from the leguminous weed *Desmodium* sp. (Fabaceae); and ocimum yellow vein virus (OcYVV), ocimum

mosaic virus (OcMV) and ocimum golden mosaic virus (OcGMV) infecting *Ocimum gratissimum* (Lamiaceae). With the exception of DesMoV, the new begomoviruses are grouped in different clusters in the main phylogenetic begomovirus group, the Old World begomoviruses. DesMoV clustered within the legumovirus phylogenetic group, being the first legumovirus described for East Africa (**Figure 22**). Phylogenetic analyses also showed that some of the new begomoviruses are distantly related to all previously characterized begomoviruses.

In regard to VeCrB, according to the accepted betasatellite species demarcation threshold of 91% (https://talk.ictvonline.org/files/proposals/taxanomy_proposals_plant1/m/plant02/6357), it represents a new species. This is the first fully molecular characterization of begomoviruses infecting wild plants in Uganda and the first identification of a betasatellite in the same country.

Collectively, these findings indicate that a high species diversity of begomoviruses infecting wild plants seems to exist in Uganda and probably in the rest of East Africa.

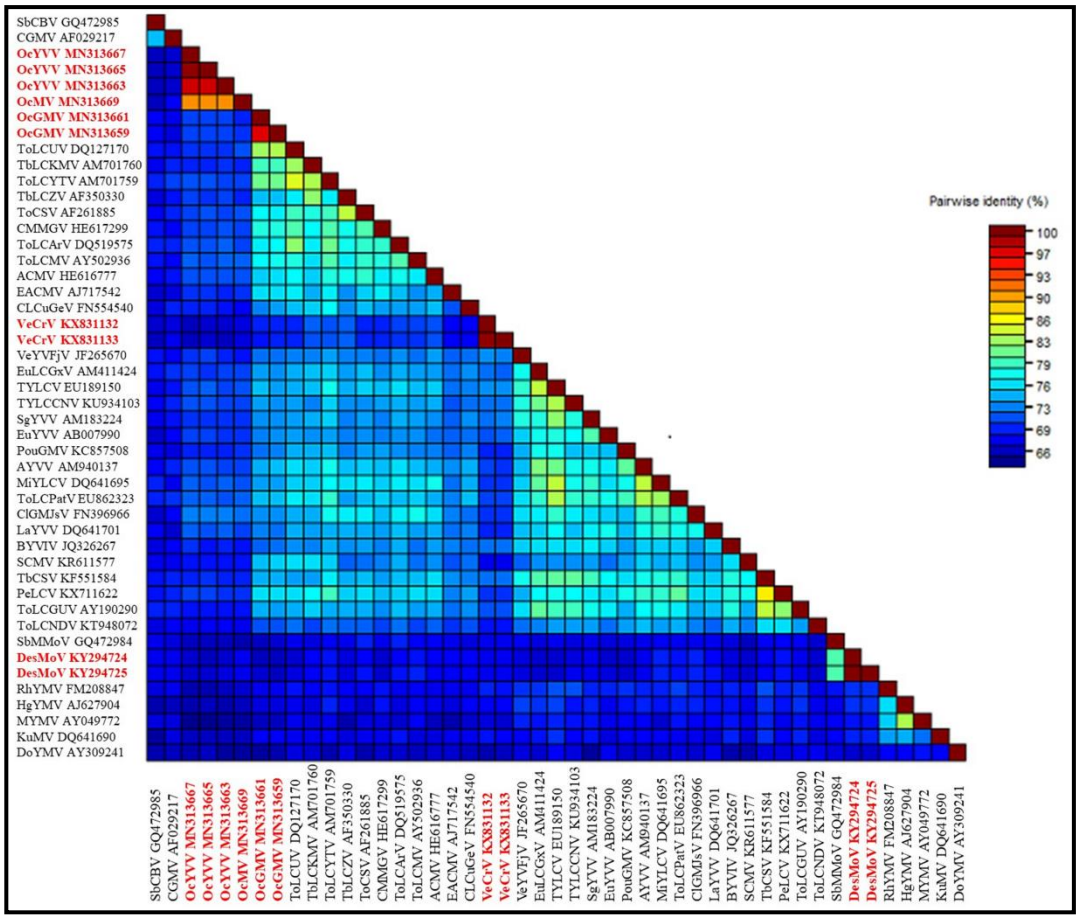

Figure 21. Colour-coded matrix of pairwise similarity scores of begomoviruses identified in this study (in bold red colour) and other Old World begomoviruses retrieved from GenBank.

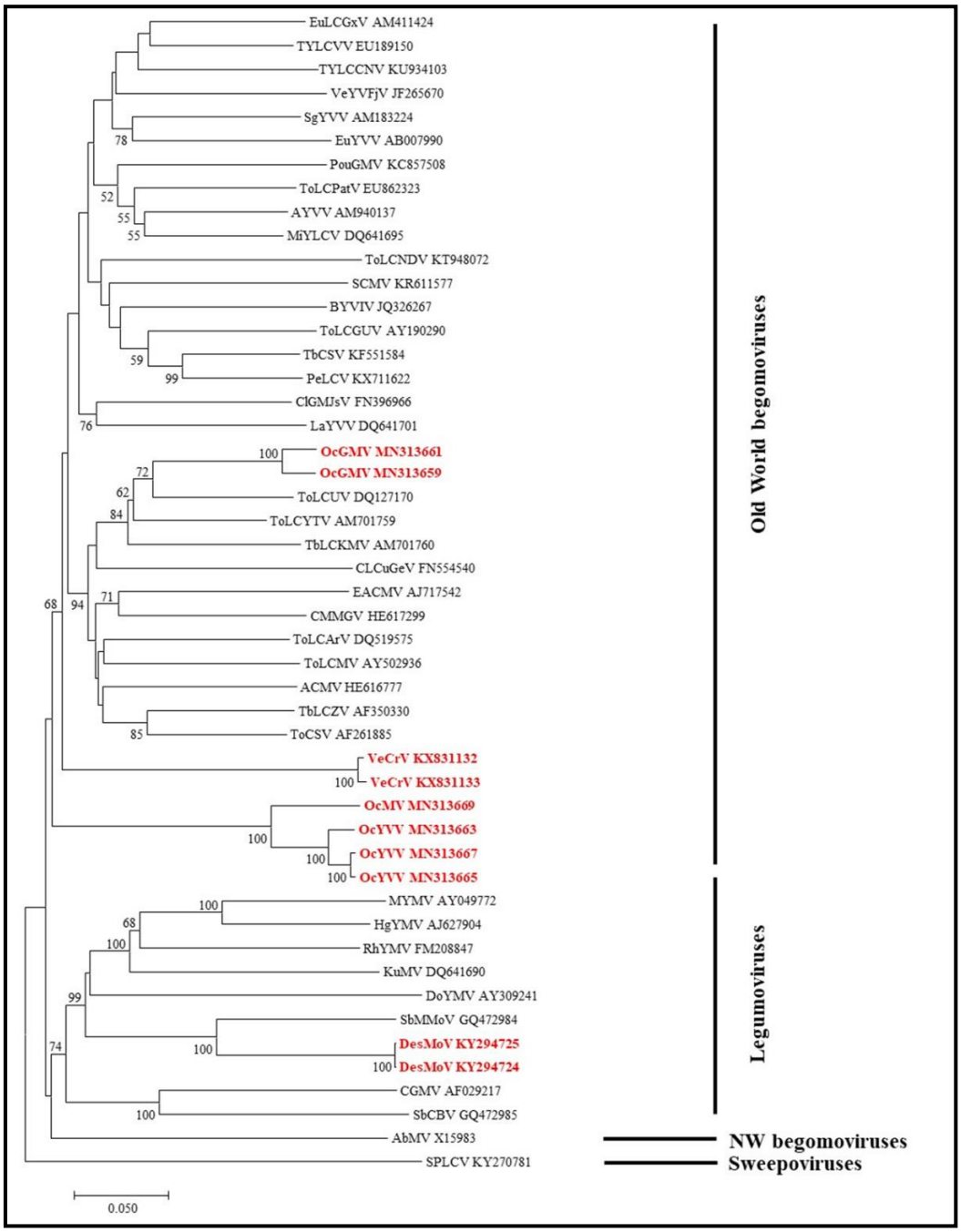

Figure 22. Phylogenetic relationships between begomoviruses identified in this study (in bold in red colour) and other Old Word begomoviruses. Complete genomes (for monopartite begomoviruses) or DNA-A components (for bipartite begomoviruses) were used to construct the phylogenetic tree by the Neighbour-joining method using the MEGA 7 program. The tree was rooted by using Abutilon mosaic virus (a New World begomovirus) as an outgroup. Only bootstrap values >50% (1000 replicates) are shown. The bar below the tree indicates nucleotide substitutions per site.

Begomoviruses infecting non-cultivated plants have barely received attention from researchers in East Africa, it should be taken in account that they are a potential threat for important crops.

Related to this, it should be stressed that, based on sequence pairwise identity analysis, some of the new begomoviruses are related to begomoviruses infecting crops in Africa. For instance, DesMoV and OcGMV are related to soybean mild mottle virus infecting soybean in Nigeria and tomato leaf curl Uganda virus infecting tomato in Uganda, respectively (Alabi *et al*., 2010; Shih *et al*., 2006a). Also, the begomovirus West Africa asystasia virus 1, initially described infecting the wild plant *Asystasia* sp., has been also reported to infect cassava plants in the field in Cameroon (Leke *et al*., 2016a). Furthermore, in other parts of the world, weed-infecting begomoviruses have been shown to infect crops naturally. For example, sida micrantha mosaic virus (SimMV) and macroptilium yellow spot virus (MaYSV) isolated from common bean plants in Brazil (Fernandes-Acioli *et al*., 2011; Lima *et al*., 2013), ageratum enation virus (AEV) isolated from tomato plants in India (Swarnalatha *et al*., 2013) and Euphorbia mosaic virus isolated from commercial tobacco crops in Cuba (Fiallo-Olivé *et al*., 2010b) and sweet pepper (*Capsicum annuum*) in Mexico (GenBank DQ520942). Moreover, some studies demonstrated the transmission of begomoviruses by the insect vector, *B. tabaci* from non-cultivated hosts to cultivated species (Barreto *et al.*, 2013; Castillo-Urquiza *et al*., 2007; Rocha *et al*., 2013; Silva *et al.*, 2010). Hence, these results reinforce the importance of considering non-cultivated plants when developing sustainable control strategies for viral infection diseases within a given region.

Elsewhere in the world, different begomoviruses have been isolated from the plant genera analyzed in this work. *Desmodium* sp. has been shown to be infected by Rhynchosia yellow mosaic Yucatan virus and Desmodium leaf distortion virus in Mexico, and Macroptilium yellow spot virus in Brazil (Fontenele *et al.*, 2016; Hernández-Zepeda *et al*., 2009, 2010). *Vernonia* sp., for instance, *Vernonia cinerea* has been found to be infected by vernonia yellow vein Fujian virus and associated cognate alpha- and beta-satellites in China (Zulfiqar *et al*., 2012), and vernonia yellow vein virus and its associated betasatellite in India (Packialakshmi *et al*., 2010). *Ocimum sanctum* has been found to be infected by tomato leaf curl virus and its associated betasatellite in India (Gaur *et al*., 2012; Nehra *et al*., 2018). *Ocimum basilicum* cultivated in Oman has been reported to host the begomoviruses chilli leaf curl virus and tomato yellow leaf curl virus, as well as the betasatellite tomato leaf curl betasatellite (Ammara *et al*., 2015).

