HAL
open science

Épidémiologie moléculaire des géminivirus responsables de maladies émergentes sur les cultures maraîchères au Burkina Faso et en Côte d'Ivoire

Alassane Ouattara

To cite this version:

Alassane Ouattara. Épidémiologie moléculaire des géminivirus responsables de maladies émergentes sur les cultures maraîchères au Burkina Faso et en Côte d'Ivoire. Biologie végétale. Université de la Réunion; Université Ouaga 1 Professeur Joseph Ki-Zerbo (Ouagadougou, Burkina Faso), 2017. Français. NNT : 2017LARE0052 . tel-02408328v1

HAL Id: tel-02408328

https://theses.hal.science/tel-02408328v1

Submitted on 13 Dec 2019 (v1), last revised 6 Feb 2020 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE OUAGA I
PR JOSEPH KI-ZERBO

Unité de Formation et de Recherche
Sciences de la Vie et de la Terre

UNIVERSITÉ DE LA RÉUNION

Faculté des Sciences et Technologies

UMR Peuplements Végétaux et Bio-agresseurs en Milieu Tropical
CIRAD – Université de La Réunion

INERA – LMI Patho-Bios

THÈSE EN COTUTELLE
Pour obtenir le diplôme de Doctorat en Sciences

Epidémiologie moléculaire des géminivirus responsables de maladies émergentes sur les cultures maraîchères au Burkina Faso

par

Alassane Ouattara

Soutenance le 14 Décembre 2017, devant le jury composé de

Stéphane POUSSIER	Professeur, Université de La Réunion	Président
Justin PITA	Professeur, Université Houphouët-Boigny, Côte d'Ivoire	Rapporteur
Philippe ROUMAGNAC	Chercheur HDR, CIRAD, UMR BGPI, France	Rapporteur
Fidèle TIENDREBEOGO	Chercheur, INERA, Burkina Faso	Examinateur
Nathalie BECKER	Maître de conférences HDR MNHN, UMR ISYEB, France	Examinatrice
Jean-Michel LETT	Chercheur HDR, CIRAD, UMR PVBMT, La Réunion	Co-Directeur de thèse
Nicolas BARRO	Professeur, Université Ouagadougou, Burkina Faso	Co-Directeur de thèse

DEDICACES

*A mon épouse **Dadjata** et à mon fils **Jaad Kaamil**: merci pour l'amour, les encouragements et la compréhension tout au long de ces trois années.*

*A mon père **Kassoum,** à ma mère **Salimata,** à ma belle-mère **Bila,** à mon oncle **Soumaïla** et son épouse **Abibata** : merci pour votre amour, j'ai toujours reçu soutien, encouragements et bénédictions de votre part. Puisse Dieu vous garder en bonne santé !*

REMERCIEMENTS

Mes remerciements vont à l'endroit du personnel des Universités Ouaga I Pr Joseph KI-ZERBO et de La Réunion pour avoir accepté mon inscription. Je remercie les différents financeurs de mes travaux : l'AIRD (Projet PEERS-EMEB), l'Union Européenne (ERDF), le Conseil Régional de La Réunion et le Cirad (Bourse Cirad-Sud).

Je voudrais remercier les membres du jury : Prof. **Stéphane POUSSIER** pour m'avoir fait l'honneur de présider ce jury ; Dr **Fidèle TIENDREBEOGO**, Dr **Nathalie BECKER**, Dr HDR **Jean-Michel LETT**, Dr HDR **Philippe ROUMAGNAC**, Dr **HDR Oumar TRAORE**, Prof. **Justin PITA,** Prof. **Nicolas BARRO** pour avoir accepté de juger ce travail. Je vous remercie tous pour votre disponibilité.

Je voudrai également remercier toutes les personnes qui ont été membres des comités de suivi de ma thèse.

Mes plus vifs remerciements vont à :

- **Jean Michel LETT** qui m'a ouvert les portes du Cirad depuis 2013. Merci pour avoir guidé mes premiers pas dans le monde de la virologie. Collègue de bureau, tu étais disponible à tout moment que ce soit dans le cadre du travail ou non. J'ai encore en mémoire ce film dans lequel tu m'aides à nettoyer le labo NS3 à l'issu des manipes. Au-delà d'un chef d'équipe, j'ai vu en toi un père et un époux exemplaire. Par ces lignes, je profite dire merci à ta famille pour avoir souvent sacrifié leurs week-ends afin qu'on puisse arriver à bout de cette thèse.
- **Pierre LEFEUVRE** qui a été mon guide en matière d'analyses bioinformatiques. Ton caractère perfectionniste a fortement contribué à améliorer mon travail. Merci pour toute la patience avec laquelle tu m'as enseigné. Tu as créé en moi l'amour pour la bioinformatique. Tu étais également présent lorsque j'étais confronté à des difficultés aussi bien au labo que dans la vie courante.
- **Murielle HOAREAU** qui a été une mère pour moi dans le monde de la biologie moléculaire en m'accompagnant dans l'exécution de mes premiers pipetages. Merci pour cet agréable sourire que tu as toujours exprimé lors des manipes.

- **Fidèle TIENDREBEOGO** pour tes enseignements et les encouragements tout au long de ces cinq dernières années. Je tiens à te témoigner ma plus vive reconnaissance pour ta rigueur scientifique et pour tout le soutien.
- **Oumar TRAORE**, qui m'a permis d'intégrer le Laboratoire de Virologie et Biotechnologie Végétale (LVBV), INERA Ouagadougou. Un grand merci pour tes conseils avisés, ton enthousiasme et ton humanité, mais aussi pour le sacrifice. Sans ton soutien je n'aurais peut-être pas eu la chance de faire cette thèse.
- **Koutoua SEKA** qui m'a permis de travailler sur des échantillons de la Côte d'Ivoire. Merci pour cette collaboration qui a abouti à la publication de deux articles.

Je tiens également à remercier fortement l'inépuisable patron **Nicolas BARRO**, pour avoir accepté de codiriger cette thèse, mais surtout pour son soutien infaillible.

Un immense merci à la **Bégomo-Team**. J'ai quitté ma famille pour La Réunion et vous avez été pour moi une seconde famille. Merci à **Nathalie** pour sa gentillesse et son aide au labo et à **Martial** pour avoir mis à ma disposition les plantules et les aleurodes. Merci à **Jérémy** pour avoir été notre "Mumu" durant plusieurs mois. Merci à **Sohini** pour avoir été mon collègue apprenti chercheur. Bon courage à toi pour la suite. Surtout n'abandonnes pas. Ça va aller !

Un grand merci à **Bernard REYNAUD**, Directeur de l'UMR PVBMT Cirad-Université de la Réunion ainsi qu'à tout le personnel Cirad. Ce fut un plaisir de partager ces cinq années de ma vie avec vous.

Je ne saurai terminer cette épopée sans remercier **B James NEYA,** responsable de notre LVBV et le personnel INERA, Kamboinsé, Burkina Faso. Merci à mes deux ainés **Ezéchiel** et **Monique** mes collègues de bureau ainsi qu'à tous les doctorants de la LVBV. J'ai vraiment apprécié votre compagnie et vos soutiens multiformes.

Sommaire

Liste des acronymes viraux selon le comité international de taxonomie virale i

Liste des abréviations ii

Liste des figures vi

Liste des tableaux ix

Introduction 1

Chapitre I : Epidémiologie moléculaire des geminivirus infectant les cultures maraîchères au Burkina Faso et en Côte d'Ivoire 45

Article 1 48
Molecular epidemiology of vegetable-infecting geminiviruses in Burkina Faso

Article 2 85
Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso

Article 3 88
New strains of chickpea chlorotic dwarf virus discovered on diseased papaya and tomato plants in Burkina Faso

Article 4 92
First report of Pepper yellow vein Mali virus associated with pepper yellow vein disease in Cote d'Ivoire

Article 5 93
First reports of Cotton leaf curl Gezira virus and Okra yellow crinkle virus associated with okra leaf curl disease in Côte d'Ivoire

Chapitre II : Evaluation des principaux paramètres épidémiologiques associés à l'émergence du PepYVMLV au Burkina Faso 94

Article 6 96
Stronger together: recruitment of a DNA-B component by an emerging monopartite begomovirus

Chapitre III : Diversité et dynamiques des populations de geminivirus à l'échelle des systèmes agroécologiques : approche metagénomique ciblée 121

Article 7 125
Metagenomics approach to uncover geminiviruses diversity and community structure at the scale of agro-ecological system in two localities of Burkina Faso

Discussion générale 163

Références bibliographiques 174

Annexes 211

Liste des acronymes viraux selon le comité international de taxonomie virale

ACMBFV :	*African cassava mosaic Burkina Faso virus*
ACMV :	*African cassava mosaic virus*
ALCCMA :	*Ageratum leaf curl Cameroon alphasatellite*
BDMV :	*Bean dwarf mosaic virus*
ChYMBJB :	*Chayote yellow mosaic Benin betasatellite*
CLCuGA :	*Cotton leaf curl Gezira alphasatellite*
CLCuGB :	*Cotton leaf curl Gezira betasatellite*
CLCuGV :	*Cotton leaf curl Gezira virus*
CpCAV :	*Chickpea chlorosis Australia virus*
CpCDV :	*Chickpea chlorotic dwarf virus*
CpCV :	*Chickpea chlorosis virus*
CpRLV :	*Chickpea redleaf virus*
CpYDV :	*Chickpea yellow dwarf virus*
CpYV :	*Chickpea yellows virus*
EACMKV :	*East African cassava mosaic Kenya virus*
EACMV :	*East African cassava mosaic virus*
EACMV-UG :	*East African cassava mosaic virus-Uganda*
EACMZV :	*East African cassava mosaic Zanzibar virus*
HoLCrV :	*Hollyhock leaf crumple virus*
MeAV :	*Melochia associated virus*
MSV :	*Maize streak virus*
OLCCMV :	*Okra leaf curl Cameroon virus*
OLCuMLA:	*Okra leaf curl Mali alphasatellite*
OLCuMLB :	*Okra leaf curl Mali betasatellite*
OLCV :	*Okra leaf curl virus*
OYCrCMA :	*Okra yellow crinkle Cameroon alphasatellite*
OYCrV :	*Okra yellow crinkle virus*
PepYVMLV :	*Pepper yellow vein Mali virus*

SPLCV :	*Sweet potato leaf curl virus*
TbLCZV :	*Tobacco leaf curl Zimbabwe virus*
ToLCBFV :	*Tomato leaf curl Burkina Faso virus*
ToLCCMV :	*Tomato leaf curl Cameroon virus*
ToLCGHB :	*Tomato leaf curl Ghana betasatellite*
ToLCGHV :	*Tomato leaf curl Ghana virus*
ToLCGV :	*Tomato leaf curl Gujarat virus*
ToLCMLV :	*Tomato leaf curl Mali virus*
ToLCNDV :	*Tomato leaf curl New Dehli virus*
ToLCNGV :	*Tomato leaf curl Nigeria virus*
ToLCTGB :	*Tomato leaf curl Togo betasatellite*
ToLCV :	*Tomato leaf curl virus*
ToMoV :	*Tomato mottle virus*
TYDV :	*Tobacco yellow dwarf virus*
TYLCMLV :	*Tomato yellow leaf curl Mali virus*
TYLCTHV :	*Tomato yellow leaf curl Thailand virus*
TYLCV :	*Tomato yellow leaf curl virus*
TYLCV-IL :	*Tomato yellow leaf curl virus-Israel*
WAAV1 :	*West Africa Asystasia begomovirus 1*

Liste des abréviations

a) Organismes et institutions

3P :	Pôle de Protection des Plantes
AIRD :	Agence Inter-établissements de Recherche pour le Développement
CIRAD :	Centre de coopération Internationale en Recherche Agronomique pour le Développement
GenBank :	Banque de données génomiques des Etats-Unis d'Amérique
ICTV :	*International Committee on Taxonomy of Viruses*

INERA :	Institut de l'Environnement et de Recherches Agricoles
IRD :	Institut de Recherche pour le Développement
LVBV :	Laboratoire de Virologie et Biotechnologie Végétale
NCBI :	National Center of Biotechnology Information
UFR/SVT :	Unité de Formation et de Recherche/Sciences de la Vie et de la Terre
UMR/PVBMT :	Unité Mixte de Recherche/Peuplements Végétaux et Bio-agresseurs en Milieu Tropical

b) Autres abréviations

AAP :	Acquisition Access Period
ADN :	Acide Désoxyribonucléique
ADN-A :	Composant A du génome des bégomovirus
ADN-B :	Composant B du génome des bégomovirus
ADNdb :	Acide désoxyribonucléique double brin
ADNsb :	Acide désoxyribonucléique simple brin
AnSL :	Biotype *Sub-Saharan Africa non-Silver Leafing* (mouches blanches)
ARN :	Acide Ribonucléique
ARNm :	Acide Ribonucléique messager
ASL :	Biotype *Sub-Saharan Africa Silver Leafing* (mouches blanches)
***BamH*I :**	Enzyme de restriction produit par la souche bactérienne *Bacillus amyloliquefaciens*
cccDNA :	Forme super-enroulée de l'ADN viral
CMD :	*Cassava mosaic disease*
CMGs :	*Cassava mosaic geminiviruses*
COI :	Cytochrome oxydase I
CP :	Protéine de Capside virale
CR :	Common region
dNTP :	désoxynucléotide triphosphate

***EcoR*I :**	Enzyme de restriction produit par la souche bactérienne *Escherichia coli*
ELISA :	*Enzyme Linked Immuno Sorbent Assay*
Genn. :	Gennadius
IAP :	*Inoculation Access Period*
IR :	*Intergenic Region* ou Région intergénique
LIR :	*Long intergenic region*
MEAM :	Espèce cryptique Middle East-Asia Minor (mouches blanches)
MED :	Espèce cryptique Mediterranean (mouches blanches)
$MgCl_2$:	Dichlorure de magnésium
ML :	Maximum-likelihood
MP :	Protéine de Mouvement
MSD :	*Maize streak disease*
NGS :	New Generation Sequencing
NSP :	*Nuclear Shuttle Protein*
OLCD :	*Okra leaf curl disease*
ORF :	*Open Reading Frames* ou Cadre ouverte de lecture
Ori :	Origine de réplication
pb :	Paire de bases
PCR :	Réaction de Polymérisation en Chaine
PDR :	*Pathogen-derived resistance*
***Pst*I :**	Enzyme de restriction produit par la souche bactérienne *Providencia stuartii*
PTGS :	*Post-Transcriptional Gene Silencing*
PYVD :	*Pepper yellow vein disease*
RA :	*Random Amplification*
RCA :	*Rolling Circle Amplification*
RCR :	*Rolling Circle Replication* ou Réplication en Cercle Roulant
RDR :	*Recombination-Dependent Replication* ou Réplication dépendante de la Recombinaison
REn :	***R**eplication **En**hancer* ou Protéine activatrice de la réplication
Rep :	Protéine de Réplication

SCR :	*Satellite Conserved Region*
SDT :	Strain Demarcation Tools
SGS :	*Second Generation Sequencing*
SIR :	*Short intergenic region*
siRNA :	*Small interfering Ribonucleic Acid*
SRAS :	Syndrome respiratoire aigu sévère
ßC1 :	Protéine multifonctionnelle intervenant dans la sévérité de la maladie
ssDNA :	Single-stranded circular DNA
TAS-ELISA :	*Triple Antibody Sandwich-Enzyme Linked Immuno Sorbent Assay*
TGS :	*Transcriptional Gene Silencing*
ToLCD :	*Tomato leaf curl disease*
TrAP :	***T**ranscriptional **A**ctivator **P**rotein* ou Protéine activatrice de la transcription
TYLCD :	*Tomato yellow leaf curl disease*
Vα :	brin viral

Liste des figures

Introduction générale

Figure 1 : (a) Reconstruction par cryo-microscopie électronique de la particule virale du maize streak virus (MSV). (b) Particules virales purifiées de MSV montrant leur forme icosaédrique associée en doublets **3**
Figure 2 : Organisation génomique des géminivirus. Les ORFs sont colorés conformément à la fonction des protéines correspondantes. **5**
Figure 3 : Schématisation du processus de réplication des géminivirus. **10**
Figure 4 : Couple d'individus adultes de Bemisia tabaci en alimentation sur une feuille. . **14**
Figure 5 : Arbre phylogénétique des protéines Rep associées à la réplication en cercle roulant basée sur la méthode du maximum de vraisemblance. **14**
Figure 6 : Structuration géographique des bégomovirus. **16**
Figure 7 : Phylogéographie des bégomovirus décrits en Afrique de l'Ouest et Centrale ... **20**
Figure 8 : Graphe biparti des liens entre les mastrévirus et leurs plantes hôtes monocotylédones. **22**
Figure 9 : Comparaison entre taux de mutation et de substitution des virus selon la nature de leur génome. **24**
Figure 10 : Phylogénie de Bemisia tabaci sur la base des séquences uniques du gène mitochondial du cytochrome oxydase I. **27**
Figure 11 : Schéma des principaux groupes de virus de plantes transmis par les aleurodes, en fonction de la structure de leurs particules virales, de leur mode de transmission et de leur organisation génomique. **29**
Figure 12 : Distribution géographique des biotypes de *Bemisia tabaci* collectés sur des plantes cultivées (coton, courge, gombo, tabac, tomate) et non cultivées (*Lantana camara*) en Afrique de l'Ouest. **30**
Figure 13 : Vue dorsale d'un adulte **(a)** et d'une nymphe **(b)** d'*Orosius albicinctus*..... **31**
Figure 14 : Symptômes **(a)** de la maladie de l'enroulement foliaire du gombo (*okra leaf curl disease*, OLCD), **(b)** de la maladie du jaunissement des nervures du piment (*pepper yellow vein disease*, PYVD) et **(c)** de la maladie de l'enroulement (et du jaunissement) foliaire de la tomate (*tomato (yellow) leaf curl disease*, ToLCD-TYLCD) observés au Burkina Faso. **33**
Figure 15 : Photographie en cryo-microscopie électronique de trichomes sur la surface foliaire de la tomate. Les différents types de trichomes présents sont indiqués par des flèches. **36**
Figure 16 : Diversité des formes d'acides nucléiques présentes dans la cellule végétale infectée par un phytovirus pouvant servir de cible aux approches métagénomiques. **40**

Chapitre I

Article 1

Figure 1 : (a) Representation of the 48 locations surveyed from 2013 to 2016 in the three agro-ecological areas. **(b)** Barplots showing variable proportions of geminivirus infecting vegetables and uncultivated plants in the all surveyed locations in Burkina Faso. **53**
Figure 2: Three predominant symptom phenotypes observed on tomato plants surveyed between 2013-2016 ..... **56**
Figure 3: Distribution of tomato-infecting geminivirus according to the agroclimatic areas and the ten surveyed locations..... **58**
Figure 4: Maximum-likelihood phylogenetic tree inferred from alignments of complete **(a)** DNA-A and **(b)** DNA-B sequences of begomovirus groups..... **61**
Figure 5: Maximum-likelihood phylogenetic tree inferred from alignments of complete alphasatellite (**a, b**) and betasatellite (**c**) sequences associated with begomovirus groups.. **64**
Supplementary figure 1: Color-coded matrix of pairwise nucleotide identity inferred from alignments of complete DNA-B sequences associated with PepYVMLV and selected DNA-B from NCBI. **71**

Article 2

Figure 1 : Symptoms of severe leaf curling observed on the tomato plant from which the isolate of tomato leaf curl Burkina Faso virus was collected. **86**
Figure 2: Maximum-likelihood phylogenetic tree and colour-coded matrix of pairwise nucleotide identity inferred from alignments of representing the complete DNA-A-like sequences of ToLCBFV and selected Old World monopartite begomoviruses originating from Africa **86**

Article 3

Figure 1 : Maximum-likelihood phylogenetic tree of full genome sequences of two new strains of chickpea chlorotic dwarf virus from tomato and papaya in Burkina Faso with representative sequences of 16 strains of CpCDV. **90**

Article 4

Figure 1 : Severe yellowing, leaf curling and deformation symptoms observed on pepper (Capsicum spp.) across Côte d'Ivoire. **92**
Figure 2: Maximum-likelihood phylogenetic tree of full genome sequences of Côte d'Ivoire isolates of PepYVMLV with publicly available begomovirus genome sequences **92**

Article 5

Figure 1 : Severe symptoms of leaf curling, deformation, and vein thickening observed on okra in south-eastern Côte d'Ivoire..... **93**
Figure 2: Maximum-likelihood phylogenetic tree of full genome sequences of Côte d'Ivoire isolates of CLCuGeV and OYCrV with publicly available begomovirus genome sequences..... **93**

Chapitre II

Article 6

Figure 1: Disease symptoms on **(a)** Nicotiana benthamiana and **(b)** tomato (Solanum lycopersicum, cv. Farmer 209, Known-you seed) agroinoculated with PepYVMLV in single (DNA-A) or mixed (DNA-A and -B) infection at 30 days post inoculation..... **104**

Figure 2: Kinetics of estimated symptom severity (solid line) of tomato (yellow) leaf curl disease following agroinoculation of tomato plants with PepYVMLV DNA-A and TYLCV-IL DNA-A in single or mixed infection with PepYVMLV DNA-B. **106**

Figure 3: Mean size (in centimeters) of tomato plants after agroinfection with PepYVMLV DNA-A or TYLCV-IL in single or mixed infection with PepYVMLV DNA-B at 32 days post inoculation. **107**

Figure 4: (a) Average of normalized viral DNA-A accumulation of PepYVMLV in single (DNA-A) or mixed (DNA-A and -B) infection of 10 tomato plants (Solanum lycopersicum, cv. Farmer 209, Known-You Seed). **(b)** Graph representing the linear correlation between PepYVMLV DNA-A and -B loads. **108**

Chapitre III

Article 7

Figure 1: Map of locations surveyed in Burkina Faso based on grid sampling strategy and diversity of collected plant samples..... **129**

Figure 2: (a) Venn diagram illustrating the comparison of number of plants detected as host of geminivirus using classical PCR and RCA-RA-SGS method. Bipartite diagram showing a complex of begomoviruses detected from **(b)** the 87/372 positive plants by PCR and **(c)** the 154/372 positive plants by RCA-RA-SGS. **133**

Figure 3: (a) Genome organization of MeAV comprising the six ORFs (V1, V2, C1, C2, C3, C4), the long intergenic region (LIR) including the origin of replication, the short intergenic region (SIR) and the common region (CR). **(b)** Colour-coded matrix of pairwise nucleotide identity inferred from alignments of complete DNA-A-like sequences of MeAV and selected Old World monopartite begomoviruses originating from Africa. **140**

Figure 4: Network showing geminivirus community structure in Goué **(a)** and Léguéma **(b)** in Burkina Faso. **(c)** Interconnection between plants sharing similar geminiviruses. . **142**

Supplementary figure 1: Metagenomic data analysis pipeline. **147**

Supplementary figure 2: Bipartite graph showing the diversity of geminivirus from Goué and the corresponding plant genera from which viral hits were obtained. **148**

Supplementary figure 3: Bipartite graph showing the diversity of geminivirus from Léguéma and the corresponding plant genera from which viral hits were obtained. **150**

Supplementary figure 4: Geographical distribution of distinct genera of plants collected from Goué and Léguéma sampling grids. **152**

Supplementary figure 5: Maximum-likelihood phylogenetic tree inferred from alignments of **(a)** begomovirus DNA-A/DNA-A like sequences, **(b)** alphasatellites and **(c)** betasatellites associated with begomovirus groups. **155**

Liste des tableaux

Introduction

Tableau 1 : Liste des 50 espèces de bégomovirus décrites en Afrique (ICTV, 2017) .. **18**

Chapitre I

Article 1

Table 1: Prevalence of tomato leaf curl disease and geminiviruses in tomato fields sampled in ten locations in Burkina Faso from 2015 to 2016. ..... **57**
Supplementary Table 1: Specific primers designed for geminivirus detection **72**
Supplementary table 2: Results of the PCR diagnosis performed with the universal begomovirus primers on the samples collected between 2013 and 2016..... **73**
Supplementary table 3: List of complete sequences obtained using RCA-RFLP protocol for geminivirus amplification, cloning and sequencing **74**

Article 3

Table 1: Geographical origin of leaf samples collected in Burkina Faso and description of chickpea chlorotic draft virus (CpCDV) isolates. **89**

Article 4

Table 1: Origine of pepper samples sampled in Côte d'Ivoire and results of PCR testing. **92**

Article 5

Table 1: Geographical origin of okra samples collected in Côte d'Ivoire and description of Cotton leaf curl Gezira virus and Okra yellow crinkle virus isolates. **93**

Chapitre II

Article 6

Table 1: Infectivity of PepYVMLV and TYLCV-IL DNA-As in single or mixed association with PepYVMLV DNA-B after agroinoculation of *N. benthamiana* and tomato plants. **105**
Table 2: Estimated parameters [95% confidence intervals] of the logit growth models governing symptom progression for tomato plants agroinoculated with PepYVMLV and TYLCV-IL in single or mixed inoculation with PepYVMLV DNA-B. **106**
Table 3: Transmission rates of PepYVMLV by *Bemisia tabaci* MEAM1..... **109**
Supplementary table 1: Primers, probes and standards developed and used in quantitative PCR..... **115**

Chapitre III

Article 7

Table 1: Percentage of infected plants collected from Goué and Léguéma after metagenomic virus detection. **136**

Supplementary table 1: Table presenting main virus(es) associated with each assignation profiles-based cluster obtained from samples collected from Goué and Léguéma using grid sampling strategy. **156**

Annexes

Supp. table 1: Summary of the 2063 collected samples from 2013 to 2016 in Burkina Faso **211**

Supp. table 2: General information about the 1068 samples collected in Goué and Léguéma using sampling grids in 2016. ..... **217**

Supp. table 3: General information about the 372 samples randomly collected in Burkina Faso from 2013 to 2015. ..... **220**

Supp. table 4: List of the 270 positive samples collected from Goué and Léguéma using grid sampling strategy. **223**

Introduction

Introduction

1. Contexte général

"Les maladies infectieuses ne disparaîtront jamais. Il en naîtra toujours de nouvelles; il en disparaîtra lentement quelques-unes; celles qui subsisteront ne se montreront plus sous la forme que nous connaissons aujourd'hui"

Charles Nicolle – Destin des maladies infectieuses, 1933

Lauréat du Prix Nobel de Médecine

Les maladies associées à des populations virales émergentes font l'actualité de ce début de siècle. On peut citer en exemple quelques maladies virales émergentes de l'Homme (Syndrome respiratoire aigu sévère, SRAS, Holmes & Rambaut, 2004 ; dengue, Farrar *et al.*, 2007; chikungunya, Pialoux *et al.*, 2007; Ebola, Coltart *et al.*, 2017; Zika, Saiz *et al.*, 2017), des animaux (fièvre aphteuse Grubman & Baxt, 2004) et des plantes (enroulement foliaire et jaunissement de la tomate, Navas-Castillo *et al.*, 2011; striure brune du manioc, Patil *et al.*, 2015; nécrose létale du maïs, Mahuku *et al.*, 2015). Cependant, même si les virus ont longtemps été regardés comme des parasites ou des pathogènes, les dernières avancées en métagénomique virale ont permis de démontrer que les virus font partie intégrante du vivant et qu'ils représentent le groupe d'organisme vivant le plus abondant et diversifié au Monde (Suttle, 2005; Roossinck, 2013).

L'émergence d'une nouvelle maladie virale apparaît aujourd'hui comme un processus naturel associé aux déplacements des équilibres écologiques aux seins des communautés hôtes/virus. Ainsi, parmi le large spectre d'association entre virus et hôtes, incluant la symbiose et le mutualisme, l'association antagoniste aboutissant à l'expression de maladie serait minoritaire. En effet, durant des millions d'années, de nombreux phytovirus ont co-évolué avec leurs hôtes sauvages (Cooper & Jones, 2006). Ce processus évolutif progressif a permis aux virus de s'adapter de façon très fine aux évolutions de leurs hôtes, en intégrant les aspects les plus complexes de leur biologie. La domestication des plantes a profondément perturbé cette coévolution, en modifiant

rapidement les relations entre les plantes cultivées, les virus et leurs vecteurs (Jones, 2009). Ces perturbations de l'équilibre des pathosystèmes ont été exacerbées par le transport des plantes cultivées de leurs centres de domestication vers des régions lointaines ou d'autres continents. Ces **invasions biologiques** ont favorisé de nouvelles rencontres entre les plantes cultivées et les virus. Associées aux pratiques agricoles modernes, tels que les monocultures, ils ont favorisé **l'émergence d'épidémies fréquentes et dommageables**, généralement considérées comme rares dans les communautés végétales indigènes.

La plupart des recherches *in natura* en évolution virale ont porté sur l'évolution de virus à ARN, qui représentent les plus importants agents pathogènes viraux humains et animaux en général. Leur succès en tant que virus émergent a souvent été attribué à des taux d'évolution très élevés, que ce soit par mutation ou recombinaison. En effet, même si il est toujours difficile de lier taux d'erreur et taux de substitution, le peu de fidélité des ARN polymérases et des transcriptases inverses serait un atout majeur pour la création de diversité et l'adaptation à de nouveaux hôtes et nouvelles conditions virus-hôte-environnement. Au contraire, chez les plantes la plupart des émergences virales sont liées aux virus à ADN simple brin (ADNsb) du genre *Begomovirus et Mastrevirus* de la famille des *Geminiviridae* (Seal *et al.*, 2006). Malgré leur réplication par les polymérases cellulaires de leur hôte, ayant un niveau de fidélité élevé, ces virus présentent un taux élevé de mutation et de recombinaison et affichent à l'échelle de la population ou de l'hôte des niveaux de diversité comparables à ceux de certains virus à ARN (Isnard *et al.*, 1998; Ge *et al.*, 2007; Duffy & Holmes, 2008). Le potentiel important d'évolution de ces virus à ADNsb, l'importante gamme de plantes hôtes monocotylédones et dicotylédones et la très large distribution géographique et gamme d'hôte de leurs insectes vecteurs ont fait de ces virus une menace majeure pour la sécurité alimentaire de nombreux pays dans le monde et d'Afrique en particulier. Outre l'impérieuse nécessité de lutter contre ces maladies, elles nous offrent l'opportunité de mieux comprendre les conditions et les facteurs associés à leur émergence et les mécanismes biologiques et moléculaires sous-jacents.

2. Géminivirus

2.1. Généralités et classification des géminivirus

La famille des *Geminiviridae* comprend des virus dont le génome est constitué d'une molécule (monopartite) ou de deux molécules (bipartite) d'ADN simple brin (ADNsb) protégées par une capside icosaédrique en doublet de 18-20 nm de diamètre et d'environ 30 nm de long (Figure 1) qui a donné son nom à cette famille virale (du latin *Geminus*, jumeau) (Zhang *et al.*, 2001; Jeske, 2009).

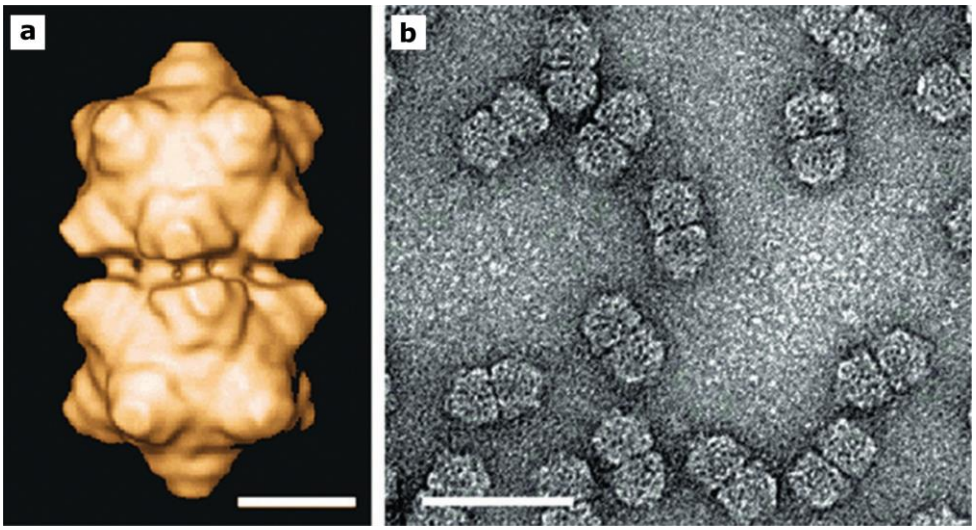


Figure 1 : (a) Reconstruction par cryo-microscopie électronique de la particule virale du maize streak virus (MSV). L'échelle représente 10 nm. D'après Zhang et al., 2001. (b) Particules virales purifiées de MSV montrant leur forme icosaédrique associée en doublets. L'échelle représente 50 nm. D'après Brown *et al.*, 2011.

Historiquement, ces virus avaient été classés en quatre genres, en fonction de leur gamme d'hôte, leur insecte vecteur et leur organisation génomique (Brown *et al.*, 2011), avec le genre ***Mastrevirus***, infectant principalement des plantes monocotylédones et transmis par des cicadelles, le genre ***Begomovirus***, infectant principalement des plantes dicotylédones et transmis exclusivement par l'aleurode *Bemisia tabaci*, le genre ***Curtovirus*** infectant des plantes dicotylédones et aussi transmis par des cicadelles et enfin le genre ***Topocuvirus*** infectant des plantes dicotylédones et transmis par un membracide (pour revue Rojas *et al.*, 2005). Grâce à l'essor des nouvelles méthodes de biologie moléculaire, et notamment le développement de protocoles de clonage et de

séquençage adaptés aux géminivirus (Inoue-Nagata *et al.*, 2004; Haible *et al.*, 2006; Shepherd *et al.*, 2008), de nombreuses nouvelles espèces ont été caractérisées, et cinq nouveaux genres ont été récemment approuvés dans la famille *Geminiviridae* par *l'International Committee on Taxonomy of Viruses* (ICTV, Zerbini *et al.*, 2017) : *Becurtovirus, Eragrosvirus* et *Turncurtovirus* (Varsani *et al.*, 2014) puis *Capulavirus, Grablovirus* (Varsani *et al.*, 2017). En outre, quatre autres espèces de virus très divergents qui sont étroitement liés aux géminivirus ont été acceptées par le Comité exécutif de l'ICTV pour inclusion dans la famille *Geminiviridae* et resteront non attribuées à un genre, en attendant l'identification des espèces vectrices et la confirmation la morphologie de la particule virale (Varsani *et al.*, 2017). Il s'agit des espèces *Apple geminivirus* (Liang *et al.*, 2015), *Citrus chlorotic dwarf associated virus* (Loconsole *et al.*, 2012; Guo *et al.*, 2015), *Grapevine geminivirus A* (Al Rwahnih *et al.*, 2017), *Mulberry mosaic dwarf associated virus* (Lu *et al.*, 2015; Ma *et al.*, 2015) et *Tomato associated geminivirus 1* (Fontenele *et al.*, 2017). Ces récentes découvertes ne font que confirmer la très grande diversité et plasticité d'organisation génomique des géminivirus et des virus à ADN circulaire simple brin en général (Krupovic, 2013; Labonté & Suttle, 2013; Roux *et al.*, 2013).

Parmi les neuf genres actuellement définis qui composent la famille des *Geminiviridae*, le genre *Begomovirus* dont l'espèce type est le ***Be**an **go**lden **mo**saic **virus*** comporte le plus grand nombre d'espèces décrites, selon le seuil de distinction taxonomique de 91 % d'identité nucléotidique (Brown *et al.*, 2015), avec plus de 320 espèces reconnues par l'ICTV (nombre en augmentation permanente). Il est suivit du genre *Mastrevirus* dont l'espèce type est le ***Ma**ize **stre**ak* **virus** (MSV) avec plus de 30 espèces reconnues (Zerbini *et al.*, 2017) sur la base d'un seuil distinction de 78 % d'identité nucléotidique (Muhire *et al.*, 2013).

2.2. Organisation génomique des genres *Begomovirus* et *Mastrevirus*

2.2.1. Organisation génomique des genres *Begomovirus*

Sur la base de leur origine, les bégomovirus ont été divisés en deux groupes : les bégomovirus du Nouveau Monde regroupant les pays du continent américain et de l'Ancien Monde impliquant le reste du monde (Abhary *et al.*, 2007). La plupart des bégomovirus originaires de l'Ancien Monde sont monopartites, alors que ceux originaires du Nouveau Monde sont presque exclusivement bipartites (Figure 2) (Duffy & Holmes, 2007; Melgarejo *et al.*, 2013; Sánchez-Campos *et al.*, 2013; Brown *et al.*, 2015).

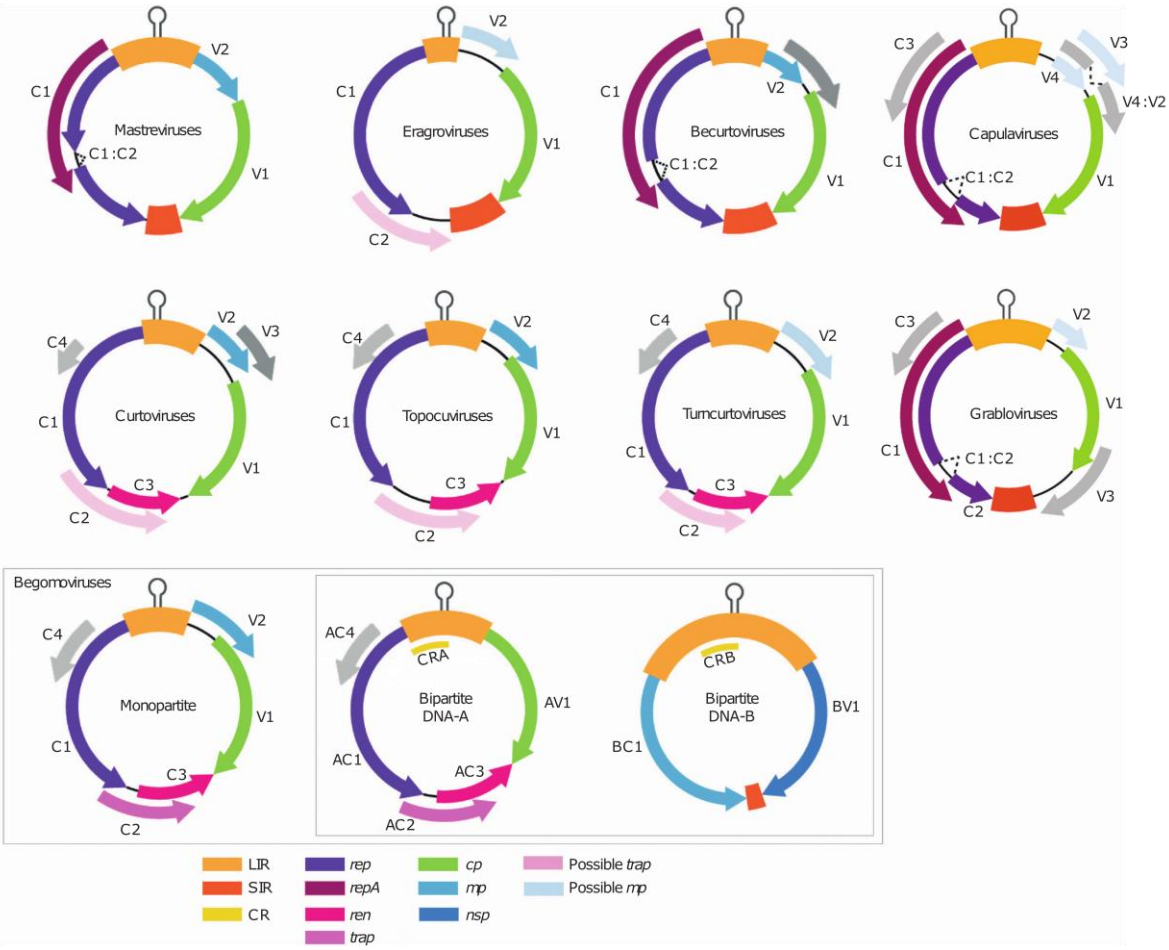


Figure 2 : Organisation génomique des géminivirus. Les ORFs sont colorés conformément à la fonction des protéines correspondantes. LIR, *long intergenic region*; SIR, *short intergenic region*; CR, *common region*. D'après Zerbini *et al.*, 2017.

Les bégomovirus monopartites possèdent un unique composant génomique, appelé ADN-A like, d'une taille d'environ 2800 pb tandis que le génome des bégomovirus bipartites est constitué de deux composants, appelé ADN-A et ADN-B, de taille similaire. Chacun des composants possède des gènes codant pour plusieurs protéines et une région intergénique (*Intergenic Region*, IR) d'environ 200 pb comprenant l'origine de réplication (ori), les itérons (courtes séquences répétées), la région commune (*Common Region*, CR) et la tige boucle (5' TAATATT↓AC 3') (Brown *et al.*, 2011; Zerbini *et al.*, 2017 ; Figure 2).

L'ADN-A porte une IR et six cadres ouverts de lecture (régions codantes) ou ORFs (*Open Reading Frames*) dont certains sont chevauchants (Brown *et al.*, 2011; Fondong, 2013). Dans le sens viral, l'ORF **V1** (**AV1** chez les bipartites) code pour la **P**rotéine de **C**apside **CP** qui représente l'unité de base dans la constitution de la particule virale en doublet des géminivirus. L'ORF **V2** (**AV2** chez les bipartites) code pour la **P**rotéine de **M**ouvement **MP** qui intervient dans les mécanismes de diffusion du virus dans la plante hôte (Brown *et al.*, 2011; Fondong, 2013). Les bégomovirus bipartites du Nouveau Monde ne possèdent pas l'ORF AV2 (Rojas *et al.*, 2001; Bull *et al.*, 2007; Fondong, 2013; Hanley-Bowdoin *et al.*, 2013). Dans le sens complémentaire, l'ORF **C1**/**AC1** code pour la protéine associée à la réplication (**Rep**). L'ORF **C2**/**AC2** code pour la **TrAP** (***T**ranscriptional **A**ctivator **P**rotein*) une protéine activatrice de la transcription et responsable de la transactivation des ORFs AV1 et BC1 (Sunter & Bisaro, 1997; Wezel *et al.*, 2001). La TrAP est également impliquée dans la suppression du mécanisme de *Post-Transcriptional Gene Silencing* (PTGS) (Voinnet *et al.*, 1999; Vanitharani *et al.*, 2005). La protéine **REn** (***R**eplication **En**hancer*) (ORF **C3**/**AC3**) est une protéine activatrice de la réplication (Hanley-Bowdoin *et al.*, 1999). La protéine **C4** codée par l'ORF **C4**/**AC4** semble être un déterminant majeur dans l'expression des symptômes, et également un suppresseur du PTGS (Krake *et al.*, 1998; Vanitharani *et al.*, 2005).

L'ADN-B porte deux ORFs qui codent pour deux protéines. La protéine **MP** (ORF **BC1**) est impliquée dans le mouvement de cellule à cellule de l'ADN viral tandis que la protéine **NSP** (*Nuclear Shuttle Protein*, ORF **BV1**) est impliquée dans le transport de l'ADN viral du noyau vers le cytoplasme à travers la membrane nucléaire (Sanderfoot *et al.*, 1996; Gafni & Epel, 2002).

Trois types de molécule d'ADNsb ont principalement été décrits en association avec les bégomovirus monopartites : les **alphasatellites** et les **betasatellites** d'une taille d'environs 1350 nucléotides chacun (Briddon & Stanley, 2006; Zhou, 2013) et les **deltasatellites** d'une taille d'environs 700 nucléotides (Lozano *et al.*, 2016). Les **alphasatellites** (précédemment connus sous le nom d'ADN-1, Briddon *et al.*, 2004) ne sont pas des satellites stricts, puisqu'ils sont capables d'autoréplication dans les cellules végétales, mais dépendent de leurs bégomovirus assistants, également nommés « *helper* », pour le mouvement intra-plante et la transmission inter-plantes via l'insecte vecteur (Mansoor *et al.*, 1999; Saunders & Stanley, 1999). Ces satellites qui codent pour une seule protéine associée à la réplication (alpha-Rep, Zhou, 2013), semblent moduler la virulence du complexe bégomovirus-betasatellite (Idris *et al.*, 2011) en diminuant, voir en supprimant l'expression des symptômes de certaines maladies à bégomovirus. Si les alphasatellites ont d'abord été identifiés dans l'Ancien Monde en association avec des bégomovirus monopartites (Saunders *et al.*, 2000; Briddon *et al.*, 2001), des données récentes décrivent des alphasatellites en association avec des bégomovirus bipartites dans l'Ancien (Patil & Fauquet, 2010; Harimalala *et al.*, 2013). et le Nouveau Monde (Paprotka *et al.*, 2010; Romay *et al.*, 2010).

Les **betasatellites** (précédemment connus sous le nom d'ADN-β) contiennent trois régions conservées : une séquence riche en Adénine (*A-rich sequence*), une séquence conservée connus sous le nom de région conservée des satellites (*satellite conserved region*, SCR) et un seul gène dans le sens complémentaire qui code pour une petite protéine (~118 acides aminés) connue sous le nom de βC1. La protéine βC1 est une protéine multifonctionnelle qui intervient dans la sévérité de la maladie et la suppression de certains mécanismes de résistance de la plante comme l'extinction génique transcriptionnelle (TGS pour *transcriptional gene silencing*) et l'extinction génique post-transcriptionnelle (PTGS pour *post-transcriptional gene silencing* ; Briddon et al., 2001; Cui et al., 2005; Zhou, 2013). Les betasatellites augmentent l'accumulation de leurs bégomovirus associés dans les plantes hôtes (Briddon, 2003; Zhou *et al.*, 2003). La seule séquence de similarité entre les betasatellites et leurs bégomovirus est le nanonucléotide TAATATTAC dans la tige-boucle faisant partie du SCR. Les betasatellites ont quasi-exclusivement étés identifiés en association avec des virus monopartites en Afrique et en Asie (Zhou, 2013) même si leur association avec un

virus bipartite a été décrite en Inde (Sivalingam & Varma, 2012).

Les **deltasatellites,** troisième classe d'ADN satellites en association avec des bégomovirus, a récemment été définit (Fiallo-Olivé *et al.*, 2012; Zhou, 2013; Lozano *et al.*, 2016). Ces satellites non codant, auparavant qualifiés de défectifs, comprennent entre autres le satellite associé au *Tomato leaf curl virus* un bégomovirus monopartite d'Australie (ToLCV-sat ; Dry *et al.*, 1997), des satellites identifiés à partir d'une collection de *Bemisia tabaci* en Floride (Ng *et al.*, 2011) et les satellites trouvés à Cuba (Fiallo-Olivé *et al.*, 2012). Cependant, ils sont distincts des alphasatellites et des betasatellites en raison de la petite taille de leur génome (~ 700 nucléotides) et de l'absence d'ORF. Ces satellites contiennent une tige boucle portant le nanonucléotide TAATATTAC, une TATA box et des itérons des bégomovirus dont une région riche en Adénine (A-rich) et une petite région partageant un pourcentage d'identité nucléotidique élevé avec des SCRs de betasatellites (Zhou, 2013). D'autres études sont nécessaires pour confirmer si ces molécules d'ADN satellites sont répliquées par leurs virus associés et si elles ont des rôles similaires à celles des betasatellites ou des alphasatellites dans la pathogénicité.

2.2.2. Organisation génomique du genre *Mastrevirus*

Les mastrévirus ont un génome monopartite (ADN-A) d'une taille d'environ 2700 pb qui comprend quatre ORFs, **V1** et **V2** codés par le brin viral et, **C1** et **C2** (ORFs chevauchants), portés par le brin complémentaire. Les ORFs V1 et V2 codent respectivement pour la **MP** et la **CP**. Le transcrit qui porte les ORFs C1 et C2 codent, après épissage, pour une **Rep** (Mullineaux *et al.*, 1990; Dekker *et al.*, 1991; Wright *et al.*, 1997). Deux IRs séparent les deux paires d'ORFs, une grande appelée LIR (*Large Intergenic Region*) qui porte l'ori et le promoteur bidirectionnel et une petite appelée SIR (*Small Intergenic Region*) qui contient les deux signaux de terminaison des transcrits (Figure 2). Récemment, un alphasatellite et un betasatellite ont été associés à un mastrévirus chez le blé. Ce fut la première fois que l'un de ces composants a été identifié chez une monocotylédone (Kumar *et al.*, 2014).

2.3. Réplication et transcription

Les géminivirus sont dépendants de la machinerie cellulaire de l'hôte pour leur réplication dont des facteurs comme l'ADN polymérase cellulaire. Les géminivirus se multiplient dans le noyau de cellules différenciées en phase G qui ont terminé leur activité de réplication propre. L'infection virale réactive la réplication en convertissant la cellule en phase S du cycle cellulaire (Hanley-Bowdoin *et al.*, 1999). Pour les mastrévirus, la synthèse du brin complémentaire après décapsidation de l'ADN viral se fait à partir d'une amorce d'ADN déjà présente et hybridée dans la région SIR (Figure 3). Pour les bégomovirus, il y a synthèse d'une amorce ARN complémentaire à la IR après décapsidation (Saunders & Stanley, 1999). La synthèse du brin complémentaire est initiée à partir de ces amorces (Figure 3). L'ADN db circulaire des deux types de virus en s'associant avec les histones de la cellule hôte, forme des minichromosomes viraux. Cette première étape est réalisée avec la seule utilisation des protéines de l'hôte (Pilartz & Jeske, 1992). Cet ADN db est alors capable de se répliquer en cercle roulant (*Rolling Circle Replication* ou RCR, Figure 3) de manière analogue aux phages à ADN circulaire simple brin (Novick, 1998). Après liaison à l'ADN au niveau des iterons, la protéine Rep initie le cycle de réplication en introduisant une ouverture (↓) au sein de la séquence conservée de la tige boucle (TAATATT↓AC) (Hanley-Bowdoin *et al.*, 1999). Après le clivage, la protéine Rep se retrouve liée de façon covalente au bout 5' de l'ADN clivé. La synthèse de l'ADN simple brin (ADNsb) est régulée par l'activité synergique des protéines TrAp et REn qui agissent respectivement sur l'activation de la transcription et l'augmentation de la réplication (Sunter & Bisaro, 1991). Selon Xie *et al.*, (1999), les géminivirus dépendent des facteurs cellulaires pour compléter leur cycle de réplication. Ce modèle de réplication en cercle roulant a été confirmé par microscopie électronique (Jeske *et al.*, 2001). Ces expériences ainsi que d'autres utilisant l'électrophorèse bidimensionnelle, ont par ailleurs permis l'identification d'intermédiaires additionnels de réplication compatibles avec un modèle de réplication dépendante de la recombinaison (*Recombination-Dependent Replication* ou RDR). Ce mécanisme utilise des intermédiaires réactionnels et est analogue à celui du bactériophage T4 (Mosig, 1998; Mosig *et al.*, 2001; Preiss & Jeske, 2003).

La transcription a aussi lieu dans le noyau de la cellule. Elle est bidirectionnelle depuis les séquences promotrices situées au sein de la zone intergénique pour tous les ORFs, excepté pour les ORFs C2 et C3 où la région promotrice est intégrée dans l'ORF C1. La transcription des bégomovirus est complexe, aboutissant fréquemment à des ARN messagers polycistroniques (Hanley-Bowdoin *et al.*, 1989; Sunter & Bisaro, 1989; Shivaprasad *et al.*, 2005).

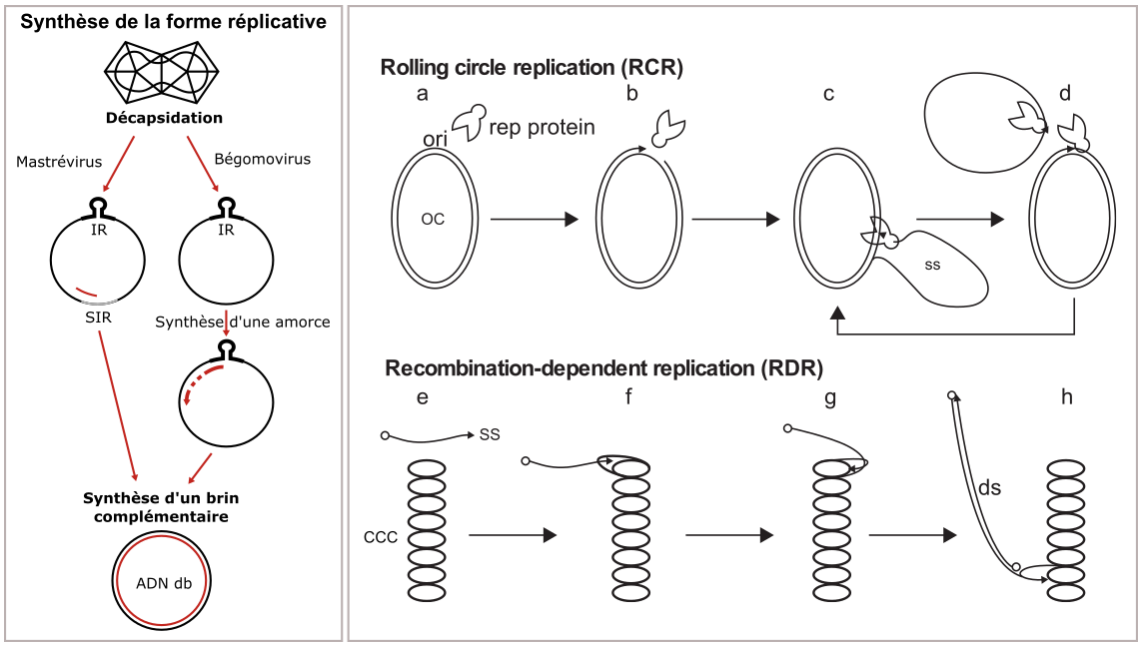


Figure 3 : Schématisation du processus de réplication des géminivirus.

Synthèse de la forme réplicative ADN db : Pour les mastrévirus après décapsidation de l'ADN viral (noir), la synthèse du brin complémentaire se fait à partir d'une amorce d'ADN (rouge) déjà présente et hybridée dans la région SIR (indiquée). Pour les bégomovirus, après décapsidation, la synthèse d'une amorce ARN (rouge) complémentaire à la région intergénique IR indiquée. La synthèse du brin complémentaire est initiée à partir de ces amorces. Les ADN db circulaires serviront dans la réplication. D'après Bernardi & Timchenko, 2008.

Réplication en cercle roulant (RCR) : Etape a : accrochage de la protéine associée à la réplication (Rep) à l'origine de réplication (ori). Etape b : ouverture de l'ADN et liaison covalente de la Rep à l'extrémité 5'. Etape c : déplacement et réplication. Etape d : nouvelle ouverture de l'ADN, fermeture des ADN simple brin et relargage de la Rep.

Réplication dépendante de la recombinaison (RDR) : Etape e : interaction entre un ADN simple brin incomplet et la forme super-enroulée de l'ADN viral (cccDNA) à des sites homologues. Etape f : recombinaison homologue. Etape g : élongation de l'ADN simple brin. Etape h : synthèse de l'ADN complémentaire et obtention d'un ADN double brin. D'après Jeske et *al.*, 2001.

3. Transmission circulante non propagative des géminivirus par des hémiptères

Les pièces buccales de la plupart des insectes vecteurs de virus témoignent d'un mode d'alimentation de type piqueur-suceur (Gray & Banerjee, 1999). Ce type morphologique est commun aux insectes qui s'alimentent de contenus cellulaires et surtout de sève phloémienne, tels que les aleurodes, cicadelles et pucerons (pour revue Blanc *et al.*, 2014). Leurs stylets leurs permettent de pénétrer dans les tissus et de traverser les parois cellulaires et les membranes cytoplasmiques par l'intermédiaire de forces mécaniques et/ou avec l'aide d'enzymes salivaires et intestinales. C'est au cours des périodes d'alimentation comportant des phases d'ingestion et de salivation qu'ils acquièrent et inoculent respectivement les particules virales. Les géminivirus sont considérés comme des virus circulants non multipliant, c'est à dire qu'ils circulent dans le corps de l'insecte et traversent plusieurs membranes cellulaires, celles de l'épithélium intestinal pour se retrouver dans le corps de l'insecte, et celles des glandes salivaires qui permettent leur évacuation par salivation dans une plante hôte (pour revue Ghanim, 2014).

Les bégomovirus sont transmis aux plantes par l'intermédiaire de l'insecte vecteur, *Bemisia tabaci* (Hemiptera : Aleyrodidae), selon le mode circulant persistant (Czosnek *et al.*, 2001). L'existence de cette spécificité de transmission suggère une association « intime » dans laquelle sont impliquées à la fois les protéines codées par le virus et le vecteur (pour revue Gray *et al.*, 2014). *B. tabaci* est un aleurode mesurant 1 à 1,5 millimètre de long qui arbore une couleur blanche-jaunâtre **(**Figure 4**).** Il appartient à un complexe d'espèces cryptiques communément appelées biotypes (Dinsdale *et al.*, 2010). Insecte phytophage, il semble se nourrir essentiellement de sève élaborée. Une fois ingéré par l'insecte via le bol alimentaire, le virus va pénétrer dans le tractus intestinal. Il va ensuite traverser une première barrière spécifique qui est la paroi de l'épithélium intestinal pour se retrouver dans l'hémolymphe de l'insecte (Ohnishi *et al.*, 2009; Uchibori *et al.*, 2013). Les particules virales vont ensuite s'associer puis traverser la deuxième barrière à la transmission que représentent les glandes salivaires. Le virus circulant peut ainsi être transmis à une nouvelle plante par salivation lors de

l'alimentation de l'insecte (Czosnek, 2007; Becker *et al.*, 2015). De nombreuses zones d'ombres persistent toutefois sur le mouvement du virus dans son vecteur et notamment sur les interactions cellulaires et moléculaires entre les bégomovirus et *B. tabaci*, la forme de circulation du virus dans l'insecte vecteur (particule virale, nucléoprotéine...) et l'existence de morphes viraux particuliers impliqués dans le mouvement à travers le vecteur et adaptés à la transmission (Blanc *et al.*, 2011).

Figure 4 : Couple d'individus adultes de *Bemisia tabaci* en alimentation sur une feuille. Source : CIRAD.

Les virus du genre *Mastrevirus* sont transmis par plusieurs espèces de cicadelles du genre *Cicadulina* (Hemiptera : Cicadellidae), qui se nourrissent et se reproduisent sur une large gamme d'espèces de *Poaceae* comprenant des poacées sauvages et cultivées. Ce genre est très homogène pour la structure du corps, de la tête et des ailes ainsi que pour de nombreux caractères de l'appareil reproducteur mâle, mais se distingue généralement par des variations de la coloration de la robe (Webb, 1987). Parmi ces cicadelles, *Cicadulina mbila* représente un vecteur particulièrement efficace pour la transmission du MSV sur maïs avec des taux de transmission individuelle de 100% (Reynaud, 1988). L'étude de son comportement alimentaire et de ses capacités d'acquisition et d'inoculation du MSV ont montré que *C. mbila* à la capacité d'ingérer massivement des particules virales à partir des cellules non phloémiennes (épiderme et

mésophylle), contrairement à *B. tabaci*. Cette particularité est compatible avec le fait que *C. mbila* est capable d'acquérir le MSV suite à une période d'alimentation de quelques secondes sur plantes virosées puis de le transmettre durant toute sa durée de vie (Storey, 1938; Reynaud, 1988; Lett *et al.*, 2001, 2002).

4. Origine, diversité, évolution et dissémination des géminivirus

4.1. Origine des géminivirus

Il existe un certain nombre de preuves suggérant que les géminivirus modernes auraient évolué à partir de réplicons d'ADN extra-chromosomiques présents chez les ancêtres procaryotes ou eucaryotes primitifs des plantes modernes (Koonin & Ilyina, 1992); Figure 5 d'après Rojas *et al*, 2005). Dans ce scénario, le géminivirus ancestral est envisagé sous la forme d'un plasmide extra-chromosomique circulaire d'ADN sb qui se serait répliqué par un mécanisme de cercle roulant impliquant une forme réplicative d'ADN db. Les relations phylogénétiques étroites des protéines Rep des géminivirus avec celles des protéines Rep des plasmides des phytoplasmes, ainsi que la ressemblance structurelle forte entre la CP des géminivirus et celle des virus icosaédriques à ARNsb ont permis de proposer un deuxième scénario évolutif dans lequel l'acquisition du gène codant pour la protéine de capside d'un virus de plante à ARNsb par un plasmide de phytoplasme a donné naissance à l'ancêtre des géminivirus (Figure 5, Krupovic *et al.*, 2009). Les principales preuves d'une origine très ancienne des géminivirus sont (i) la nature conservée des gènes protéiques de Rep parmi les réplicons d'ADN procaryotes et eucaryotes modernes et (ii) les caractéristiques des procaryotes présentes chez les géminivirus modernes (comme par exemples, les propriétés du promoteur de la CP, les ARNm polycistroniques et l'existence d'une capacité de réplication dans les cellules procaryotes). Ainsi, le ou les ancêtre(s) des géminivirus serai(en)t apparu(s), il y a des milliards d'années, sous une forme primitive ne codant que pour des protéines destinées à la réplication de l'ADN (pour revues Rojas *et al.*, 2005; Hull, 2014).

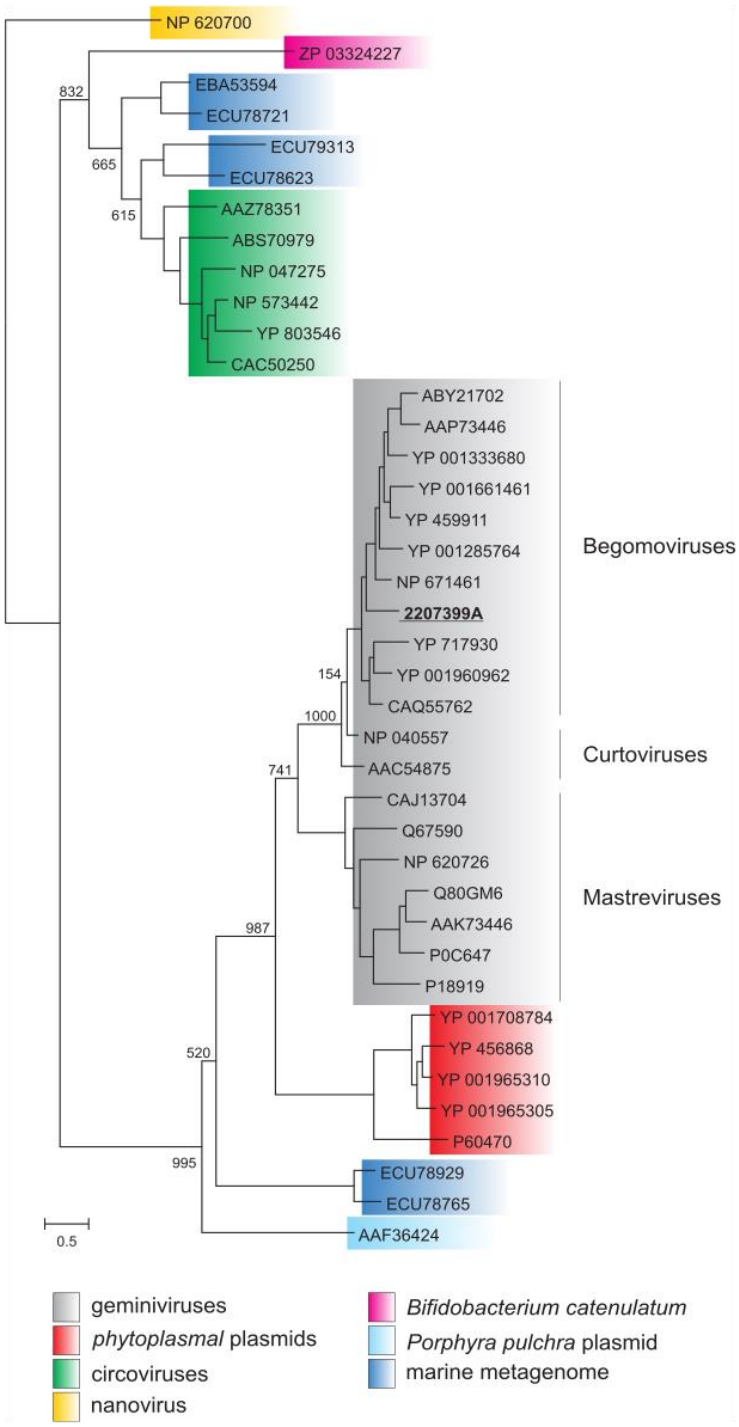


Figure 5 : Arbre phylogénétique des protéines Rep associées à la réplication en cercle roulant basée sur la méthode du maximum de vraisemblance. La Rep des nanovirus a été choisie comme un groupe externe pour enraciner l'arbre. La Rep de *Nicotiana tabacum* dérivée de *Geminivirus* est soulignée. La barre d'échelle représente le nombre de substitutions d'acides aminés par site. D'après Krupovic *et al*., 2009.

4.2. Diversité génétique des bégomovirus et mastrévirus

En raison de leur impact majeur sur le rendement de nombreuses cultures, l'intérêt pour les géminivirus n'a cessé d'augmenter depuis ces dernières décennies, aboutissant à la caractérisation fréquente de nouvelles espèces. Néanmoins, les efforts de caractérisation se sont principalement focalisés sur les plantes d'intérêt agronomique, fortement impactées par les maladies émergentes, et la diversité connue des bégomovirus, bien qu'impressionnante, ne représente probablement qu'une infime partie de leur diversité réelle. En effet, on considère que la majeure partie des maladies causées par les bégomovirus sont liées au passage de virus indigènes, ayant co-évolué avec leur(s) hôte(s) naturel(s), sur des plantes introduites notamment pour la culture, non adaptés à ces virus et donc souvent plus sensibles à leur infection. Les exemples les plus emblématiques sont les maladies de la mosaïque du manioc (Patil & Fauquet, 2009) et de la striure du maïs (Shepherd *et al.*, 2010b), toutes deux endémiques d'Afrique mais absente d'Amérique du Sud, l'aire d'origine des plantes infectées. De plus, de manière globale, la diversité des bégomovirus est structurée géographiquement (Briddon *et al.*, 2010 ; Figure 6), et différentes lignées virales semblent s'être adaptées indépendamment aux même plantes hôtes. Aussi, la diversité des espèces virales infectant un hôte est souvent très différente dans les zones géographiques d'introduction secondaire, de celle de son aire d'origine (Nawaz-ul-Rehman & Fauquet, 2009). Tous ces éléments suggèrent qu'une grande diversité de géminivirus reste encore à décrire à partir de plantes non-cultivées des écosystèmes naturels (Bernardo *et al.*, 2017).

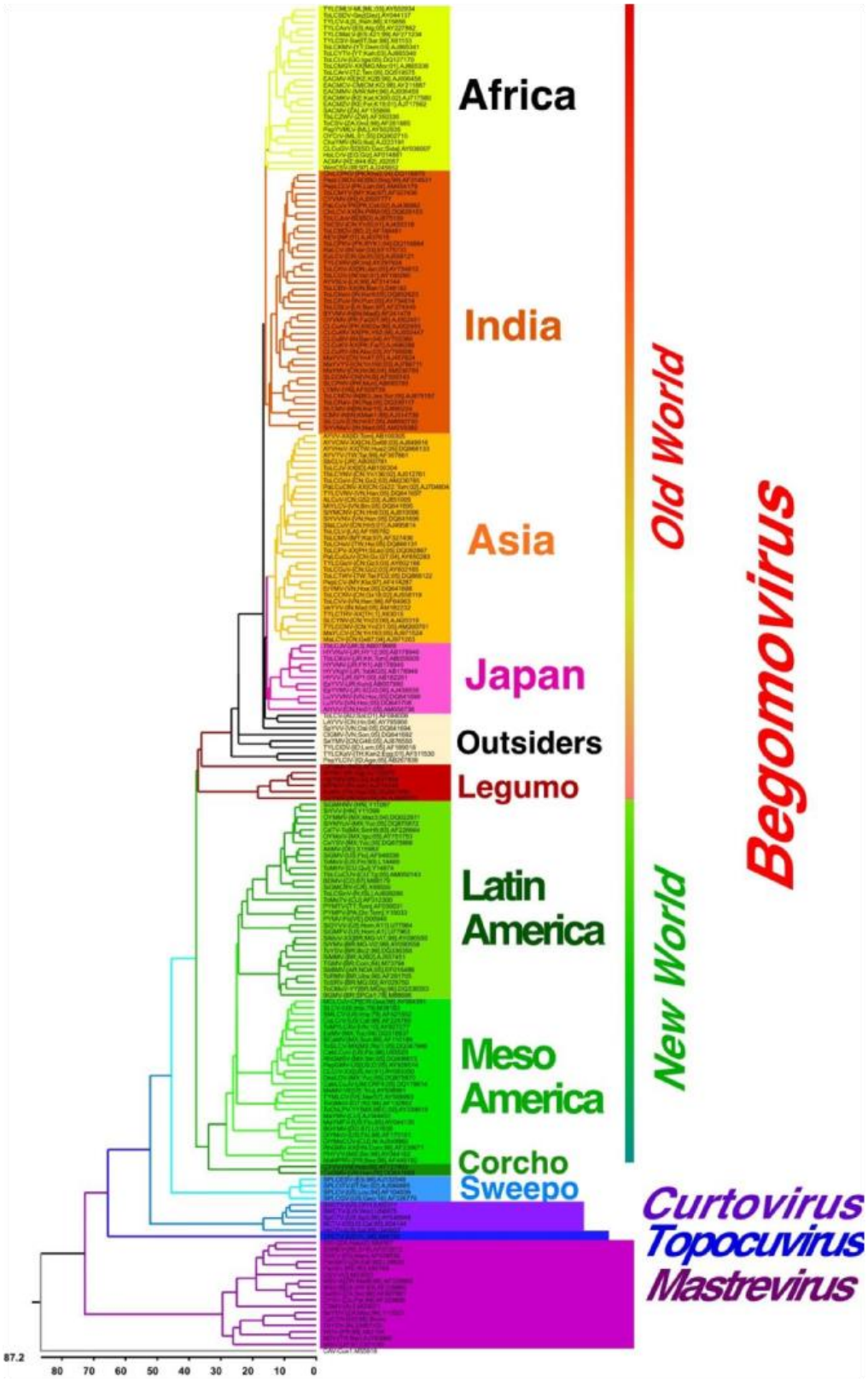


Figure 6 : Structuration géographique des bégomovirus. D'après Briddon *et al*. 2010.

4.2.1. *Bégomovirus*

Le biais dans la découverte de la diversité virale est particulièrement prégnant en Afrique. En effet, sur 322 espèces de bégomovirus décrites actuellement dans le monde (ICTV), les 50 espèces décrites en Afrique (Tableau 1) ont été caractérisées essentiellement à partir d'un nombre restreint de plantes cultivées, comme le manioc (*Manihot esculenta*), la tomate (*Solanum esculentum*), le gombo (*Abelmoschus esculentus)*, le tabac (*Nicotiana tabacum*), le haricot (*Phaseolus vulgaris*) et la patate douce (*Ipomoea batatas*) (Figure 7). Les bégomovirus africains les plus étudiés et ayant le plus important impact économique sont sans nul doute ceux associés à la maladie de la mosaïque du manioc (*cassava mosaic disease*, CMD; (Legg *et al.*, 2014b). Cette maladie, décrite dans toutes les régions d'Afrique sub-saharienne, est causée par un complexe d'au moins neuf espèces bipartites de bégomovirus dénommés « *cassava mosaic geminiviruses* » (CMGs, Legg et al., 2014a). La tomate est aussi l'une des cultures les plus fortement impactée à l'échelle mondiale par des maladies causées par des bégomovirus. Cette plante qui présente une très forte susceptibilité aux bégomovirus s'est révélée être un biomarqueur remarquable pour leur détection. Ainsi, plus de 90 espèces de bégomovirus ont été identifiés sur tomate à travers le monde. En Afrique, pas moins de 25 espèces ont été décrites, et de cette plante sont continuellement caractérisées de nouvelles espèces de bégomovirus (Rey *et al.*, 2012; Leke *et al.*, 2015). La majorité des espèces africaines infectant la tomate forment un groupe monophylétique, bien distinct des espèces retrouvées par ailleurs dans les autres aires d'introduction de la tomate ou de son aire d'origine, suggérant une nouvelle fois l'adaptation indépendante de lignées virales africaines à cette plante sensible (Rey *et al.*, 2012).