Furthermore, the detection of three distinct begomovirus species from *Ocimum gratissimum* suggest that this plant species may act as 'mixing vessel' in which recombinant viruses may arise. In west Africa, other wild plant, *Asystasia gangetica* has been reported to have mixed infections of distinct begomovirus species (asystasia begomovirus 1 and asystasia begomovirus 2) (Wyant *et al*.,

2015). Mixed infections by different begomoviruses are common in non-cultivated plants (Alabi *et al*., 2008; Ferro *et al*., 2017; Mohammed *et al*., 2018; Rodríguez-Negrete *et al*., 2019), facilitating recombination events even among distantly related begomoviruses which may contribute to the frequent emergence of new species or strains.

Betasatellites have been found in association with begomoviruses from both cultivated and non-cultivated plants in different parts of the world (Idris *et al*., 2012; Khatri *et al*., 2014; Kumar *et al*., 2013; Nawaz-ul-Rehman and Fauquet, 2009; Saunders *et al*., 2000; Yasmin *et al*., 2017; Zubair *et al*., 2017). VeCrB is the fourth betasatellite to be detected in wild plants in Africa. Previously, three betasatellites had been identified: cotton leaf curl Gezira betasatellite (CLCuGeB) (Idris *et al*., 2005), ageratum leaf curl Cameroon betasatellite (ALCCMB) (Leke *et al*., 2012) and tomato leaf curl Togo betasatellite (ToLCTGB) (Kon and Gilbertson, 2012). Detection of a fourth betasatellite indicates a growing complexity of betasatellites in Africa.

In the present study, recombination events have been detected in both begomoviruses and betasatellite genomes. Recombination events detected from VeCrV and its associated VeCrB involve begomoviruses that infect pepper, tomato, and *Siegesbeckia glabrescens* and *Vernonia cinerea* respectively, in Asia, and tomato, cassava, tobacco and cotton in Africa (Leke *et al.*, 2016b; Packialakshmi *et al*., 2010; Wu *et al.*, 2007). Begomoviruses involved in recombination events of DNA-A of OcYVV, OcMV and OcGMV infect cassava, tomato and soybean in Asia, tomato in the Indian Ocean Islands, and cassava, tomato, and *Desmodium* sp. in Africa (Lefeuvre *et al*., 2007; Mollel *et al*., 2017b; Shih *et al*., 2006b). Natural recombinant begomoviruses have been directly involved in the emergence of new diseases and epidemics on crops in many countries (García-Andrés *et al*., 2006, 2007a, b; Lefeuvre *et al*., 2010; Monci *et al*., 2002). The results obtained in this work support the idea from previous studies that non-cultivated crops are melting pots for recombination and emerging of novel begomoviral and DNA satellite genomes.

ii. Species boundaries in the *Bemisia tabaci* complex: the case of the Sub-Saharan African group

An objective of this thesis was to determine the species boundaries in the Sub-Saharan African high-level group of *Bemisia tabaci* through reproductive compatibility studies. The work was focused on a population of Sub-Saharan African 2 from Spain (SSA2-Sp) and populations colonizing cassava in Sub-Saharan Africa: SSA1-SG1-Ug, SSA1-SG3-Tz, SSA2-Ng, SSA2-Ug and SSA3-Ng.

Whiteflies, including *B. tabaci*, are haplodiploid insects, with haploid male progeny being produced from unfertilized eggs and diploid female progeny from fertilized eggs (Byrne and Bellows,

1991). To ascertain whether two whitefly populations behave like distinct biological species, reproductive compatibility studies need to be carried out that should reveal the proportion of female progeny obtained from intra- and inter-population reciprocal crosses (Xu *et al.*, 2010). This study reports the production of female progeny among SSA2-Sp and SSA2-Ug, SSA2-Ng and SSA3-Ng; few female progeny between SSA2-Sp and SSA1-SG3-Tz and no female progeny among SSA2-Sp and SSA1-SG1-Ug in both direction of the crosses.

The production of only-male progeny in reciprocal crosses carried out between SSA1-SG1-Ug and SSA2-Sp populations demonstrated mating incompatibility between these whiteflies. The pairing of male and female parents, which is a mating behaviour characteristic in *B. tabaci* (Li *et al.*, 1989; Luan *et al.*, 2008), was observed in all of these inter-population crosses, however, no females progeny were generated. This provides strong evidence that the SSA1-SG1-Ug and SSA2-Sp populations studied belonged to different species in the *B. tabaci* complex. Another study conducted by Mugerwa (2018) also showed that SSA1-SG1-Ug do not interbreed with SSA2-Ug. Additionally, using a STRUCTURE analysis (Pritchard *et al.*, 2000), no gene flow was seen when field populations of SSA1 from East and Central Africa and SSA2 from Kenya and Cameroon were analyzed (Chen *et al*., 2019; Wosula *et al*., 2017).

The few female progeny reported here between SSA1-SG3-Tz and SS2-Sp populations could probably be due to forced copulation between these two populations when confined under laboratory conditions (Sun *et al*., 2011), hence demonstrating a high level of reproductive incompatibility. Few female progeny, less than 30%, in inter-population crosses was also observed between SSA1-SG3 (TzCas-Mtw) and SSA2 (UgCas-Nam), and MEAM1 and other species (Costa *et al.*, 1993a; De Barro and Hart, 2000; Maruthi *et al.*, 2004) that was attributed to the use of already mated females (Luan *et al*., 2008). However, the experimental setup used in this study allows only virgin adult whiteflies to participate in the reciprocal crossing experiments, thus avoiding false positive results. This method has also been used previously for crosses among populations of different genetic groups of *B. tabaci* (Mugerwa, 2018; Qin *et al*., 2016; Vyskočilová *et al*., 2018). The female progeny produced in this cross was confirmed molecularly using a nuclear marker that demonstrated the true nature of the hybrids and verified that there was no contamination during the experiment.

The production of female progeny from two generations, F_1 and F_2, observed between SSA2-Sp and SSA2-Ug, SSA2-Ng and SSA3-Ng was a good indication of successful mating between them, and hence showed that they belong to the same biological species. Hybrids among populations of SSA1-SG1-Ug and SSA1-SG2-Ug of *B. tabaci* produced under laboratory conditions have also been reported (Mugerwa, 2018). Under field conditions, the presence of gene flow has been detected between SSA2 and SSA4 (Chen *et al*., 2019; Wosula *et al*., 2017).

In accordance with a species delimitation of 3.5% partial mtCOI sequence divergence among different *B. tabaci*, Dinsdale *et al.* (2010) suggested that populations should be considered as separate species based on this threshold. Based on the crossing data obtained from different populations used in this study, the 3.5% nt at mtCOI level was a good indicator of species delimitation with the exception of a cross between SSA2-Sp and SSA3-Ng that diverged from one another by 5.78%, yet were reproductively compatible. This is not the first report of compatibility observed between populations that diverged by >3.5% nt at mtCOI level. Qin *et al*. (2016) reported nearly complete mating compatibility in one direction and partial compatibility in other direction between Asia II-3 and Asia II-9, that diverged from each other by 4.57%. In addition, MED and MED-ASL are reported to be different species (Vyskočilová *et al*., 2018), yet they were classified as one species based on the 3.5% at mtCOI level. Other studies also supported the 3.5% threshold for biological species, involving for example Q1 and Q2 belonging to MED species (Vyskočilová *et al*., 2018); MEAM1 and MED (Wang *et al*., 2011); SSA1-SG1-Ug and SSA1-SG2-Ug (Mugerwa, 2018); and Asia I and Asia II -7, Asia II-6 and Asia II -1 (Qin *et al*., 2016). The results from this study add to the growing number of examples that show that the 3.5% mtCOI sequence divergence may still be used as a rule of thumb for species boundaries, but it needs to be confirmed by crossing experiments when considering biological species.

iii. Concluding remarks and prospects

The findings of this study indicated that non-cultivated plants are common hosts of novel begomoviruses in Uganda, in some cases associated to DNA satellites. In addition, as it has been shown before, some wild plant species are vulnerable to infections by different begomoviruses that may later facilitate the emergence of novel species or strains by recombination. This may result in more virulent and destructive variants than their parents. Furthermore, the identification of the first betasatellite in East Africa adds additional complexity to the pathosystem. The results obtained with samples from Uganda suggest that very probably there are many more undiscovered begomoviruses and DNA satellites associated infecting weeds and other wild plants surrounding farmers' fields, thus menacing their crops.

It was not demonstrated in this study whether the begomoviruses characterized could infect and cause diseases in cultivated plants, with the exception of some African basil plants. However, there are examples in the literature that viruses in wild plants do spread to infect crops. For instance, cassava, a crop that originated in Latin America, became infected with several begomoviruses upon its introduction to Africa strongly suggesting that these viruses were transmitted from plants native to the region. The potential threat that begomoviruses infecting weeds and other wild plants could be

worse to the cultivated crops compared to their natural hosts, if the viruses become more virulent and are readily transmissible by the whitefly vector, *B. tabaci*.

Further research on diversity of begomoviruses and DNA satellites infecting non-cultivated plants in East Africa is important in order to understand their active role in the epidemiology of begomoviruses in crop plants, either by acting as primary inoculum sources or as a continuous source of novel viruses, which could disrupt management strategies based on the deployment of resistant varieties. Furthermore, there is a need to carry out transmission studies to find out if the detected begomoviruses can be transmitted to important crops in the continent.

According to the mating data obtained from the reciprocal crossing experiments of the current study together with available genetic data and the biological species concept, Spanish populations of SSA2 belong to the same species with sub-Saharan African populations SSA2-Ug, SSA2-Ng and SSA3-Ng colonising cassava plants. These allopatric populations have retained the ability to interbreed and produce fertile offspring in experimental conditions. Also, the results have shown that SSA2-Sp is a distinct species from cassava whiteflies SSA1-SG1-Ug and SSA1-SG3-Tz.

Although SSA2-Sp interbred with SSA2-Ug, SSA2-Ng and SSA3-Ng and produced viable offspring, further research is required in order to understand whether it could colonise cassava plants and it is able to transmit the major cassava viral diseases, and other begomoviruses of economic importance. It would also be interesting to study the host range of SSA2-Sp and other SSA populations to determine the importance that it could have to a range of important crops.

The complexity of emerging begomovirus problems and their associated DNA satellites together with the explosion of whitefly *B. taba*ci populations belonging to species of the complex, therefore, require a joint effort by virologists as well as entomologist and plant breeders to develop sustainable control measures for minimizing the damage caused by begomovirus diseases and their vector, *B. tabaci* to allow sustainable crop production.

CONCLUSIONS / CONCLUSIONES

CONCLUSIONS

1. Vernonia crinkle virus and vernonia crinkle betasatellite constitute the first monopartite begomovirus-betasatellite complex identified from Uganda.
2. Desmodium mottle virus, characterized in this work, is the first legumovirus reported from East Africa.
3. African basil (*Ocimum gratissimum*) was infected by three new divergent begomoviruses in Uganda. These are the first viruses identified for this species and the first in *Ocimum* spp. in Africa.
4. Begomoviruses were frequently found infecting wild plants in Uganda.
5. Sub-Saharan Africa 2 whitefly populations infecting *Ipomoea indica* and other wild plants in Spain interbreed with Sub-Saharan Africa 2 and Sub-Saharan Africa 3 populations from Africa, thus showing that they belong to the same species of the *Bemisia tabaci* complex.
6. Sub-Saharan Africa 2 whitely populations from Spain are isolated reproductively from Sub-Saharan Africa 1 subgroups 1 and 3 from Africa, thus belonging to different species of the *Bemisia tabaci* complex.
7. Reciprocal crossing experiments between *Bemisia tabaci* populations provide biological evidence confirming, or challenging in some cases, the definition of species based on partial sequences of the mitochondrial CO1 gene.