Six espèces de bégomovirus ont été associées aux maladies de l'enroulement et/ou du jaunissement foliaire de la tomate en Afrique de l'Ouest et Centrale. Il s'agit des espèces virales suivantes : *Pepper yellow vein Mali virus* (PepYVMLV ; Zhou *et al.*, 2008), *Tomato leaf curl Cameroon virus* (ToLCCMV, Leke *et al.*, 2011), *Tomato leaf curl Ghana virus* (ToLCGHV ; Osei *et al.*, 2008); *Tomato leaf curl Mali virus* (ToLCMLV ; Zhou *et al.*, 2008), *Tomato leaf curl Nigeria virus* (ToLCNGV ; Kon and

Gilbertson, 2012), *Tomato yellow leaf curl Mali virus* (TYLCMLV ; Chen *et al.*, 2009; Lett *et al.*, 2009; Sattar *et al.*, 2015).

Tableau 1 : Liste des 50 espèces de bégomovirus décrites en Afrique (ICTV, 2017)

Espèce virale	Pays	Acronyme	Numéro d'accession d'une séquence type
African cassava mosaic Burkina Faso virus #	Burkina Faso **	ACMBFV	HE616777
African cassava mosaic virus	Cameroun *****	ACMV	J02057; J02058
Ageratum leaf curl Cameroon virus #	Cameroun *****	ALCCMV	FR873230
Cassava mosaic Madagascar virus	Madagascar *	CMMGV	HE617299; HE617300
Chayote yellow mosaic virus	Nigeria **	ChaYMV	AJ223191
Cotton leaf curl Gezira virus	Burkina Faso **	CLCuGV-BF	FN554540
	Cameroun *****	CLCuGV-CM	FM210276
	Egypte ***	CLCuGV-EG	AF155064
	Madagascar *	CLCuGV-MG	AM701757
	Mali **	CLCuGV-ML	EU024120
	Niger **	CLCuGV-NE	EU432373
	Soudan ***	CLCuGV-SD	AF260241
Cowpea golden mosaic virus	Nigeria **	CPGMV	AF029217
East African cassava mosaic Cameroon virus	Cameroun *****	EACMCMV	AF112354; AF112355
East African cassava mosaic Kenya virus	Kenya *	EACMKV	AJ717572; AJ704971
East African cassava mosaic Malawi virus	Malawi *	EACMMV	AJ006460
East African cassava mosaic virus	Kenya *	EACMV-KE	AJ717542; AJ704949
	Malawi *	EACMV-MW	JX473582
	Tanzanie *	EACMV-TZ	AY795983; AY795989
	Uganda *	EACMV-UG	AF126806; AF126807
East African cassava mosaic Zanzibar virus	Tanzanie *	EACMZV	AF422174; AF422175
Hollyhock leaf crumple virus	Egypte ***	HoLCrV	AJ542539
Jatropha mosaic Nigeria virus	Nigeria **	JMNV	JX025358
Okra leaf curl Cameroon virus #	Cameroun *****	OLCuCMV	HE793427
Okra yellow crinkle virus	Cameroun *****	OYCrV	HE793424
	Mali **	OYCrV-ML	DQ875879
Pepper yellow vein Mali virus	Mali **	PepYVMLV	AY502935
South African cassava mosaic virus	Afrique du Sud ****	SACMV	AF155806; AF155807
Soybean chlorotic blotch virus	Nigeria **	SbCBV	GQ472985; GQ472986
Soybean mild mottle virus	Nigeria **	SbMMoV	GQ472984
Squash leaf curl virus	Egypte ***	SLCV	DQ285019
Sweet potato leaf curl Sao Paulo virus	Afrique du Sud ****	SPLCSPV	JQ621844
Sweet potato leaf curl Uganda virus	Uganda *	SPLCUV	FR751068
Sweet potato mosaic virus	Afrique du Sud ****	SPMaV	JQ621843
Tobacco leaf curl Comoros virus	Comores *	TbLCKMV	AM701760
Tobacco leaf curl Zimbabwe virus	Zimbabwe *	TbLCZV	AF350330
Tomato curly stunt virus	Afrique du Sud ****	ToCSV	AF261885
Tomato leaf curl Anjouan virus	Comores *	ToLCAnV	AM701758
Tomato leaf curl Antsiranana virus #	Madagascar *	ToLCAntV	AM701766
Tomato leaf curl Arusha virus	Tanzanie *	ToLCArV	EF194760
Tomato leaf curl Cameroon virus	Cameroun *****	ToLCCMV	HE659516
Tomato leaf curl Comoros virus	Madagascar *	ToLCKMV-MG	AM701761
	Mayotte *	ToLCKMV	AJ865340
Tomato leaf curl Diana virus	Madagascar *	ToLCDiV	AM701765
Tomato leaf curl Ghana virus	Cameroun *****	ToLCGHV-Ag	FR873228
	Ghana **	ToLCGHV	EU350585
Tomato leaf curl Madagascar virus	Madagascar *	ToLCMGV-Ats	AJ865339
	Madagascar *	ToLCMGV-Men	AJ865338
Tomato leaf curl Mali virus	Mali **	ToLCMLV	AY502936
Tomato leaf curl Mayotte virus	Mayotte *	ToLCYTV	AJ865340
Tomato leaf curl Moheli virus	Comores *	ToLCMohV	AM701763
Tomato leaf curl Namakely virus	Madagascar *	ToLCNaV	AM701764
Tomato leaf curl Nigeria virus	Nigeria **	ToLCNGV	FJ685621
Tomato leaf curl Seychelles virus	Seychelles *	ToLCSCV	AM491778

Tableau 1 : suite et fin

Espèce virale	Pays	Acronyme	Numéro d'accession d'une séquence type
Tomato leaf curl Sudan virus	Soudan ***	ToLCSDV	AY044137
Tomato leaf curl Toliara virus	Madagascar *	ToLCToV	AM701768
Tomato leaf curl Uganda virus	Uganda *	ToLCUV	DQ127170
Tomato yellow leaf curl Mali virus	Burkina Faso **	TYLCMLV	LM651400
	Cameroun *****	TYLCMLV	FM212662
	Ethiopie *	TYLCMLV-ET	DQ358913
	Ghana **	TYLCMLV	EU847740
	Mali **	TYLCMLV-ML	AY502934
Tomato yellow leaf curl Sardinia virus	Maroc ***	TYLCSV	AY702650
Tomato yellow leaf curl virus	Egypte ***	TYLCV	AY594174
	Maroc ***	TYLCV	EF060196
	Maurice *	TYLCV	HM448447
	Réunion *	TYLCV	AJ865337
	Soudan ***	TYLCV	AY044138
	Tunisie ***	TYLCV	EF101929
Watermelon chlorotic stunt virus	Soudan ***	WmCSV	AJ245650
West African Asystasia virus 1	Afrique de l'Ouest **	WAAV1	JF694484; JF694485
West African Asystasia virus 2	Afrique de l'Ouest **	WAAV2	JF694486

: Espèces décrites en Afrique absentes de la liste de l'ICTV
* : Pays d'Afrique de l'Est
** : Pays d'Afrique de l'Ouest
*** : Pays d'Afrique du Nord
**** : Pays d'Afrique du Sud
***** : Pays d'Afrique Centrale

La maladie de l'enroulement foliaire a également été décrite sur le gombo dans de nombreux pays d'Afrique sahélienne et ses environs (Burkina Faso Tiendrébéogo *et al.*, 2010 ; Cameroun, Leke *et al.*, 2013 ; Côte d'Ivoire, (N'Guessan *et al.*, 1992) ; Ghana Swanson & Harrison, 1993 ; Mali Kon *et al.*, 2009 ; Niger, Shih *et al.*, 2009 ; Nigéria, Tchad, Swanson & Harrison, 1993) et de la vallée du Nil (Egypte, Bigarré *et al.*, 2001). Plusieurs espèces de bégomovirus ont été associés à cette maladie, l'*Okra leaf curl virus* (OLCV) au Cameroun (Leke *et al.*, 2007), le *Cotton leaf curl Gezira virus* (CLCuGV) au Soudan, Niger, Côte d'Ivoire et Burkina Faso (Idris & Brown, 2002, 2005; Shih *et al.*, 2009; Tiendrébéogo *et al.*, 2010), l'*Okra yellow crinkle virus* (OYCrV) au Mali (Shih *et al.*, 2007; Kon *et al.*, 2009), et le *Hollyhock leaf crumple virus* (HoLCrV) en Egypte et au Soudan (Bigarré *et al.*, 2001; Idris & Brown, 2002).

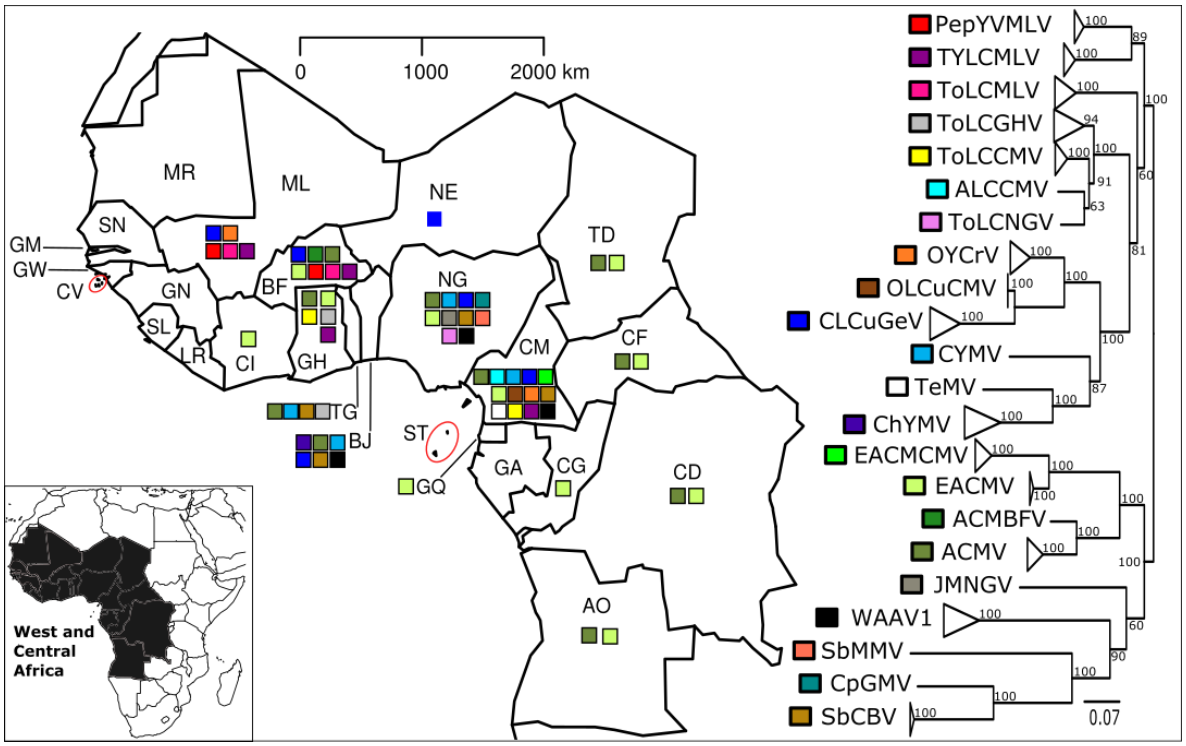


Figure 7 : Phylogéographie des bégomovirus décrits en Afrique de l'Ouest et Centrale. La répartition géographique des bégomovirus et la construction de l'arbre phylogénétique sont basées sur les séquences complètes d'ADN-A disponibles sur GenBank/EMBL/DDBJ (Octobre, 2017).

Mis à part le manioc, la tomate et le gombo, des bégomovirus ont été détectés dans d'autres plantes cultivées en Afrique. Une vingtaine d'espèces ont ainsi été isolées sur des plants de tabac, de coton (*Gossypium barbadense*), de piment (*Capsicum annuum*), de haricot (*Phaseolus vulgaris*), de chayotte (*Sechium edule*), de pignons d'Inde (*Jatropha curcas*) et sur plusieurs légumineuses (pour revues Rey *et al.*, 2012; Leke *et al.*, 2015). Enfin, le groupe divergent des bégomovirus monopartites décrits sur patate douce (*sweepovirus*) a également été décrit en Afrique du Sud (Esterhuizen *et al.*, 2012), au Kenya (Miano *et al.*, 2006), en Tanzanie (Mbanzibwa *et al.*, 2014), et en Ouganda (Wasswa *et al.*, 2011).

4.2.2. *Mastrévirus*

Les mastrévirus infectant les monocotylédones comptent de nos jours 25 espèces ayant été toutes caractérisées dans l'Ancien Monde (Figure 8). Douze d'entre elles ont été caractérisées en Afrique et dans les îles de l'océan Indien et neuf en Australie (Kraberger *et al.*, 2017). Trois espèces ont été caractérisées en Eurasie et une espèce au Vanuatu. Parmi les douze espèces de mastrévirus infectant les monocotylédones en Afrique, le MSV a été le plus étudié à cause de son effet dévastateur sur les cultures de maïs très importantes pour l'agriculture de subsistance.

Sept espèces de mastrévirus qui infectent les plantes dicotylédones ont été décrites (pour revue Kraberger *et al*., 2015). Le *Chickpea chlorotic dwarf virus* (CpCDV) a été caractérisé en Afrique, au Moyen-Orient, en Turquie et en Inde (Horn *et al.*, 1993; Liu *et al.*, 1997; Nahid *et al.*, 2008; Kraberger *et al.*, 2013) et le *Chickpea yellow dwarf virus* (CpYDV) au Pakistan (Kraberger, 2014). Les cinq autres espèces à savoir le *Chickpea chlorosis virus* (CpCV) (Thomas *et al.*, 2010; Hadfield *et al.*, 2012; Kraberger *et al.*, 2013), le *Chickpea chlorosis Australia virus* (CpCAV) (Hadfield *et al.*, 2011), le *Tobacco yellow dwarf virus* (TYDV) (Morris *et al.*, 1992; Hadfield *et al.*, 2012), le *Chickpea redleaf virus* (CpRLV) (Thomas *et al.*, 2010), et le *Chickpea yellows virus* (CpYV) (Hadfield *et al.*, 2012), pour l'instant, n'ont été décrits qu'en Australie (Kraberger, 2014). Parmi ces matrévirus infectant les dicotylédones, le CpCDV est le plus largement caractérisé avec 16 souches identifiées de nos jours (Kraberger, 2014; Kraberger *et al.*, 2017).

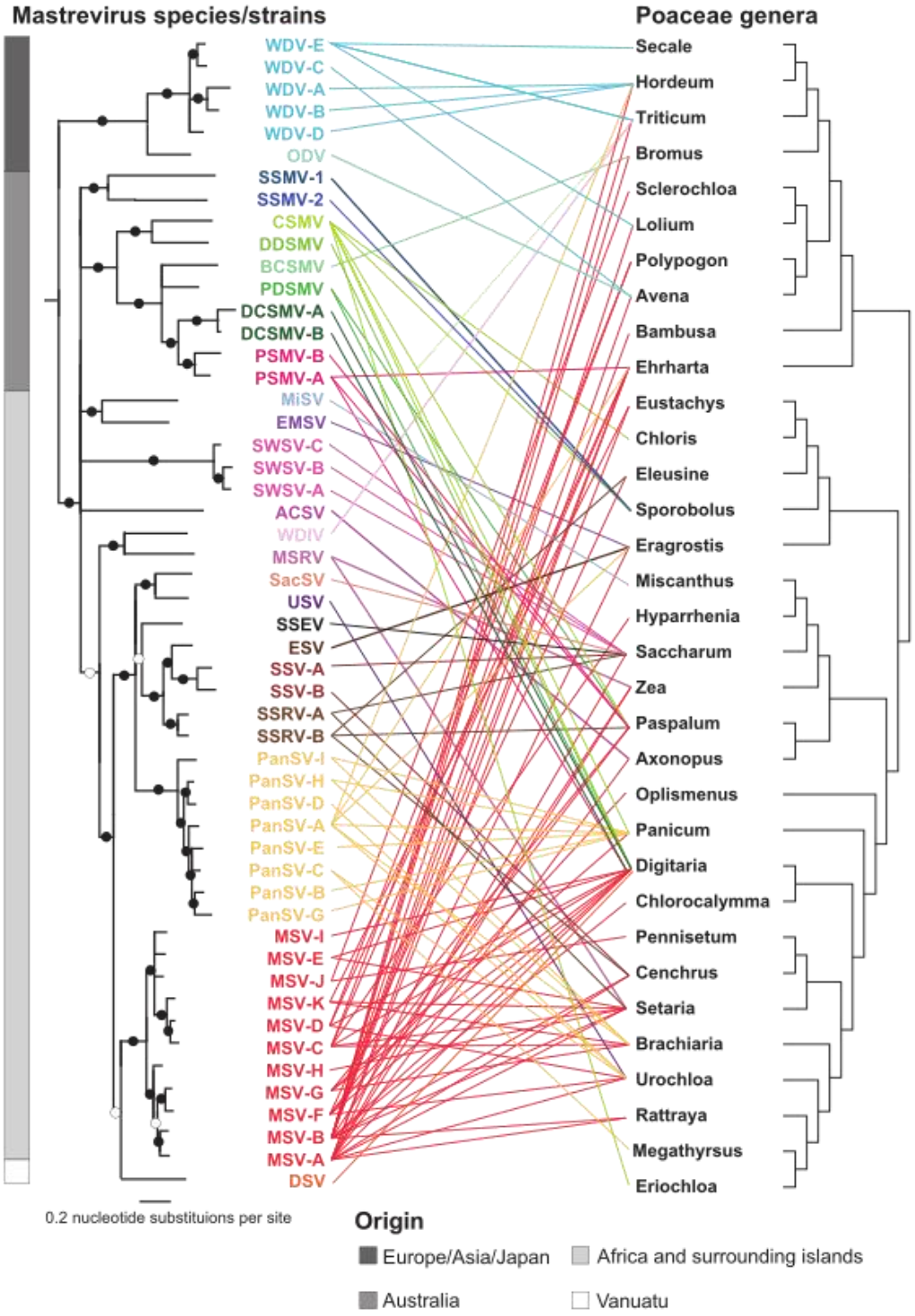


Figure 8 : Graphe biparti des liens entre les mastrévirus et leurs plantes hôtes monocotylédones. D'après Kraberger *et al.*, 2014.

4.3. Principaux processus évolutifs impliqués dans la diversification des géminivirus : Mutation et Recombinaison/Réassortiment

Le taux de substitution d'un virus est exprimé en nombre de substitutions par position nucléotidique et par unité de temps. Il représente le taux auquel les mutations se fixent dans les génomes. Il dépend du taux de mutation, qui correspond au nombre d'erreurs introduites par position nucléotidique et par évènement de réplication. Le taux de substitution pouvant être influencé par plusieurs processus en aval de la mutation, tels que par exemple la sélection, le lien entre ces deux valeurs n'est pas toujours direct. Si le taux de mutation est difficile à estimer et s'il dépend de plusieurs paramètres biochimiques, il est largement lié au taux d'erreur de la polymérase utilisée pour la réplication virale. Ces taux d'erreur varient grandement entre les ADN polymérases, présentant divers mécanismes de correction et des taux d'erreurs relativement faibles (entre 10^{-6} et 10^{-8} erreurs par nucléotide et par réplication; Garcia-Diaz and Bebenek, 2007) et les ARN polymérases aux taux d'erreurs nettement plus élevés (entre 10^{-3} et 10^{-5} erreurs par nucléotide et par réplication; Jenkins *et al.*, 2002). En conséquence, les virus à ADN, qui utilisent la polymérase de la cellule hôte, évolueraient moins rapidement que les virus à ARN, pouvant présenter des taux de substitution entre 10^{-2} et 10^{-5} substitutions par position nucléotidique et par année (Duffy *et al.*, 2008; Figure 9). Néanmoins, les estimations des taux de substitution de plusieurs espèces virales à ADN simple brin se sont révélées étonnement proches des taux de substitution des virus à ARN. Ainsi, chez les géminivirus, plusieurs études ont estimé des taux de substitution de l'ordre de 10^{-3} à 10^{-4} substitutions par position nucléotidique et par année (Isnard *et al.*, 1998; Ge *et al.*, 2007; van der Walt *et al.*, 2008; Duffy & Holmes, 2009). Plusieurs hypothèses ont été avancées pour expliquer ces taux élevés, tels que l'incompatibilité entre les enzymes de correction et le génome viral, la présence de structures secondaires favorisant les erreurs, ou bien le rôle potentiel de protéines virales dans l'altération de la fidélité de la polymérase (Gutierrez, 1999; Arguello-Astorga *et al.*, 2007).

Le taux de substitution élevé des bégomovirus en tant que moteur de leur diversification est associé à un deuxième processus évolutif majeur chez les géminivirus : la recombinaison. Elle correspond à la création de génomes hybrides à partir de portions

d'ADN issus de génomes parentaux (Lefeuvre, 2008). La recombinaison présente l'énorme avantage pour un virus de permettre l'acquisition en une seule étape d'une grande variabilité génétique et ainsi de permettre de créer de nouveaux types d'arrangements au sein du génome. Conjointement, mutation et recombinaison peuvent fournir l'accès à un polymorphisme beaucoup plus important que par mutation seule (Stemmer, 1994; Crameri *et al.*, 1998).

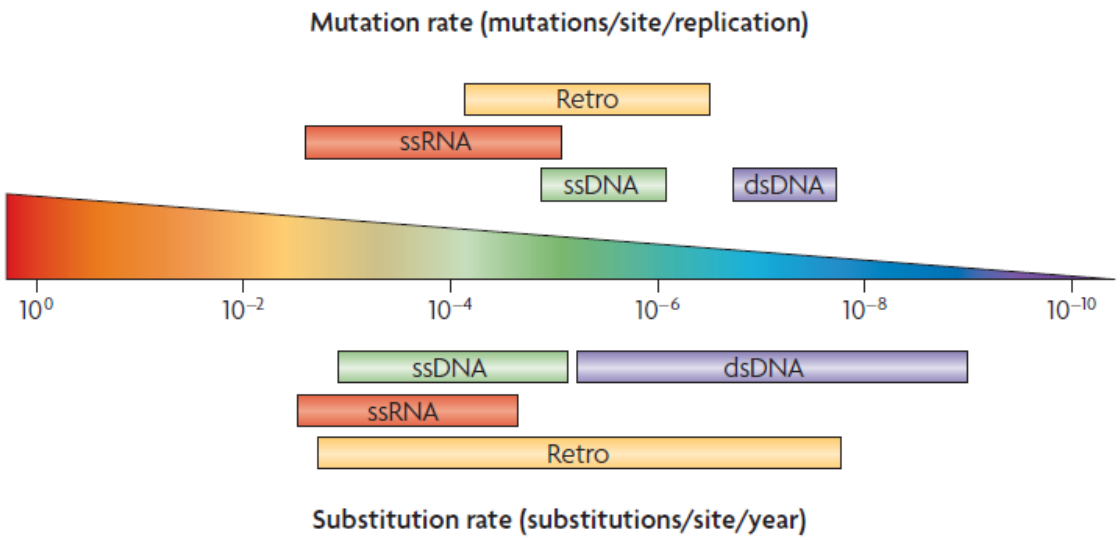


Figure 9 : Comparaison entre taux de mutation et de substitution des virus selon la nature de leur génome. D'après la revue de Duffy *et al.*, 2008.

Plusieurs types de recombinaison peuvent être distingués, tel que la recombinaison homologue durant laquelle une portion d'un génome est remplacée par une portion homologue d'un autre génome, la recombinaison non-homologue qui implique le réarrangement génomique par duplication, insertion ou délétion, ou le réassortiment (pseudo-recombinaison) durant lequel des composants entiers de génomes multipartites sont échangés. Dans le cas de la recombinaison homologue, différents mécanismes ont étés identifiés, tous étant intimement liés à la réplication du virus. La recombinaison peut ainsi être issue de cassures au niveau de la molécule d'ADN matrice ou de conflit entre les enzymes de réplication et de transcription, qui causeraient le détachement prématuré des complexes de réplication. Le ré-attachement de ces complexes de réplication et du brin en cours de synthèse à des matrices distinctes provoquerait la création de molécules nucléotidiques chimères (Owor *et al.*, 2007). La

recombinaison peut également être issue des processus de réparation des ADN double brin de la cellule hôte, induits par des cassures des molécules virales double brin et impliquant des mécanismes de recombinaison dépendante de séquences homologues. Ce mécanisme de réplication dépendante de la recombinaison (RDR) représente un des deux mécanismes de la réplication des géminivirus, avec la réplication en cercle roulant (pour revue Martin *et al.*, 2011).

Depuis longtemps, le phénomène de recombinaison a été identifié chez les géminivirus avec des recombinaisons entre virus de même espèce (Zhou *et al.*, 1997), d'espèces différentes (Briddon *et al.*, 1996; Klute *et al.*, 1996; Monci *et al.*, 2002) et de genre différents (Saunders & Stanley, 1999). Suite à l'obtention d'un nombre croissant de séquences génomiques complètes et au développement de méthodes de détection de la recombinaison, il a été montré que ce processus joue un rôle primordial dans la génération de la diversité chez les géminivirus (Padidam *et al.*, 1999; Lefeuvre *et al.*, 2009). Si la détection fréquente de variants recombinants semble impliquer que la recombinaison est susceptible de produire des descendants mieux adaptés que leurs parents, rares sont les exemples vérifiés expérimentalement (Lefeuvre & Moriones, 2015). Chez les bégomovirus, l'exemple type est celui de la souche sévère Ougandaise de *l'East African cassava mosaic virus* (EACMV-UG). Ce virus est issu d'une recombinaison entre l'EACMV et l'ACMV. En infection mixte et par synergie virale, l'EACMV-UG a été responsable d'une épidémie très sévère sur le manioc en Afrique de l'Est avec des conséquences humanitaires et économiques dramatiques (Zhou *et al.*, 1997; Pita *et al.*, 2001b). Aussi une autre espèce de CMG nommé ACMBFV, a récemment été décrite au Burkina Faso comme étant le résultat d'une recombinaison interspécifique entre un ancêtre des isolats de l'ACMV d'Afrique de l'Ouest (parent majoritaire), et un ancêtre des isolats du *Tomato leaf curl Cameroun virus* et du *Cotton leaf curl Gezira virus* (parents minoritaires) (Tiendrébéogo *et al.*, 2012). Chez les mastrévirus, on peut citer le cas du MSV, dont l'unique souche causant des symptômes sévères sur maïs, la souche MSV-A, est issue de la recombinaison de génomes parentaux apparentés aux souches actuelles du MSV-B et MSV-G/-F décrites sur des poacées sauvages (Varsani *et al.*, 2008).

Les phénomènes de réassortiments de composants d'une espèce ou souche avec un composant d'une autre espèce ou souche de virus ont été décrits à plusieurs reprises dans la littérature. Ainsi l'étude de la capacité d'adaptation à l'hôte des bégomovirus bipartites du Nouveau Monde, *Bean dwarf mosaic virus* (BDMV) et *Tomato mottle virus* (ToMoV), ont permis de démontrer que les réassortiments entre les composants d'ADN-A et -B des deux virus étaient infectieux sur plusieurs plantes hôtes comme la tomate, le haricot et *Nicotiana benthamiana,* mais que l'expression de symptômes de maladie et les taux d'infection élevés étaient variables selon les compositions (Hou *et al.*, 1998). Les travaux réalisés sur les bégomovirus bipartites de l'Ancien Monde associés à la mosaïque du manioc ont montré des réassortiments intra-spécifiques entre l'ADN-A de la souche UG2 de l'EACMV et l'ADN-B de la souche UG3 du même virus, et l'expression de symptômes plus sévères de la maladie (Pita *et al.*, 2001b).

4.4. Dissémination spatiale et temporelle des géminivirus

Si l'impressionnante capacité de recombinaison et d'adaptation des géminivirus a sans nul doute largement participé à leur émergence en tant que pathogènes de nombreuses cultures, deux acteurs majeurs semblent avoir contribués aux émergences des maladies à géminivirus : les **insectes vecteurs** et les **activités humaines**.

4.4.1. *Insectes vecteurs*

B. tabaci a été décrit pour la première fois en 1889 en Grèce (Gennadius, 1889). Depuis, il a été décrit sur tous les continents excepté l'Antarctique (Perring, 2001). Du fait de sa plasticité écologique, la caractérisation taxonomique de l'espèce *B. tabaci* a posé et pose encore de nombreuses difficultés. Pour cette raison, les taxonomistes ont d'abord proposé de subdiviser cette espèce en biotypes, correspondant à des populations biologiquement et génétiquement distinctes (Brown & Bird, 1995; Frohlich *et al.*, 1999; Perring, 2001). La description de ces biotypes est principalement basée sur le polymorphisme enzymatique, le polymorphisme génétique, la gamme de plantes hôtes, la fécondité, le comportement de dispersion, et la capacité à transmettre des virus de plantes (Perring, 2001). A ce jour, au moins 37 biotypes ont été identifiés sur ces critères. Par la suite, en se basant sur le gène mitochondrial du cytochrome oxydase I

(COI), *B. tabaci* a été reclassé en 24 espèces cryptiques, appartenant à 11 groupes génétiques (Dinsdale *et al.*, 2010; De Barro & Ahmed, 2011; De Barro *et al.*, 2011) (Figure 10).

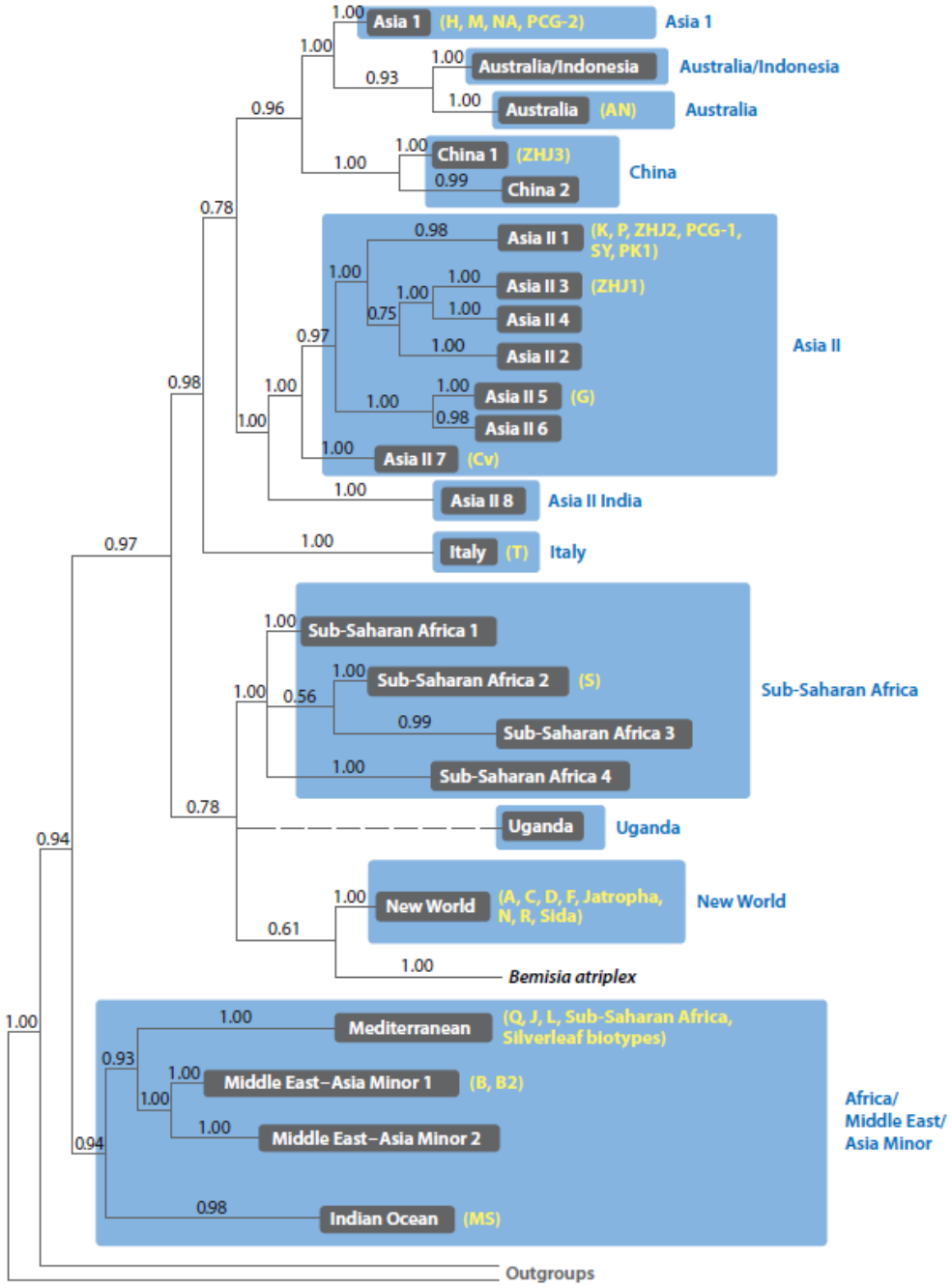


Figure 10 : Phylogénie de *Bemisia tabaci* sur la base des séquences uniques du gène mitochondial du cytochrome oxydase I. Les séquences sont réparties dans 11 groupes génétiques (boites bleues) et 24 espèces cryptiques (boites noires). Les biotypes (en jaune) ont été assignés aux différents groupes génétiques. D'après De Barro *et al*., 2011.

B. tabaci est reconnu comme l'une des espèces les plus invasives au monde principalement pour ses dégâts directs et indirects. Les dégâts trophiques directs, liés à la prise de nourriture dans le phloème, sont d'autant plus préjudiciables que les densités de population de l'insecte sont élevées. Les dégâts indirects sont quant à eux premièrement associés à l'excrétion de miellat produit par les aleurodes et le développement ultérieur de fumagine qui peut affecter la fonction photosynthétique de l'hôte et déprécier la valeur commerciale des produits récoltés (Byrne & Bellows, 1991). Toutefois, les dégâts indirects principaux sont associés en majorité à de nombreuses maladies virales. En effet, parmi les nombreuses espèces virales transmises par les aleurodes et responsables de maladies sur les plantes cultivées, la grande majorité est transmise par *B. tabaci* (Jones, 2003). Ils appartiennent à cinq genres viraux : les bégomovirus qui sont largement majoritaire (90%), les crinivirus, les ipomovirus, les torradovirus et certains carlavirus (Figure 11 ; pour revue Navas-Castillo et al., 2011).

De nombreuses études ont établi un lien direct entre l'étiologie de maladies à bégomovirus et la pullulation de populations de deux espèces cryptiques de *B. tabaci* MEAM1 et MED (communément appelées biotypes B et Q ; Conférer paragraphe 3 ; Figure 10). En effet, à partir des années 1990, le biotype B vraisemblablement originaire du Moyen-Orient se serait rapidement propagé aux dépens des autres biotypes locaux (Brown & Bird, 1995). Sa propagation à travers le monde a directement été associée à la dissémination de l'espèce *Tomato yellow leaf curl virus* (TYLCV) responsable d'épidémies virales très dommageables sur les cultures de tomate notamment dans les pays du pourtour méditerranéen (Moriones & Navas-Castillo, 2000), les Caraïbes (Polston *et al.*, 1994), l'Amérique du Nord (Etats-Unis d'Amérique, Polston *et al.*, 1995) et l'Amérique Centrale (Mexico, Ascencio-Ibáñez *et al.*, 1999), l'Asie (Japon, Kato *et al.*, 1998; Chine, Yongping *et al.*, 2008), les îles du Sud-ouest de l'océan Indien (La Réunion, Peterschmitt *et al.*, 1999 ; Maurice, Lobin *et al.*, 2010), et en dernier lieu l'Océanie (Australie, Van Brunschot *et al.*, 2010 ; Nouvelle-Calédonie, Péréfarres *et al.*, 2012 ; Polynésie Française, Wong *et al.*, 2017) (pour revue Mabvakure *et al.*, 2016).

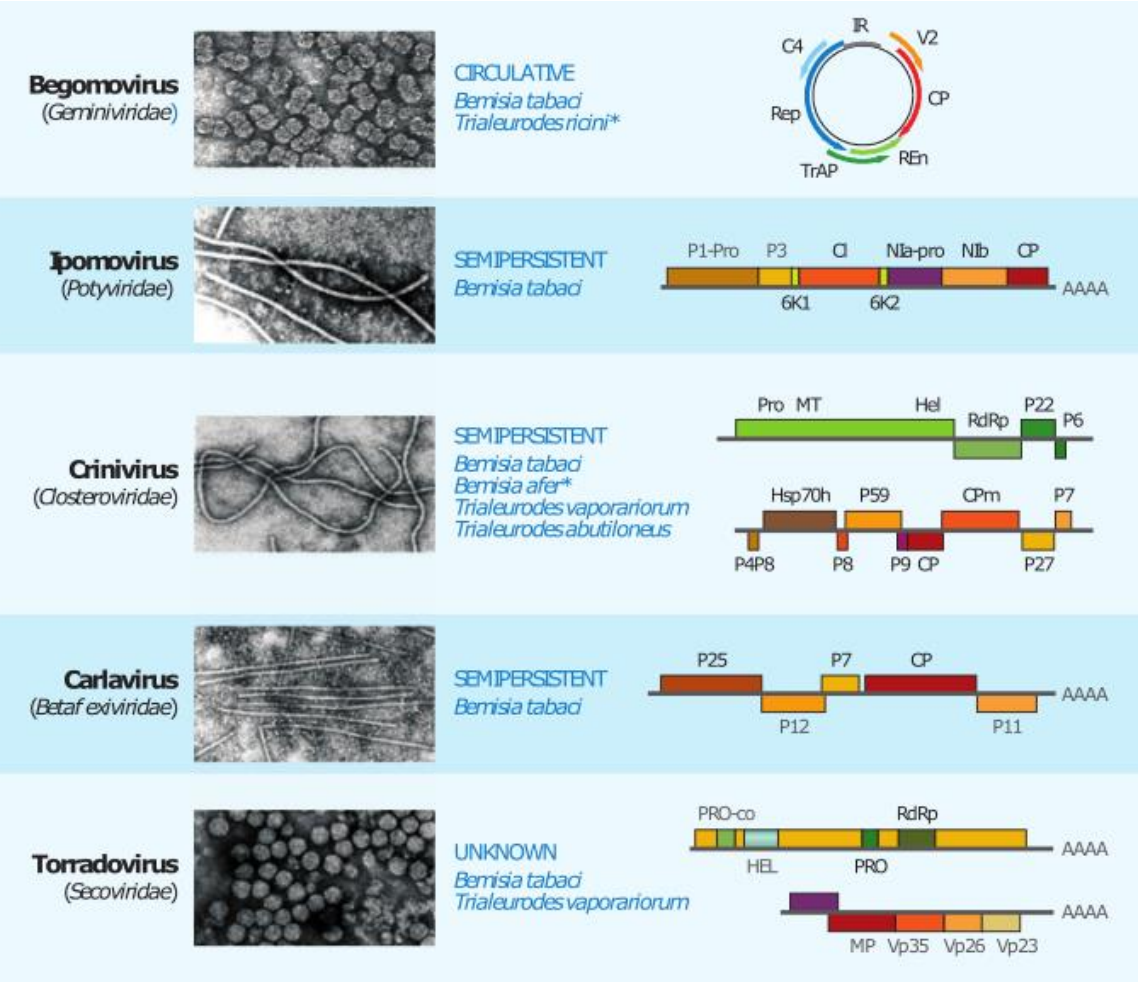


Figure 11 : Schéma des principaux groupes de virus de plantes transmis par les aleurodes, en fonction de la structure de leurs particules virales, de leur mode de transmission et de leur organisation génomique. D'après Navas-Castillo *et al*, 2011.

Ces biotypes invasifs présentent des caractéristiques biologiques particulières : (1) un taux de fertilité accru (Bethke *et al.*, 1991; Delatte *et al.*, 2008), (2) une mobilité qui leur permet de se déplacer à longue distance (Blackmer *et al.*, 1995), (3) une gamme de plantes hôtes très importante (Brown & Bird, 1995), et (4) une résistance à certains insecticides (Brown & Bird, 1995; Elbert & Nauen, 2000; Horowitz *et al.*, 2005). La pullulation d'autres biotypes de *B. tabaci* a également été associée à l'émergence de certaines maladies virales comme la pandémie sévère de la mosaïque africaine du manioc en Afrique de l'Est (pour revue Legg and Fauquet, 2004; Legg and Ogwal, 1998).

Des études récentes ont permis de décrire trois biotypes de *B. tabaci* en Afrique de l'Ouest : le biotype subsaharien *Africa non-Silver Leafing* (AnSL), le biotype subsaharien *Africa Silver Leafing* (ASL), et le biotype Q (Gueguen *et al.*, 2010). Ces trois biotypes ont été regroupés au sein de la même espèce cryptique MED (De Barro *et al.*, 2011 ; Figure 10). Au Burkina Faso, le biotype Q semble prédominant avec deux sous-groupes Q1 et Q3 (Gnankiné *et al.*, 2013 ; Figure 12).

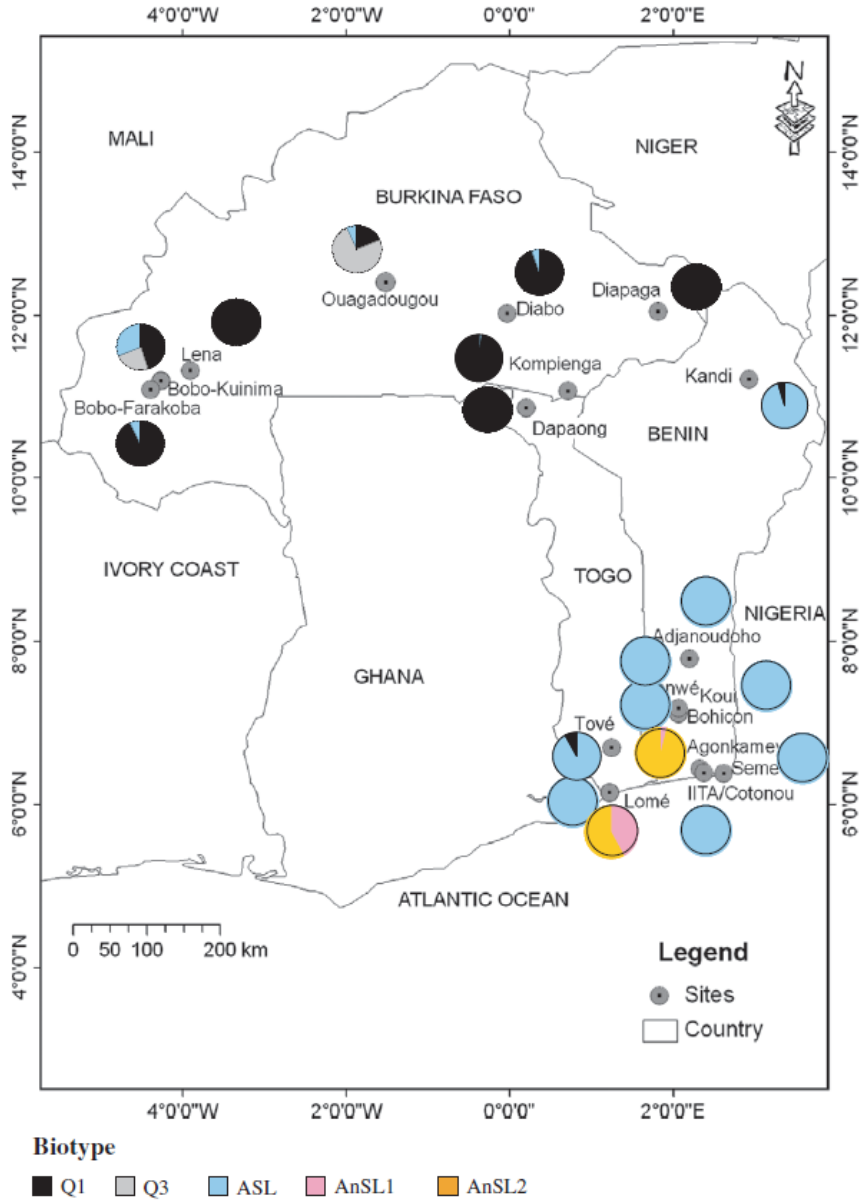


Figure 12 : Distribution géographique des biotypes de *Bemisia tabaci* collectés sur des plantes cultivées (coton, courge, gombo, tabac, tomate) et non cultivées (*Lantana camara*) en Afrique de l'Ouest. D'après Gnankiné *et al*, 2013.

La diversité génétique ou plasticité écologique importante de *B. tabaci* et l'existence de nombreux biotypes lui permet d'accéder à une très large gamme de plantes hôtes. Plus de 900 espèces de plantes hôtes de *B. tabaci*, appartenant à 74 familles botaniques, ont ainsi été recensées (Mound, 1978; Oliveira *et al.*, 2001). L'ensemble de ces espèces représente 73% des plantes cultivées (Servin *et al.*, 1999).

Les cicadelles du genre *Cicadulina* sont les plus importants vecteurs de pathogènes sur le maïs et la canne à sucre dans l'Ancien Monde. Sur les 22 espèces reconnus, 14 sont endémiques d'Afrique et quatre sont présente en Afrique et au Moyen Orient (Webb, 1987). Sur les huit espèces décrites comme vecteur du MSV (Webb, 1987), *C. mbila* semble être à la fois l'espèce vectrice la plus efficace et la plus largement distribuée en Afrique subsaharienne avec plus de 80 espèces de poacées hôtes (Markham et al., 1984 ; Reynaud. 1988). Au Burkina Faso, cinq espèces de *Cicadulina* ont été identifiées : *C. mbila, C. similis, C. arachidis, C. triangula* et très rarement *C. hartmansi* (Traoré *et al.*, 2003). Les espèces de cicadelles *Orosius orientalis* (Horn *et al.*, 1993) et *Orosius albicinctus* (Figure 13 ; Kumari *et al.*, 2004; Akhtar *et al.*, 2011) ont été décrits comme vecteurs du CpCDV, mastrévirus inféodé aux dicotylédones, en Inde (Horn *et al.*, 1993), en Syrie (Kumari *et al.*, 2004) et au Pakistan (Akhtar *et al.*, 2011).

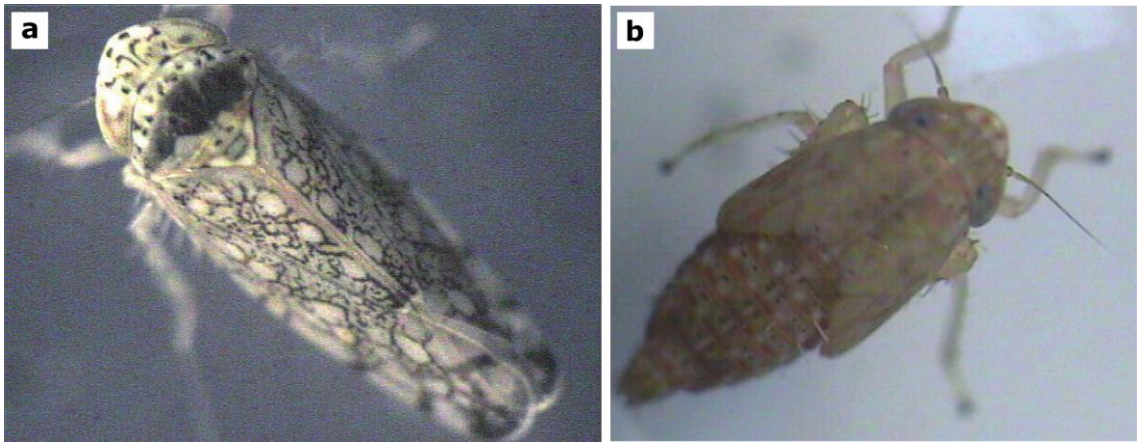


Figure 13 : Vue dorsale d'un adulte **(a)** et d'une nymphe **(b)** d'*Orosius albicinctus*. D'après Akhtar et *al.*, 2011.

4.4.2. *Activités humaines*

Comme pour la plupart des maladies émergentes, le rôle de l'activité humaine a également été largement reconnu dans l'émergence des bégomovirus. Le transfert d'espèces végétales hors de leur centre d'origine, et l'agriculture intensive basée sur la monoculture, ont fortement participé à la mise en place de conditions favorables à la transmission par *B. tabaci* de virus indigènes vers les plantes exotiques introduites (Jones, 2009). De la même façon, il a été démontré que la présence de certains bégomovirus hors de leur aire géographique d'origine est la conséquence de l'introduction par l'Homme de matériel végétal contaminé. On peut citer en exemple le cas emblématique du TYLCV, pour lequel les analyses phylogéographiques ont démontré que certaines souches invasives ont été disséminées indépendamment et dans certains cas à plusieurs reprises, en Asie, à La Réunion, en Amérique du Nord et dans les Caraïbes, à partir du Moyen-Orient et du bassin Méditerranéen (Duffy & Holmes, 2007; Lefeuvre *et al.*, 2010). Enfin, les échanges de matériel végétal contaminé favorisent à la fois l'élargissement des aires de distributions des virus, les infections mixtes et l'évolution par recombinaison. C'est en particulier le cas pour les plantes à multiplication végétative comme le manioc dont les boutures sont fréquemment transportées et échangées (Swanson & Harrison, 1994).

5. Hôtes et symptômes causés par les bégomovirus et les mastrévirus

Les plantes hôtes des bégomovirus les plus répertoriées dans la littérature appartiennent aux familles dicotylédones des *Amaranthaceae*, *Asteraceae*, *Brassicaceae*, *Convolvulaceae*, *Cucurbitaceae*, *Euphorbiaceae*, *Labiaceae*, *Fabaceae*, *Malvaceae*, *Solanaceae* et *Verbenaceae* (Silva *et al*., 2017).

Les principaux symptômes (Figure 14) de maladie associés aux bégomovirus sont (1) des enroulements foliaires en forme de cuillère associés ou non à du jaunissement et du nanisme, comme la maladie de l'enroulement et/ou du jaunissement foliaire de la tomate (*tomato (yellow) leaf curl disease*, ToLCD-TYLCD) (pour revues Navas-Castillo *et al.*, 2011; Péréfarres *et al.*, 2014) et la maladie de l'enroulement foliaire du gombo (*okra leaf curl disease*, OLCD ; Tiendrébéogo *et al.*, 2010), (2) des déformations foliaires associées à des mosaïques, comme la mosaïque du manioc (*cassava mosaic*

disease, CMD; (Legg *et al.*, 2014a,b), et (3) des jaunissements des nervures, comme la maladie du jaunissement des nervures du piment (*pepper yellow vein disease*, Tiendrébéogo *et al.*, 2008).

Figure 14 : Symptômes **(a)** de la maladie de l'enroulement foliaire du gombo (*okra leaf curl disease*, OLCD), **(b)** de la maladie du jaunissement des nervures du piment (*pepper yellow vein disease*, PYVD) et **(c)** de la maladie de l'enroulement (et du jaunissement) foliaire de la tomate (*tomato (yellow) leaf curl disease*, ToLCD-TYLCD) observés au Burkina Faso (Tiendrébéogo, 2010).

Les plantes hôtes des mastrévirus actuellement décrites concernent une soixantaine d'espèces végétales appartenant à 44 genres (Muhire *et al.*, 2013; Kraberger *et al.*, 2015b). Parmi ces 44 genres, onze sont des dicotylédones se répartissant entre les familles *Amaranthaceae* (Farzadfar *et al.*, 2008), *Fabaceae* (Muhire *et al.*, 2013; Kraberger *et al.*, 2015b), *Cucurbitaceae* (Fahmy *et al.*, 2015), *Malvaceae* (Manzoor *et al.*, 2014) et *Solanaceae* (Akhtar *et al.*, 2014). Les 33 autres genres font tous parties de la famille des poacées et sont en grande majorité des poacées sauvages (Figure 8 ; Muhire *et al.*, 2013; Kraberger, 2014). Les mastrévirus inféodés aux monocotylédones (poacées) provoquent des maladies avec des symptômes de striure très caractéristiques (Kraberger, 2014). Au contraire, les mastrévirus inféodées aux dicotylédones (familles citées dans le paragraphe précédent) provoquent des symptômes variés : jaunissement, déformation et nanisme foliaire (Kraberger *et al.*, 2015b).

6. Principales méthodes de lutte contre les géminivirus

Les maladies virales sont responsables de pertes importantes à la fois sur la quantité et la qualité des productions agricoles. Peu de méthodes de lutte sont disponibles dans le cas d'une infection virale et surtout, aucun traitement curatif ne peut être appliqué. Les moyens de contrôle des maladies virales peuvent se décliner en deux catégories

principales : (1) les stratégies qui permettent de limiter ou de réduire au minimum l'arrivée du virus sur la parcelle et (2) les stratégies qui tendent à limiter les effets du virus sur la plante une fois qu'elle a été infectée.

6.1. Pratiques culturales, prophylaxie

La première des pratiques consiste à empêcher ou réduire fortement l'arrivée du virus sur la parcelle. Empêcher les introductions et les mouvements de plants contaminés vers les zones encore indemnes est une priorité dans la lutte contre les géminivirus (Anderson *et al.*, 2004). Dans les zones contaminées, l'emploi de plants sains et certifiés ainsi que la suppression des plantes adventices hôtes de bégomovirus à proximité de la parcelle constituent les premières mesures qui tendent à repousser l'infestation de la parcelle (Sánchez-Campos *et al.*, 1999; Papayiannis *et al.*, 2011). De même, la rotation des cultures ainsi que l'élimination des résidus de culture sont également recommandés (Morilla *et al.*, 2005).

Le contrôle de l'insecte vecteur est un élément important de la lutte contre les viroses associées aux géminivirus. Elle peut se baser sur l'utilisation de barrières physiques (filets *insect-proof*, Cohen & Antignus, 1994), de méthodes qui perturbent le comportement de l'insecte (modification du rayonnement UV dans les serres, (Antignus *et al.*, 2001), de la lutte biologique avec des insectes auxiliaires des cultures (insectes des genres *Eretmocerus* et *Encarsia* par exemple pour la lutte contre *B. tabaci*; Picó *et al.*, 1996) et bien sûr de la lutte chimique. Toutefois, les emplois répétés d'insecticide favorisent le développement de résistances chez les insectes vecteurs, comme ce fut notamment le cas des biotypes invasifs B et Q de *B. tabaci* (Byrne *et al.*, 2003) et n'ont montré qu'une faible efficacité pour limiter la dissémination des maladies virales avec l'inconvénient des effets indésirables des produits chimiques sur l'environnement (Elbert & Nauen, 2000; Otoidobiga *et al.*, 2003; Horowitz *et al.*, 2005).

6.2. Résistance variétale

Les pratiques culturales peuvent aider à contrôler le développement de la maladie mais sont en général insuffisantes (Antignus, 2007; Polston & Lapidot, 2007). Développer des cultivars résistants reste la meilleure stratégie pour assurer une production

acceptable en quantité et en qualité. Les programmes d'amélioration variétale pour la résistance aux virus, comme pour les autres pathogènes, doivent répondre à trois objectifs majeurs : le niveau de résistance (efficacité de contrôle de la quantité de virus), le spectre d'action de la résistance (efficacité à contrôler différentes souches ou espèces virales), et la durabilité de la résistance face à la dynamique évolutive des populations virales (Moury *et al.*, 2011). Toutefois, le contournement possible de la résistance ou la non acceptation des variétés résistantes par les populations restent un frein à leur utilisation.

Une plante est qualifiée de résistante à un virus quand l'infection, la réplication et/ou le mouvement du virus dans la plante est/sont restreinte(s) (Cooper *et al.*, 1983). La sélection de cultivars résistants repose sur la disponibilité de caractères de résistance dans le germplasme (ensemble des ressources génétiques végétales) et leur introgression dans des lignées élites aux fortes qualités agronomiques. Plusieurs gènes de résistance majeure et mineure, dominants ou récessifs ont été identifiés et utilisés avec succès dans la lutte contre les géminivirus que ce soit sur la tomate (Caro *et al.*, 2015; Li *et al.*, 2017), le manioc (Fondong, 2017) ou le maïs (Martin & Shepherd, 2009; Nair *et al.*, 2015).

Une autre possibilité est l'utilisation des plantes transgéniques résistantes. Bien que peu acceptées par la société, les techniques de transgénèse contribuent à élargir la diversité des mécanismes de résistance utilisable par les sélectionneurs. Différentes méthodologies basées sur la résistance dérivée du pathogène (*Pathogen-derived resistance*, PDR), avec ou sans expression de protéines virales, et l'expression d'agents antiviraux non issus d'agents pathogènes ont été utilisées principalement à des fins expérimentales (pour revues voir Prins *et al.*, 2008; Shepherd *et al.*, 2009). En ce qui concerne la résistance aux géminivirus, notons l'efficacité des expériences de résistance dérivée du pathogène pour le MSV (Shepherd *et al.*, 2009), les CMGs (Fondong, 2017) et le TYLCV (Fuentes *et al.*, 2006). Il est à noter que toute activité relative à l'utilisation en milieu confiné ou en milieu ouvert, à la dissémination, à la mise sur le marché et aux mouvements transfrontière, incluant le transit d'organismes génétiquement modifiés fait l'objet d'un accord préalable au Burkina Faso. L'accord est

donné par l'Agence Nationale de Biosécurité (Loi n° 064-2012/AN du 20 décembre 2012 portant régime de sécurité en matière de biotechnologie).

En dehors de la résistance au virus *sensu stricto*, la résistance au vecteur *B. tabaci* est une option intéressante mais encore peu exploitée pour la lutte contre les maladies causées par les géminivirus. Des sources naturelles de résistance à *B. tabaci* ont été décrites dans des accessions originaires de *Solanum habrochaites* (Momotaz *et al.*, 2010), *Solanum pennellii* (Blauth *et al.*, 1998) et *Solanum pimpinellifolium* (Delatte *et al.*, 2006; Rodríguez-López *et al.*, 2011). Ces résistances reposent sur la sécrétion de métabolites secondaires par des trichomes glandulaires de type IV et VI présents sur les feuilles (Blauth *et al.*, 1998; Muigai *et al.*, 2002 ; Figure 15). Ces composés ont des propriétés à la fois d'antibiose (effet négatif sur la physiologie de l'insecte) et d'antixénose (perturbation du comportement de l'insecte). L'efficacité de ces mécanismes de résistance a été démontrée pour limiter l'infection (Delatte *et al.*, 2006) et la dissémination du virus (Rodríguez-López *et al.*, 2011).

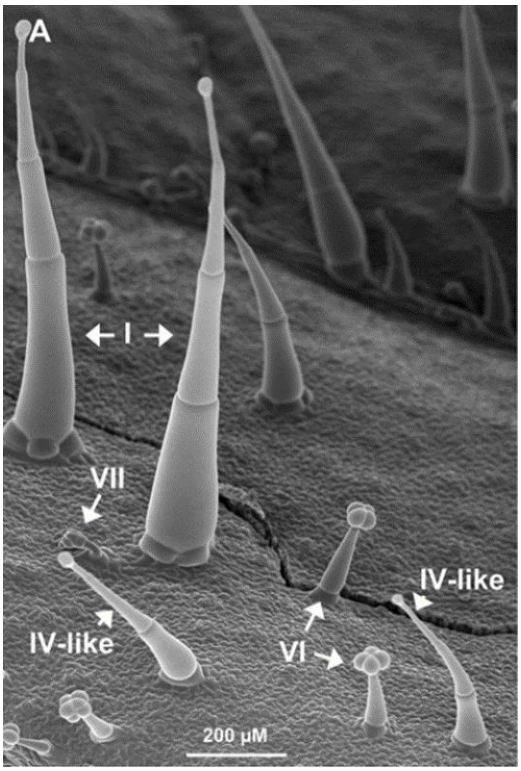


Figure 15 : Photographie en cryo-microscopie électronique de trichomes sur la surface foliaire de la tomate. Les différents types de trichomes présents sont indiqués par des flèches. D'après Kang *et al.*, 2010.

7. Principales méthodes de diagnostics et d'analyses

7.1. Méthodes de diagnostics

Comme nous l'avons décrit dans le paragraphe 5, les symptômes provoqués par les bégomovirus et les mastrévirus sont souvent particuliers et constituent à ce titre un bon diagnostic en vue de suspecter une infection par ces groupes de virus. Cependant, plusieurs facteurs peuvent influencer la forme des feuilles et l'état général de la plante, comme l'action des insectes défoliateurs, des insectes piqueurs-suceurs et du stress hydrique (Jones, 2003). Face à cette difficulté voir impossibilité de relier de façon formelle symptômes et agent causal, plusieurs méthodes de diagnostic ont été développées. Dans le cas des bégomovirus et des mastrévirus, des méthodes de diagnostic sérologique (Al Bitar & Luisoni, 1995; Wu *et al.*, 2012) et moléculaire (Patel *et al.*, 1993; Inoue-Nagata *et al.*, 2004) sont couramment utilisées.

7.1.1. *Les tests sérologiques*

Les techniques immuno-enzymatiques (ELISA) sont utilisés pour la détection courante des bégomovirus ainsi que des mastrévirus. Les anticorps utilisant les tests ELISA ont permis la détection des bégomovirus dans différents extraits végétaux (Konaté *et al.*, 1995; Harrison & Robinson, 1999). La variante TAS-ELISA utilisant les anticorps monoclonaux a été utilisée pour la détection et la classification de ce groupe de virus (Barro, 1994; Konaté *et al.*, 1995).

Cependant, du fait des fortes similitudes structurales et antigéniques entre les géminivirus et de l'observation de réactions sérologiques croisées, il est difficile voire impossible de distinguer la plupart des espèces virales par ces méthodes (Brown *et al.*, 2001).

7.1.2. *Les tests moléculaires*

Méthodes séquences dépendantes

Les techniques moléculaires de détection et d'identification des virus ont été développées avec la réaction de polymérisation en chaîne (PCR) (Duncan & Torrrance,

1992). Comprenant différentes variantes, cette technique nécessite bien souvent la mise au point d'amorces consensus ou dégénérées à partir de séquences virales connues (Rector *et al.*, 2004; Mokili *et al.*, 2012). Cette approche se cantonne au diagnostic des espèces virales connues ou aux espèces virales génétiquement proches. L'identification taxonomique nécessite le séquençage des amplicons PCR et une recherche de séquences similaires dans des bases de données (*e.g.* BLASTn, NCBI). Malgré la rapidité de mise en œuvre de cette approche, l'identification taxonomique est le plus souvent basée sur une portion du génome. Du fait de la fréquence de recombinaison élevée de ces virus, l'identification taxonomique définitive nécessite d'avoir la connaissance de la séquence complète de leur génome voire de la molécule d'ADN-A pour les bégomovirus bipartites.

Méthodes séquences indépendantes

La mise au point d'une technique d'amplification en cercle roulant (*Rolling-circle amplification*, RCA) a révolutionné l'étude des virus à ADN circulaire (Inoue-Nagata *et al.*, 2004; Shepherd *et al.*, 2008). Tirant parti des propriétés de la polymérase *Phi*29, une enzyme de bactériophage à haute-fidélité, active et stable à 30°C, et en utilisant des amorces hexamèriques aléatoires non spécifiques, cette méthode permet de produire de multiples copies répétées en tandem des ADN circulaires de petite taille, comme ceux des génomes des géminivirus. Il est à noter que l'amplification n'est pas spécifique aux petits ADN circulaires mais que ce type de matrice est très efficacement répliqué aboutissant à un enrichissement de l'échantillon en ADN viral en comparaison des autres ADN présents dans l'échantillon comme l'ADN de plante. Une fois linéarisé par restriction sur un site de coupure unique, le composant d'ADN viral double brin est cloné et séquencé. Cette technique présente le grand avantage de ne pas nécessiter de connaissances préalables sur la séquence virale à amplifier. Elle s'est révélée particulièrement efficace et peu couteuse pour la caractérisation des géminivirus (Haible *et al.*, 2006; Shepherd *et al.*, 2008), et a été largement adoptée par la communauté scientifique. Elle est à l'origine de la forte augmentation du nombre d'espèces de bégomovirus découverts ces dernières années, ainsi que de la caractérisation des différentes molécules impliquées dans ces complexes viraux comme les ADN satellites.

Métagénomique virale

Le développement des méthodes de séquençage à haut-débit (post-Sanger) et la transposition des protocoles de métagénomique (génomique environnementale) aux communautés virales ont permis de réaliser des études de la diversité virale d'échantillons complexes (Breibart *et al*., 2002). L'absence de gène canonique universel chez les virus ne permet pas de développer des approches classiques de *barcoding* comme chez d'autres groupes d'organismes. Plusieurs protocoles d'isolation de différentes cibles virales (pour revue Roossinck, 2015; Figure 16) et de séquençage sans a priori à partir d'échantillon environnementaux ont vu le jour. Ces protocoles sont actuellement basés sur (1) les siRNA qui nécessitent l'extraction des petits ARNs générés lors du mécanisme de défense de la plante vis-à-vis des virus (également appelé *silencing*) (Kreuze *et al.*, 2009), (2) l'extraction des acides nucléiques viraux après une étape d'enrichissement en particules virales (*virion-associated nucleic acid*, VANA; Candresse *et al.*, 2014), (3) l'extraction des ARN bicaténaires (Roossinck *et al.*, 2010), (4) l'extraction des ARN totaux de la plante avec ou sans déplétion des ARN ribosomiques (Gu *et al.*, 2014) et (5) l'amplification spécifique d'ARN ou d'ADN viraux ou encore le traitement direct des acides nucléiques totaux (De Bruyn, 2014). Si l'usage routinier de la métagénomique comme test de diagnostic est encore confronté à de nombreuses difficultés (Martin *et al.*, 2016; Massart *et al.*, 2017), les possibilités offertes par ces méthodes pour la découverte de nouveaux virus ou la caractérisation précise des communautés virales sont très importantes. Plusieurs travaux pionniers en métagénomique virale ont permis de démontrer l'immense diversité virale présente dans les écosystèmes naturels (Breitbart *et al.*, 2002; Venter *et al.*, 2004) et d'espérer pouvoir étudier le fonctionnement de ces communautés virales (Roossinck et al., 2010). En ce qui concerne les virus à ADN circulaire en général et les géminivirus en particulier, la combinaison d'amplification en cercle roulant avec des méthodes de séquençage haut-débit a été appliquée avec succès sur des échantillons de plantes (Wyant *et al.*, 2012; Idris *et al.*, 2014), d'insectes vecteurs (Rosario *et al.*, 2015, 2016) ou environnementaux (Kraberger *et al.*, 2015a; Dayaram *et al.*, 2016).

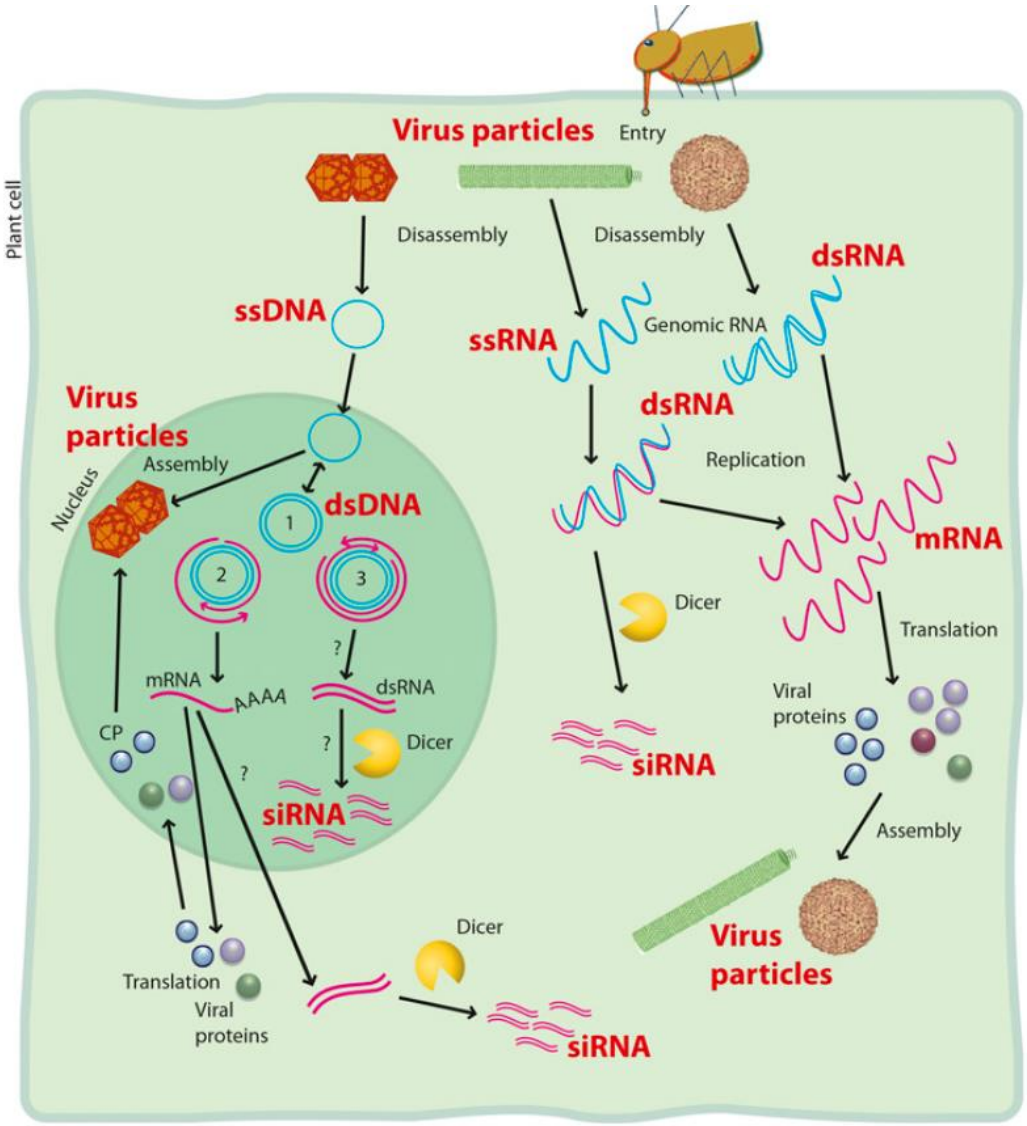


Figure 16 : Diversité des formes d'acides nucléiques présentes dans la cellule végétale infectée par un phytovirus pouvant servir de cible aux approches métagénomiques. D'après Roossinck *et al.*, 2015.

8. Problématique et objectifs

Les épidémies virales causent des dégâts considérables aux cultures maraîchères, vivrières et industrielles. Elles sont particulièrement dommageables dans les régions tropicales où les populations dépendent étroitement de la production agricole pour leur subsistance.

Depuis la fin des années 1980s, le continent Africain et sa population ont dû faire face à l'émergence des géminivirus, notamment sur le maïs avec le MSV (pour revues Martin & Shepherd, 2009; Shepherd *et al.*, 2010), le manioc avec le complexe d'espèces des CMGs (pour revue Legg *et al.*, 2014a; Legg and Fauquet, 2004), les solanées cultivées avec un complexe de bégomovirus monopartites africains communément appelés les ToLCVs (Leke *et al.*, 2015; Rey *et al*, 2012).

En Afrique de l'Ouest et Centrale, outre le fait de diagnostiquer la présence d'espèces de bégomovirus préalablement décrites en Afrique subsaharienne, comme le CLCuGeV responsable de l'OLCD (Tiendrébéogo *et al.*, 2010) et la souche Ouganda de l'EACMV responsable de la forme sévère de la CMD (Tiendrébéogo *et al.*, 2009; Zinga *et al.*, 2012), les travaux récents d'épidémiologie moléculaire ont permis de caractériser une **multitude de nouvelles espèces de bégomovirus** (Leke *et al.*, 2015) notamment sur les solanées cultivées, comme les PepYVMLV (Zhou *et al.*, 2008; Kon & Gilbertson, 2011; Tiendrébéogo *et al.*, 2011), ToLCCMLV (Leke *et al.*, 2011), ToLCGHV (Osei *et al.*, 2008), ToLCMLV (Zhou *et al.*, 2008), ToLCNGV (Kon & Gilbertson, 2011), et TYLCMLV (Chen *et al.*, 2009; Lett *et al.*, 2009; Sattar *et al.*, 2015). Malgré la description de toutes ces nouvelles espèces de géminivirus, les connaissances sur leur épidémiologie restent néanmoins incomplètes, notamment concernant la diversité et l'identification précise des espèces virales incriminées dans l'étiologie des maladies, et les gammes de plantes hôtes réservoirs ou alternatives. De plus, de nombreux symptômes de maladies à bégomovirus (enroulement ou déformation foliaire, nanisme, jaunissement) ont été décrits à la fois sur les cultures maraîchères, dont notamment la tomate et certaines plantes adventices et ornementales en Afrique de l'Ouest et Centrale (Konaté *et al.*, 1995), pour lesquelles l'agent causal ou les agents causaux restent à être déterminés.