CONCLUSIONES

1. Vernonia crinkle virus y vernonia crinkle betasatellite constituyen el primer complejo begomovirus-betasatellite identificado en Uganda.
2. Desmodium mottle virus, caracterizado en este trabajo, es el primer legumovirus aislado en el este de África.
3. Se han caracterizado tres nuevos virus divergentes infectando la albahaca de clavo o albahaca africana (*Ocimum gratissimum*) en Uganda, que son los primeros virus identificados en esta especie y en *Ocimum* spp. en África.
4. Los resultados obtenidos en esta tesis demuestran que los begomovirus infectan frecuentemente plantas silvestres en Uganda.
5. Las poblaciones de mosca blanca Sub-Saharan Africa 2 que infectan *Ipomoea indica* y otras plantas silvestres en España pueden reproducirse con las poblaciones de Sub-Saharan Africa 2 y Sub-Saharan Africa 3 de África, lo que demuestra que pertenecen a la misma especie del complejo *Bemisia tabaci*.
6. Las poblaciones de Sub-Saharan Africa 2 de España están sin embargo aisladas reproductivamente de los subgrupos 1 y 3 de Sub-Saharan Africa 1, por lo que pertenecen a diferentes especies del complejo *Bemisia tabaci*.
7. Los experimentos de cruzamientos recíprocos entre poblaciones de *Bemisia tabaci* proporcionan evidencia biológica que confirma, o desafía en algunos casos, la definición de especies basada en secuencias parciales del gen mitocondrial CO1.

REFERENCES

REFERENCES

Alabi OJ, Kumar PL, Mgbechi-Ezeri JU, Rayapati AN (2010) Two new 'legumoviruses' (genus *Begomovirus*) naturally infecting soybean in Nigeria. Arch Virol 155:643-656

Alabi OJ, Ogbe FO, Bandyopadhyay R, Kumar PL, Dixon AGO, Hughes JA, Naidu RA (2008) Alternate hosts of African cassava mosaic virus and East African cassava mosaic Cameroon virus in Nigeria. Arch Virol 153:1743-1747

Alemandri V, Vaghi Medina CG, Dumón AD, Argüello Caro EB, Mattio MF, García Medina S, López Lambertini PM, Truol G (2015) Three members of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex occur sympatrically in Argentine horticultural crops. J Econ Entomol 108:405-413

Ammara UE, Al-Ansari M, Al-Shihi A, Amin I, Mansoor S, Al-Maskri AY, Al-Sadi A (2015) Association of three begomoviruses and a betasatellite with leaf curl disease of basil in Oman. Can J Plant Pathol 37:506-513

Andou T, Yamaguchi A, Kawano S, Kawabe K, Ueda S, Onuki M (2010) Ageratum yellow vein virus isolated from tomato plants with leaf curl on Ishigaki Island, Okinawa, Japan. J. Gen Plant Pathol 76:287-291

Antignus Y, Mor N, Ben-Joseph R, Lapidot M, Cohen S (1996) UV-absorbing plastic sheets protect crops from insect pests and from virus diseases vectored by insects. Envir Entomol 25:919-924

Argüello-Astorga GR, Guevara-González RG, Herrera-Estrella LR, Rivera-Bustamante R (1994) Geminivurs replication origins have a group specific organization of iterative element: a model for replication. Virology 203:90-100

Argüello-Astorga GR, Ruiz-Medrano (2001) An iteron-related domain is associated to motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by comparative approach. Arch Virol 146:1465-1485

Avidov, Z (1956) Bionomics of the tobacco whitefly (*Bemisia tabaci* Gennad.) in Israel. Ktavim, 7:25-41

Avise JC (2000) Phylogeography: The history and formation of Species. Harvard University Press, Cambridge, Massachusetts

Banks GK, Bedford ID, Beitia FJ, Rodriguez-Cerezo E, Markham PG (1999) A novel geminivirus of *Ipomoea indica* (Convolvulaceae) from southern Spain. Plant Dis 83:486-486

Banks GK, Markham PG (2000) *Bemisia tabaci*-How many biotypes are there? In: Gerling D, Jones W, eds., *Bemisia*. New letters 13:2-3

REFERENCES

Barbosa JC, Barreto SS, Inoue-Nagata AK, Reis MS, Firmino AC, Bergamin Filho A, Rezende JAM (2009) Natural infection of *Nicandra physaloides* by tomato severe rugose virus in Brazil. J Gen Plant Pathol 75:440-443

Barnabas AD, Radhakrishnan GK, Usha R (2010) Characterization of a begomovirus causing horsegram yellow mosaic disease in India. Eur J Plant Pathol 127:41-51

Barreto SS, Hallwass M, Aquino OM, Inoue-Nagata AK (2013) A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. Phytopathology 103:436-444

Bedford ID, Briddon RW, Brown JK, Rossel RC, Markham PG (1994) Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. Ann Appl Biol 125:311-325

Bedford ID, Kelly A, Banks GK, Briddon RW, Cenis JL, Markham, PG (1998) *Solanum nigrum*: an indigenous weed reservoir for tomato yellow leaf curl geminivirus in southern Spain. Eur J Plant Pathol 104:221-222

Bellows TSJr, Arakawa K (1988) Dynamics of preimaginal populations of *Bemisia tabaci* (Homoptera: Aleyrodidae) and *Eretmocerus* sp. (Hymenoptera: Aphelinidae) in southern California cotton. Environ Entomol 17:483-487

Berry SD, Fondong VN, Rey C, Rogan D, Fauquet CM, Brown JK (2004) Molecular evidence for five distinct *Bemisia tabaci* (Homoptera: Aleyrodidae) geographic haplotypes associated with cassava plants in sub-Saharan Africa. Systematics 97:852-859

Bird J (1957) A whitefly transmitted mosaic of *Jatropha gossypiifolia*. Technical Paper. Agricultural Experiment Station, Puerto Rico 22:1-35

Bird J, Maramorosch K (1978) Viruses and virus diseases associated with whiteflies. Adv Virus Res 22:55-110

Bird J, Sánchez J, Rodríguez RL, Juliá FJ (1975) Rugaceous (whitefly-transmitted) viruses in Puerto Rico. In: Tropical diseases of legumes. Bird J, Maramorosch K, eds. Academic Press, New York, pp 3-25

Bisaro DM (1996) Geminivirus DNA replication. In: DNA replication in Eukaryotic cells. Edited by De Pamphilis ML, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, pp 833-854

Boykin LM, Armstrong KF, Kubatko L, De Barro P (2012) Species delimitation and global biosecurity. Evol Bioinform 8:1-37

Boykin LM, De Barro PJ (2014) A practical guide to identifying members of the *Bemisia tabaci* species complex: and other morphologically identical species. Front Ecol Evol 2:1-5

Boykin LM, Shatters RGJr, Rosell RC, McKenzie CL, AnnBagnall R, De Barro P, Frohlich DR (2007) Global relationships of *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. Mol Phylogenet Evol 44:1306-1319

Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M (2008) Recommendations for the classification and nomenclature of the DNA-βsatellites of begomoviruses. Arch Virol 153:763-781

Briddon RW, Bull SE, Amin I, Idris AM, Mansoor S, Bedford IA, Dhawan P, Rishi N, Siwatch SS, Abdel-Salam AM (2003) Diversity of DNA β, a satellite molecule associated with some monopartite begomoviruses. Virology 312:106-121

Briddon RW, Bull SE, Amin I, Mansoor S, Bedford ID, Rishi N, Siwatch SS, Zafar Y, Abdel-Salam AM, Markham PG (2004) Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus-DNA betasatellite complexes. Virology 324:462-474

Briddon RW, Mansoor S, Bedford ID, Pinner MS, Saunders K, Stanley J, Zafar Y, Malik KA, Markham PG (2001) Identification of DNA components required for induction of cotton leaf curl disease. Virology 285:234-243

Briddon RW, Markham PG (2000) Cotton leaf curl virus disease. Virus Res 71:151-159

Briddon RW, Martin DP, Roumagnac P, Navas-Castillo J, Fiallo-Olivé E, Moriones E, Lett JM, Zerbini FM, Varsani A (2018) "*Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites". Arch Virol 163:2587-2600

Briddon RW, Patil BL, Bagewadi B, Nawaz-ul-Rehman MS, Fauquet CM (2010) Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evol Biol 10:97

Briddon RW, Stanley J (2006) Subviral agents associated with plant single-stranded DNA viruses. Virology 344:198-210

Brown JK (1994) Current status of *Bemisia tabaci* as a pest and virus vector in world agroecosystems. FAO Plant Protect B 42:3-32

Brown JK (2000) Molecular markers for the identification and global tracking of whitefly vector begomovirus complexes. Virus Res 71:233-260

Brown JK, Coats SA, Bedford ID, Markham PG, Bird J, Frohlich DR (1995b) Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). Biochem Genet 33:205-214

REFERENCES

Brown JK, Czosnek H (2002) Whitefly transmission of plant viruses. In: Advances in Botanical Research. Edited by Plumb RT, Callow JA. Academic Press, New York, pp 65-100

Brown JK, Fauquet CM, Briddon RW, Zerbini M, Moriones E, Navas-Castillo J (2012) Family *Geminiviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (Ed), Virus Taxonomy: Classification and Nomenclature of Viruses-Ninth Report of the International Committee on Taxonomy of Viruses. In: Elsevier Academic Press, USA, pp 351-373

Brown JK, Frohlich DR, Rosell RC (1995a) The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or species complex? Ann Rev Entomol 40:511-534

Brown JK, Rogan D, Legg JP (2004) Hybridization between the 'invader' and 'local' *Bemisia tabaci* haplotypes associated with severe mosaic disease of cassava in East Africa. 2nd European Whitefly Symposium, Croatia

Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JC, Fiallo-Olivé E, Briddon RW, Hernández-Zepeda C, Idris A, Malathi VG, Martin DP, Rivera-Bustamante R, Ueda S, Varsani A (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol 160:1593-1619

Brown JK, Idris AM, Ostrow KM, Goldberg N, French R, Stenger DC (2005) Genetic and phenotypic variation of the Pepper golden mosaic virus complex. Phytopathology 95:1217-1224

Buckner JS, Freeman TP, Ruud RL, Chu CC, Henneberry TJ (2002) Characterization and functions of the whitefly egg pedicel. Arch Insect Biochem 49:22-33

Bull SE, Briddon RW, Sserubombwe WS, Ngugi K, Markham PG, Stanley J (2006) Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. J Gen Virol 87:3053-3065

Burban C, Fishpool LDC, Fauquet C, Fargette D, Thouvenel JC (1992) Host associated biotypes within West African populations of the whitefly *Bemisia tabaci* (Genn), (Hom, Aleyrodidae). J Appl Entomol 113:416-423

Butler GDJR, Hennebery TJ, Clayton TE (1983) *Bemisia tabaci* (Hom., Aleyrodidae): Development, oviposition, and longevity in relation to temperature. Ann Entomol Soc Am 76:310-313

Byrne DN, Bellows Jr TS, Parrella MP (1990) Whiteflies in Agriculture systems. In. Gerling D (ed.) Whiteflies: their bionomics, pest status and management. Intercept Ltd., Andover, Hants, UK, pp 227-261

Byrne DN, Bellows TSJ (1991) Whitefly biology. Annu Rev Entomol 36:431-457

Byrne FJ, Devonshire AL (1996) Biochemical evidence of haplodiploidy in the whitefly *Bemisia tabaci*. Biochem Genet 34:93-107

Campbell BC, Duffus JE, Baumann P (1993) Determining whitefly species. Science 261:1333-1335

Campbell BC, Steffen-Campbell JD, Sorensen JT, Gill RJ (1994) Evolutionary origin of whitefly (Hemiptera: Sternorrhyncha: Aleyrodidae) inferred from 18S rDNA sequences. Insect Mol Biol 3:73-88

Castillo-Urquiza GP, Beserra Junior JEA, Alfenas-Zerbini P, Varsani A, Lima ATM, Barros DR, Zerbini FM (2007) Genetic diversity of begomoviruses infecting tomato in Paty do Alferes, Rio de Janeiro state, Brazil. Virus Rev Res 12:233