Aussi, les données récentes de métagénomique suggèrent que la grande majorité des virus caractérisés sur les cultures ne représentent qu'une faible fraction des virus infectant les plantes en général (Malmstrom *et al.*, 2011; Alexander *et al.*, 2014). Le manque d'exhaustivité des études actuelles sur la diversité des populations virales s'explique à la fois par le fait que la majorité des travaux de recherche s'est focalisée sur les maladies des plantes cultivées et également par les difficultés à identifier les éventuels symptômes associés à l'infection de plantes sauvages par des virus. En effet, les virus ayant longuement co-évolués avec leurs plantes hôtes, provoquent très rarement des maladies dommageables dans leurs écosystèmes naturels (Jones, 2009). On considère que dans ces systèmes écologiques non perturbés par l'activité humaine les interactions entre les virus-vecteurs-plantes hôtes sont évolutivement anciennes et stables (Malmstrom *et* al. 2011). Cependant, du fait des perturbations environnementales associées à l'agriculture moderne, les équilibres virus-plantes hôtes sont souvent fortement perturbés, entraînant de nouvelles interactions virus-vecteur-hôte-environnement, qui peuvent provoquer l'apparition de maladies qualifiées d'émergentes (Fargett*e et al*., 2006; Lefeuvre & Moriones, 2015). En outre, ces phénomènes sont exacerbés par les activités humaines qui favorisent la proximité entre les écosystèmes naturels et cultivés au sein « d'interfaces agroécologiques » composées de plantes cultivées, d'adventices des systèmes de cultures et de plantes non cultivées exotiques, indigènes ou endémiques (Alexander *et al.*, 2014).

En conséquence, outre les lacunes de nos connaissances de la diversité des populations virales naturelles, de leurs répartitions géographiques, de leurs liens phylogénétiques avec les populations virales décrites sur les plantes cultivées et de leurs pathogénicités, on constate un manque de connaissances fondamentales sur la dynamique des communautés virales à l'échelle des écosystèmes, l'écologie évolutive des virus en général et leur capacité à changer d'hôte en particulier (Alexander *et al.*, 2014; Roossinck *et al.*, 2015).

Ainsi, l'objectif général de cette thèse a été de contribuer à sécuriser durablement les productions maraîchères du Burkina Faso en améliorant les connaissances d'épidémiologie moléculaire des géminivirus responsables de maladies des cultures

maraîchères afin de proposer des méthodes de gestion durable.

Afin d'approfondir nos connaissances sur la diversité des géminivirus infectant les cultures maraîchères et de leur répartition géographique au Burkina Faso, le **premier chapitre** de cette thèse a été consacrée (1) à la caractérisation moléculaire des géminivirus, (2) à l'étude de leur diversité génétique, et (3) à l'évaluation des principaux paramètres épidémiologiques en plein champ dont notamment la prévalence des géminivirus sur les cultures de tomate et leur gamme de plantes hôtes. Cette activité a été réalisée à l'échelle nationale, avec des prospections étalées de 2013 à 2016 sur les trois zones agroclimatiques du pays. Ce travail a été conduit en utilisant deux approches. La première dite **séquence dépendante** est basée sur l'utilisation d'amorces spécifiques et universelles des bégomovirus monopartites de l'Ancien Monde, confirmée systématiquement par le séquençage des amplicons PCR. La deuxième dite **séquence indépendante** est basée sur la RCA-RFLP, le clonage et le séquençage de génomes complets. Les principaux résultats obtenus dans le cadre de ce chapitre ont permis de confirmer à la fois la non-exhaustivité de notre connaissance de la diversité des géminivirus présent sur les cultures maraîchères au Burkina Faso, l'existence d'un gradient Nord-Sud de complexité des espèces de géminivirus associés à la maladie du ToLCD-TYLCD, la prédominance du PepYVMLV sur les solanées cultivées et la présence de nombreuses plantes alternatives ou réservoirs. Ces résultats ont également permis la première caractérisation, en Côte d'Ivoire, de trois bégomovirus dont deux sur le gombo et un sur le piment.

Au regard de la découverte d'un nouvel ADN-B en association avec l'ADN-A du PepYVMLV, au Burkina Faso, et afin d'élucider l'énigme de leur association, le **deuxième chapitre** de cette thèse s'est intéressé à décortiquer les principaux paramètres du pouvoir pathogène du PepYVMLV en association ou non avec le composant d'ADN-B. En conditions contrôlées, nous avons évalué le **pouvoir pathogène** des deux molécules d'ADN en association simple ou mixte, les cinétiques de **virulence** (symptômes de maladie), l'**abondance virale** (*fitness* relative) et le **taux de transmission individuelle** par l'aleurode *B. tabaci*. L'ensemble de nos résultats démontrent que même si le PepYVMLV n'est pas un bégomovirus bipartite strict,

l'ADN-B représente à la fois un activateur fort de la virulence, de l'accumulation de l'ADN-A dans la plante et de la transmission individuelle par son insecte vecteur *B. tabaci*, qui semblent concourir à sa prédominance sur le terrain. Avec une sévérité, en conditions contrôlées, supérieure à celle du TYLCV-IL, considéré actuellement comme le bégomovirus le plus sévère sur les cultures de tomate dans le monde, et des symptômes ultimes de la maladie du ToLCD-TYLCD aboutissant à la mort de la plante en conditions contrôlées, on peut considérer que le PepYVMLV représente une menace majeure pour la culture de la tomate au Burkina Faso, en Afrique et plus largement dans le monde.

Le **troisième chapitre** de cette thèse a été consacré à l'étude de la **diversité** et de la **dynamique des populations de géminivirus à l'échelle des systèmes agroécologiques** sur les plantes cultivées et non cultivées par une approche métagénomique. Dans le but de détecter et identifier les espèces virales présentes dans nos échantillons sans a priori sur leurs séquences nucléotidiques, j'ai participé à la mise au point d'une approche métagénomique ciblée basée sur l'amplification des acides nucléiques viraux, le marquage et le séquençage à haut débit (RCA-RA-SGS). Cette approche a non seulement permis de confirmer la présence d'une majorité des virus décrits préalablement avec les approches classiques, mais également la détection de plusieurs nouvelles espèces. L'étude de la diversité des gammes d'hôtes et de leurs chevauchements permet d'établir les possibles réseaux d'échange de virus entre l'ensemble des plantes hôtes décrites. La description de la structure de ces réseaux apporte une vision non-biaisée et inédite de la structure des communautés de géminivirus au sein des agro-écosystèmes du Burkina Faso.

CHAPITRE I

EPIDEMIOLOGIE MOLECULAIRE DES GEMINIVIRUS INFECTANT LES CULTURES MARAÎCHÈRES AU BURKINA FASO ET EN COTE D'IVOIRE

CHAPITRE I : EPIDEMIOLOGIE MOLECULAIRE DES GEMINIVIRUS INFECTANT LES CULTURES MARAÎCHÈRES AU BURKINA FASO ET EN COTE D'IVOIRE

Ces dernières décennies, le continent Africain et sa population ont dû faire face à l'émergence de nombreuses maladies à géminivirus, sur les cultures vivrières, comme la mosaïque du manioc (CMD ; Legg *et al.*, 2014), la striure du maïs (MSD ; Shepherd *et al.*, 2010), le jaunissement et l'enroulement foliaire du gombo (OLCD) et de la tomate (ToLCD-TYLCD ; pour revues Leke *et al.*, 2015 ; Rey *et al.*, 2012). En Afrique Occidentale et du Centre, une **multitude de nouvelles espèces de bégomovirus** a été décrite comme responsable des épidémies virales observées dans les périmètres maraîchers (Leke *et al.*, 2015). Ainsi, trois espèces de bégomovirus ont été décrites comme responsables de l'OLCD (Tiendrébéogo *et al.*, 2010a; Leke *et al.*, 2015) et cinq autres espèces seraient impliquées dans les maladies de l'enroulement et du jaunissement foliaire des solanées cultivées, notamment la tomate avec le ToLCMLV et le TYLCMLV (Kon & Gilbertson, 2011; Sattar *et al.*, 2015), le ToLCNGV (Leke *et al.*, 2015), et le ToLCGHV (Osei *et al.*, 2008) ainsi que le piment et le poivron avec le PepYVMLV (Kon & Gilbertson, 2011; Tiendrébéogo *et al.*, 2011).

Malgré la description d'un grand nombre de nouvelles espèces de géminivirus sur les cultures maraîchères en Afrique Occidentale et Centrale, les connaissances sur leurs aspects épidémiologiques majeurs (diversité des virus impliqués dans une maladie, gamme de plantes hôtes réservoirs ou alternatives (cultivées et/ou sauvages), prévalence, répartition géographique et capacité de transmission) restent largement incomplètes. Par ailleurs, de nombreux symptômes de maladies virales (enroulement ou déformation foliaire, nanisme, jaunissement) ont été décrits à la fois sur les cultures maraîchères, dont notamment la tomate et certaines plantes adventices et ornementales de cette région (Konaté *et al.*, 1995) pour lesquelles le ou les agents viraux restent à être déterminés. Par conséquent, la caractérisation fine de la diversité des géminivirus, l'identification précise des espèces virales incriminées dans l'étiologie des maladies sur les cultures maraîchères en Afrique de l'Ouest, et la description précise de leur répartition géographique et exhaustive de leurs gammes de plantes hôtes réservoirs ou

alternatives devraient nous permettre d'approfondir nos connaissances et d'améliorer notre compréhension de l'épidémiologie de ces virus.

Dans ce but, le **premier chapitre** de cette thèse a été consacré (1) à la caractérisation moléculaire des géminivirus, (2) à l'étude de leur diversité génétique, et (3) à l'évaluation des principaux paramètres épidémiologiques en plein champ dont notamment la prévalence des géminivirus sur les cultures de tomate et leur gamme de plantes hôtes. Cette activité a été réalisée au Burkina Faso, à l'échelle nationale sur les trois zones agroclimatiques avec des prospections réalisées de 2013 à 2016 (**Articles 1-3**), ainsi qu'en Côte d'Ivoire dans le cadre d'une collaboration avec l'Université de Nangui-Abrogoua (**Articles 4 et 5**). Ce travail a été conduit en utilisant deux approches.

La première approche, dite **séquence dépendante,** a été basée sur l'utilisation d'amorces universelles et spécifiques des géminivirus, dont le diagnostic a été confirmé systématiquement par le séquençage des amplicons PCR. Cette activité a permis premièrement de confirmer la non-exhaustivité de notre connaissance de la diversité des géminivirus présente sur les cultures maraîchères au Burkina Faso, avec la description de **trois bégomovirus** préalablement décrits en Afrique de l'Ouest (ToLCMLV, ToLCGHV, PepYVMLV; **Article 1**), d'une **nouvelle espèce de bégomovirus** nommée provisoirement *Tomato leaf curl Burkina Faso virus* (ToLCBFV, **Article 2**), et enfin d'un **mastrévirus** décrit pour la première fois en Afrique sur la tomate (CpCDV, **Article 3**), tous associés à des symptômes de la maladie du ToLCD-TYLCD. Elle a également permis d'estimer la prévalence des géminivirus associés à la maladie du ToLCD-TYLCD sur 10 sites répartis dans les trois zones agroclimatiques du Burkina Faso (**Figure 3**), et de montrer que les géminivirus sont présents sur toutes les parcelles échantillonnées, avec néanmoins des prévalences très variables (de 2 à 78%). De plus, ces travaux ont permis de montrer l'existence d'un gradient Nord-Sud de complexité croissante des géminivirus, avec la prédominance du PepYVMLV sur les solanées cultivées, et l'étendue de leur gamme d'hôte.

La deuxième approche, dite **séquence indépendante,** a été basée sur l'amplification en cercle roulant (RCA), le clonage et le séquençage de génomes complets. Cette activité a permis de confirmer la présence, sur les cultures de tomate, de cinq géminivirus

distincts (ToLCMLV, ToLCGHV, PepYVMLV, ToLCBFV et CpCDV), dont trois diagnostiqués pour la première fois au Burkina Faso (ToLCGHV, **Article 1** ; ToLCBFV, **Article 2** ; CpCDV, **Article 3**). Parallèlement, dans le cadre d'une collaboration avec une équipe ivoirienne, nous avons réalisé la première description du PeYVMLV associé à la maladie du PepYVD sur piment et poivron (**Article 4**) et des CLCuGeV et OYCrV associés à l'OLCD en Côte d'Ivoire (**Article 5**). Le replacement phylogénétique des isolats burkinabés de bégomovirus décrits au sein de l'ensemble des bégomovirus africains a révélé l'existence de plusieurs groupes monophylétiques suggérant l'adaptation indépendante de plusieurs lignées virales à la tomate. De manière inattendue, un composant génomique d'ADN-B a été caractérisé très majoritairement en association avec le PepYVMLV, un virus préalablement décrit comme un virus monopartite (Zhou *et al.*, 2008; Tiendrébéogo *et al.*, 2011). Cette découverte ouvre de nouvelles questions sur l'implication de cette molécule d'ADN-B dans l'étiologie des maladies du ToLCD-TYLCD sur tomate et du PYVD sur piment-poivron associées au PepYVMLV. Enfin, cette activité a permis l'identification de six espèces de plantes sauvages potentiellement réservoirs ou alternatives des bégomovirus, dont certaines plantes très communes autour des parcelles cultivées comme *Sida acuta* (Malvaceae) et *Amaranthus* spp. (Amaranthaceae), également utilisées comme légume feuille; ces plantes pourraient jouer un rôle essentiel pendant les périodes d'inter-cultures et plus largement dans la dynamique virale au sein des agroécosystèmes.

Globalement, les résultats de ce premier chapitre ont permis de confirmer à la fois la non-exhaustivité de notre connaissance de la diversité des géminivirus présents sur les cultures maraîchères au Burkina Faso et en Côte d'Ivoire. La description de plusieurs groupes viraux distincts sur les cultures de tomates dans toutes les zones géographiques échantillonnées confirme l'intérêt du couple géminivirus/tomate en tant que modèle d'étude de l'émergence virale et la grande sensibilité de cette plante à l'infection par les géminivirus. Pour finir, l'association du nouveau composant d'ADN-B avec le PepYVMLV et la détection de ce virus sur, au moins, cinq plantes sauvages connues comme adventices des cultures pourraient jouer un rôle essentiel dans la forte prévalence du PepYVMLV observée sur les solanées cultivées au Burkina Faso, et plus largement en Afrique Occidentale et Centrale.

Article 1

Molecular epidemiology of vegetable-infecting geminiviruses in Burkina Faso

Article 1

Molecular epidemiology of vegetable-infecting geminiviruses in Burkina Faso

Alassane Ouattara[1,2,3,4,5] , Fidèle Tiendrébéogo[1,5] , Pierre Lefeuvre[2], Murielle Hoareau[2], Sohini Claverie[2,3], Agathe Allibert[2], Edgar Valentin Traoré[1,5], Nicolas Barro[4], Oumar Traoré[1,5], Jean-Michel Lett[2]

[1] INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

[2] CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

[3] Université de La Réunion, UMR PVBMT, 97410 Saint-Pierre, Ile de La Réunion, France

[4] Université Ouaga I Professeur Joseph Ki-Zerbo, 03 BP 7021, Ouagadougou 03, Burkina Faso

[5] Laboratoire Mixte International Patho-Bios, IRD-INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

Abstract

Background: In West and Central Africa, as in many regions of the world, vegetables are severely affected by geminivirus diseases which represent a major threat to tomato, pepper and okra production. In Burkina Faso, the observation of various virus-like symptoms, especially on tomato, suggests the involvement of several geminiviruses and associated satellites, and the pressing need for additional information on their diversity, distribution, prevalence and host plant reservoirs.

Findings: Surveys conducted in 48 localities of Burkina Faso confirmed the presence of tomato (yellow) leaf curl diseases (ToLCD-TYLCD) and geminiviruses in all localities with prevalence means of 25 and 44%, respectively. The highest values of ToLCD-TYLCD prevalence (42%), geminivirus prevalence (78%) and mixed infection (11%) were observed in the Sudan-Sahel and Sudan areas. Five geminiviruses including four begomovirus species *Pepper yellow vein Mali virus*, *Tomato leaf curl Burkina Faso virus*, *Tomato leaf curl Mali virus* and *Tomato leaf curl Ghana virus*, and a dicot-infecting mastrevirus species *Chickpea chlorotic dwarf virus* were characterized on tomato. In addition, pepper yellow vein Mali virus (PepYVMLV) and cotton leaf curl Gezira virus (CLCuGeV) were characterized on pepper and okra, respectively, in combination or not with alphasatellites and betasatellites for CLCuGeV. The most severe, prevalent and widely distributed virus on vegetables was PepYVMLV which was characterized for the first time in combination with a newly described DNA-B species. Finally, of the eight weeds identified as potential reservoir hosts of begomoviruses. Four of them host PepYVMLV, exclusively.

Conclusion: Our results confirm the importance of geminivirus diseases on vegetable crops in Burkina Faso and highlight the complex association of geminiviruses and satellites. The association between PepYVMLV and a newly described DNA-B component could be an important epidemiological factor for the emergence of PepYVMLV as the most prevalent and severe plant virus disease on tomato and pepper in Burkina Faso. The detection of begomoviruses in weeds close to crops points towards the increased necessity of considering reservoir plants in the control of virus diseases.

1. Introduction

Emerging plant viruses cause major economic losses on food and fiber crops, and pose a continuous threat to global food security (Anderson *et al.*, 2004; Jones, 2009). Since the late 1980s, begomoviruses (*Geminiviridae* family), transmitted by the whitefly *Bemisia tabaci*, draw particular attention due to its emergence on a large set of monocots and dicots worldwide and especially in Africa (Rojas *et al.*, 2005; Mansoor *et al.*, 2006; Navas-Castillo *et al.*, 2011). Beside the well-characterized cassava mosaic geminiviruses (Fauquet *et al.*, 2008; Patil & Fauquet, 2009) and African streak viruses (Martin & Shepherd, 2009), a flurry of distinct begomovirus species has been described over a large range of vegetable and staple crops in Africa.

Begomovirus genomes consist of one (monopartite) or two (bipartite) single-stranded circular DNA (ssDNA) components of ~2.7 kb that conjointly encode four to eight proteins involved in viral replication, movement, transmission and pathogenesis (Mansoor *et al.*, 2006; Jeske, 2009; Scholthof *et al.*, 2011; Zerbini *et al.*, 2017). Interestingly, except for a small number of bipartite begomoviruses (presenting with the so-called DNA-A and DNA-B components) described in the Old World (De Bruyn *et al.*, 2012; Leke *et al.*, 2016a,b), most African begomoviruses are monopartite with a unique DNA-A-like component (Brown *et al.*, 2015). Nevertheless, an increasing body of evidence has demonstrated that most monopartite begomoviruses are associated with satellites DNA. Betasatellites and alphasatellites are about half the size of a begomovirus component (~1.4 kb) and require the helper begomovirus for systemic spread and insect transmission (Saunders *et al.*, 2000; Briddon & Stanley, 2006). Betasatellite DNAs are required for induction of typical disease symptoms and depend on the helper begomovirus for replication and movement (Briddon & Stanley, 2006; Mansoor *et al.*, 2006). Alphasatellites are also commonly associated with Old World monopartite begomoviruses. These alphasatellites encode for a replication associated protein and are able to replicate autonomously (Mansoor *et al.*, 1999; Saunders & Stanley, 1999; Briddon *et al.*, 2004). In contrast with betasatellites, alphasatellites have not been shown to play a role in the etiology of begomovirus-induced diseases.

In recent years, vegetable crops and in particular tomato crops have been negatively impacted by epidemics of leaf curling and yellowing diseases in many countries of West and Central Africa (e.g., Benin, Burkina Faso, Cameroon, Ghana, Côte d'Ivoire, Mali, Niger, Nigeria, Senegal and Togo; for review Leke *et al*, 2015). A complex of at least 10 species of begomoviruses have been characterized in association with okra leaf curl disease (OLCD), tomato leaf curl or yellow leaf curl diseases (ToLCD-TYLCD) and pepper yellow vein disease (PYVD) in this region. *Cotton leaf curl Gezira virus* (CLCuGV; Shih *et al.*, 2009; Tiendrébéogo *et al.*, 2010a,b; Séka *et al.*, 2016), *Okra leaf curl Cameroon virus* (OLCCMV; Leke et al, 2013) and *Okra yellow crinkle virus* (OYCrV; Shih *et al.*, 2007; Kon *et al.*, 2009) have been associated with OLCD, in combination or not with *Cotton leaf curl Gezira betasatellite* (CLCuGeB; Tiendrébéogo *et al.*, 2010a), *Cotton leaf curl Gezira alphasatellite* (CLCuGeA), *Okra leaf curl Burkina Faso alphasatellite* (OLCBFA), *Okra leaf curl Mali alphasatellite* (OLCuMLA) and *Okra yellow crinkle alphasatellite* (OYCrA; Tiendrébéogo *et al.*, 2010a; Leke *et al.*, 2013).

Tomato leaf curl Burkina Faso virus (ToLCBFV; Ouattara *et al.*, 2017a), *Tomato leaf curl Cameroon virus* (ToLCCMV; Leke *et al.*, 2011), *Tomato leaf curl Nigeria virus* (ToLCNGV; Kon & Gilbertson, 2011), *Tomato leaf curl Mali virus* (ToLCMLV; Osei *et al.*, 2008; Zhou *et al.*, 2008; Leke *et al.*, 2011), *Tomato leaf curl Ghana virus* (ToLCGHV; Kon & Gilbertson, 2012), *Tomato yellow leaf curl Mali virus* (TYLCMLV; Chen *et al.*, 2009; Lett *et al.*, 2009; Sattar *et al.*, 2015b) and recently the dicot-infecting *Mastrevirus* species *Chickpea chlorotic dwarf virus* (CpCDV; Ouattara *et al.*, 2017b) have been associated with ToLCD-TYLCD, in combination or not with CLCuGeB (TYLCMLV; Chen *et al.*, 2009), *Tomato leaf curl Togo betasatellite* (ToLCTGB; Kon & Gilbertson, 2012), *Ageratum leaf curl Cameroon alphasatellite* (ALCCMA; Leke & Kvarnheden, 2014), and *Tomato leaf curl Cameroon alphasatellite* (ToLCCMA; Leke *et al.*, 2011). Using agroinoculation systems, Chen *et al.* (2009) demonstrated that TYLCMLV in association with CLCuGeB causes a severe symptom phenotype of ToLCD-TYLCD, similar to that observed in tomato fields in Mali and neighboring countries. *Pepper yellow vein Mali virus* (PepYVMLV) has been

associated with PepYVD and ToLCD-TYLCD (Zhou *et al.*, 2008; Tiendrébéogo *et al.*, 2011; Séka *et al.*, 2017).

Several lines of evidence have suggested a complex begomovirus etiology for these diseases with several begomovirus and satellite species involved in association or not with characteristic disease symptoms, and high populations of whiteflies (Konaté *et al.*, 1995; Zhou *et al.*, 2008; Chen *et al.*, 2009). In Burkina Faso, the observation of various virus-like symptoms suggests the involvement of several viruses possibly associated with satellites and the pressing need for additional information on their diversity and distribution. This study was undertaken in Burkina Faso to inventory geminiviruses infecting vegetable crops and weeds serving as potential reservoirs or alternative hosts and to study their main epidemiological parameters such as disease and virus prevalence in the field.

2. Materials and Methods

2.1. Samples collection

Surveys were conducted from 2013 to 2016 in the three agroclimatic areas (Figure 1a; the humid Sudan [annual rainfall between 900 and 1200 mm], the sub-humid Sudan-Sahel [annual rainfall between 600 and 900 mm] and the dry Sahel [annual rainfall between 400 and 600 mm] areas) of Burkina Faso. Leaves were sampled form a total of 1650 cultivated and 413 uncultivated plant in 48 localities (Figure 1a). The plant species, the name of the locality and the geographical coordinates (latitude, longitude) of samples were recorded. All samples were collected in envelopes and dried using an oven. Of the 1650 vegetable samples, 450 samples (tomato, okra, sweet and hot peppers) were collected based on their virus-like symptoms from 2013 to 2016, and 1200 tomato samples were collected regardless of symptomatology (blind random sampling) in 2015 and 2016, to study the prevalence of ToLCD-TYLCD and geminiviruses on this host. This sampling was performed from ten localities (three localities in the Sahel area, four in the Sudan-Sahel area and three in the Sudan area). One and three fields per locality were surveyed in 2015 and 2016, respectively. Leaves from 30 tomato plants were randomly collected per field.

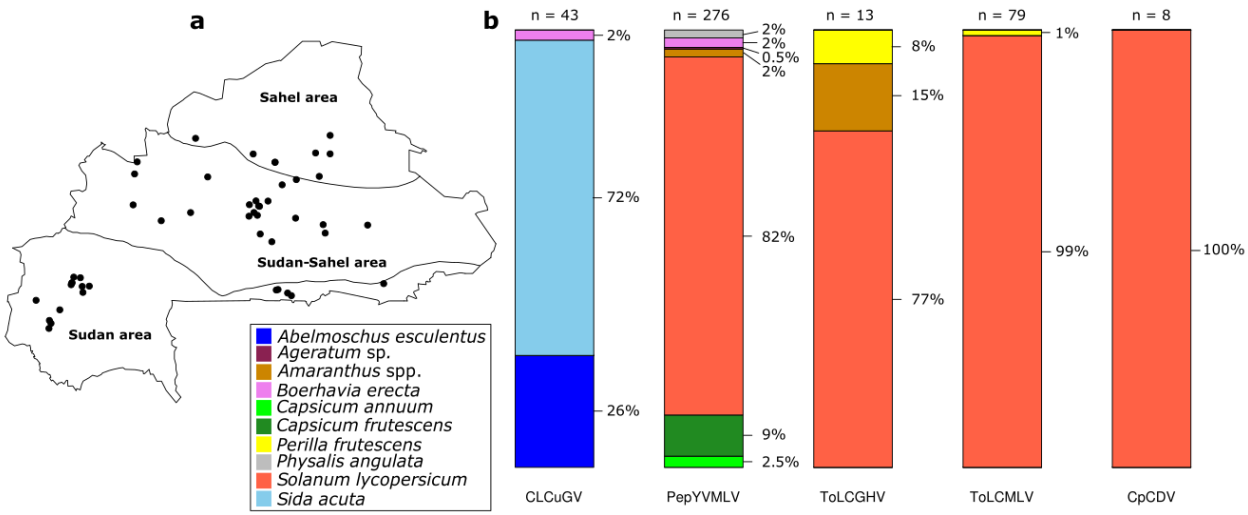


Figure 1: (a) Representation of the 48 locations surveyed from 2013 to 2016 in the three agro-ecological areas. **(b)** Barplots showing variable proportions of geminivirus infecting vegetables and uncultivated plants in the all surveyed locations in Burkina Faso

2.2. DNA extraction and PCR detection

Total DNA was extracted using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. The degenerate primer pair Clust4CP-F342 and Clust4CP-R1032 was used for the universal detection of begomoviruses by PCR (De Bruyn, 2014). Six sets of primer pairs were designed for the specific detection of tomato leaf curl Mali virus (ToLCMLV), tomato leaf curl Ghana virus (ToLCGHV), tomato leaf curl Burkina Faso (ToLCBFV), pepper yellow vein Mali virus (PepYVMLV) (DNA-A or -B) and chickpea chlorotic dwarf virus (CpCDV) (Supplementary table 1). The PCR screening of samples were carried out in 25 µL volumes with 5 µL of 5X buffer, 2.5 µL of dNTPs (2 mM), 1.5 µL of $MgCl_2$, 1 µL each of forward and reverse primers (10 mM) and 0.2 µL of GoTaq® Flexi DNA Polymerase. After an initial denaturation of 5 min at 94 °C, 30 cycles at 94 °C for 30 sec, 50 to 62 °C for 30 sec (see Supplementary table 1 for details) and 72 °C for 1 min were conducted before a final elongation step at 72 °C for 5 min. Amplifications were checked by electrophoresis onto 1% agarose gels.

2.3 Cloning and sequencing of full viral genomes

Complete sequences were obtained using a previously described RFLP-RCA method (Inoue-Nagata et al., 2004). Full genome amplicons were digested using *BamH*I, *EcoR*I or *Pst*I, before ligation into similarly linearized pGem-3zf (Promega, USA) or pJET1.2 (Thermo Fisher Scientific, USA) cloning vectors and transformation into *Escherichia coli*. The resulting clones and PCR amplicons obtained with the degenerate primers (Clust4CP) were sequenced by Sanger sequencing methods using primer-walking when needed (Macrogen, Europe).

2.4. Phylogenetic analysis

Sequences were edited and assembled using DNA Baser v.2.91 (Heracle BioSoft S.R.L., Romania) and Geneious R8 (Biomatters Ltd, New Zealand) softwares. Nucleotide sequences were subjected to a BLAST search for preliminary species assignment. Alignments of the request sequences along with representative geminivirus sequences available from GenBank were performed using the MUSCLE alignment tool (Edgar, 2004) available in MEGA 6 (Tamura *et al.*, 2013) and edited by eye. In order to assign sequences to known virus species and strains, pairwise identity comparisons of nucleotide sequences were performed using SDT v1.2 with pairwise deletion of gaps (Muhire *et al.*, 2013). Maximum-likelihood (ML) phylogenetic trees were constructed using FastTree v2.1.7 (Price *et al.*, 2010) and edited using FigTree v1.3 (available at http://tree.bio.ed.ac.uk/software/figtree/).

2.4. Field data analysis

Disease and virus prevalences were calculated as the percentage of ToLCD-TYLCD symptomatic and PCR positive tomato plants, respectively, among 30 plants for each site in the ten localities. Data processing and analysis were performed with R v3.3.2 (R Core Team, 2016).

3. Results

3.1. Survey for virus-like symptoms in vegetables and weeds and detection of geminiviruses DNA

Surveys conducted in the 48 localities of Burkina Faso confirmed the presence of ToLCD-TYLCD, PepYVD and OLCD. PCR diagnosis of begomoviruses based on universal primers revealed that 21.5% (355/1650) of the cultivated plants (tomato, pepper and okra) and 13% (52/413) of the uncultivated plants taken near the crops are infected by begomoviruses (Figure 1b). Direct sequencing of PCR amplicons and search for sequence similarity in the GenBank database (NCBI, BLASTn), revealed the presence of at least four species of begomoviruses in cultivated plants: CLCuGeV on okra (55%, n = 11/20), PepYVMLV on sweet and hot pepper (41%, n = 33/80) and tomato (15%, n = 226/1520), ToLCMLV on tomato (5%, n = 78/1520) and ToLCGHV on tomato (1%, n = 10/1520) (Figure 1b, Supplementary table 2). The same viruses were detected on uncultivated plants: CLCuGeV on *Sida acuta* (66%, n = 31/47); PepYVMLV on *Amaranthus hybridus* (24%, n = 4/17), *Amaranthus* sp. (3%, n = 1/33), *Ageratum* sp. (50%, n = 1/2), *Physalis angulata* (9%, n = 5/55) and *Boerhavia erecta* (22%, n = 6/27), ToLCMLV on *Perilla frutescens* (100%, n = 1/1) and ToLCGHV on *Amaranthus spinosus* (9%, n = 2/22) (for details see the Supplementary table 2).

Tomato plants showed three predominant symptom phenotypes (Figure 2): (phenotype 1) typical symptoms of leaf upcurling with or without stunting and yellowing, (phenotype 2) unusual symptoms of distortion and crumpling of leaves with yellowing, and (phenotype 3) interveinal yellowing and stunting of leaves without upcurling.

Figure 2: Three predominant symptom phenotypes observed on tomato plants surveyed between 2013-2016: **(a)** typical symptoms of leaf upcurling with or without stunting and yellowing (phenotype 1), **(b)** unusual symptoms of distortion and crumpling of leaves with yellowing **(**phenotype 2), and **(c)** interveinal yellowing and stunting of leaves without upcurling (phenotype 3).

3.2. Prevalence of ToLCD-TYLCD and geminiviruses on tomato

The prevalence of ToLCD-TYLCD and associated geminiviruses was measured with blind random sampling of 30 tomato samples per field, in 40 fields surveyed in the three agroecological areas of Burkina Faso (Table 1, Figure 3). Symptoms of ToLCD-TYLCD were observed in all localities with a mean disease prevalence of 25% (Table 1). The highest values of disease prevalence were observed in the Sudan-Sahel (average prevalence 30 ± 13%) and Sudan (average prevalence 30 ± 10%) areas for which no significant difference was observed (p = 0.98). In contrast, significant differences were observed between the Sahel area (average prevalence 13 ± 9%) and the both Sudan-Sahel and Sudan areas ($p < 10^{-6}$). However, disease prevalences varied widely according to localities (2 ± 2% to 42 ± 15%). The highest disease prevalences were observed in the Sudan-Sahel localities of Loumbila (42 ± 15%) and Korsimoro (40 ± 26%). The locality least affected by ToLCD-TYLCD was Dakiri in the Sahel area with a prevalence of only 2 ± 2%.

Table 1: Prevalence of tomato leaf curl disease and geminiviruses in tomato fields sampled in ten locations in Burkina Faso from 2015 to 2016.

Area	Localities	Disease prevalence ± SD (%) [symptomatic samples/total]	Geminivirus DNA prevalence ± SD (%) [PCR positive samples/total]
Sahel	Dakiri	2 ± 2 [2/120]	2 ± 2 [2/120]
	Dem	18 ± 13 [22/120]	23 ± 5 [28/120]
	Ouahigouya	18 ± 10 [21/119]	45 ± 43 [53/119]
		13 ± 9 [45/359]¤	23 ± 22 [83/359]¤
Soudan-Sahel	Guiédougou	19 ± 18 [22/119]	53 ± 43 [63/119]
	Korsimoro	40 ± 26 [48/120]	78 ± 37 [94/120]
	Loumbila	42 ± 15 [51/120]	62 ± 22 [74/120]
	Tandjari	17 ± 5 [20/120]	32 ± 13 [38/120]
		30 ± 13 [141/479]¤	56 ± 19 [269/479]¤
Soudan	Karangasso-sambla	20 ± 10 [24/120]	31 ± 16 [37/120]
	Leguéma	31 ± 18 [37/120]	63 ± 44 [76/120]
	Toussiana	39 ± 20 [47/120]	62 ± 33 [74/120]
		30 ± 10 [108/360]¤	52 ± 18 [187/360]¤

Based on PCR diagnosis using virus specific primers, geminiviruses were detected on tomato in all localities surveyed but with very variable prevalences (Table 1). In agreement with the localities with the highest disease prevalences, the highest geminivirus prevalences were observed in the Sudan-Sahel localities of Korsimoro (78 ± 37%) and Loumbila (62 ± 22%), and the Sudan localities of Leguéma (63 ± 44%) and Toussiana (63 ± 44%). Four geminiviruses were detected in the 1200 tomato samples evaluated: PepYVMLV, ToLCMLV, ToLCGHV and CpCDV (Figure 3). PepYVMLV was the most prevalent begomovirus species on tomatoes in the three agro-climatic areas of Burkina Faso with prevalence ranging between 0 and 45%, 31 and 78%, 31 and 48% in the Sahel, Sudan-Sahel and Sudan areas, respectively. The second most prevalent begomovirus was the ToLCMLV detected only in the Sudan area with prevalence varying between 10 and 53%. The least prevalent begomovirus is ToLCGHV which was detected only in four tomato plants from the same locality Toussiana in the Sudan area. The mastrevirus CpCDV was detected in the three agroecological areas but in a very low prevalence, below 3% (Figure 3). The use of

virus specific primers revealed the occurrence of mixed infection in 11% of the samples tested from the Sudan area (38/360). Mixed infections consisted of 37 cases of double infection (PepYVMLV-ToLCMLV, 36/360; PepYVMLV-CpCDV, 1/360) and a case of triple infection (PepYVMLV-ToLCMLV-CpCDV; 1/360).