Castle SJ, Palumbo JC, Prabhaker N, Horowitz AR, Denholm I (2010) Ecological determinants of *Bemisia tabaci* resistance to insecticides. In: Stansly PA, Naranjo SE (Eds) *Bemisia*: bionomics and management of a global pest. Springer, Dordrecht, the Netherlands

Cervera MT, Cabezas JA, Simon B, Martinez-Zapater JM, Beitia F, Cenis JL (2000) Genetic relationships among biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on AFLP analysis. B Entomol Res 90:391-396

Chant SR (1958) Studies on the transmission of cassava mosaic virus by *Bemisia* spp. (Aleyrodidae). Ann Appl Biol 46:210-215

Chen W, Wosula EN, Hasegawa DK, Casinga C, Shirima RR, Fiaboe KKM, Hanna R, Fosto A, Goergen G, Tamò M, Mahuku G, Murithi HM, Tripathi L, Mware B, Kumar LP, Ntawuruhunga P, Moyo C, Yomeni M, Boahen S, Edet M, Awoyale W, Wintermantel WM, Ling KS, Legg JP, Fei Z (2019) Genome of the African cassava whitefly *Bemisia tabaci* and distribution and genetic diversity of cassava-colonizing whiteflies in Africa. Insect Bioch Mol Biol 110:112-120

Cheng X, Wang X, Wu J, Briddon RW, Zhou X (2011) βC1 encoded by tomato yellow leaf curl China betasatellite forms multimeric complexes in vitro and in vivo. Virology 409:156-162

Choudhury NR, Malik PS, Singh DK, Islam MN, Kaliappan K, Mukherjee SK (2006) The oligomeric Rep protein of mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. Nucleic Acids Res 34:6362-6377

Chowda-Reddy RV, Kirankumar M, Seal SE, Muniyappa V, Valand GB, Govindappa MR, Colvin J, (2012) *Bemisia tabaci* phylogenetic groups in India and the relative transmission efficacy of tomato leaf curl Bangalore virus by an indigenous and an exotic population. J Integr Agr 11:235-248

Clark MA, Baumann L, Munson MA, Baumann P, Campbell BC, Duffus JE, Osborne LS, Moran NA (1992) The eubacterial endosymbionts of whiteflies (Homoptera: Aleyrodoidea) constitute a lineage distinct from the endosymbionts of aphids and mealybugs. Curr Microbiol 25:199-123

Cook SM, Khan ZR, Pickett JA (2007) The use of push-pull strategies in integrated pest management. Annu Rev Entomol 52:375-400

REFERENCES

Cooper I, Jones RAC (2006) Wild plants and viruses: underinvestigated ecosystems. Adv Virus Res 67:1-47

Costa AS, Russell, LM (1975) Failure of *Bemisia tabaci* to breed on cassava plants in Brazil (Homoptera: Aleyrodidae). Ciência e Cultura 27:388-390

Costa HS, Brown J, Sivasupramaniam S, Bird J (1993a) Regional distribution, insecticide resistance and reciprocal crosses between the A-biotype and B-biotype of *Bemisia tabaci*. Insect Sci Appl 14:255-266

Costa HS, Brown JK (1990) Variability in biological characteristics, isozyme patterns and virus transmission among populations of *Bemisia tabaci* Genn. in Arizona. Phytopathology 80:888

Costa HS, Brown JK (1991) Variation in biological characteristics and in esterase patterns among populations of *Bemisia tabaci* (Genn.) and the association of one population with silverleaf symptom development. Entomol Exp Appl 61:211-219

Costa AS, Bennett CW (1950) Whitefly transmitted mosaic of *Euphorbia prunifolia*. Phytopathology 40:266-283

Coudriet DL, Prabhaker N, Kishaba AN, Meyerdirk DE (1985) Variation in developmental rate on different hosts and overwintering of the sweetpotato whitefly, *Bemisia tabaci* (Homeptera: Aleyrodidae). Environ Entomol 14:516-519

Cui X, Li G, Wang D, Hu D, Zhou X (2005) A begomovirus DNAβ-encoded protein binds DNA functions as a suppressor of RNA silencing and targets the cell nucleus. J Virol 79:10764-10775

Davino S, Accotto GP, Masenga V, Torta L, Davino M (2009) Basil (*Ocimum basilicum*), a new host of Pepino mosaic virus. Plant Pathol 58:407

De Barro PJ (2011**)** *Bemisia tabaci*, the capacity to invade. In the whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae) interaction with geminivirus-infected host plants; Thompson WMO Ed.; Springer: Dordrecht, The Netherlands, pp 181-204

De Barro PJ, Coombs MT (2009) Post-release evaluation of *Eretmocerus hayati* Zolnerowich and Rose in Australia. B Entomol Res 99:193-206

De Barro PJ, Driver F (1997) Use of RAPD PCR to distinguish the B biotype from other biotypes of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae). Aust J Entomol 36:149-152

De Barro PJ, Driver F, Trueman JWH, Curran J (2000) Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. Mol Phylogenet Evol 16:29-36

De Barro PJ, Hart PJ (2000) Mating interactions between two biotypes of the whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Australia. B Entomol Res 90:103-112

De Barro PJ, Liu SS, Boykin LM, Dinsdale AB (2011**)** *Bemisia tabaci*: A statement of species status. Annu Rev Entomol 56:1-19

De Barro PJ (2012) The *Bemisia tabaci* species complex: questions to guide future research. J Integr Agric 11:187-196

De Barro P, Ahmed MZ (2011) Genetic networking of the *Bemisia tabaci* cryptic species complex reveals pattern of biological invasions. PLoS One 6:e25579

De Barro PJ, Trueman JWH, Frohlich DR (2005) *Bemisia argentifolii* is a race of *B. tabaci* (Hemiptera: Aleyrodidae): The molecular genetic differentiation of *B. tabaci* populations around the world. B Entomol Res 95:1-11

De Bruyn A, Harimalala M, Zinga M, Mabvakure BM, Hoareau M, Ravigne V, Walters M, Reynaud B, Varsani A, Harkins GW, Martin DP, Lett JM, Lefeuvre P (2016) Divergent evolutionary and epidemiological dynamics of cassava mosaic geminiviruses in Madagascar. BMC Evol Biol 16:182

De la Rúa P, Simón B, Cifuentes D, Martinez-Mora C, Cenis JL (2006) New insights into the mitochondrial phylogeny of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in the Mediterranean Basin. J Zool Syst Evol Res 44:25-33

Debrot E, Centeno F (1986) Ocurrencia del virus del mosaico de las euforbiáceas infectando a *Euphorbia heterophylla* L. en Venezuela. Agron Trop 35:5-12

Delatte H, David P, Granier M, Lett JM, Goldbach R, Peterschmitt M, Reynaud B (2006) Microsatellites reveal extensive geographical, ecological and genetic contacts between invasive and indigenous whitefly biotypes in an insular environment. Genet Res 87:109-124

Demichelis S, Arno B, Bosco D, Marina D, Caciagli P (2005) Characterization of biotype T of *Bemisia tabaci* associated with *Euphorbia characias* in Sicily. Phytoparasitica 33:196-208

Dinsdale A, Cook L, Riginos C, Buckley, YM, De Barro P (2010) Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. Annu Entomol Soc Am 103:196-208

Domingo E, Holland EE (1997) RNA virus mutations and fitness for survival. Annu Rev Microbiol 51:151-178

Dry IB, Krake LR, Rigden JE, Rezaian MA (1997) A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. P Natl Acad Sci USA 94:7088-7093

Duffy S, Holmes E (2009) Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J Gen Virol 90:1539-1547

REFERENCES

Duffy S, Holmes EC (2008) Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus (TYLCV). J Virol 82:957-965

Eagle PA, Orozco BM, Hanley-Bowdoin L (1994) A DNA sequence required for geminivirus replication also mediates transcriptional regulation. Plant Cell 6:1157-1170

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797

Eini O, Dogra S, Selth LA, Dry IB, Randles JW, Rezaian MA (2009) Interaction with a host ubiquitin-conjugating enzyme is required for the pathogenicity of a geminiviral DNA beta satellite. Mol Plant Microbe In 22:737-746

Elbaz M, Lahav N, Morin S (2010) Evidence for pre-zygotic reproductive barrier between the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). B Entomol Res 100:581-591

Esterhuizen LL, Mabasa KG, van Heerden SW, Czosnek H, Brown JK, van Heerden H, Rey MEC (2013) Genetic identification of members of the *Bemisia tabaci* cryptic species complex from South Africa reveals native and introduced haplotypes. J Appl Entomol 137:1-14

Farooq J, Farooq A, Riaz M, Shahid MR, Saeed F, Iqbal MS, Hussain T, Batool A, Mahmood A (2014) Cotton leaf curl virus disease a principle cause of decline in cotton productivity in Pakistan (a mini review). Can J Plant Prot 2:9-16

Fauquet C, Stanley J (2003) Geminivirus classification and nomenclature: progress and problems. Ann Appl Biol 142:165-189

Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X (2008) Geminivirus strain demarcation and nomenclature. Arch Virol 153:783-821

Fernandes-Acioli NAN, Pereira-Carvalho RC, Fontenele RS, Lacorte C, Ribeiro SG, Fonseca MEN, Boiteux LS (2011) First report of sida micrantha mosaic virus in *Phaseolus vulgaris* in Brazil. Plant Dis 9:1196

Ferro CG, Silva JP, Xavier CAD, Godinho MT, Lima ATM, Mar TB, Lau D Zerbini FM (2017) The ever-increasing diversity of begomoviruses infecting non-cultivated hosts: new species from *Sida* spp. and *Leonurus sibiricus*, plus two New World alphasatellites. Ann Appl Biol 170:204-218

Fiallo-Olivé E, Hamed A, Navas-Castillo J, Moriones E (2013) Cotton leaf curl Gezira alphasatellite associated with tomato leaf curl Sudan virus approaches the expected upper size limit in the evolution of alphasatellites. Virus Res 178:506-510

Fiallo-Olivé E, Martínez-Zubiaur Y, Moriones E, Navas-Castillo J (2012) A novel class of DNA satellites associated with New World begomoviruses. Virology 426:1-6

Fiallo-Olivé E, Navas-Castillo J, Moriones E, Martínez-Zubiaur Y (2010a) Two novel begomoviruses belonging to different lineages infecting *Rhynchosia minima*. Arch Virol 155:2053-2058

Fiallo-Olivé E, Tovar R, Navas-Castillo J (2016) Deciphering the biology of deltasatellites from the New World: maintenance by New World begomoviruses and whitefly transmission. New Phytol 212:680-692

Fiallo-Olivé E, Louro D, Trenado HP, Navas-Castillo J (2019a) Recurrent speciation of a tomato yellow leaf curl geminivirus in Portugal by recombination. Sci Rep 9:1332

Fiallo-Olivé E, Chirinos DT, Geraud-Pouey F, Moriones E, Navas-Castillo J (2013) Complete genome sequences of two begomoviruses infecting weeds in Venezuela. Arch Virol 158:277-280

Fiallo-Olivé E, Rivera-Bustamante RF, Martínez-Zubiaur Y (2010b) First report of tobacco as a natural host of Euphorbia mosaic virus in Cuba. Plant Pathol. 59:795

Fiallo-Olivé E, Pan L-L, Liu S-S, Navas-Castillo J (2019b) Transmission of begomoviruses and other whitefly-borne viruses: dependence on the vector species. Phytopathology. doi: 10.1094/PHYTO-07-19-0273-FI.

Firdaus S, Vosman B, Hidayati N, Jaya Supena ED, Visser RG, van Heusden AW (2013) The *Bemisia tabaci* species complex: additions from different parts of the world. Insect Sci 20:723-733

Fishpool LDC, Burban C (1994) *Bemisia tabaci*: the whitefly vector of African cassava mosaic geminivirus. Trop Sci 34:55-72

Fishpool LDC, Fauquet C, Fargette D, Thouvenel JC, Burban C, Colvin J (1995) The phenology of *Bemisia tabaci* (Homoptera: Aleyrodidae) populations on cassava in southern Côte d'Ivoire. B Entomol Res 85:197-207

Fondong VN, Chen K (2011) Genetic variability of East African cassava mosaic Cameroon virus under field and controlled environment conditions. Virology 413:275-282

Fontenele RS, Poppiel R, Matos VORL, Costa F, Faria JC, Ribeiro SG (2016) First Report of macroptilium yellow spot virus in *Desmodium glabrum* in Brazil. Plant Dis 100:657

Fontes EP, Eagle P A, Sipe PS, Luckow VA, Hanley-Bowdoin L (1994) Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J Biol Chem 269:8459-8465

Fourment M, Gibbs AJ, Gibbs MJ (2008) SWeBLAST: a sliding window web-based BLAST tool for recombinant analysis. J Virol Methods 152:98-101

Frohlich DR, Torres-Jerez I, Bedford ID, Markham PG, Brown JK (1999) A phylogeographic analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. Mol Ecol 8:1683-1691

REFERENCES

García-Andrés S, Accotto GP, Navas-Castillo J, Moriones E (2007a) Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. Virology 359:302-312

García-Andrés S, Monci F, Navas-Castillo J, Moriones E (2006) Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: evidence for the presence of a new virus species of recombinant nature. Virology 350:433-442

García-Andrés S, Tomas DM, Sánchez-Campos S, Navas-Castillo J, Moriones E (2009) Resistance-driven selection of begomoviruses associated with the tomato yellow leaf curl disease. Virus Res 146:66-72

García-Andrés S, Tomas DM, Sanchez-Campos S, Navas Castillo J, Moriones E (2007b) Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. Virology 365:210-219

García-Arenal F, Zerbini FM (2019) Life on the edge: geminiviruses at the interface between crops and wild plant hosts. Annu. Rev. Virol. (in press) doi: 10.1146/annurev-virology-092818-015536

Gaur RK (2012) First report of begomovirus infecting two ornamental plants: *Ocimum Sanctum* and *Alternanthera variegata* in India. J Clin Exp Pathol 2:46

Gawel NJ, Bartlett AC (1993) Characterisation of differences between whiteflies using RAPD PCR. Insect Mol Biol 2:33-38

Ge L, Zhang J, Zhou X, Li H (2007) Genetic structure and population variability of tomato yellow leaf curl China virus. J Virol 81:5902-5907

Gerling D, Alomar O, Arno J (2001) Biological control of *Bemisia tabaci* using predators and parasitoids. Crop Prot 20:779-799

Gerling D, Horowitz AR (1984) Yellow traps for evaluating the population levels and dispersal patterns of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). Ann Entomol Soc Am 77:753-759

Gerling D, Horowitz AR, Baumgaertner J (1986) Autecology of *Bemisia tabaci*. Agr Ecosyst Environ 17:5-19

Gerling D, Mayer RT (1996) (Eds). *Bemisia*: 1995. Taxonomy, Biology, Damage, Control and Management. Intercept, Andover, UK

Gershoni M, Templeton AR, Mishmar D (2009) Mitochondrial bioenergetics as a major motive force of speciation. BioEssays 31:642-650

Ghosh S, Kanakala S, Lebedev G, Kontsedalov S, Silverman D, Alon T, Mor N, Sela N, Luria N, Dombrovsky A, Mawassi M, Haviv S, Czosnek H, Ghanim M (2019) Transmission of a new

polerovirus infecting pepper by the whitefly *Bemisia tabaci*. J Virol 93:e00488-19

Gibson RW, Legg JP, Otim-Nape GW (1996) Unusually severe symptoms are a characteristic of the current epidemic of mosaic virus disease of cassava in Uganda. Ann Appl Biol 128:479-490

Gilbertson RL, Batuman O, Webster CG, Adkins S (2015) Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. Annu Rev Virol 2:67-93

Gill RJ (1990) The morphology of whiteflies Gerling D (Ed.), Whiteflies: their bionomics, pest status and management, Intercept, Andover, UK, pp 13-46

Gopal P, Kumar P, Sinilal B, Jose J, Kasin Yadunandam A, Usha R (2007) Differential roles of C4 and βC1 in mediating suppression of posttranscriptional gene silencing: evidence for transactivation by the C2 of Bhendi yellow vein mosaic virus, a monopartite begomovirus. Virus Res 123:9-18

Gould JR, Hoelmer KA, Goolsby JA (2008) Classical biological control of *Bemisia tabaci* in the United States: a review of interagency research and implementation (co-editors). In: Hokkanen HMT (ed) Vol. 4, progress in biological control, Springer, Dordrecht, the Netherlands

Gregorio-Jorge J, Bernal-Alcocer A, Bañuelos-Hernández B, Alpuche-Solís AG, Hernández-Zepeda C, Moreno-Valenzuela O, Frías-Treviño G, Argüello-Astorga G (2010) Analysis of a new strain of Euphorbia mosaic virus with distinct replication specificity unveils a lineage of begomoviruses with short Rep sequences in the DNA-B intergenic region. Virol J 7:275

Guirao P, Beitia F, Cenis JL (1997) Biotype determination in Spanish populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae). B Entomol Res 87:587-593

Gunning RV, Byrne FJ, Devonshire, AL (1997) Electrophoretic analysis of non-B and B-biotype *Bemisia tabaci* (Gennaduis) (Hemiptera: Aleyrodidae) in Australia. Aust J Entomol 36:245-249

Ha C, Coombs S, Revill P, Harding R, Vu M, Dale J (2008) Molecular characterization of begomoviruses and DNA satellites from Vietnam: additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J Gen Virol 89:312-326

Hadjistylli M, Roderick GK, Gauthier N (2015) First report of the sub-Saharan Africa 2 species of the *Bemisia tabaci* complex in the Southern France. Phytoparasitica 43:679-687

Hanley-Bowdoin L, Settlage SB, Orozoo BM, Nagar S, Robertson D (2000) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Biochem Mol 35:105-140

Hanley-Bowdoin L, Settlage SB, Robertson D (2004) Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. Mol Plant Pathol 5:149-156

REFERENCES

Hassan I, Orílio AF, Fiallo-Olivé E, Briddon RW, Navas-Castillo J (2016) Infectivity, effects on helper viruses and whitefly transmission of the deltasatellites associated with sweepoviruses (genus *Begomovirus*, family *Geminiviridae*). Sci Rep 6:30204

Hernández-Zepeda C, Brown JK, Moreno-Valenzuela OA, Argüello-Astorga G, Idris AM, Carnevali G, Rivera-Bustamante RF (2010) Characterization of Rhynchosia yellow mosaic Yucatan virus, a new recombinant begomovirus associated with two fabaceous weeds in Yucatan, Mexico. Arch Virol 155:1571-1579

Heywood VH (1993) Flowering plants of the World. Oxford University Press, New York

Hilje L, Costa, HS, Stansly PA (2001) Cultural practices for managing *Bemisia tabaci* and associated viral diseases. Crop Prot 20:801-812

Hillocks RJ (1998) Integrated pest management reviews 3:155-167

Hollingsworth PM, Forrest LL, Spouge JL (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794-12797

Holm L, Pancho JV, Herberger JP, Plucknett DL (1979) A geographical Atlas of World weeds. John Wiley, Sons New York

Horowitz AR, Antignus Y, Gerling D (2011) Management of *Bemisia tabaci* Whiteflies. In: Thompson WMO (ed.), The whitefly, *Bemisia tabaci (*Homoptera: *Aleyrodidae)* Interaction with geminivirus-infected host plants, Springer Dordrecht, pp 293-322

Horowitz AR, Podoler H, Gerling D (1984) Life table analysis of the tobacco whitefly *Bemisia tabaci* (Gennadius) in cotton fields in Israel. Acta Oecol 5:221-233

Horowitz R, Denholm I, Morin S (2007) Resistance to insecticides in the tomato yellow leaf curl virus vector, *Bemisia tabaci*. In: Czosnek H (ed) Tomato yellow leaf curl virus disease, Springer Dordrecht, the Netherlands, pp 305-325

Horowitz AR, Ishaaya, I (2014) Dynamics of biotypes B and Q of the whitefly *Bemisia tabaci* and its impact on insecticide resistance. Pest Manag Sci 70:1568-1572

Horowitz, AR, (1986) Population dynamics of *Bemisia tabaci* (Gennadius): with special emphasis on cotton fields. Agric Ecosystems Environ 17:37-47

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biometrical J. 50:346-363

Hou YM, Gilbertson RL (1996) Increased pathogenicity in a pseudo-recombinant bipartite geminivirus correlates with intermolecular recombination. J Virol 70:5430-5436

Hsieh CH, Ko CC, Chung CH, Wang HY (2014) Multilocus approach to clarify species status and the divergence history of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. Mol Phylogenet Evol 76:172-180

Hu J, De Barro P, Zhao H, Wang J, Nardi F, Liu SS (2011) An extensive field survey combined with a phylogenetic analysis reveals rapid and widespread invasion of two alien whiteflies in China. PLoS One 6:e16061

Hussain M, Mansoor S, Iram S, Fatima AN, Zafar Y (2005) The nuclear shuttle protein of tomato leaf curl New Delhi virus is a pathogenicity determinant. J Virol 79:4434-4439

Hussain M, Mansoor S, Iram S, Zafar Y, Briddon RW (2007) The hypersensitive response to tomato leaf curl New Delhi virus nuclear shuttle protein is inhibited by transcriptional activator protein. Mol Plant Microbe In 20:1581-1588

ICTV database. https://talk.ictvonline.org/ICTV/proposals/2016.021a-kP.A.v2. *Tolecusatellitidae*

ICTV database. https://talk.ictvonline.org/taxonomy/

Idris AM, Briddon RW, Bull SE, Brown JK (2005) Cotton leaf curl Gezira virus-satellite DNAs represent a divergent, geographically-isolated Nile Basin lineage: predictive identification of a satDNA REP binding motif. Virus Res 109:19-32

Idris AM, Shahid MS, Briddon RW, Khan AJ, Zhu JK, Brown JK (2011) An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. J Gen Virol 92:706-717

Idris AM, Abdullah NM, Brown JK (2012) Leaf curl diseases of two solanaceous species in Southwest Arabia are caused by a monopartite begomovirus evolutionarily most closely related to a species from the nile basin and unique suite of betasatellites. Virus Res 169:296-300

Ilyas M, Qazi J, Mansoor S, Briddon RW (2009) Molecular characterization and infectivity of a ‘‘Legumovirus’’ (genus *Begomovirus*: family *Geminiviridae*) infecting the leguminous weed *Rhynchosia minima* in Pakistan. Virus Res 145:279-284

Ilyas M, Qazi J, Mansoor S, Briddon RW (2010) Genetic diversity and phylogeography of begomoviruses infecting legumes in Pakistan. J Gen Virol 91:2091-2101

Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage φ29 DNA polymerase. J Virol Methods 116:209-211

Jeffrey JL, Pooma W, Petty ITD (1996) Genetic requirements for local and systemic movement of tomato golden mosaic virus in infected plants. Virology 223:208-218

REFERENCES

Jeske H (2007) Replication of geminiviruses and the use of rolling circle amplification for their diagnosis. In: tomato yellow leaf curl virus disease. Edited by Czosnek H Springer, Dordrecht, pp 141-156

Jeske H (2018) Barcoding of plant viruses with circular single-stranded DNA based on rolling circle amplification. Viruses 10:469