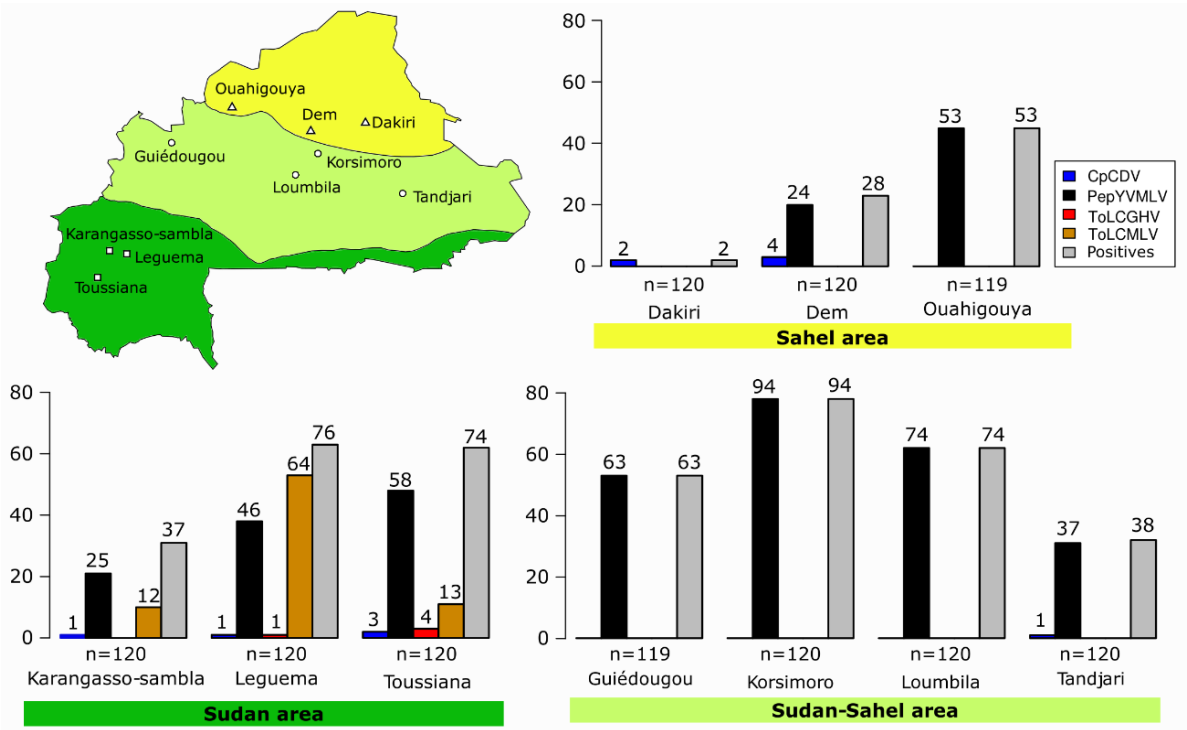


Figure 3: Distribution of tomato-infecting geminivirus according to the agroclimatic areas and the ten surveyed locations. The bars were colored according to the four geminivirus species. The height of each bar corresponds to the prevalence of each virus species-detected (%) calculated by the ratio of the number of infected plants to the total of tested plants. The grey bar represents the proportion of infected plants regardless virus species calculated by the ratio of number of infected plants to the number of tested plants per location. The number of infected plants is given at the top of each bar and the values under bars correspond to the number of tested plants par location.

3.3. Molecular characterization of a complex of five begomoviruses species and a mastrevirus

Using RCA-RFLP protocol for geminivirus amplification, cloning and sequencing (Shepherd *et al.*, 2008), 134 DNA sequences were obtained form 117 out of the 2063 leaf samples collected in Burkina Faso from 2013 to 2016 (Supplementary table 3).

These sequences consisted of 84 begomovirus DNA-A, 33 begomovirus DNA-B, eight alphasatellites, nine betasatellites and two mastreviruses.

The 84 complete DNA-A sequences of begomovirus (2732-2822 nucleotides) were obtained from tomato (n = 57), peppers (n = 10), okra (n = 07), S*ida acuta* (n = 08), *Amaranthus* sp. (n = 1) and *Amaranthus hybridus* (n = 1). All sequences present the typical genome organization of Old World monopartite begomoviruses with the presence of six open reading frames (ORFs) corresponding to V1 and V2 on the virion strand and C1, C2, C3 and C4 on the complementary strand (Rojas *et al.*, 2005). In the IR region located between the start codons of the C1 and V2, all components present a typical replication origin (↓), including an inverted repeat sequence containing the highly conserved nonanucleotide sequence TAATATT↓AC (Zhou *et al.*, 2008; Kon *et al.*, 2009).

Sequence analysis revealed that the ten sequences isolated from peppers, one from *S. acuta* and the sequence obtained from *Amaranthus* sp. were all isolates of PepYVMLV with similarities ranging from 96 to 100% (Figure 4a).

Sequences isolated from okra shared 99-100% similarities with Burkinabe isolates of the Niger strain of CLCuGeV (CLCuGeV-NE). Conversely, the seven CLCuGeV genomes isolated from *S. acuta* were only 91-92% similar to CLCuGV-NE isolates and form a distinct clade on the phylogeny (Figure 4a). According to the International Committee on Taxonomy of Viruses (ICTV) guidelines for species and strain demarcation in the genus *Begomovirus* (Brown *et al.*, 2015), these sequences are members of a new strain of CLCuGeV provisionally named Sida (CLCuGeV-Si).

Begomoviruses isolated from tomato belong to four distinct clades (West African tomato group 1 to 4) (Figure 4a). The 19 DNA-A sequences from the West African tomato group 1 shared 99 to 100% similarities between them and the highest nucleotide similarities of 96-97% with isolates of the Burkina Faso strain of ToLCMLV (ToLCMLV-BF). The 2 DNA-A sequences from the West African tomato group 2 shared 94% similarities between them and the highest nucleotide similarities of 92-94 % with isolates of ToLCGHV. The DNA-A sequence from the West African tomato group 3 shared the highest nucleotide similarities of 85% with isolates of PepYVMLV,

suggesting the presence of a new tomato-infecting begomovirus named ToLCBFV (Ouattara *et al.*, 2017a). Finally, 35 DNA-A sequences from the West African tomato group 4 shared 95 to 100% similarities between them and the highest nucleotide similarities of 96-99% with isolates of PepYVMLV.

Unexpectedly, 33 complete sequences of DNA-B component were obtained from cultivated (tomato, n = 19; pepper, n = 10) and uncultivated (*Amaranthus* sp., n = 1; *Amaranthus hibrydus,* n = 1; *Boerhavia erecta,* n = 1; *Sida acuta,* n = 1) plants (Figure 4b, Supplementary table 3). A DNA-B sequence found from tomato and the other one found from *Amaranthus hibrydus* were not associated with geminivirus. In contrast, DNA-B sequence characterized from *Sida acuta* (n = 1) and the others from tomato (n = 18) were associated with CLCuGeV and PepYVMLV (based on PCR or cloning), respectively. These components were 2657 to 2700 nucleotides long and ranged 92-100% similarity amongst themselves. All presented the typical genome organization of begomovirus DNA-B component with the two ORFs BV1 and BC1 and a common region (CR) (Padidam *et al.*, 1995). BLAST searches and pairwise comparison to reference sequences using SDT revealed that these sequences shared the highest nucleotide similarity of 64-66% with the DNA-B component of East African cassava mosaic Kenya virus and East African cassava mosaic Zanzibar virus *(*EACMKV, AJ704965; EACMZV, AJ704942; Supplementary table 3, Supplementary figure 1). The results of phylogenetic analysis grouped these 33 DNA-B sequences in a unique clade distantly related to the DNA-B sequences of the African cassava-infecting begomoviruses and the recently characterized West Africa Asystasia begomovirus 1 (WAAV1, Figure 4b) (Wyant *et al.*, 2015). Based on the recommendations of Briddon et *al.* (2010), theses sequences could be assigned to a new DNA-B species tentatively named PepYVMLV associated DNA-B.

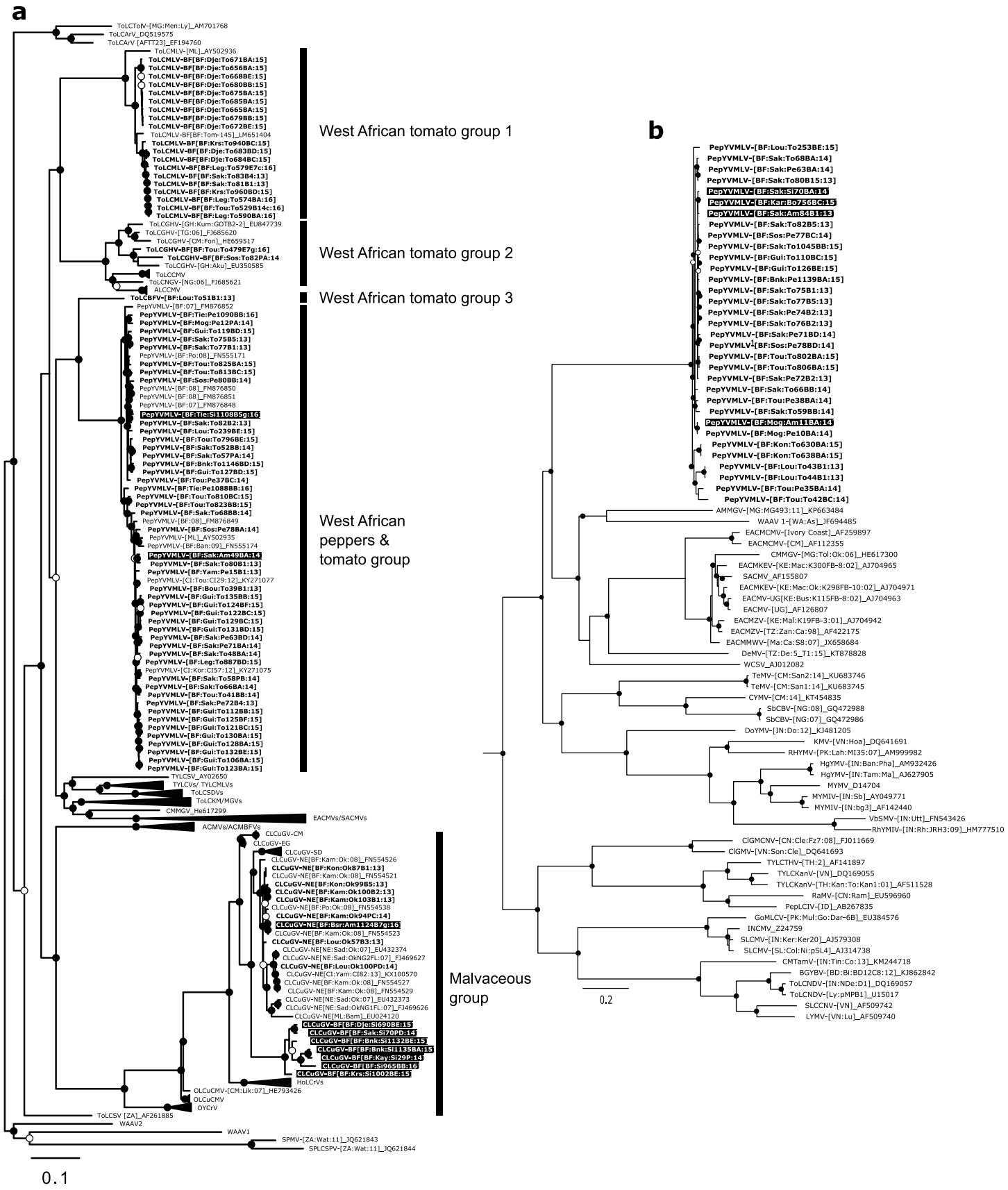

a
ToLCToIV-[MG:Men:Ly]_AM701768
ToLCArV_DQ519575
ToLCArV [AFTT23]_EF194760
ToLCMLV-[ML]_AY502936
ToLCMLV-BF[BF:Dje:To671BA:15]
ToLCMLV-BF[BF:Dje:To656BA:15]
ToLCMLV-BF[BF:Dje:To668BE:15]
ToLCMLV-BF[BF:Dje:To680BB:15]
ToLCMLV-BF[BF:Dje:To675BA:15]
ToLCMLV-BF[BF:Dje:To685BA:15]
ToLCMLV-BF[BF:Dje:To665BA:15]
ToLCMLV-BF[BF:Dje:To679BB:15]
ToLCMLV-BF[BF:Dje:To672BE:15]
ToLCMLV-BF[BF:Tom-145]_LM651404
ToLCMLV-BF[BF:Krs:To940BC:15]
ToLCMLV-BF[BF:Dje:To683BD:15]
ToLCMLV-BF[BF:Dje:To684BC:15]
ToLCMLV-BF[BF:Leg:To579E7c:16]
ToLCMLV-BF[BF:Sak:To83B4:13]
ToLCMLV-BF[BF:Sak:To81B1:13]
ToLCMLV-BF[BF:Krs:To960BD:15]
ToLCMLV-BF[BF:Leg:To574BA:16]
ToLCMLV-BF[BF:Tou:To529B14c:16]
ToLCMLV-BF[BF:Leg:To590BA:16]
ToLCGHV-[GH:Kum:GOTB2-2]_EU847739
ToLCGHV-[TG:06]_FJ685620
ToLCGHV-[CM:Fon]_HE659517
ToLCGHV-BF[BF:Tou:To479E7g:16]
ToLCGHV-BF[BF:Sos:To82PA:14
ToLCGHV-[GH:Aku]_EU350585
ToLCCMV
ToLCNGV-[NG:06]_FJ685621
ALCCMV
ToLCBFV-[BF:Lou:To51B1:13]
PepYVMLV-[BF:07]_FM876852
PepYVMLV-[BF:Tie:Pe1090BB:16]
PepYVMLV-[BF:Mog:Pe12PA:14]
PepYVMLV-[BF:Gui:To119BD:15]
PepYVMLV-[BF:Sak:To75B5:13]
PepYVMLV-[BF:Sak:To77B1:13]
PepYVMLV-[BF:Po:08]_FN555171
PepYVMLV-[BF:Tou:To825BA:15]
PepYVMLV-[BF:Tou:To813BC:15]
PepYVMLV-[BF:Sos:Pe80BB:14]
PepYVMLV-[BF:08]_FM876850
PepYVMLV-[BF:08]_FM876851
PepYVMLV-[BF:07]_FM876848
PepYVMLV-[BF:Tie:Si1108B5g:16
PepYVMLV-[BF:Sak:To82B2:13]
PepYVMLV-[BF:Lou:To239BE:15]
PepYVMLV-[BF:Tou:To796BE:15]
PepYVMLV-[BF:Sak:To52BB:14]
PepYVMLV-[BF:Sak:To57PA:14]
PepYVMLV-[BF:Bnk:To1146BD:15]
PepYVMLV-[BF:Gui:To127BD:15]
PepYVMLV-[BF:Tou:Pe37BC:14]
PepYVMLV-[BF:Tie:Pe1088BB:16]
PepYVMLV-[BF:Tou:To810BC:15]
PepYVMLV-[BF:Tou:To823BB:15]
PepYVMLV-[BF:Sak:To68BB:14]
PepYVMLV-[BF:08]_FM876849
PepYVMLV-[BF:Sos:Pe78BA:14]
PepYVMLV-[ML]_AY502935
PepYVMLV-[BF:Ban:09]_FN555174
PepYVMLV-[BF:Sak:Am49BA:14
PepYVMLV-[BF:Sak:To80B1:13]
PepYVMLV-[BF:Yam:Pe15B1:13]
PepYVMLV-[CI:Tou:CI29:12]_KY271077
PepYVMLV-[BF:Bou:To39B1:13]
PepYVMLV-[BF:Gui:To135BB:15]
PepYVMLV-[BF:Gui:To124BF:15]
PepYVMLV-[BF:Gui:To122BC:15]
PepYVMLV-[BF:Gui:To129BC:15]
PepYVMLV-[BF:Gui:To131BD:15]
PepYVMLV-[BF:Sak:Pe63BD:14]
PepYVMLV-[BF:Sak:Pe71BA:14]
PepYVMLV-[BF:Sak:To48BA:14]
PepYVMLV-[BF:Leg:To887BD:15]
PepYVMLV-[CI:Kor:CI57:12]_KY271075
PepYVMLV-[BF:Sak:To58PB:14]
PepYVMLV-[BF:Sak:To66BA:14]
PepYVMLV-[BF:Tou:To41BB:14]
PepYVMLV-[BF:Sak:Pe72B4:13]
PepYVMLV-[BF:Gui:To112BB:15]
PepYVMLV-[BF:Gui:To125BF:15]
PepYVMLV-[BF:Gui:To121BC:15]
PepYVMLV-[BF:Gui:To130BA:15]
PepYVMLV-[BF:Gui:To128BA:15]
PepYVMLV-[BF:Gui:To132BE:15]
PepYVMLV-[BF:Gui:To106BA:15]
PepYVMLV-[BF:Gui:To123BA:15]
West African tomato group 1
West African tomato group 2
West African tomato group 3
West African peppers & tomato group
TYLCSV_AY02650
TYLCVs/ TYLCMLVs
ToLCSDVs
ToLCKM/MGVs
CMMGV_He617299
EACMVs/SACMVs
ACMVs/ACMBFVs
CLCuGV-CM
CLCuGV-EG
CLCuGV-SD
CLCuGV-NE[BF:Kam:Ok:08]_FN554526
CLCuGV-NE[BF:Kon:Ok87B1:13]
CLCuGV-NE[BF:Kam:Ok:08]_FN554521
CLCuGV-NE[BF:Kon:Ok99B5:13]
CLCuGV-NE[BF:Kam:Ok100B2:13]
CLCuGV-NE[BF:Kam:Ok103B1:13]
CLCuGV-NE[BF:Po:Ok:08]_FN554538
CLCuGV-NE[BF:Kam:Ok94PC:14]
CLCuGV-NE[BF:Bsr:Am1124B7g:16]
CLCuGV-NE[BF:Kam:Ok:08]_FN554523
CLCuGV-NE[BF:Lou:Ok57B3:13]
CLCuGV-NE[NE:Sad:Ok:07]_EU432374
CLCuGV-NE[NE:Sad:OkNG2FL:07]_FJ469627
CLCuGV-NE[BF:Lou:Ok100PD:14]
CLCuGV-NE[CI:Yam:CI82:13]_KX100570
CLCuGV-NE[BF:Kam:Ok:08]_FN554527
CLCuGV-NE[BF:Kam:Ok:08]_FN554529
CLCuGV-NE[NE:Sad:Ok:07]_EU432373
CLCuGV-NE[NE:Sad:OkNG1FL:07]_FJ469626
CLCuGV-NE[ML:Bam]_EU024120
CLCuGV-BF[BF:Dje:Si690BE:15
CLCuGV-BF[BF:Sak:Si70PD:14
CLCuGV-BF[BF:Bnk:Si1132BE:15
CLCuGV-BF[BF:Bnk:Si1135BA:15
CLCuGV-BF[BF:Kay:Si29P:14
CLCuGV-BF[BF:Si965BB:16
CLCuGV-BF[BF:Krs:Si1002BE:15
HoLCrVs
OLCuCMV-[CM:Lik:07]_HE793426
OLCuCMV
OYCrV
Malvaceous group
ToLCSV [ZA]_AF261885
WAAV2
WAAV1
SPMV-[ZA:Wat:11]_JQ621843
SPLCSPV-[ZA:Wat:11]_JQ621844
0.1
b
PepYVMLV-[BF:Lou:To253BE:15]
PepYVMLV-[BF:Sak:To68BA:14]
PepYVMLV-[BF:Sak:Pe63BA:14]
PepYVMLV-[BF:Sak:To80B15:13]
PepYVMLV-[BF:Sak:Si70BA:14
PepYVMLV-[BF:Kar:Bo756BC:15
PepYVMLV-[BF:Sak:Am84B1:13
PepYVMLV-[BF:Sak:To82B5:13]
PepYVMLV-[BF:Sos:Pe77BC:14]
PepYVMLV-[BF:Sak:To1045BB:15]
PepYVMLV-[BF:Gui:To110BC:15]
PepYVMLV-[BF:Gui:To126BE:15]
PepYVMLV-[BF:Bnk:Pe1139BA:15]
PepYVMLV-[BF:Sak:To75B1:13]
PepYVMLV-[BF:Sak:To77B5:13]
PepYVMLV-[BF:Sak:Pe74B2:13]
PepYVMLV-[BF:Sak:To76B2:13]
PepYVMLV-[BF:Sak:Pe71BD:14]
PepYVMLV-[BF:Sos:Pe78BD:14]
PepYVMLV-[BF:Tou:To802BA:15]
PepYVMLV-[BF:Tou:To806BA:15]
PepYVMLV-[BF:Sak:Pe72B2:13]
PepYVMLV-[BF:Sak:To66BB:14]
PepYVMLV-[BF:Tou:Pe38BA:14]
PepYVMLV-[BF:Sak:To59BB:14]
PepYVMLV-[BF:Mog:Am11BA:14
PepYVMLV-[BF:Mog:Pe10BA:14]
PepYVMLV-[BF:Kon:To630BA:15]
PepYVMLV-[BF:Kon:To638BA:15]
PepYVMLV-[BF:Lou:To43B1:13]
PepYVMLV-[BF:Tou:To44B1:13]
PepYVMLV-[BF:Tou:Pe35BA:14]
PepYVMLV-[BF:Tou:To42BC:14]
AMMGV-[MG:MG493:11]_KP663484
WAAV 1-[WA:As]_JF694485
EACMCMV-[Ivory Coast]_AF259897
EACMCMV-[CM]_AF112355
CMMGV-[MG:Tol:Ok:06]_HE617300
EACMKEV-[KE:Mac:K300FB-8:02]_AJ704965
SACMV_AF155807
EACMKEV-[KE:Mac:Ok:K298FB-10:02]_AJ704971
EACMV-UG[KE:Bus:K115FB-8:02]_AJ704963
EACMV-[UG]_AF126807
EACMZV-[KE:Mal:K19FB-3:01]_AJ704942
EACMZV-[TZ:Zan:Ca:98]_AF422175
EACMMWV-[Ma:Ca:S8:07]_JX658684
DeMV-[TZ:De:5_T1:15]_KT878828
WCSV_AJ012082
TeMV-[CM:San2:14]_KU683746
TeMV-[CM:San1:14]_KU683745
CYMV-[CM:14]_KT454835
SbCBV-[NG:08]_GQ472988
SbCBV-[NG:07]_GQ472986
DoYMV-[IN:Do:12]_KJ481205
KMV-[VN:Hoa]_DQ641691
RHYMV-[PK:Lah:MI35:07]_AM999982
HgYMV-[IN:Ban:Pha]_AM932426
HgYMV-[IN:Tam:Ma]_AJ627905
MYMV_D14704
MYMIV-[IN:Sb]_AY049771
MYMIV-[IN:bg3]_AF142440
VbSMV-[IN:Utt]_FN543426
RhYMIV-[IN:Rh:JRH3:09]_HM777510
ClGMCNV-[CN:Cle:Fz7:08]_FJ011669
ClGMV-[VN:Son:Cle]_DQ641693
TYLCTHV-[TH:2]_AF141897
TYLCKanV-[VN]_DQ169055
TYLCKanV-[TH:Kan:To:Kan1:01]_AF511528
RaMV-[CN:Ram]_EU596960
PepLCIV-[ID]_AB267835
GoMLCV-[PK:Mul:Go:Dar-6B]_EU384576
INCMV_Z24759
SLCMV-[IN:Ker:Ker20]_AJ579308
SLCMV-[SL:Col:Ni:pSL4]_AJ314738
CMTamV-[IN:Tin:Co:13]_KM244718
BGYBV-[BD:Bi:BD12C8:12]_KJ862842
ToLCNDV-[IN:NDe:D1]_DQ169057
ToLCNDV-[Ly:pMPB1]_U15017
SLCCNV-[VN]_AF509742
LYMV-[VN:Lu]_AF509740
0.2

Figure 4: Maximum-likelihood phylogenetic tree inferred from alignments of complete **(a)** DNA-A and **(b)** DNA-B sequences of begomovirus groups. The new sequences characterized in our study are shown in bold. Sequences obtained from uncultivated plants are highlighted by using a black background. For better visibility, the bootstrap values between 50 and 70 are represented by white dots and those higher than 70 are represented by black dots. The horizontal scale indicates the genetic distance.

Seven alphasatellite sequences were characterized from okra samples for which CLCuGeV were detected by PCR and three complete DNA-A sequences were obtained ([BF:Kam:Ok100B2:13], [BF:Lou:Ok57B3:13] and [BF:Lou:Ok100PD:14], Supplementary table 3). In addition, another alphasatellite was described from *Amaranthus hybridus* for which no begomovirus was detected by PCR. Alphasatellite sequences consisted of 1346-1387 nucleotides and showed typical features consisting of the presence of a single *alpha-Rep* ORF in viral sense and *A-rich* region. All alphasatellites had a predicted alphasatellite hairpin structure with a loop containing a nonanucleotide, TAGTATT↓AC or CAGTATT↓AC, common to nanoviruses (Tiendrébéogo *et al.*, 2010a; Zhou, 2013). After sequence comparison and based on the established alphasatellite species demarcation threshold of 83% (Mubin *et al.*, 2009), it was apparent that four of the alphasatellites isolated from okra in association with CLCuGeV and from *Amaranthus hybridus* (89 to 100% similarity between them) were isolates of the *Okra leaf curl alphasatellite* with 97-98% similarity (Figure 5a). The remaining alphasatellites isolated from okra (96 to 100% similarity between them) were isolates of the *Cotton leaf curl Gezira alphasatellite* (CLCuGeA) with 96 to 98% similarities (Figure 5b).

Eight complete nucleotide sequences of betasatellites were obtained from okra (n = 5) and *Sida acuta* (n = 3) plants (Figure 5c, Supplementary table 3). Three of them obtained from okra (n = 2) and *Sida acuta* (n = 1) were associated with CLCuGeV. Another betasatellite sequence of 1203 nucleotides was identified from tomato in association with ToLCGHV-BF. Except this tomato-infecting betasatellite which had a partial ORF βC1, all the malvaceous-infecting betasatellites count 1344-1360 nucleotides and had the characteristic features including the βC1 ORF on the complementary strand, an adenine-rich (A-rich) sequence region and a stem-loop

structure containing the ubiquitous nonanucleotide TAATATTAC sequence that serves as the origin of virion-strand DNA replication (Briddon *et al.*, 2008; Zhou, 2013). The betasatellite sequences from *S. acuta* were 96-100% similar and shared 74-75% nucleotide similarities with *Chayote yellow mosaic Benin betasatellite* (ChYMBJB, KT454828) from Benin. These similarities fall below the established species demarcation threshold for betasatellites (78% nucleotide sequence identity; Briddon et al 2008), and thus form a new species tentatively named *Cotton leaf curl Gezira betasatellite 1* (CLCuGeB1). The betasatellite isolated from okra were 97-100% similar and shared 96-98% nucleotides similarities with CLCuGeB and OLCuMLB isolates. The proximity of these species on the phylogenetic tree (Figure 5c), suggest that they both belong to the same species. The two identical sequences of betasatellite, characterized on tomato and associated with ToLCGHV, shared only 75% sequences similarities with *Chayote yellow mosaic Benin betasatellite* (ChYMBJB, KT454828) from Nigeria and may constitute a new species provisionally named *Tomato leaf curl Ghana betasatellite* (ToLCGHB).

Along with the description of begomoviruses associated with ToLCD-TYLCD, two mastrevirus genomes of 2562 nucleotides were characterized from tomato (98.5% similarity with one another). The two genomes shared ~93% genome-wide pairwise identity with the Sudanese CpCDV-M strain (CpCDV-M[SD:SD231:13]_KM229909). Following the guidelines for mastrevirus classification (Muhire *et al.*, 2013), these isolates were assigned to a new strain named CpCDV-R (Ouattara *et al.*, 2017b).

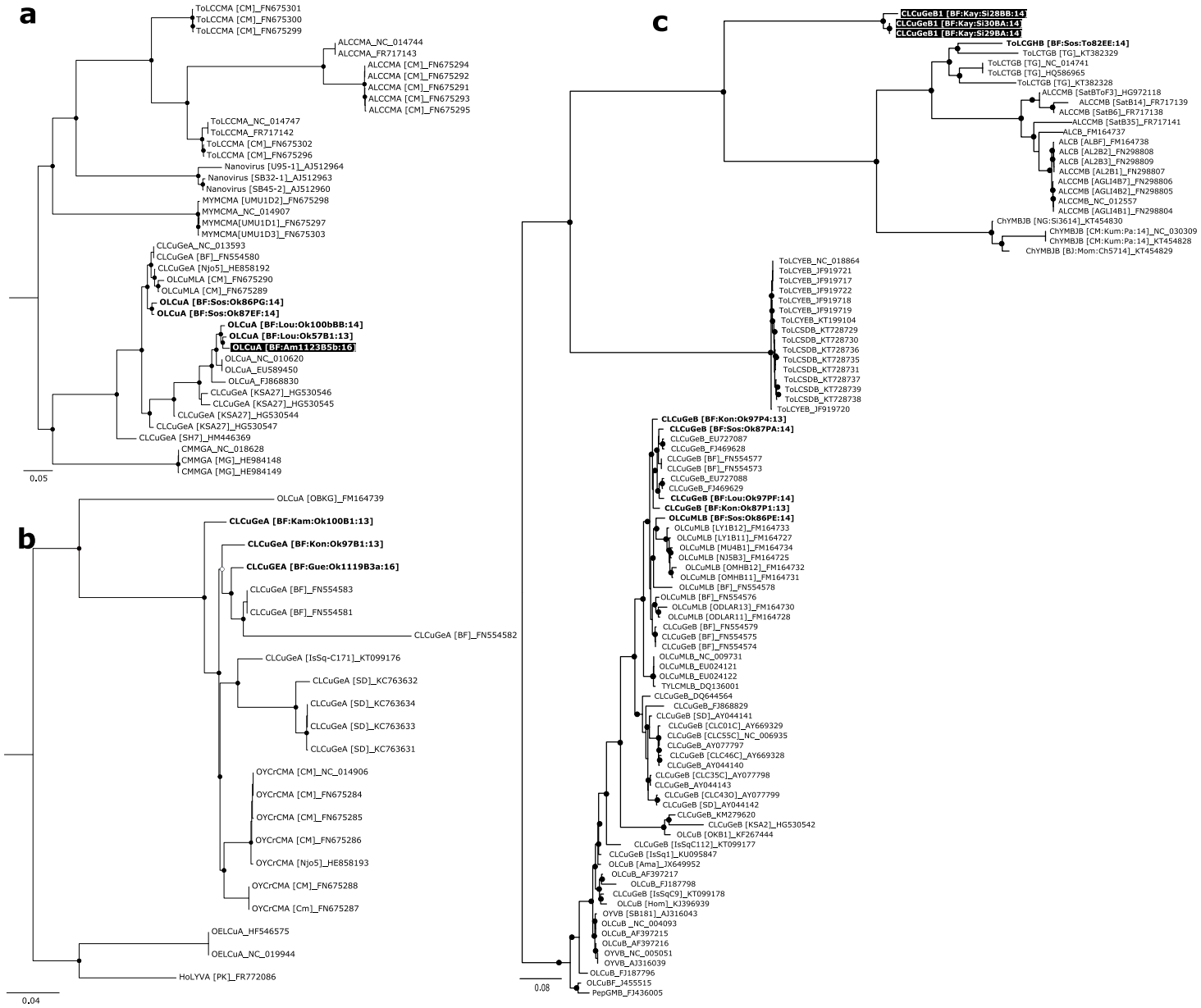


Figure 5: Maximum-likelihood phylogenetic tree inferred from alignments of complete alphasatellite (**a, b**) and betasatellite (**c**) sequences associated with begomovirus groups. The new sequences characterized in our study are shown in bold. Sequences obtained from uncultivated plants are highlighted by using a black background. For better visibility, the bootstrap values between 50 and 70 are represented by white dots and those higher than 70 are represented by black dots. The horizontal scale indicates the genetic distance.

4. Discussion

In West Africa, diseases caused by whitefly-transmitted begomoviruses have emerged as major constraints on production of vegetables, especially tomato, pepper, and okra (Chen *et al.*, 2009; Kon *et al.*, 2009; Tiendrébéogo *et al.*, 2010a, 2011; Leke *et al.*, 2012; Sattar *et al.*, 2015). Considerable progress have been made on the identification of begomoviruses circulating in that region (Kon & Gilbertson, 2011). Studies have revealed a diversity of genetically distinct and locally evolved viruses, thereby establishing West and Central Africa as a center of begomovirus diversity (Chen *et al.*, 2009; Kon *et al.*, 2009; Kon & Gilbertson, 2011; Leke *et al.*, 2015). In this study, we improve our understanding of diversity and etiology of geminiviruses associated with vegetable crops in Burkina Faso.

4.1. High prevalence of geminivirus-like diseases on vegetables in Burkina Faso

The large surveys conducted in the three agroclimatic areas of Burkina Faso confirmed the omnipresence of geminivirus-like diseases: ToLCD-TYLCD, PepYVD and OLCD. Our study revealed that on average one quarter (25%) of tomato plants randomly evaluated in fields present symptoms of ToLCD-TYLCD, and that almost half (44%) are infected by geminiviruses. Our results confirm the initial report of the existence of geminiviruses infecting crops in West and Central Africa in the 90s, based on serology and hybridization (N'Guessan *et al.*, 1992; Konaté *et al.*, 1995), and more recently based on advanced molecular techniques (for review Leke *et al.*, 2015). However, an increasing gradient of ToLCD-TYLCD prevalence and geminivirus prevalence from North to South were observed, with the highest values observed in the Sudan-Sahel and Sudan areas. The prevalence of begomoviruses on crops is directly related to the abundance and type of their insect vector *Bemisia tabaci* and/or the presence of virus reservoir or alternative plants in the agroecosystems. Gnankiné *et al.* (2013) showed that the polyphagous MED cryptic species of *B. tabaci* (formely Q biotype) was dominant on tomato crops in Burkina Faso. The presence of an abundant population of the MED cryptic species of *B. tabaci* is correlated with the both high diversity and prevalence of begomovirus diseases on vegetable crops. Moreover, the lower prevalence of diseases and the lower viral diversity recorded in the Sahel area seems to correlate

with the arid climate to an environment less favorable to weeds and potential virus plant reservoirs and to phytophagous insects.

4.2. Occurrence of geminiviruses and DNA satellites on vegetables in Burkina Faso

A complex of five begomoviruses, including CLCuGV on okra, ToLCMLV, ToLCGHV, ToLCBFV on tomato, PepYVMLV on peppers and tomato, and two species of betasatellites and three species of alphasatellites, mostly associated with the CLCuGV on okra, were identified on vegetables in Burkina Faso. These results are in agreement with previous reports suggesting a complex etiology for ToLCD-TYLCD (Sattar *et al.*, 2015; Ouattara *et al.*, 2017a), PepYVD (Tiendrébéogo *et al.*, 2008) and OLCD (Tiendrébéogo *et al.*, 2010a,b) with several begomovirus and DNA satellite species (Tiendrébéogo *et al.*, 2010a), and the recent and surprisingly description of a dicot-infecting mastrevirus CpCDV on tomato in Burkina Faso (Ouattara *et al.*, 2017b).

Even if CpCDV has already been detected on tomato in 2012 in Pakistan (GenBank accession number KP881605; unpublished data), the occurrence of this virus at a country level and its potential involvement as a causative agent of ToLCD-TYLCD disease represents a novelty of great significance for managing the “health” of tomato in Africa and more largely in the Old World, since CpCDV has also been found infecting leguminous hosts in Middle East and the Indian-subcontinent (Kraberger *et al.*, 2015). The presence of this dicot-infecting virus in Africa on two new hosts, tomato and papaya, confirms its wide occurrence and plant host range which reopens the question of the geographical origin of this highly interesting group of dicot-infecting viruses.

First reported in Mali, Benin, Ghana, Niger, Senegal, and Togo (Zhou *et al.*, 2008), our results confirmed the recent description of ToLCMLV on symptomatic tomato plants in Burkina Faso (Sattar *et al.*, 2015). Unlike ToLCMLV, this is the first description of ToLCGHV outside Ghana where it was first described (Osei *et al.*, 2008). The description of a new betasatellite named ToLCGHB in association with ToLCGHV on a tomato yellow leaf curl diseased plant opens the question of the impact of this betasatellite on symptoms induction. Chen *et al.* (2009) demonstrated that even if TYLCMLV causes TYLCD throughout West Africa, the association of TYLCMLV and CLCuGeB infrequently detected in the field, represent a reassortant that causes a severe

phenotype of TYLCD. Inversely, tomato leaf curl Togo betasatellite (ToLCTGB), described in association with ToLCGHV and ToLCNGV in Togo, seems to have no effect on disease symptoms on tomato induced by either begomovirus, but increased symptom severity by both viruses in *Nicotiana* spp. and *Datura stramonium* (Kon & Gilbertson, 2011).