John P, Sivalingam PN, Haq QM, Kumar N, Mishra A, Briddon RW, Malathi VG (2008) Cowpea golden mosaic disease in Gujarat is caused by a mungbean yellow mosaic India virus isolate with a DNA B variant. Arch Virol 153:1359-1365

Jones DR (2003) Plant viruses transmitted by whiteflies. Eur J Plant Pathol 109:195-219

Khan AA, Sharma R, Afreen B, Naqvi QA, Kumar S, Snehi SK, Raj SK (2011) Molecular identification of a new isolate of Cucumber mosaic virus subgroup II from basil (*Ocimum sanctum*) in India. Phytoparasitica 39:199-203

Khatri S, Nahid N, Fauquet CM, Mubin M, Nawaz-ul-Rehman MS (2014) A betasatellite-dependent begomovirus infects ornamental rose: Characterization of begomovirus infecting rose in Pakistan. Virus Genes 49:124-131

Kokwaro JO (2009) Medicinal plants of East Africa, 3rd ed. University of Nairobi Press, Nairobi

Kon T, Gilbertson RL (2012) Two genetically related begomoviruses causing tomato leaf curl disease in Togo and Nigeria differ in virulence and host range but do not require a betasatellite for induction of disease symptoms. Arch Virol 157:107-120

Kumar PP Usha R, Zrachya A, Levy Y, Spanov H, Gafni Y (2006) Protein-protein interactions and nuclear trafficking of coat protein and βC1 protein associated with bhendi yellow vein mosaic disease. Virus Res 122:127-136

Kumar S, Stecher G, Tamura K (2015) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874

Kumar J, Gunapati S, Singh SP, Gadre R, Sharma NC, Tuli R (2013) Molecular characterization and pathogenicity of a carrot (*Daucus carota*) infecting begomovirus and associated betasatellite from India. Virus Res 178:478-485

Laarif A, Saleh D, Clouet C, Gauthier N (2015) Regional co-occurrence between distinct *Bemisia tabaci* species in Tunisia with new insights into the role of host plants. Phytoparasitica 13:135-150

Lawrence BM (1992) Chemical components of Labiatae oils and their explotation. In: Advances in Labiatae Science. Harley KM, Reynolds T, eds. Royal Botanic Gardens, Kew, pp 399-436

Lee W, Park J, Lee GS, Lee S, Akimoto S (2013) Taxonomic status of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) and reassessment of the number of its constituent species. PLoS One 8:e63817

Lefeuvre P, Martin DP, Harkins G, Lemey P, Gray AJA (2010) The spread of tomato yellow leaf curl virus from the Middle East to the World. PLoS Pathog 6:e1001164

Lefeuvre P, Martin DP, Hoareau M, Naze F, Delatte H, Thierry M, Varsani A, Becker N, Reynaud B, Lett JM (2007) Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. J Gen Virol 88:3458-3468

Legg JP (1994) *Bemisia tabaci*: The whitefly vector of cassava mosaic geminiviruses in Africa: An ecological perspective. Afr Crop Sci J 2:437-448

Legg JP, French R, Rogan D, Okao-Okuja G, Brown JK (2002) A distinct *Bemisia tabaci* (Gennadius) (Hemiptera: Sternorrhyncha: Aleyrodidae) genotype cluster is associated with the epidemic of severe cassava mosaic virus disease in Uganda. Mol Ecol 11:1219-1229

Legg JP, Gibson RW, Otim-Nape GW (1994) Genetic polymorphism amongst Ugandan populations of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae), vector of African cassava mosaic geminivirus. Trop Sci 34:73-81

Legg JP, Lava Kumar P, Makeshkumar T, Tripathi L, Ferguson M, Kanju E, Ntawuruhunga P, Cuellar W (2015) Cassava virus diseases: biology, epidemiology, and management. Adv Virus Res 91:85-142

Legg JP (1999) Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. Crop Prot 18:627-637

Leke WN, Brown JK, Ligthart ME, Sattar MN, Njualem DK, Kvarnheden A (2012) *Ageratum conyzoides*: a host to a unique begomovirus disease complex in Cameroon. Virus Res 163:229-237

Leke WN, Kvarnheden A, Ngane EB, Titanji VP, Brown JK (2011) Molecular characterization of a new begomovirus and divergent alphasatellite from tomato in Cameroon. Arch Virol 156:925-928

Leke WN, Mignouna DB, Brown JK, Fondong VN (2016a) First report of soybean chlorotic blotch virus and West African asystasia virus 1 infecting cassava and a wild cassava relative in Cameroon and Togo. New Dis Rep 33:24

Leke WN, Khatabi B, Mignouna DB, Brown JK, Fondong VN (2016b) Complete genome sequence of a new bipartite begomovirus infecting cotton in the Republic of Benin in West Africa. Arch Virol 161:2329-2333

REFERENCES

Li J, Zhou X (2010) Molecular characterization and experimental host-range of two begomoviruses infecting *Clerodendrum cyrtophyllum* in China. Virus Genes 41:250-259

Li TY, Vinson SB, Gerling D (1989) Courtship and mating behaviour of *Bemisia tabaci* (Homoptera, Alyerodidae). Environ Entomol 18:800-806

Lima ATM, Silva JCF, Silva FN, Castillo-Urquiza GP, Silva FF, Seah YM, Mizubuti ESG, Duffy S, Zerbini FM (2017) The diversification of begomovirus populations is predominantly driven by mutational dynamics. Virus Evol 3:vex005

Lima ATM, Sobrinho RR, Gonzalez-Aguilera J, Rocha CS, Silva SJC, Xavier CAD, Silva F, Duffy S, Zerbini FM (2013) Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. J Gen Virol 94:418-431

Liu J, Zhao H, Jiang K, Zhou XP, Liu SS (2009) Differential indirect effects of two plant viruses on an invasive and an indigenous whitefly vector: implications for competitive displacement. Ann Applied Biol 155:439-448

Liu SS, Colvin J, De Barro PJ (2012) Species concepts as applied to the whitefly *Bemisia tabaci* systematics: how many species are there? J Integr Agr 11:176-186

Liu SS, De Barro PJ, Xu J, Luan JB, Zang LS, Ruan YM, Wan FH (2007) Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. Science 318:1769-1772

Lozano G, Trenado HP, Fiallo-Olivé E, Chirinos D, Geraud-Pouey F, Briddon RW, Navas-Castillo J (2016) Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, *Geminiviridae*)-definition of a distinct class of begomovirus-associated satellites. Front Microbiol 7:162

Lozano G, Trenado HP, Valverde RA, Navas-Castillo J (2009) Novel begomovirus species of recombinant nature in sweet potato (*Ipomoea batatas*) and *Ipomoea indica*: taxonomic and phylogenetic implications. J Gen Virol 90:2550-2562

Luan JB, Ruan YM, Zhang L, Liu SS (2008) Pre-copulation intervals, copulation frequencies, and initial progeny sex ratios in two biotypes of whitefly, *Bemisia tabaci*. Entomol Exp Appl 129:316-324

Macedo MA, Albuquerque LC, Maliano MR, Souza JO, Rojas MR, Inoue-Nagata AK, Gilbertson RL (2018) Characterization of tomato leaf curl purple vein virus, a new monopartite New World begomovirus infecting tomato in Northeast Brazil. Arch Virol 163:737-743

Mansoor S, Briddon RW, Bull SE, Bedford ID, Bashir A, Hussain M, Saeed M, Zafar MY, Malik KA, Fauquet C, Markham PG (2003) Cotton leaf curl disease is associated with multiple monopartite begomoviruses supported by single DNA β. Arch Virol 148:1969-1986

Mansoor S, Khan SH, Bashir A, Saeed M, Zafar Y, Malik KA (1999) Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259:190-199

Mar TB, Mendes IR, Lau D, Fiallo-Olivé E, Navas-Castillo J, Alves MS, Zerbini FM (2017) Interaction between the New World begomovirus Euphorbia yellow mosaic virus and its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. J G Virol 98:1552-1562

Márquez-Martín B, Aragon-Caballero L, Fiallo-Olivé E, Navas-Castillo J, Moriones E (2011) Tomato leaf deformation virus, a novel begomovirus associated with a severe disease of tomato in Peru. Eur J Plant Pathol 129:1-7

Martin JA, Mound LA (2007) "An annotated check list of the world's whiteflies (Insecta: Hemiptera: Aleyrodidae)"Zootaxa 1492:1-84

Martin JH (1987) An identification guide to common whitefly pest species of the world (Homoptera, Aleyrodidae). Trop Pest Manag 33:298-322

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol 1:1-5

Martin JH, Mifsud D, Rapisarda C (2000) The whiteflies (Hemiptera: Aleyrodidae) of Europe and the Mediterranean basin. B Entomol Res 90:407-448

Maruthi MN, Colvin J, Gibson G, Thwaites RM, Banks GK, Seal S (2004) Reproductive incompatibility and cytochrome oxidase I gene sequence variability amongst host-adapted and geographically separate *Bemisia tabaci* populations (Hemiptera: Aleyrodidae). Syst Entomol 29:560-568

Maruthi MN, Hillocks RJ, Mtunda K (2005b) Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). Phytopathology 153:307-312

Maruthi MN, Colvin J, Seal S (2001) Mating compatibility, life-history traits, and RAPD-PCR variation in *Bemisia tabaci* associated with the cassava mosaic disease pandemic in East Africa. Entomol Exp Appl 99:13-23

Maruthi MN, Manjunatha B, Rekha AR, Govindappa MR, Colvin J, Muniyappa V (2006) Dolichos yellow mosaic virus belongs to a distinct lineage of Old World begomoviruses; its biological and molecular properties. Ann Appl Biol 149:187-195

Maruthi MN, Colvin J, Seal S, Gibson G, Cooper J (2002) Co-adaptation between cassava mosaic geminiviruses and their local vector populations. Virus Res 86:71-85.

Masaba SC (2000) The antimalarial activity of *Vernonia amygdalina* Del. (Compositae). Trans R Soc Trop Med Hyg 94:694-695

REFERENCES

Mayo MA, Leibowitz MJ, Palukaitis P, Scholthof KBG, Simon AE, Stanley J Taliansky MB, (2005) Satellites. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, Eds. VIIIth Report of the International Committee on Taxonomy of Viruses Virus Taxonomy. London: Elsevier/Academic Press, pp 1163-1169

McKenzie CL, Bethke JA, Byrne FJ, Chamberlin JR, Dennehy TJ, Dickey AM, Gilrein D, Hall PM, Ludwig S, Oetting RD, Osborne LS, Schmale L, Shatters RG Jr (2012) Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in North America after the Q invasion. J Econ Entomol 105:753-766

Melgarejo TA, Kon T, Rojas MR, Paz-Carrasco L, Zerbini FM, Gilbertson RL (2013) Characterization of a New World monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. J Virol 87:5397-5413

Merrell DJ (1981) Ecological genetics. Longman, London, pp 358-399

Moffat AS (1999) Plant pathology-geminiviruses emerge as serious crop threat. Science 286:1835

Mohammed HS, El Siddig MA, El Hussein AA, Navas-Castillo J, Fiallo-Olivé E (2018) Complete genome sequence of datura leaf curl virus, a novel begomovirus infecting *Datura innoxia* in Sudan, related to begomoviruses causing tomato yellow leaf curl disease. Arch Virol 163:273-275

Mollel HG, Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017a) A novel East African monopartite begomovirus-betasatellite complex that infects *Vernonia amygdalina*. Arch Virol 162:1079-1082

Mollel HG, Ndunguru J, Sseruwagi P, Alicai T, Colvin J, Navas-Castillo J, Fiallo-Olivé E (2017b) Desmodium mottle virus, the first legumovirus (genus *Begomovirus*) from East Africa. Arch. Virol. 162:1799-1803

Monci F, Sánchez-Campos S, Navas-Castillo J, Moriones E (2002) A natural recombinant between the geminiviruses tomato yellow leaf curl Sardinia virus and tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303:317-326

Morales FJ, Anderson, PK (2001) The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Arch Virol 146:415-441

Moriones E, Navas-Castillo J, Díaz-Pendón J (2011) Emergence of begomovirus diseases. In: Recent Advance Plant Virology. Edited by Caranta C, Aranda MA, Tepfer M, Lopez-Moya JJ Caister Academic Press, UK, pp 301-320

Mound LA, Halsey SH (1978) Whitefly of the World. A systematic catalogue of the Aleyrodidae (Homoptera) with host plant and natural enemy data. British Museum (Natural History) and John Wiley and Sons, Chichester (Great Britain), pp 340

Mubin M, Amin I, Amrao L, Briddon RW, Mansoor S (2010) The hypersensitive response induced by the V2 protein of a monopartite begomovirus is countered by the C2 protein. Mol Plant Pathol 11:245-254

Mubin M, Briddon RW, Mansoor S (2009) Diverse and recombinant DNA betasatellites are associated with a begomovirus disease complex of *Digera arvensis*, a weed host. Virus Res 142:208-212

Mugerwa (2018) Molecular characterisation of cassava-colonising populations of the whitefly, *Bemisia tabaci*, PhD Thesis, University of Greenwich, UK.

Mugerwa H, Rey MEC, Alicai T, Ateka E, Atuncha H, Ndunguru J, Sseruwagi P (2012) Genetic diversity and geographic distribution of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) genotypes associated with cassava in East Africa. Ecol Evol 2:2749-2762

Mugerwa H, Seal S, Wang HL, Patel MV, Kabaalu R, Omongo CA, Alicai T, Tairo F, Ndunguru J, Sseruwagi P, Colvin J (2018) African ancestry of New World, *Bemisia tabaci*-whitefly species. Sci Rep 8:2734

Muhire B, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9:e108277

Naranjo SE, Ellsworth, PC (2005) Mortality dynamics and population regulation in *Bemisia tabaci*. Entomol Exp Appl 116:93-108

Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibáñez JT, Hanley-Bowdoin L (2011) Functional analysis of a novel motif conserved across geminivirus Rep proteins. J Virol 85:1182-1192

Navas-Castillo J, Fiallo-Olivé E, Sánchez-Campos S (2011) Emerging virus diseases transmitted by whiteflies. Annu Rev Phytopathol 49:219-248

Navas-Castillo J, Sánchez-Campos S, Noris E, Louro D, Accotto GP, Moriones E (2000) Natural recombination between tomato yellow leaf curl virus and tomato leaf curl virus. J Gen Virol 81:2797-2801

Nawaz-ul-Rehman MS, Fauquet CM (2009) Evolution of geminiviruses and their satellites. FEBS Lett 583:1825-1832

Ndunguru J, Legg JP, Aveling TAS, Thompson G, Fauquet CM (2005) Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. J Virol 2:21

REFERENCES

Nehra C, Marwa A, Kumar VR, Gaur RK (2019) Molecular characterization of begomoviruses DNA-A and associated betasatellites with new host *Ocimum sanctum* in India. P Natl A Sci India B 89:903-910

Oerke EC (2006) Crop losses to pests. J Agr Sci 144:31-43

Omondi BA, Sseruwagi P, Obeng-Ofori D, Danquah EY, Kyerematen RA (2005) Mating interactions between okra and cassava biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on eggplant. Int J Trop Insect Sci 25:159-167

Otim M, Kyalo G, Kyamanywa S, Asiimwe P, Legg JP, Guershon M, Gerling D (2008) Parasitism of *Bemisia tabaci* (Homoptera: Aleyrodidae) by *Eretmocerus mundus* (Hymenoptera: Aphelinidae) on cassava. Int J Trop Insect Sci 28:158-167

Packialakshmi RM, Srivastava N, Girish KR, Usha R (2010) Molecular characterization of a distinct begomovirus species from *Vernonia cinerea* and its associated DNA-betasatellite using the bacteriophage Phi 29 DNA polymerase. Virus Genes 41:135-143

Padidam M, Beachy RN, Fauquet CM (1996) The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. Virology 224:390-404

Palumbo JC, Horowitz AR, Prabhaker N (2001) Insecticidal control and resistance management for *Bemisia tabaci*. Crop Prot 20:739-765

Pandey P, Choudhury N, Mukherjee S (2009) A geminiviral amplicon (VA) derived from tomato leaf curl virus (ToLCV) can replicate in a wide variety of plant species and also acts as a VIGS vector. Virol J 6:152

Pant V, Gupta D, Choudhury NR, Malathi VG, Varma A, Mukherjee SK (2001) Molecular characterization of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus. J Gen Virol 82:2559-2567

Paprotka T, Metzler V, Jeske H (2010) The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. Virology 404:148-157

Pasumarthy KK, Mukherjee SK, Choudhury, NR (2011) The presence of tomato leaf curl Kerala virus AC3 protein enhances viral DNA replication and modulates virus induced gene-silencing mechanism in tomato plants. Virol J 8:178

Paton A (1992) A synopsis of *Ocimum* L. (Labiatae) in Africa. Kew Bull 47:405-437

Paton A, Harley RM, Harley MM (1999) *Ocimum*: An overview of classification and relationships. In: The genus *Ocimum*. Hiltunen R, Holm Y, eds. Harwood Academic, Amsterdam, pp 1-38

Pedersen TJ, Hanley-Bowdoin L (1994) Molecular characterization of the AL3 protein encoded by a bipartite geminivirus. Virology 202:1070-1075

Permingeat HR, Romagnoli MV, Vallejos RH (1998) A simple method for isolating high yield and quality DNA from cotton (*Gossypium hirsutum* L.) leaves. Plant Mol Biol Rep 16:1-6

Perring TM, Cooper AD, Rodriguez RJ, Farrar CA, Bellows TS (1993) Identification of a whitefly species by genomic and behavioral studies. Science 259:74-77

Perring T (2001) The *Bemisia tabaci* complex. Crop Prot 20:725-737

Perring TM, Symmes EJ (2006) Courtship behavior of *Bemisia argentifolii* (Hemiptera: Aleyrodidae) and whitefly mate recognition. Ann Entomol Soc Am 99:598-606

Pita JS, Fondong VN, Sangare A, Otim-Nape GW, Ogwal S, Fauquet CM (2001) Recombination, pseudo-recombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. J Gen Virol 82:655-665

Polston JE, De Barro P, Boykin LM (2014) Transmission specificities of plant viruses with the newly identified species of the *Bemisia tabaci* species complex. Pest Manag Sci 70:1547-1552

Poojari S, Naidu RA (2013) First report of Impatiens necrotic spot virus (INSV) infecting basil (*Ocimum basilicum*) in the United States. Plant Dis 97:850

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Qazi J, Ilyas M, Mansoor S, Briddon R (2007) Legume yellow mosaic viruses: genetically isolated begomoviruses. Mol Plant Pathol 8:343-348

Qin L, Pan LL, Liu SS (2016) Further insight into reproductive incompatibility between putative cryptic species of the *Bemisia tabaci* whitefly complex. Insect Sci 23:215-224

R Core Team (2013) R: A language and environment for statistical computing. (R Foundation for Statistical Computing)

Rey C, Vanderschuren H (2017) Cassava mosaic and brown streak diseases: Current perspectives and beyond. Annu Rev Virol 4:429-452

Rocha CS, Castillo-Urquiza GP, Lima ATM, Silva FN, Xavier CAD, Hora-Junior BT, Beserra-Junior JEA, Malta AWO, Martin DP, Varsani A, Alfenas-Zerbini P, Mizubuti ESG, Zerbini FM (2013) Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. J Virol 87:5784-5799

Rodríguez-Negrete EA, Morales-Aguilar JJ, Domínguez-Duran G, Torres-Devora, G, Camacho-Beltrán E, Leyva-López NE, Voloudakis AE, Bejarano ER, Méndez-Lozano J (2019) High-Throughput Sequencing reveals differential begomovirus species diversity in non-cultivated plants in Northern-Pacific Mexico. Viruses 11:594

Rojas MR, Gilbertson RL, Russell DR, Maxwell DP (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis 77:340-347

REFERENCES

Rojas MR, Hagen C, Lucas WJ, Gilbertson RL (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu Rev Phytopathol 43:361-394

Rojas MR, Macedo MA, Maliano MR, Soto-Aguilar M, Souza JO (2018) World management of geminiviruses. Annu Rev Phytopathol 56:637-677

Romay G, Chirinos D, Geraud-Pouey F, Desbiez C (2010) Association of an atypical alphasatellite with a bipartite New World begomovirus. Arch Virol 155:1843-1847

Ronda MA, Adan A, Cifuentes D, Cenis JL, Bietia F (1999) Laboratory evidence of interbreeding between biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) present in Spain. In VIIth International Plant Virus Epidemiology Symposium-Plant Virus Epidemiology: Current status and future prospects. Aguadulce, pp 83-84

Russell LM (1957) Synonyms of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). B Brooklyn Entomol Soc 52:122-123

Saeed S, Zafar YY, Randles JW, Rezaian MA (2007) A monopartite begomovirus-associated DNA β satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. J Gen Virol 88:2881-2889

Samretwanich K, Kittipakorn K, Chiemsombat P, Ikegami M (2001) Complete nucleotide sequence and genome organization of soybean crinkle leaf virus. J Phytopathol 149:333-336

Sánchez-Campos S, Martínez-Ayala A, Márquez-Martín B, Aragón-Caballero L, Navas-Castillo J, Moriones E (2013) Fulfilling Koch's postulates confirms the monopartite nature of tomato leaf deformation virus: a begomovirus native to the New World. Virus Res 173:286-293

Sanderfoot AA, Ingham DJ, Lazarowitz SC (1996) A viral movement protein as a nuclear shuttle. Plant Physiol 11:22-33

Sanderfoot AA, Lazarowitz SG (1996) Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. Trends in Cell Biol 6:353-358

Saunders K, Bedford ID, Briddon RW, Markham PG, Wong SM, Stanley J (2000) A unique virus complex causes ageratum yellow vein disease. P Natl A Sci USA 97:6890-6895

Saunders K, Briddon RW, Stanley J (2008) Replication promiscuity of DNA-β satellites associated with monopartite begomoviruses; deletion mutagenesis of the ageratum yellow vein virus DNA-β satellite localizes sequences involved in replication. J Gen Virol 89:3165-3172

Shankarappa KS, Rangaswamy KT, Narayana DSA, Rekha AR, Raghavendra N (2007) Development of silverleaf assay, protein and nucleic acid-based diagnostic techniques for the quick and reliable detection and monitoring of biotype B of the whitefly, *Bemisia tabaci* (Gennadius). B Entomol Res 97:503-513

Sharaf NS, Al-Musa AM, Batta Y (1985) Effect of different host plants on population development of the whitefly *Bemisia tabaci* Genn. (Homoptera: Aleyrodidae). Dirasat 12:89-100

Sharma P, Gaur RK, Ikegami M (2011) Subcellular localization of V2 protein of tomato leaf curl Java virus by using green fluorescent protein and yeast hybrid system. Protoplasma 248:281-288

Sharma P, Ikegami M (2009) Characterization of signals that dictate nuclear/nucleolar and cytoplasmic shuttling of the capsid protein of Tomato leaf curl Java virus associated with DNA β satellite. Virus Res 144:145-153

Shih SL, Green SK, Tsai WS, Ssekyewa C (2006a) Molecular characterization of a begomovirus associated with tomato leaf curl disease in Uganda. Plant Dis 90:246

Shih SL, Tsai WS, Green SK, Lee LM (2006b) Molecular characterization of a distinct begomovirus associated with tomato leaf curl disease in Arusha of Tanzania. Plant Dis 90:1550

Silva AKF, Santos CDG, Nascimento AKQ (2010) Begomovirus transmission from weeds to tomato by the whitefly. Planta Daninha 28:507-514

Silva F, Lima A, Rocha C, Castillo-Urquiza G, Alves-Junior M, Zerbini F (2014) Recombination and pseudo-recombination driving the evolution of the begomoviruses tomato severe rugose virus (ToSRV) and tomato rugose mosaic virus (ToRMV): two recombinant DNA-A components sharing the same DNA-B. Virol J 11:66

Silva SJC, Castillo-Urquiza GP, Hora-Junior BT, Assuncao IP, Lima GSA, Pio-Ribeiro G, Mizubuti ESG, Zerbini FM (2012) Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. Plant Pathol 61:457-46

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87:651-701

Snehi SK, Raj SK, Prasad V, Singh V (2015) Recent research findings related to management strategies of begomoviruses. J Plant Pathol Microb 6:273

Sseruwagi P, Maruthi G, Colvin J, Rey C, Brown J, Legg J (2006) Colonization of non-cassava plant species by whiteflies (*Bemisia tabaci*) in Uganda. Entomol Exp App 119:145-153

Sseruwagi P, Legg JP, Maruthi MN, Colvin J, Rey MEC, Brown J (2005) Genetic diversity of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) populations and presence of the B biotype and a non-B biotype that can induce silverleaf symptoms in squash, in Uganda. Ann Appl Biol 147:253-265

Stanley J (1995) Analysis of African cassava mosaic virus recombinants suggest strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. Virology 206:707-712

REFERENCES

Stanley J, Bisaro DM, Briddon RW, Brown JK, Fauquet CM, Harrison BD, Rybicki EP, Stenger, DC (2005) Family *Geminiviridae*. In Virus Taxonomy: Edited by Fauquet, CM, Mayo MA, Maniloff J, Desselberger U, Ball LA Elsevier Academic Press, London. Eighth Report of the International Committee on Taxonomy of Viruses, pp 301-326

Stepp JR (2004) The role of weeds as sources of pharmaceuticals. J Ethnopharmacol 92:163-166

Stepp JR, Moerman DE (2001) The importance of weeds in ethnopharmacology. J Ethnopharmacol 75:25-31

Sun DB, Xu J, Luan JB, Liu SS (2011) Reproductive incompatibility between the B and Q biotypes of the whitefly *Bemisia tabaci* in China: genetic and behavioural evidence. B Entomol Res 101:211-220

Swarnalatha P, Mamatha M, Manasa M, Singh RP, Krishnareddy M (2013) Molecular identification of Ageratum enation virus (AEV) associated with leaf curl disease of tomato (*Solanum lycopersicum*) in India. Australasian Plant Dis Notes 8:67-71

Tahiri A, Halkett F, Granier M, Gueguen G, Peterschmitt M (2013) Evidence of gene flow between sympatric populations of the Middle East-Asia Minor 1 and Mediterranean putative species of *Bemisia tabaci.* Ecol Evol 3:2619-2633

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725-2729

Tan PHN, Wong SM, Wu M, Bedford ID, Saunders K, Stanley J (1995) Genome organization of ageratum yellow vein virus, a monopartite whitefly-transmitted geminivirus isolated from a common weed. J Gen Virol. 76:2915-2922

Terraz G, Gueguen G, Arnó J, Fleury F, Mouton L (2014) Nuclear and cytoplasmic differentiation among Mediterranean populations of *Bemisia tabaci*: testing the biological relevance of cytotypes. Pest Manag Sci 70:1503-1513

Thresh JM, Cooter RJ (2005) Strategies for controlling cassava mosaic virus disease in Africa. Plant Pathol 54:587-614

Tiendrébéogo F, Lefeuvre P, Hoareau M, Harimalala MA, De Bruyn A, Villemot J, Traore VSE, Konate G, Traore AS, Barro N, Reynaud B, Traore O, Lett JM (2012) Evolution of African cassava mosaic virus by recombination between bipartite and monopartite begomoviruses. Virol J 9:67

Tiendrébéogo F, Lefeuvre P, Hoareau M, Villemot J, Konate G, Traore AS, Barro N, Traore VS, Reynaud B, Traore O, Lett, JM (2010) Molecular diversity of cotton leaf curl Gezira virus isolates, and their satellite DNAs associated with okra leaf curl disease in Burkina Faso. Virol J 7:48

Trenado HP, Orílio AF, Márquez-Martín B, Moriones E, Navas-Castillo J (2011) Sweepoviruses cause disease in sweet potato and related *Ipomoea* spp.: Fulfilling Koch's postulates for a divergent group in the genus *Begomovirus*. PLoS One 6: e27329

Trinks D, Rajeswaran R, Shivaprasad PV, Akbergenov R, Oakeley EJ, Veluthambi K, Hohn T, Pooggin MM (2005) Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J Virol 79:2517-2527

Tukey JW (1949) Comparing individual means in the analysis of variance. Biometrics 5:99-114

Van der Walt E, Martin DP, Varsani A, Polston JE, Rybicki EP (2008) Experimental observations of rapid maize streak virus evolution reveal a strand-specific nucleotide substitution bias. Virol J 5:104

Vanitharani R, Chellappan P, Pita JS, Fauquet CM (2004) Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of post-transcriptional gene silencing. J Virol 78:9487-9498

Varma A, Malathi VG (2003) Emerging geminivirus problems. A serious threat to crop production. Ann Appl Biol 142:145-164

Varsani A, Navas-Castillo J, Moriones E, Hernández-Zepeda C, Idris A, Brown JK, Zerbini FM, Martin DP (2014) Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. Arch Virol 159:2193-2203

Venables WN, Ripley BD (2002) Modern applied statistics with *S*. (Springer-Verlag)

Vyskočilová S, Tay WT, van Brunschot S, Seal S, Colvin J (2018) An integrative approach to discovering cryptic species within the *Bemisia tabaci* whitefly species complex. Sci Rep 8:10886

Wang Z, Yan H, Yang Y, Wu Y (2010) Biotype and insecticide resistance status of the whitefly *Bemisia tabaci* from China. Pest Manag Sci 66:1360-1366

Wang P, Crowder DW, Liu, SS (2012) Roles of mating behavioral interactions and life history traits in the competition between alien and indigenous whiteflies. B Entomol Res 102:395-405

Wang P, Sun DB, Qiu BL, Liu SS (2011) The presence of six cryptic species of the whitefly *Bemisia tabaci* complex in China as revealed by crossing experiments. Insect Sci 18:67-77

Wasswa P, Otto B, Maruthi M N, Mukasa SB, Monger W, Gibson RW (2011) First identification of a sweet potato begomovirus (sweepovirus) in Uganda: characterization, detection and distribution. Plant Pathol 60:1030-1039

White JA, Kelly SE, Perlman SJ, Hunter MS (2009) Cytoplasmic incompatibility in the parasitic wasp *Encarsia inaron*: disentangling the roles of *Cardinium* and *Wolbachia* symbionts. Heredity 102:483-489

REFERENCES

Wintermantel WM, Natwick ET (2012) First report of alfalfa mosaic virus infecting basil (*Ocimun basilicum*) in California. Plant Dis 96:295

Wool D, Gerling D, Nolt BL, Constantino LM, Bellotti AC, Morales FJ (1989) The use of electrophoresis for identification of adult whiteflies (Homoptera: Aleyrodidae) in Israel and Colombia. J Appl Entomol 107:344-350

Wosula EN, Chen W, Fei Z, Legg JP (2017) Unravelling the genetic diversity among cassava *Bemisia tabaci* whiteflies using NextRAD sequencing. Genome Biol Evol 9:2958-2973

Wu JB, Cai JH, Zhou XP (2007) A new begomovirus associated with yellow vein disease of *Siegesbeckia glabrescens*. Plant Pathol 56:343

Wyant P, Strohmeiera S, Fischera A, Schäfera B, Briddon RW, Krenz B, Jeske H (2015) Light-dependent segregation of begomoviruses in *Asystasia gangetica* leaves. Virus Res 195:225-235

Wyant PS, Gotthardt D, Schafer B, Krenz B, Jeske H (2011) The genomes of four novel begomoviruses and a new Sida micrantha mosaic virus strain from Bolivian weeds. Arch Virol 156:347-352

Xu J, De Barro PJ, Liu SS (2010) Reproductive incompatibility among genetic groups of *Bemisia tabaci* supports the proposition that the whitefly is a cryptic species complex. B Entomol Res 100:359-366

Xu J, Lin KK, Liu SS (2011) Performance on different host plants of an alien and an indigenous *Bemisia tabaci* from China. J Appl Entomol 135:771-779

Yadava P, Suyal G, Mukherjee SK (2010) Begomovirus DNA replication and pathogenicity. Curr Sci-India 98:360-368

Yang X, Baliji S, Buchmann RC, Wang H, Lindbo JA, Sunter G, Bisaro DM (2007) Functional modulation of the geminivirus AL2 transcription factor and silencing suppressor by self-interaction. J Virol 81:11972-11981

Yasmin S, Raja NI, Hameed S, Brown JK (2017) First association of pedilanthus leaf curl virus, papaya leaf curl virus, cotton leaf curl Kokhran virus, and papaya leaf curl betasatellite with symptomatic chilli pepper in Pakistan. Plant Dis 101:2155

Zang LS, Chen WQ, Liu SS (2006) Comparison of performance on different host plants between the B biotype and a non-B biotype of *Bemisia tabaci* from Zhejiang, China. Entomol Exp Appl 121:221-227

Zang LS, Liu SS (2007) A comparative study on mating behaviour between the B biotype and a non-B biotype of *Bemisia tabaci* (Hemiptera: Aleyrodidae) from Zhejiang, China. J Insect Behav 20:157-171

Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas-Castillo J, Rivera-Bustamante R, Roumagnac P, Varsani A, ICTV Report Consortium (2017) ICTV Virus Taxonomy Profile: *Geminiviridae*. J Gen Virol 98:131-133

Zhou X (2013) Advances in understanding begomovirus satellites. Annu Rev Phytopathol 51:357-381

Zhou X, Liu Y, Robinson DJ, Harrison BD (1998) Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. J Gen Virol 79:915-923

Zhou YC, Noussourou M, Kon T, Rojas M, Jiang H, Chen LF, Gamby K, Foster R, Gilbertson R (2008) Evidence of local evolution of tomato-infecting begomovirus species in West Africa: characterization of tomato leaf curl Mali virus and tomato yellow leaf crumple virus from Mali. Arch Virol 153:693-706

Zhou X (2013) Advances in understanding begomovirus satellites. Annu Rev Phytopathol 51:357-381

Zinga I, Chiroleu F, Legg J, Lefeuvre P, Komba EK, Semballa S, Yandia SP, Mandakombo NB, Reynaud B, Lett JM (2013) Epidemiological assessment of cassava mosaic disease in Central African Republic reveals the importance of mixed viral infection and poor health of plant cuttings. Crop Prot 44:6-12

Zrachya A, Glick E, Levy Y, Arazi, T, Citovsky V, Gafni Y (2007) Suppressor of RNA silencing encoded by tomato yellow leaf curl virus-Israel. Virology 358:159-165

Zubair M, Zaidi SSEA, Shakir S, Amin I, Mansoor S (2017) An insight into cotton leaf curl Multan betasatellite, the most important component of cotton leaf curl disease complex. Viruses 9:280

Zulfiqar A, Zhang J, Cui X (2012) A new begomovirus associated with alpha- and beta-satellite molecules isolated from *Vernonia cinerea* in China. Arch Virol 157:189-191

Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora"
(IHSM-UMA-CSIC)

Programa de Doctorado de Biotecnología Avanzada
Facultad de Ciencias
Universidad de Málaga