TYLCMLV was reported on tomato in Burkina Faso in 2010 (Sattar *et al.*, 2015). But our study could not confirm its presence in several localities spread across the country including the region surveyed by Sattar *et al.* (2015), five years earlier. Although this result might suggest a low prevalence of TYLCMLV in Burkina Faso or a dynamic of viral communities over time, it highlights the importance of continuous epidemiological surveillance.

The description of a new tomato-infecting begomovirus ToLCBFV (Ouattara *et al.*, 2017a) confirms that West and Central Africa represents a center of tomato-infecting begomovirus diversity in that these diseases have emerged relatively recently and are caused by genetically distinct locally evolved species (Kon & Gilbertson, 2011). It reflects also the genetic diversity and wide distribution of indigenous begomoviruses present in weeds or other plants in inter-tropical regions and the capacity of these viruses to evolve into tomato-infecting variants following introduction or migration events mediated by high populations of *B. tabaci* (Zhou *et al.*, 2008).

The lack of genetic differentiation between PepYVMLV isolates described in our study on tomato, pepper and weeds and the isolates described elsewhere in West Africa on tomato and pepper (Zhou *et al.*, 2008; Tiendrébéogo *et al.*, 2011; Sattar *et al.*, 2015), suggests that there is no genetic structuring of this virus species depending on the host plant (cultivated or wild) and that the same viral population migrates from one host plant species to another. PepYVMLV was named on the basis of symptoms observed on pepper fields in the early 2000s (Tiendrébéogo *et al.*, 2008; Zhou *et al.*, 2008). This virus was subsequently reported in China (AM691548, AM691555), Burkina Faso (Tiendrébéogo *et al.*, 2011) and recently in Côte d'Ivoire (Séka *et al.*, 2017). Based on the infectivity of the agroinfectious clone generated from the virus, Zhou *et al.* (2008) concluded that this virus was more infectious on *Nicotiana benthamiana* and tomato

(*Solanum lycopersicum* cv. Glamor) than on pepper (*Capsicum annuum* cv. Cayenne Long and cv. Carolina). The work done by Tiendrebeogo *et al.* (2011) also revealed that PepYVMLV was virulent on tomato, but no symptoms were surprisingly observed on pepper in controlled chambers suggesting the absence of a component implicated in the pathogenicity on pepper. However, the present work revealed a new species of DNA-B distantly related with African begomoviruses involved in the cassava mosaic disease. Its discovery, mainly in association with PepYVMLV DNA-A components, suggests that PepYVMLV described until now as a monopartite begomovirus could be associated with a DNA-B component. As bottom lines of evidence, the similarity of the common region (CR) between the DNA-A and DNA-B components of PepYVMLV (90%, whereas DNA-B CR share less than 80% similarity with ToLGHV, ToLCBFV and ToLCMLV DNA-As; Supplementary figure 2) and the presence of similar iteron sequences (GGGTAC for both PepYVMLV DNAs whereas GGTCTC, GGGTCT and GGGTTC are used by ToLCGHV, ToLCBFV and ToLCMLV DNA-As, respectively) may indicate that these DNA components are coadapted. Moreover, the large distribution and the high prevalence of PepYVMLV observed in our study could also be the result of this positive association. However, the strict or unstrict association of these DNAs and the involvement of this DNA-B in the pathogenicity and more broadly in the epidemiology of PepYVMLV need further investigations.

Molecular characterization of tomato-begomoviruses associated with typical symptoms of TYLCD-ToLCD in Burkina Faso (this study and Sattar *et al.*, 2015), demonstrated that these symptoms of leaf upcurling with or without stunting and yellowing are associated with a complex of at least four begomoviruses: ToLCBFV, ToLCMLV, ToLCGHV, TYLCMLV. The unusual distortion and crumpling symptoms with yellowing observed in the field on tomato leaves have been associated with PepYVMLV. This unusual symptom phenotype has also been observed on tomato leaves in 2004 in Mali and consistently associated with PepYVMLV (Zhou *et al.*, 2008). The authors proposed to rename this begomovirus species tomato yellow leaf crumple virus based upon the predominant tomato disease symptoms in the field and the high rate of infectivity and symptoms induced in tomato by the agroinfectious clone and the absence of symptoms on pepper. Finally, the interveinal yellowing and stunting of

leaves without upcurling have been associated with the dicot-infecting mastrevirus CpCDV. Knowing that ToLCD-TYLCD are associated with more than 70 begomovirus species in the world, only a molecular diagnosis is able to identify the causative agent. For the particular tomato symptom phenotypes associated with PepYVMLV and CpCDV, monitoring the incidence and the distribution of these two viral agents in simple infection in the field on the basis of symptoms appears feasible.

First described on cotton in Sudan (Idris & Brown, 2000), CLCuGeV has been reported on okra in several countries bordering Burkina Faso with Mali (Kon *et al.*, 2009), Niger (Shih *et al.*, 2009) and recently in Côte d'Ivoire (Séka *et al.*, 2016). CLCuGeV was also described in Egypt, Cameroon and Nigeria (Bigarré *et al.*, 2001; Leke *et al.*, 2013; Silva *et al.*, 2017). Our study confirms the presence in Burkina Faso on cultivated and uncultivated malvaceous plants of the Niger strain of CLCuGeV in association with a complex of satellites including alphasatellites and betasatellites (Tiendrébéogo *et al.*, 2010b).

4.3. Host plant reservoirs of geminiviruses

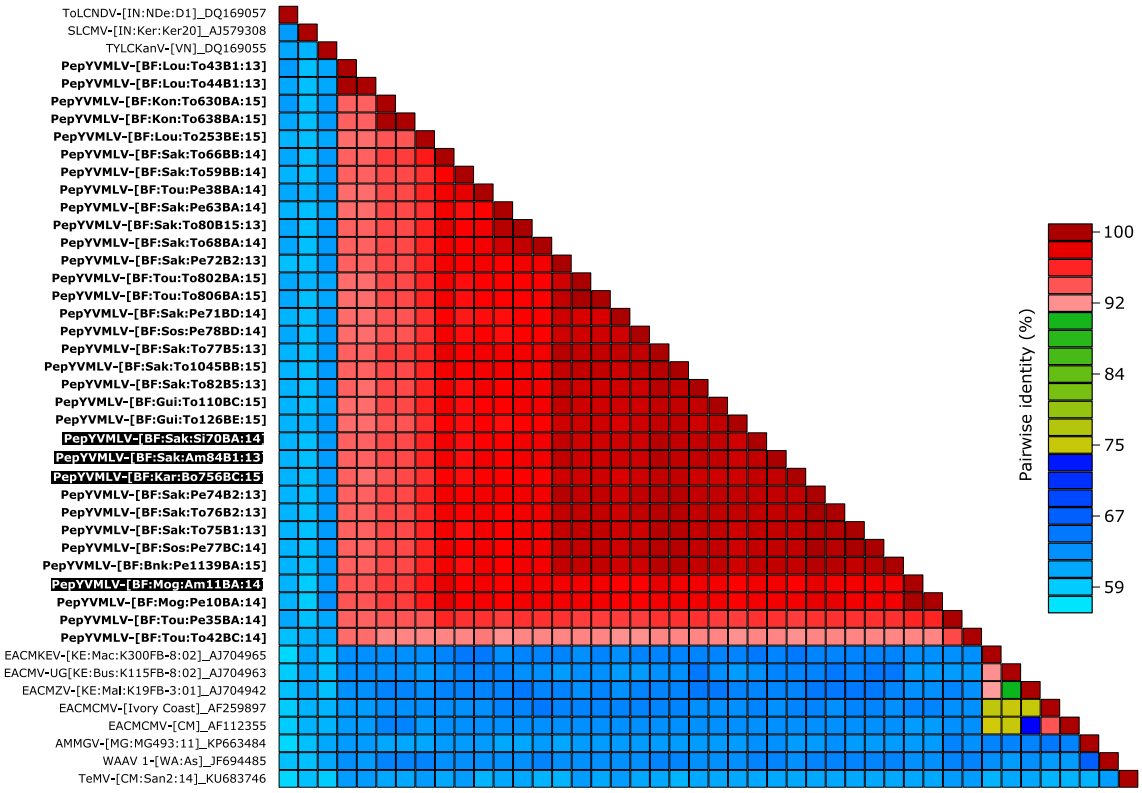
For a better understanding of the epidemiology of virus diseases and to develop appropriate control measures, the identification of alternative or reservoir hosts is an important aspect which has been somewhat neglected in the past (Leke *et al.*, 2015). In our study, we identified eight weeds as potential reservoir hosts of begomoviruses, four exclusively host PepYVMLV. Interestingly, despite their proximity and promiscuity on the fields, isolates of CLCuGeV isolated from *Sida acuta* and okra were genetically distant (90-92% nucleotides identities). In the past, work based on serological tests had reported *S. acuta* as an alternative host of okra leaf curl viruses implicated in OLCD in Burkina Faso (Konaté *et al.*, 1995). This weed has also been reported in Nigeria to be host of *Chayote yellow mosaic virus* (KT454826) and *Soybean chlorotic blotch virus* (KT454817). Due to its high resistance to drought, *S. acuta* could constitute an important reservoir of various viruses that remains uncharacterized. Conversely, PepYVMLV isolates found on *Amarantus* spp. and *S. acuta* are variants of the same strain Niger (PepYVMLV-NG) also characterized on pepper. This suggests that these weeds found around the pepper fields directly impacts PepYVMLV prevalence on

pepper. Based solely on PCR diagnosis, the detection of PepYVMLV on *Ageratum* sp., *Boerhavia erecta* and *Physalis angulata*, and ToLCGHV on *Amaranthus spinosus* and ToLCMLV and ToLCGHV on *Perilla frutescens* highlight further the existence of a large variety of uncultivated plant reservoirs.

Besides identifying a large set of geminivirus circulating within different agro-systems in Burkina Faso, our results highlights (1) the complex association of DNA components (monopartite and/or bipartite viruses, and satellites) in the field, and (2) the existence of weeds as potential reservoir or alternative hosts to geminiviruses encountered on cultivated plants. Whereas some degree of host association is observed for some virus species (e.g. CLCuGeV/okra and CLCuGeV/*S. acuta*); a more relax transition between cultivated and uncultivated plants seems to be the rule for others (e.g. PepYVMLV and cultivated solanaceous, *Amaranthus* spp., *Ageratum* sp., *B. erecta* and *P. angulata*). Understanding how these transitions participate to the global ToLCD-TYLCD in Burkina Faso and more largely in West and Central Africa is primordial to implement proper pest management. Concomitant use of metagenomic-based screening and molecular characterization could represent a powerful strategy.

Acknowledgements

This study was co-funded by the AIRD (PEERS, 2013-2014), the European Union (ERDF), the Conseil Régional de la Réunion and by the CIRAD. AO is a recipient of a PhD fellowship from CIRAD (2015-2017).



Supplementary figure 1: Color-coded matrix of pairwise nucleotide identity inferred from alignments of complete DNA-B sequences associated with PepYVMLV and selected DNA-B from NCBI.

Supplementary Table 1: Specific primers designed for geminivirus detection

Primer name	Primer sequence (5' – 3')	Amplicon size (bp)	Hybridation temperature (°C)	Targeted virus
PepYVMLV-A-F	GCTCTTGAGTGCGTAATTC	559	55	PepYVMLV DNA-A
PepYVMLV-A-R	ATGCAGATTCCGCTGAAG			
PepYVMLV-B-F	GAGATCCAGACAGGTACTG	1290	57	PepYVMLV DNA-B
PepYVMLV-B-R	GTCGACCTTCACTACTTCTC			
ToLCBFV-F	GTCTCTATATACTTCCTCC	1156	60	ToLCBFV DNA-A like
ToLCBFV-R	GTTCTCAAGCATCTGAAGC			
ToLCGHV-F	CACTCTTGGTCACGATCTG	595	62	ToLCGHV DNA-A like
ToLCGHV-R	CACTTGATAACGGTCTCTG			
ToLCMLV-F	TGTCATGTTCTACTTGGTC	652	62	ToLCMLV DNA-A like
ToLCMLV-R	GAACCACGACATGATATCAG			
CpCDV-F	TGTCGTCACACCAACAAG	671	60	CpCDV DNA-A
CpCDV-R	AGTCACTGAACGTGCCTCT			

Supplementary table 2: Results of the PCR diagnosis performed with the universal begomovirus primers on the samples collected between 2013 and 2016

	Family	Specie	CLCuGV	PepYVMLV	ToLCGHV	ToLCMLV	CpCDV	Percentage of infected plants (infected /tested)
Cultivated plants	Malvaceae	*Abelmoschus esculentus* (okra)	11	0	0	0	0	55 (11/20)
	Solanaceae	*Capsicum annuum* (sweet pepper)	0	7	0	0	0	33 (7/21)
		Capsicum frutescens (hot pepper)	0	26	0	0	0	44 (26/59)
		Solanum lycopersicum (tomato)	0	226	10	78	8	20 (311/1520)
Uncultivated plants	Amaranthaceae	*Amaranthus hybridus*	0	4	0	0	0	24 (4/17)
		Amaranthus sp	0	1	0	0	0	3 (1/33)
		Amaranthus spinosus	0	0	2	0	0	9 (2/22)
	Asteraceae	*Ageratum sp*	0	1	0	0	0	50 (1/2)
	Lamiaceae	*Perilla frutescens*	0	0	1	1	0	100 (1/1)
	Malvaceae	*Sida acuta*	31	0	0	0	0	66 (31/47)
	Nyctaginaceae	*Boerhavia erecta*	1	6	0	0	0	26 (7/27)
	Solanaceae	*Physalis angulata*	0	5	0	0	0	9 (5/55)
Total			43	276	12	78	8	20 (407/2063)

Supplementary table 3: List of complete sequences obtained using RCA-RFLP protocol for geminivirus amplification, cloning and sequencing

Sample ID	Isolate	Virus species	Component	Size (nt)
15BF13	[BF:Yam:Pe15B1:13]	PepYVMLV	DNA-A	2786
39BF13	[BF:Bou:To39B1:13]	PepYVMLV	DNA-A	2786
57BF13	[BF:Lou:Ok57B3:13]	CLCuGeV	DNA-A	2761
51BF13	[BF:Lou:To51B1:13]	ToLCBFV	DNA-A	2784
72BF13	[BF:Sak:Pe72B4:13]	PepYVMLV	DNA-A	2786
77BF13	[BF:Sak:To77B1:13]	PepYVMLV	DNA-A	2778
80BF13	[BF:Sak:To80B1:13]	PepYVMLV	DNA-A	2785
82BF13	[BF:Sak:To82B2:13]	PepYVMLV	DNA-A	2779
81BF13	[BF:Sak:To81B1:13]	ToLCMLV	DNA-A	2804
83BF13	[BF:Sak:To83B4:13]	ToLCMLV	DNA-A	2804
75BF13	[BF:Sak:To75B5:13]	PepYVMLV	DNA-A	2783
87BF13	[BF:Kon:Ok87B1:13]	CLCuGeV	DNA-A	2761
99BF13	[BF:Kon:Ok99B5:13]	CLCuGeV	DNA-A	2761
100BF13	[BF:Kam:Ok100B2:13]	CLCuGeV	DNA-A	2761
103BF13	[BF:Kam:Ok103B1:13]	CLCuGeV	DNA-A	2761
12BF14	[BF:Mog:Pe12PA:14]	PepYVMLV	DNA-A	2779
29BF14	[BF:Kay:Si:29PA:14]	CLCuGeV	DNA-A	2752
100BF14	[BF:Lou:Ok100PD:14]	CLCuGeV	DNA-A	2761
94BF14	[BF:Kam:Ok94PC:14]	CLCuGeV	DNA-A	2762
37BF14	[BF:Tou:Pe37BC:14]	PepYVMLV	DNA-A	2785
41BF14	[BF:Tou:To41BB:14]	PepYVMLV	DNA-A	2786
70BF14	[BF:Sak:Si70PD:14]	CLCuGeV	DNA-A	2744
71BF14	[BF:Sak:Pe71BA:14]	PepYVMLV	DNA-A	2787
49BF14	[BF:Sak:Am49BA:14]	PepYVMLV	DNA-A	2785
48BF14	[BF:Sak:To48BA:14]	PepYVMLV	DNA-A	2787
52BF14	[BF:Sak:To52BB:14]	PepYVMLV	DNA-A	2784
57BF14	[BF:Sak:To57PA:14]	PepYVMLV	DNA-A	2778
58BF14	[BF:Sak:To58PB:14]	PepYVMLV	DNA-A	2788
66BF14	[BF:Sak:To66BA:14]	PepYVMLV	DNA-A	2787
68BF14	[BF:Sak:To68BB:14]	PepYVMLV	DNA-A	2785
63BF14	[BF:Sak:Pe63BD:14]	PepYVMLV	DNA-A	2782
78BF14	[BF:Sos:Pe78BA:14]	PepYVMLV	DNA-A	2787
80BF14	[BF:Sos:Pe80BB:14]	PepYVMLV	DNA-A	2780
82BF14	[BF:Sos:To82PA:14]	ToLCGHV	DNA-A	2802
106BF15	[BF:Gui:To106BA:15]	PepYVMLV	DNA-A	2786
112BF15	[BF:Gui:To112BB:15]	PepYVMLV	DNA-A	2786
119BF15	[BF:Gui:To119BD:15]	PepYVMLV	DNA-A	2778
121BF15	[BF:Gui:To121BC:15]	PepYVMLV	DNA-A	2786
122BF15	[BF:Gui:To122BC:15]	PepYVMLV	DNA-A	2787
123BF15	[BF:Gui:To123BA:15]	PepYVMLV	DNA-A	2786
124BF15	[BF:Gui:To124[BF:15]	PepYVMLV	DNA-A	2787
125BF15	[BF:Gui:To125[BF:15]	PepYVMLV	DNA-A	2786
127BF15	[BF:Gui:To127BD:15]	PepYVMLV	DNA-A	2778
128BF15	[BF:Gui:To128BA:15]	PepYVMLV	DNA-A	2786

Supplementary table 3 continued

Sample ID	Isolate	Virus species	Component	Size (nt)
129BF15	[BF:Gui:To129BC:15]	PepYVMLV	DNA-A	2787
130BF15	[BF:Gui:To130BA:15]	PepYVMLV	DNA-A	2785
131BF15	[BF:Gui:To131BD:15]	PepYVMLV	DNA-A	2787
132BF15	[BF:Gui:To132BE:15]	PepYVMLV	DNA-A	2783
135BF15	[BF:Gui:To135BB:15]	PepYVMLV	DNA-A	2787
239BF15	[BF:Lou:To239BE:15]	PepYVMLV	DNA-A	2779
164BF15	[BF:Tan:To164BG:15]	CpCDV	DNA-A	2562
656BF15	[BF:Dje:To656BA:15]	ToLCMLV	DNA-A	2799
665BF15	[BF:Dje:To665BA:15]	ToLCMLV	DNA-A	2802
668BF15	[BF:Dje:To668BE:15]	ToLCMLV	DNA-A	2802
671BF15	[BF:Dje:To671BA:15]	ToLCMLV	DNA-A	2803
672BF15	[BF:Dje:To672BE:15]	ToLCMLV	DNA-A	2803
675BF15	[BF:Dje:To675BA:15]	ToLCMLV	DNA-A	2803
679BF15	[BF:Dje:To679BB:15]	ToLCMLV	DNA-A	2804
680BF15	[BF:Dje:To680BB:15]	ToLCMLV	DNA-A	2803
683BF15	[BF:Dje:To683BD:15]	ToLCMLV	DNA-A	2801
684BF15	[BF:Dje:To684BC:15]	ToLCMLV	DNA-A	2802
685BF15	[BF:Dje:To685BA:15]	ToLCMLV	DNA-A	2803
1132BF15	[BF:Bnk:Si1132BE:15]	CLCuGeV	DNA-A	2732
1146BF15	[BF:Bnk:To1146BD:15]	PepYVMLV	DNA-A	2822
690BF15	[BF:Dje:Si690BE:15]	CLCuGeV	DNA-A	2752
1002BF15	[BF:Krs:Si1002BE:15]	CLCuGeV	DNA-A	2747
960BF15	[BF:Krs:To960BD:15]	ToLCMLV	DNA-A	2803
887BF15	[BF:Leg:To887BD:15]	PepYVMLV	DNA-A	2791
796BF15	[BF:Tou:To796BE:15]	PepYVMLV	DNA-A	2778
813BF15	[BF:Tou:To813BC:15]	PepYVMLV	DNA-A	2785
823BF15	[BF:Tou:To823BB:15]	PepYVMLV	DNA-A	2786
825BF15	[BF:Tou:To825BA:15]	PepYVMLV	DNA-A	2768
1135BF15	[BF:Bnk:Si1135BA:15]	CLCuGeV	DNA-A	2759
940BF15	[BF:Krs:To940BC:15]	ToLCMLV	DNA-A	2822
810BF15	[BF:Tou:To810BC:15]	PepYVMLV	DNA-A	2793
587BF15	[BF:Dem:To587BD:15]	CpCDV	DNA-A	2562
1124BF16	[BF:Bsr:Am1124B7g:16]	CLCuGeV	DNA-A	2759
965BF16	[BF:Kor:Si965BB:16]	CLCuGeV	DNA-A	2760
1088BF16	[BF:Tie:Pe1088BB:16]	PepYVMLV	DNA-A	2779
1090BF16	[BF:Tie:Pe1090BB:16]	PepYVMLV	DNA-A	2780
1108BF16	[BF:Tie:Si1108B5g:16]	PepYVMLV	DNA-A	2779
479BF16	[BF:Tou:To479E7g:16]	ToLCGHV	DNA-A	2793
529BF16	[BF:Tou:To529B14c:16]	ToLCMLV	DNA-A	2797
574BF16	[BF:Leg:To574BA:16]	ToLCMLV	DNA-A	2803
579BF16	[BF:Leg:To579E7c:16]	ToLCMLV	DNA-A	2804
590BF16	[BF:Leg:To590BA:16]	ToLCMLV	DNA-A	2803
43BF13	[BF:Lou:To43B1:13]	PepYVMLV	DNA-B	2700
44BF13	[BF:Lou:To44B1:13]	PepYVMLV	DNA-B	2700
72BF13	[BF:Sak:Pe72B2:13]	PepYVMLV	DNA-B	2662
74BF13	[BF:Sak:Pe74B2:13]	PepYVMLV	DNA-B	2662

Supplementary table 3 continued

Sample ID	Isolate	Virus species	Component	Size (nt)
75BF13	[BF:Sak:To75B1:13]	PepYVMLV	DNA-B	2662
76BF13	[BF:Sak:To76B2:13]	PepYVMLV	DNA-B	2661
77BF13	[BF:Sak:To77B5:13]	PepYVMLV	DNA-B	2662
80BF13	[BF:Sak:To80B15:13]	PepYVMLV	DNA-B	2662
82BF13	[BF:Sak:To82B5:13]	PepYVMLV	DNA-B	2662
84BF13	[BF:Sak:Am84B1:13]	PepYVMLV	DNA-B	2662
10BF14	[BF:Mog:Pe10BA:14]	PepYVMLV	DNA-B	2670
11BF14	[BF:Mog:Am11BA:14]	PepYVMLV	DNA-B	2661
35BF14	[BF:Tou:Pe35BA:14]	PepYVMLV	DNA-B	2668
38BF14	[BF:Tou:Pe38BA:14]	PepYVMLV	DNA-B	2669
42BF14	[BF:Tou:To42BC:14]	PepYVMLV	DNA-B	2699
59BF14	[BF:Sak:To59BB:14]	PepYVMLV	DNA-B	2661
63BF14	[BF:Sak:Pe63BA:14]	PepYVMLV	DNA-B	2662
66BF14	[BF:Sak:To66BB:14]	PepYVMLV	DNA-B	2669
68BF14	[BF:Sak:To68BA:14]	PepYVMLV	DNA-B	2662
70BF14	[BF:Sak:Si70BA:14]	PepYVMLV	DNA-B	2662
71BF14	[BF:Sak:Pe71BD:14]	PepYVMLV	DNA-B	2660
77BF14	[BF:Sos:Pe77BC:14]	PepYVMLV	DNA-B	2671
78BF14	[BF:Sos:Pe78BD:14]	PepYVMLV	DNA-B	2662
110BF15	[BF:Gui:To110BC:15]	PepYVMLV	DNA-B	2662
126BF15	[BF:Gui:To126BE:15]	PepYVMLV	DNA-B	2662
630BF15	[BF:Kon:To630BA:15]	PepYVMLV	DNA-B	2662
638BF15	[BF:Kon:To638BA:15]	PepYVMLV	DNA-B	2662
253BF15	[BF:Lou:To253BE:15]	PepYVMLV	DNA-B	2662
1139BF15	[BF:Bnk:Pe1139BA:15]	PepYVMLV	DNA-B	2668
756BF15	[BF:Kar:Bo756BC:15]	PepYVMLV	DNA-B	2663
1045BF15	[BF:Sak:To1045BB:15]	PepYVMLV	DNA-B	2670
802BF15	[BF:Tou:To802BA:15]	PepYVMLV	DNA-B	2657
806BF15	[BF:Tou:To806BA:15]	PepYVMLV	DNA-B	2657
57BF13	[BF:Lou:Ok57B1:13]	OLCuA	Alphasatellite	1387
97BF13	[BF:Kon:Ok97B1:13]	CLCuGeA	Alphasatellite	1355
100BF13	[BF:Kam:Ok100B1:13]	CLCuGeA	Alphasatellite	1349
86BF14	[BF:Sos:Ok86PG:14]	CLCuGeA	Alphasatellite	1383
87BF14	[BF:Sos:Ok87EF:14]	CLCuGeA	Alphasatellite	1346
100BF14	[BF:Lou:Ok100bBB:14]	OLCuA	Alphasatellite	1386
1119BF16	[BF:Gue:Ok1119B3a:16]	CLCuGeA	Alphasatellite	1353
1123BF16	[BF:Am1123B5b:16]	OLCuA	Alphasatellite	1386
87BF13	[BF:Kon:Ok87P1:13]	CLCuGeB	Betasatellite	1347
97BF13	[BF:Kon:Ok97P4:13]	CLCuGeB	Betasatellite	1348
28BF14	[BF:Kay:Si28BB:14]	CLCuBFB	Betasatellite	1360
29BF14	[BF:Kay:Si29BA:14]	CLCuBFB	Betasatellite	1356
30BF14	[BF:Kay:Si30BA:14]	CLCuBFB	Betasatellite	1356
82BF14	[BF:Sos:To82EE:14]	ToLCGHB	Betasatellite	1203
86BF14	[BF:Sos:Ok86PE:14]	OLCuMLB	Betasatellite	1349
87BF14	[BF:Sos:Ok87PA:14]	CLCuGeB	Betasatellite	1346
97BF14	[BF:Lou:Ok97PF:14]	CLCuGeB	Betasatellite	1344

References

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R. & Daszak, P. (2004) Emerging infectious diseases of plants: pathogen pollution , climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, **19**, 535–544.

Bigarré, L., Chazly, M., Salah, M., Ibrahim, M., Padidam, M., Nicole, M., Peterschmitt, M., Fauquet, C. & Thouvenel, J.C. (2001) Characterization of a new begomovirus from Egypt infecting hollyhock (Althea rosea). *European Journal of Plant Pathology*, **107**, 701–711.

Briddon, R. & Stanley, J. (2006) Subviral agents associated with plant single-stranded DNA viruses. *Virology*, **344**, 198–210.

Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X. & Fauquet, C.M. (2008) Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. *Archives of Virology*, **153**, 763–781.

Briddon, R.W., Bull, S.E., Amin, I., Mansoor, S., Bedford, I.D., Rishi, N., Siwatch, S.S., Zafar, Y., Abdel-Salam, A.M. & Markham, P.G. (2004) Diversity of DNA 1: A satellite-like molecule associated with monopartite begomovirus-DNA complexes. *Virology*, **324**, 462–474.

Briddon, R.W., Patil, B.L., Bagewadi, B., Nawaz-ul-Rehman, M.S. & Fauquet, C.M. (2010) Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evolutionary Biology*, **10**, 97.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C.F., Fiallo-Olivé, E., Briddon, R.W., Hernández-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S. & Varsani, A. (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of Virology*, **160**, 1593–1619.

De Bruyn, A. (2014) Diversité génétique , Evolution et Dissémination des bégomovirus africains dans les îles du sud-ouest de l'océan Indien. Thèse de doctorat, Université

de La Réunion, 240 p

De Bruyn, A., Villemot, J., Lefeuvre, P., Villar, E., Hoareau, M., Harimalala, M., Abdoul-Karime, A.L., Abdou-chakour, C., Reynaud, B., Harkins, G.W., Varsani, A., Martin, D.P. & Lett, J. (2012) East African cassava mosaic-like viruses from Africa to Indian ocean islands: molecular diversity, evolutionary history and geographical dissemination of a bipartite begomovirus. *BMC Evolutionary Biology*, **12**, 228.

Chen, L.-F., Rojas, M., Kon, T., Gamby, K., Xoconostle-ca, B. & Gilbertson, R.L. (2009) A severe symptom phenotype in tomato in Mali is caused by a reassortant between a novel recombinant begomovirus (Tomato yellow leaf curl Mali virus) and a betasatellite. *Molecular Plant Pathology*, **10**, 415–430.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M. & Zhou, X. (2008) Geminivirus strain demarcation and nomenclature. *Archives of Virology*, **153**, 783–821.

Gnankiné, O., Mouton, L., Henri, H., Terraz, G., Houndeté, T., Martin, T., Vavre, F. & Fleury, F. (2013) Distribution of Bemisia tabaci (Homoptera: Aleyrodidae) biotypes and their associated symbiotic bacteria on host plants in West Africa. *Insect Conservation and Diversity*, **6**, 411–421.

Idris, A.M. & Brown, J.K. (2000) Identification of a new, monopartite begomovirus associated with leaf curl disease of cotton in Gezira, Sudan. *Plant Disease*, **84**, 809.

Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B. & Nagata, T. (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage ?29 DNA polymerase. *Journal of Virological Methods*, **116**, 209–211.

Jeske, H. (2009) *Geminivirus*. *TT Viruses - The Still Elusive Human Pathogens* (ed. by E.-M. De Villiers) and H. Zur Hausen), pp. 185–226. Springer-Verlag Berlin

Heidelberg.

Jones, R. (2009) Plant virus emergence and evoution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, **141**, 113–130.

Kon, T. & Gilbertson, R.L. (2011) 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. *Archives of Virology*, **157**, 107–20.

Kon, T., Rojas, M.R., Abdourhamane, I.K. & Gilbertson, R.L. (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *Journal of General Virology*, **90**, 1001–1013.

Konaté, G., Barro, N., Fargette, D. & Swanson, M.M. (1995) Occurrence of whitefly-transmitted geminiviruses in crops in Burkina Faso, and their serological detection and differentiation. *Annals of Applied Biology*, **126**, 121–130.

Kraberger, S., Kumari, S.G., Hamed, A.A., Gronenborn, B., Thomas, J.E., Sharman, M., Harkins, G.W., Muhire, B.M., Martin, D.P. & Varsani, A. (2015) Molecular diversity of Chickpea chlorotic dwarf virus in Sudan: High rates of intra-species recombination - a driving force in the emergence of new strains. *Infection, Genetics and Evolution*, **29**, 203–215.

Leke, W.N., Brown, J.K., Ligthart, M.E., Sattar, N., Njualem, D.K. & Kvarnheden, A. (2012) Ageratum conyzoides: a host to a unique begomovirus disease complex in Cameroon. *Virus Research*, **163**, 229–37.

Leke, W.N., Khatabi, B., Fondong, V.N. & Brown, J.K. (2016a) Complete genome sequence of a new bipartite begomovirus infecting fluted pumpkin (Telfairia occidentalis) plants in Cameroon. *Archives of Virology*, **161**, 2347–2350.

Leke, W.N., Khatabi, B., Mignouna, D.B., Brown, J.K. & Fondong, V.N. (2016b) Complete genome sequence of a new bipartite begomovirus infecting cotton in the Republic of Benin in West Africa. *Archives of Virology*, **161**, 2329–2333.

Leke, W.N. & Kvarnheden, A. (2014) Mixed infection by two West African tomato-infecting begomoviruses and ageratum leaf curl Cameroon betasatellite in tomato in Cameroon. *Archives of Virology*, **159**, 3145–3148.

Leke, W.N., Kvarnheden, A., Ngane, E.B., Titanji, V.P.K. & Brown, J.K. (2011) Molecular characterization of a new begomovirus and divergent alphasatellite from tomato in Cameroon. *Archives of Virology*, **156**, 925–8.

Leke, W.N., Mignouna, D.B., Brown, J.K. & Kvarnheden, A. (2015) Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. *Agriculture & Food Security*, **4**, 1.

Leke, W.N., Sattar, M.N., Ngane, E.B., Ngeve, J.M., Kvarnheden, A. & Brown, J.K. (2013) Molecular characterization of begomoviruses and DNA satellites associated with okra leaf curl disease in Cameroon. *Virus Research*, **174**, 116–125.

Lett, J.M., Lefeuvre, P., Couston, L., Hoareau, M., Thierry, M., Reynaud, B., Martin, D.P. & Varsani, A. (2009) Complete genomic sequences of Tomato yellow leaf curl Mali virus isolates infecting tomato and pepper from the North Province of Cameroon. *Archives of Virology*, **154**, 535–540.

Mansoor, S., Khan, S.H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. a., Briddon, R., Stanley, J. & Markham, P.G. (1999) Identification of a Novel Circular Single-Stranded DNA Associated with Cotton Leaf Curl Disease in Pakistan. *Virology*, **259**, 190–199.

Mansoor, S., Zafar, Y. & Briddon, R.W. (2006) Geminivirus disease complexes: the threat is spreading. *Trends in Plant Science*, **11**, 209–212.

Martin, D.P. & Shepherd, D.N. (2009) The epidemiology , economic impact and control of maize streak disease. *Food Security*, **1**, 305–315.

Mubin, M., Briddon, R.W. & Mansoor, S. (2009) Complete nucleotide sequence of chili leaf curl virus and its associated satellites naturally infecting potato in Pakistan. *Archives of Virology*, **154**, 365–368.

Muhire, B., Darren, P.M., Brown, J.K., Navas-Castillo, J., Moriones, E., Zerbini, F.M., Rivera-Bustamante, R., Malathi, V.G., Briddon, R.W. & Varsani, A. (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of Virology*, **158**, 1411–1424

N'Guessan, K.P., Fargette, D., Fauquet, C. & Thouvenel, J. (1992) Aspects off the epidemiology of okra leaf curl virus in Côte d'lvoire. *Tropical Pest Management*, **38**, 122–126.

Navas-Castillo, J., Fiallo-Olivé, E. & Sanchez-Campos, S. (2011) Emerging Virus Diseases Transmitted by Whiteflies. *Annual Review of Phytopathology*, **49**, 219–248.

Osei, M., Akromah, R., Shih, S., Lee, L. & Green, S. (2008) First report and molecular characterization of three distinct Begomoviruses associated with tomato leaf curl disease in Ghana. *Plant Disease*, **92**, 1585–1585.

Ouattara, A., Tiendrébéogo, F., Lefeuvre, P., Claverie, S., Hoareau, M., Traoré, E.V., Barro, N., Traoré, O. & Lett, J.M. (2017a) Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso. *Archives of Virology*, **162**, 1427–1429.

Ouattara, A., Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Claverie, S., Traoré, E.V., Barro, N., Traoré, O., Varsani, A. & Lett, J.M. (2017b) New strains of chickpea chlorotic dwarf virus discovered on diseased papaya and tomato plants in Burkina Faso. *Archives of Virology*, **162**, 1791–1794.

Padidam, R., Beachy, R.N. & Fauquet, C.M. (1995) Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *Journal of General Virology*, **76**, 25–35.

Patil, B.L. & Fauquet, C.M. (2009) Cassava mosaic geminiviruses : actual knowledge and perspectives. *Plant Molecular Pathology*, **10**, 685–701.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2010) Approximately Maximum-Likelihood

Trees for Large Alignments. *PLoS ONE*, **5**.

R Core Team (2016) R: A Language and Environment for Statistical Computing.

Rojas, M.R., Hagen, C., Lucas, W.J. & Gilbertson, R.L. (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, **43**, 361–394.

Sattar, M.N., Koutou, M., Husseini, S., Leke, W.N., Brown, J.K. & Kvarnheden, A. (2015) First Identification of Begomoviruses Infecting Tomato With Leaf Curl Disease in Burkina Faso. *Plant Disease*, **99**, 732–732.

Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M. & Stanley, J. (2000) A unique virus complex causes Ageratum yellow vein disease. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6890–6895.

Saunders, K. & Stanley, J. (1999) A nanovirus-like DNA component associated with yellow vein disease of Ageratum conyzoides: evidence for interfamilial recombination between plant DNA viruses. *Virology*, **264**.

Scholthof, K.B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C. & Foster, G.D. (2011) Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology*, **12**, 938–954.

Séka, K., Ouattara, A., Assiri, K.P., Kra, K.D., Hoareau, M., Lefeuvre, P., Diallo, H.A. & Lett, J.M. (2017) First report of Pepper yellow vein Mali virus associated with pepper yellow vein disease in Cote d'Ivoire. *New Disease Reports*, **35**, 5197.

Séka, K., Ouattara, A., Assiri, K.P., Kra, K.D., Hoareau, M., Lefeuvre, P., Diallo, H.A. & Lett, J.M. (2016) First reports of Cotton leaf curl Gezira virus and Okra yellow crinkle virus associated with okra leaf curl disease in Côte d'Ivoire. *New Disease Reports*, **34**, 1.

Shepherd, D.N., Martin, D.P., Lefeuvre, P., Monjane, A.L., Owor, B.E., Rybicki, E.P.

& Varsani, A. (2008) A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. *Journal of Virological Methods*, **149**, 97–102.

Shih, S.L., Green, S.K., Tsai, W.S., Lee, L.M. & Levasseur, V. (2007) First report of a distinct begomovirus associated with okra yellow crinkle disease in Mali. *Plant Pathology*, **56**, 718.

Shih, S.L., Kumar, S., Tsai, W.S., Lee, L.M. & Green, S.K. (2009) Complete nucleotide sequences of okra isolates of Cotton leaf curl Gezira virus and their associated DNA- b from Niger. *Archives of virology*, **154**, 369–372.

Silva, J.C.F., Carvalho, T.F.M., Basso, M.F., Deguchi, M., Pereira, W.A., Sobrinho, R.R., Vidigal, P.M.P., Brustolini, O.J.B., Silva, F.F., Dal-Bianco, M., Fontes, R.L.F., Santos, A.A., Zerbini, F.M., Cerqueira, F.R. & Fontes, E.P.B. (2017) Geminivirus data warehouse: a database enriched with machine learning approaches. *BMC Bioinformatics*, **18**, 240.

Tamura, K., Stecher, G., Peterson, D., Filipski, a. & Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.

Tiendrébéogo, F.., Traoré, E.V.S., Barro, N., Konaté, G., Traoré, A. & Traoré, O. (2008) Characterization of Pepper yellow vein Mali virus in Capsicum sp. in Burkina Faso. *Plant Pathology Journal*, **7**, 155–161.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Traoré, V.S.E., Barro, N., Péréfarres, F., Reynaud, B., Traoré, A.S., Konaté, G., Lett, J.-M. & Traoré, O. (2011) Molecular and biological characterization of Pepper yellow vein Mali virus (PepYVMV) isolates associated with pepper yellow vein disease in Burkina Faso. *Archives of Virology*, **156**, 483–487.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Villemot, J., Konaté, G., Traoré, A.S., Barro, N., Traoré, V.S., Reynaud, B., Traoré, O. & Lett, J.-M. (2010a) Molecular diversity of cotton leaf curl Gezira virus isolates and their satellite DNAs associated with okra leaf curl disease in Burkina Faso. *Virology Journal*, **7**, 48.

Tiendrébéogo, F., Traoré, V.S.E., Lett, J., Barro, N., Konaté, G., Traoré, A.S. & Traoré, O. (2010b) Impact of okra leaf curl disease on morphology and yield of okra. *Crop Protection*, **29**, 1–5.

Wyant, P., Strohmeier, S., Fischer, A., Schäfer, B., Briddon, R.W., Krenz, B. & Jeske, H. (2015) Light-dependent segregation of begomoviruses in Asystasia gangetica leaves. *Virus Research*, **195**, 225–235.

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

Zhou, X. (2013) Advances in understanding begomovirus satellites. *Annual review of phytopathology*, **51**, 357–81.

Zhou, Y.-C., Noussourou, M., Kon, T., Rojas, M.R., Jiang, H., Chen, L.-F., Gamby, K., Foster, R. & Gilbertson, R.L. (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. *Archives of Virology*, **153**, 693–706.

Article 2

Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso

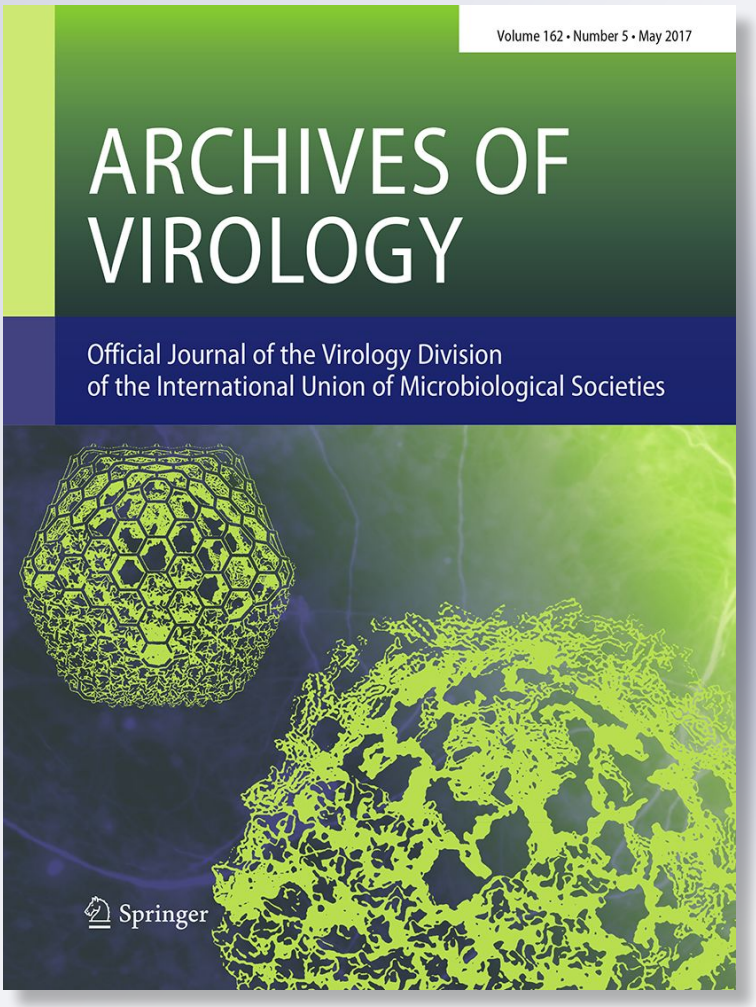
Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso

Alassane Ouattara, Fidèle Tiendrébéogo, Pierre Lefeuvre, Sohini Claverie, Murielle Hoareau, Edgard Valentin Traoré, et al.

Archives of Virology
Official Journal of the Virology Division of the International Union of Microbiological Societies

ISSN 0304-8608
Volume 162
Number 5

Arch Virol (2017) 162:1427-1429
DOI 10.1007/s00705-017-3231-6



Springer

Arch Virol
DOI 10.1007/s00705-017-3231-6

ANNOTATED SEQUENCE RECORD

Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso

Alassane Ouattara[1,2,3,4] · Fidèle Tiendrébéogo[3] · Pierre Lefeuvre[1] · Sohini Claverie[1,2] · Murielle Hoareau[1] · Edgard Valentin Traoré[3] · Nicolas Barro[4] · Oumar Traoré[3] · Jean-Michel Lett[1]

Received: 11 October 2016 / Accepted: 19 December 2016
© Springer-Verlag Wien 2017

Abstract In this report, we present the first description of the complete genome sequence of a new monopartite begomovirus isolated from tomatoes collected in Burkina Faso and presenting with symptoms of tomato leaf curl disease. We propose the tentative name "tomato leaf curl Burkina Faso virus'' (ToLCBFV). DNA-A-like nucleotide sequence of ToLCBFV shares the highest nucleotide sequence identity (85%) with the pepper yellow vein Mali virus (PepYVMLV). Phylogenetic analysis confirmed the affiliation of ToLCBFV to Old World monopartite begomoviruses. This discovery of a new species confirms the existence of high genetic diversity in monopartite begomoviruses in sub-Saharan Africa and particularly in West Africa.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-017-3231-6) contains supplementary material, which is available to authorized users.

✉ Jean-Michel Lett
lett@cirad.fr

[1] CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 Chemin de l'IRAT, 97410 Saint-Pierre, Ile de La Réunion, France

[2] Université de La Réunion, UMR PVBMT, Pôle de Protection des Plantes, 7 Chemin de l'IRAT, 97410 Saint-Pierre, Ile de La Réunion, France

[3] INERA, 01 BP 476, Ouagadougou 01, Ouagadougou, Burkina Faso

[4] Université Ouagadougou 1, Professeur Joseph KI-ZERBO, 03 BP 7021, Ouagadougou 03, Ouagadougou, Burkina Faso

Annotated sequence record

Begomoviruses (genus *Begomovirus,* family *Geminiviridae*) are a group of plant viruses that are transmitted by the whitefly *Bemisia tabaci* (family *Aleyrodidae*) and are responsible for serious diseases in a wide range of cultivated crops, including tomato (*Solanum lycopersicum*). These viruses have a circular single stranded DNA genome encapsidated in twinned icosahedral particles [9].

Since the late 1980s, begomoviruses have been responsible for many emerging vegetable and food crop diseases worldwide [6]. Except for a small number of bipartite begomoviruses described in the Old World, African begomoviruses are monopartite with a unique DNA-A-like component [3]. Today, a complex of more than ten monopartite begomoviruses has been described in tomato in Africa (for reviews see [4, 7]).

In March 2013, leaf samples from three tomato plants showing severe symptoms of leaf curling, resembling tomato leaf curl disease (ToLCD; Figure 1), were collected at Loumbila (latitude: 12.490932; longitude: −1.398806; Figure 1), located 25 km from Ouagadougou, in Burkina Faso. Total DNA was extracted using the DNeasy Plant Mini Kit (Qiagen). Viral genomes were amplified by rolling-circle amplification using *Phi29* DNA polymerase. Amplified products were digested with *Bam*HI endonuclease. The monomeric full-length DNA molecules obtained (~3 kb) were purified and ligated into the pGEM-3Zf vector (Promega). The ligated products were then cloned into *Escherichia coli* (JM109, Promega). Selected clones were completely sequenced by primer walking (Macrogen) and contigs were assembled with DNA Baser v.2.91 (Heracle BioSoft). Nucleotide sequences were subjected to a BLAST search for preliminary species assignment. Multiple sequences were aligned using

Published online: 04 February 2017

Fig. 1 Symptoms of severe leaf curling observed on the tomato plant from which the isolate of tomato leaf curl Burkina Faso virus was collected

the MUSCLE alignment method and maximum-likelihood (ML) phylogenetic trees were constructed by using FastTree v2.1.7. Pairwise identity comparisons of nucleotide sequences were performed using SDT v1.2 with pairwise deletion of gaps [5].

Three DNA-A-like sequences of 2,784 nucleotides in length and 100% identical to each other, were obtained from the same tomato sample. No DNA-B component or betasatellite DNA was detected by PCR with universal primers ([2, 8]). The DNA-A-like component presented features typical of Old World begomoviruses, with six open reading frames (ORFs): V1 (258 aa; capsid protein [CP]), V2 (116 aa; movement protein [MP]) in the virion sense, and C1 (359 aa; replication-associated protein [REP]), C2 (135 aa; transcriptional activator protein [TrAP]), C3 (134 aa; replication enhancer protein [REn]) and C4 (96 aa; C4 protein) in the complementary sense (Supplementary table 1). The DNA-A-like component presented a stem-loop structure in the origin of replication with the canonical nonanucleotide motif (5'-TAATATTAC-3'), TATA and GC boxes and two GGGGT iteron sequences (Supplementary figure 1). The corresponding iteron-related domain (IRD; Rep N-terminal domain interaction with *ori*-associated iterons) was identified as MAPPKR**FRVN** [1].

The three DNA-A-like sequences shared the highest nucleotide sequence identity (85%) with a pepper yellow vein Mali virus isolated from *Capsicum frutescens* in Burkina Faso (PepYVMV-[BF:Ban:Hpe:09], FN555174; supplementary tables 2 and 3; [10]). In accordance with the recently updated begomovirus species demarcation criteria (91% nucleotide identity of DNA-A and DNA-A-like without gaps; [3]), this isolate should be assigned to a new species of tomato-infecting monopartite begomovirus. ML tree analysis confirmed that the complete nucleotide sequence of ToLCBFV forms a distinct branch together with PepYVMLV among current African begomoviruses (Figure 2). In the absence of symptoms that differ from the other African tomato-infecting begomoviruses, we propose the name ''tomato leaf curl Burkina Faso virus'' (ToLCBFV).

Taken together, our results demonstrate that ToLCBFV is representative of a new species, genetically distinct from the currently described West African tomato-infecting monopartite begomoviruses, such as ToLCGV [4]. The infectivity and biological features of this new species should be investigated in the future.

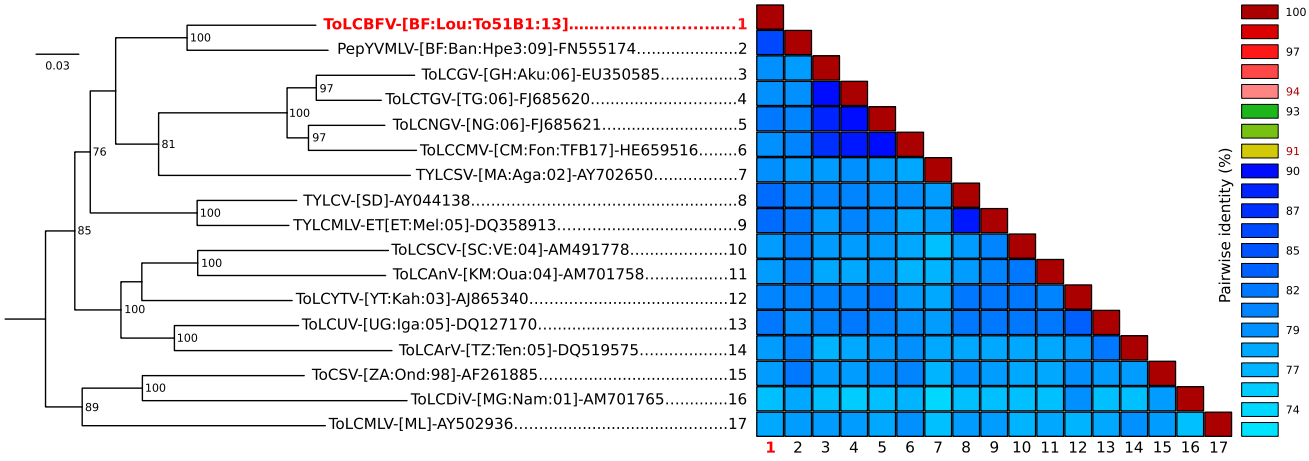


Fig. 2 Maximum-likelihood phylogenetic tree and colour-coded matrix of pairwise nucleotide identity inferred from alignments of complete DNA-A-like sequences of ToLCBFV and selected Old World monopartite begomoviruses originating from Africa. Only bootstrap values higher than 70% are shown (1000 replicates). The matrix uses a discontinuous range of three shades of colour (red, green and blue) differentiating two cut-off values representing the strain (93-94%, green-red) and the species (90-91%, blue-green) demarcation thresholds of begomoviruses. For begomovirus acronyms, see supplementary table 3

GenBank accession number

KX853168

Compliance with ethical standards

Funding This study was funded by the EMEB project (PEERS-AIRD, 2013-2014), the European Union (FEDER), the *Région Réunion* and CIRAD. AO is a recipient of a PhD fellowship from CIRAD (2015-2017).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Arguello-Astorga GR, Ruiz-Medrano R (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 a comparative approach. Arch Virol 146:1465–1485
2. Briddon RW, Bull SE, Mansoor S, Amin I, Markham PG (2002) Universal primers for the PCR-mediated amplification of DNA beta—a molecule associated with some monopartite begomoviruses. Mol Biotechnol 20:315–318
3. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, Fiallo-Olive E, Briddon RW, Hernandez-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
4. Leke WN, Mignouna DB, Brown JK, Kvarnheden A (2015) Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. Agric Food Secur 4:1–4
5. Muhire BM, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. Plos One 9:1–8
6. Navas-Castillo J, Fiallo-Olivé E, Sánchez-Campos S (2011) Emerging virus diseases transmitted by whiteflies. Ann Rev Phytopathol 49:219–248
7. Rey MEC, Ndunguru J, Berrie LC, Paximadis M, Berry S, Cossa N, Nuaila VN, Mabasa KG, Abraham N, Rybicki EP, Martin D, Pietersen G, Esterhuizen LL (2012) Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in Southern Africa, including the South-West Indian Ocean Islands. Viruses 4:1753–1791
8. Rojas MR, Gilbertson RL, Russel DR, Maxwell DP (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminivirus. Plant Dis 77:340–347
9. 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
10. Tiendrebeogo F, Lefeuvre P, Hoareau M, Traore VS, Barro N, Perefarres F, Reynaud B, Traore AS, Konate G, Lett JM, Traore O (2011) Molecular and biological characterization of pepper yellow vein Mali virus (PepYVMV) isolates associated with pepper yellow vein disease in Burkina Faso. Arch Virol 156:483–487

Article 3

New strains of chickpea chlorotic dwarf virus discovered on diseased papaya and tomato plants in Burkina Faso

Arch Virol
DOI 10.1007/s00705-017-3262-z

ANNOTATED SEQUENCE RECORD

New strains of chickpea chlorotic dwarf virus discovered on diseased papaya and tomato plants in Burkina Faso

Alassane Ouattara[1,2,3,4] · Fidèle Tiendrébéogo[1,2,4] · Pierre Lefeuvre[2] · Murielle Hoareau[2] · Sohini Claverie[2] · Edgar Valentin Traoré[1,4] · Nicolas Barro[3] · Oumar Traoré[1,4] · Arvind Varsani[5] · Jean-Michel Lett[2]

Received: 22 November 2016 / Accepted: 11 January 2017
© Springer-Verlag Wien 2017

Abstract This is the first description of full genome sequences of chickpea chlorotic dwarf virus (CpCDV; genus *Mastrevirus*; family *Geminiviridae*) identified in papaya and tomato plants sampled in Burkina Faso. The CpCDV full genome sequences from papaya and tomato share the highest pairwise sequence identity (84% and 93.5%) with Sudanese isolates of the CpCDV-K and CpCDV-M strains, respectively. Based on the strain demarcation threshold (>94% identity) for mastreviruses, we propose two new strains, CpCDV-Q and CpCDV-R, identified in papaya and tomato, respectively. Phylogenetic analysis confirmed that the sequences belong to a distinct clade of the highly diverse population of CpCDVs. Evidence of inter-strain recombination provided more support for the important role of recombination in CpCDV evolution. The discovery of CpCDV on papaya, a previously unsuspected host, raises many questions about the natural and potential host range of this dicot-infecting mastrevirus species that is reported to be emerging worldwide.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-017-3262-z) contains supplementary material, which is available to authorized users.

✉ Fidèle Tiendrébéogo
fidelet@gmail.com

✉ Jean-Michel Lett
lett@cirad.fr

1 INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

2 CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 Chemin de l'IRAT, 97410 Saint-Pierre, Ile de La Réunion, France

3 Université Ouaga I Professeur Joseph Ki-Zerbo, 03 BP 7021, Ouagadougou 03, Burkina Faso

4 Laboratoire Mixte International Patho-Bios, IRD-INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

5 The Biodesign Center for Fundamental and Applied Microbiomics, School of Life Sciences, Center for Evolution and Medicine, Arizona State University, 1001 S. McAllister Ave, Tempe, AZ 85287-5001, USA

Annotated sequence record

The genus *Mastrevirus* belongs to the family *Geminiviridae* and contains species with circular single-stranded DNA (ssDNA) genomes of approximately 2.7 kb. Mastreviruses infect a wide range of monocotyledonous and dicotyledonous hosts and are transmitted by leafhopper vectors. Seven known dicot-infecting mastreviruses have been described to date, five of which have only been documented in Australia: chickpea chlorosis virus (CpCV; [2]; [3]; [11]), chickpea chlorosis Australia virus (CpCAV; [2]), chickpea redleaf virus (CpRLV; [11]), chickpea yellows virus (CpYV; [2]) and tobacco yellow dwarf virus (TYDV; [2]). The two dicot-infecting mastreviruses found outside Australia are chickpea chlorotic dwarf virus (CpCDV), for which 16 strains are currently described ([3]; [4]; [6]) and chickpea yellow dwarf virus (CpYDV) ([5]). Whereas CpCDV has been found in the Middle East (Iran, Syria, Oman and Yemen, as well as Turkey), Africa (Egypt, Morocco, South Africa, Sudan, Tunisia) and the Indian subcontinent (India and Pakistan), so far, CpYDV has only been found in Pakistan. With the exception of TYDV, all these dicot-infecting mastrevirus species have been found to mainly infect pulses, although little is known about their potential host range. CpCDV natural hosts identified in the field include chickpea (*Cicer arietinum*), cotton (*Gossypium sp.*), lentil (*Lens culinaris*), faba bean

Published online: 22 February 2017

(*Vicia faba*), field pea (*Pisum sativum*), French bean (*Phaseolus vulgaris*), sugar beet (*Beta vulgaris*), pepper (*Capsicum annum*), *Sesbania bispinosa*, squash (*Cucurbita* sp; [1]), tomato (*Solanum lycopersicum;* [13]) and watermelon (*Citrullus lanatus*; [12]) (see [4] for specific references).

In August 2008, and from March to April 2015, leaf samples were collected from papaya (*Carica papaya*) and tomato plants showing severe symptoms of leaf dwarfing, curling and yellowing in central and eastern Burkina Faso (Table 1). Total DNA was extracted using the DNeasy Plant Mini kit (Qiagen, USA) and circular viral genomes were amplified by rolling circle amplification using *Phi29* DNA polymerase. Amplified products were digested with *BamH*I restriction endonuclease. The monomeric full-length genomes obtained (~2.6Kb) were purified and ligated into pGEM-3Zf(+) plasmid (Promega, France) restricted with *Bam*HI. The ligated products were then cloned into *Escherichia coli* (JM 109, Promega, France) and sequenced by primer walking (Macrogen Inc., Korea). Contigs were assembled with DNA Baser v.2.91 (Heracle BioSoft, Romania). Nucleotide sequences were subjected to a BLASTn search for preliminary identification. Multiple sequence alignment was performed using MUSCLE and a maximum likelihood (ML) phylogenetic tree was constructed using MEGA v.6.06 [10]. Pairwise sequence similarities between complete genome sequences were performed using SDT v1.2 with pairwise deletion of gaps [9].

Four genome sequences 2562 to 2569 nucleotides in length were obtained from the same papaya sample (Pap16A10 and Pap16A31) and from two tomato samples (To164 and To587) (Table 1). In the four genomes, we identified the four open reading frames (ORFs) that are typical of mastreviruses. These were V1 and V2 ORFs, in the viral sense, encoding the putative movement protein (MP) and capsid protein (CP), respectively and C1 and C2 ORFs in the complementary sense, encoding the putative replication-associated protein (Rep), expressed from ORFs C1 and C2 by transcript splicing, and the RepA protein from the unspliced transcript. However, we observed an unconventional nonanucleotide sequence TAAT**G**TTAC as part of a stem-loop structure in the origin of replication in the two sequences isolated from papaya (Supp. Figure 1). Interestingly, this unconventional nonanucleotide sequence has also been described in genome sequences of a Tunisian watermelon isolate of CpCDV [12] and other dicot-infecting mastreviruses characterized in Australia: CpRLV, CpCV and CpCAV ([2]; [3]; [11]; Fig. 1).

Pap16A10 and Pap16A31 genome sequences from papaya (99.7% similarity with one another) share ~84% genome-wide pairwise identity with the Sudanese CpCDV-K strain sequences (CpCDV-K [SD:SD272:13]; KM229905) (Supp. Table 2). To164 and To587 genome sequences from tomato (98.5% similarity with one another) share ~93% genome-wide pairwise identity with Sudanese CpCDV-M strain sequences (CpCDV-M [SD:SD231:13]; KM229909). Following the guidelines for mastrevirus classification [8] (i.e. genome sequences with 78-94% pairwise identity are unique strains), we assigned these isolates to two new strains we tentatively name CpCDV-Q for papaya isolates ([BF:Kou:Pap16A10:08] and [BF:Kou:Pap16A31:08]) and CpCDV-R for tomato isolates ([BF:Tan:To164:15] and [BF:Dem:To587:15]). Maximum likelihood phylogenetic analysis demonstrated that along with the CpCDV-M strain, these new CpCDV-Q and -R strains form a distinct clade within the current diversity of CpCDVs. Recombination analyses using RDP4 [7] revealed inter-strain recombination events, thereby confirming that recombination is a major feature of CpCDV evolution [4].

In conclusion, here we report the genomes of CpCDV from papaya and tomato in Burkina Faso. These represent isolates of two new strains CpCDV-Q and CpCDV-R, which form a distinct clade among the highly diverse population of CpCDV isolates. This is the first report of dicot-infecting mastreviruses in West Africa. The discovery of CpCDV on papaya, a previously unsuspected host, raises many questions concerning the natural and potential host range of this dicot-infecting mastrevirus species, which is reported to be emerging worldwide.

Table 1 Geographical origin of leaf samples collected in Burkina Faso and description of chickpea chlorotic draft virus (CpCDV) isolates

Plant	Locality	Region	Date of sampling (dd/mm/yyyy)	GPS coordinates		Mastrevirus isolates	GenBank accession N°
				Latitude	Longitude		
Tomato	Dem	Eastern	17-03-2015	13.177954	-1.159413	CpCDV-R [BF:Dem:To587:15]	KY047532
Tomato	Tandjari	Central	10-04-2015	12.190186	0.327786	CpCDV-R [BF:Tan:To164:15]	KY047533
Papaya	Koubri	Central	08-08-2008	12.193744	-1.395426	CpCDV-Q [BF:Kou:Pap16A10:08]	KY047534
						CpCDV-Q [BF:Kou:Pap16A31:08]	KY047535

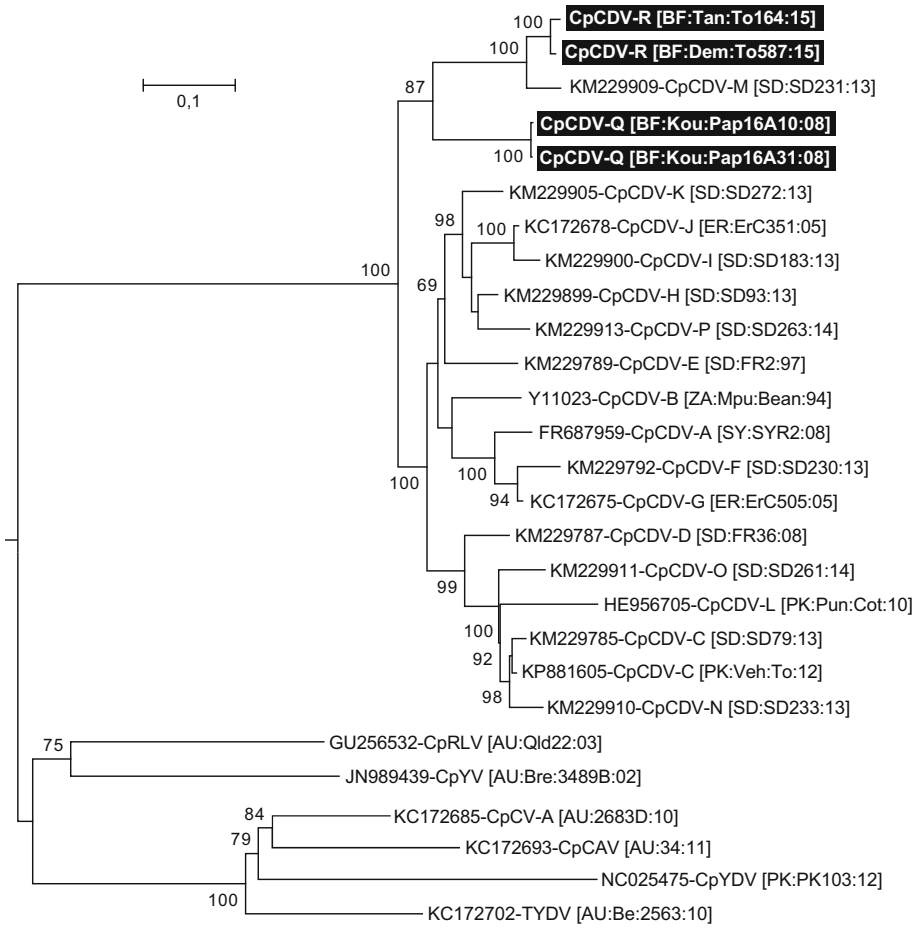


Fig. 1 Maximum-likelihood phylogenetic tree of full genome sequences of two new strains of chickpea chlorotic dwarf virus from tomato and papaya in Burkina Faso (CpCDV-Q and -R, respectively) with representative sequences of 16 strains of CpCDV (CpCDV-A to -P), as well as the six other species of previously described dicot-infecting mastreviruses. Only bootstrap values at nodes higher than 70% are displayed (1,000 replicates). The scale bar shows the genetic distance. The tree was rooted with a sequence of the African monocot-infecting maize streak virus (MSV-A, KJ4376664). For mastrevirus acronyms, see Supplementary Table 1

GenBank accession numbers

KY047532-KY047535

Compliance with ethical standards

Funding This study was funded by the European Union (FEDER), the *Région Réunion* and CIRAD (*Action incitative renforcement des compétences des chercheurs du Sud*). AO is a recipient of a PhD fellowship from CIRAD (2015-2017).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval No studies involving human participants or animals performed by any of the authors are described in this article.

References

1. Fahmy IF, Taha O, El-Ashry AN (2015) First genome analysis and molecular characterization of Chickpea chlorotic dwarf virus Egyptian isolate infecting squash. Virus Dis 26:33
2. Hadfield J, Thomas JE, Schwinghamer MW, Kraberger S, Stainton D, Dayaram A, Parry JN, Pande D, Martin DP, Varsani A (2012) Molecular characterisation of dicot-infecting mastreviruses from Australia. Virus Res 166:13–22
3. Kraberger S, Harkins GW, Kumari SG, Thomas JE, Schwinghamer MW, Sharman M, Collings DA, Briddon RW, Martin DP, Varsani A (2013) Evidence that dicot-infecting mastreviruses are particularly prone to inter-species recombination and have likely been circulating in Australia for longer than in Africa and the Middle East. Virology 444:282–291
4. Kraberger S, Kumari SG, Hamed AA, Gronenborn B, Thomas JE, Sharman M, Harkins GW, Muhire BM, Martin DP, Varsani A (2015) Molecular diversity of Chickpea chlorotic dwarf virus in Sudan: High rates of intra-species recombination - a driving force in the emergence of new strains. Infect Genet Evolut 29:203–215
5. Kraberger S, Mumtaz H, Claverie S, Martin DP, Briddon RW, Varsani A (2015) Identification of an Australian-like dicot-infecting mastrevirus in Pakistan. Arch Virol 160:825–830
6. Manzoor MT, Ilyas M, Shafiq M, Haider MS, Shahid AA, Briddon RW (2014) A distinct strain of chickpea chlorotic dwarf virus (genus *Mastrevirus*, family Geminiviridae) identified in cotton plants affected by leaf curl disease. Arch Virol 159:1217–1221
7. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolut 1:vev003
8. Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family Geminiviridae). Arch Virol 158:1411–1424
9. Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9:e108277

10. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evolut 30:2725–2729
11. Thomas JE, Parry JN, Schwinghamer MW, Dann EK (2010) Two novel mastreviruses from chickpea (*Cicer arietinum*) in Australia. Arch Virol 155:1777–1788
12. Zaagueri T, Mnari-Hattab M, Zammouri S, Hajlaoui MR, Acotto GP, Vaira AM (2016) First Report of Chickpea chlorotic dwarf virus in watermelon (*Citrullus lanatus*) in Tunisia. Plant Dis 101:392
13. Zia-Ur-Rehman M, Hameed U, Herrmann HW, Iqbal MJ, Haider MS, Brown JK (2015) First report of Chickpea chlorotic dwarf virus infecting tomato crops in Pakistan. Plant Dis 99:1287

Article 4

First report of Pepper yellow vein Mali virus associated with pepper yellow vein disease in Cote d'Ivoire

New Disease Reports

First report of *Pepper yellow vein Mali virus* associated with pepper yellow vein disease in Cote d'Ivoire

K. Séka [1,2], A. Ouattara [1,3], K.P. Assiri [2], K.D. Kra [2], M. Hoareau [1], P. Lefeuvre [1], H. Atta Diallo [2] and J.M. Lett [1]*

[1] *CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 97410 Saint-Pierre, Ile de La Réunion, France*; [2] *Université Nangui Abrogoua, 02 BP 801 Abidjan, Côte d'Ivoire*; [3] *INERA, LMI Patho-Bios, 01 BP 476, Ouagadougou 01, Burkina Faso*

*E-mail: lett@cirad.fr

Received: *27 Jan 2017.* ***Published:*** *27 Feb 2017.*

In Mali (Zhou *et al*., 2008) and Burkina Faso (Tiendrébéogo *et al*., 2008), pepper yellow vein disease (PYVD) is associated with the African monopartite begomovirus *Pepper yellow vein Mali virus* (PepYVMLV).

In January 2012 and August 2013, severe yellowing, leaf curling and deformation symptoms (Fig. 1), resembling those of PYVD were observed on pepper (*Capsicum* spp.) across Côte d'Ivoire (Table 1). Twenty leaf samples from pepper plants with symptoms were collected and tested for the presence of begomoviruses using a PCR assay with a set of degenerate primers designed to amplify the coat protein gene of Old World begomoviruses (Seka *et al*., 2016). PCR products of the expected size were obtained from seven samples, suggesting the presence of Old World monopartite begomoviruses (Table 1).

PCR-positive samples were further processed and viral genome sequences were amplified from three samples (Table 1) by rolling-circle amplification, cloned using the restriction enzyme *Bam*HI and sequenced. The three genome sequences were 2,786 nucleotides in length (GenBank Accession Nos. KY271075-KY271077) and showed the highest pairwise sequence identity, 98 to 99%, with PepYVMLV isolates from Burkina Faso (PepYVMLV-[BF:Oua:sweet pepper1:08]; FM876849) and Mali (PepYVMLV-[ML:03]; AY502935). A maximum-likelihood phylogenetic tree constructed with publicly available begomovirus genome sequences (MEGA6; Tamura *et al*., 2013), confirmed the genetic relationship of the Côte d'Ivoire isolates of PepYVMLV with isolates from neighbouring Burkina Faso and Mali (Fig. 2).

To our knowledge, this is the first report of PepYVMLV associated with PYVD in Côte d'Ivoire. Our results confirm the geographical distribution of PepYVMLV in West Africa and the close genetic relationship of this viral population (Tiendrébéogo *et al*., 2008).

Acknowledgements

This study was funded by the *Région Réunion*, the European Union (FEDER) and CIRAD. K. Séka is a recipient of a fellowship from CIRAD (*Action incitative renforcement des compétences des chercheurs du Sud*).

References

Séka K, Ouattara A, Assiri KP, Kra KD, Hoareau M, Lefeuvre P, Diallo HA, Lett J-M, 2016. First report of *Cotton leaf curl Gezira virus* and *Okra yellow crinkle virus* associated with okra leaf curl disease in Côte d'Ivoire. *New Disease Reports* **34**, 8. http://dx.doi.org/10.5197/j.2044-0588.2016.034.008

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

Tiendrébéogo F, Traoré VSE, Barro N, Traoré AS, Konaté G, Traoré O, 2008. Characterization of *Pepper yellow vein Mali virus* in *Capsicum* sp. in Burkina Faso. *Plant Pathology Journal* **7**, 155-161. http://dx.doi.org/10.3923/ppj.2008.155.161

Zhou YC, Noussourou M, Kon T, Rojas MR, Jiang H, Chen LF, Gamby K, Foster R, Gilbertson RL, 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. *Archives of Virology* **153**, 693-706 http://dx.doi.org/10.1007/s00705-008-0042-9

Figure 1

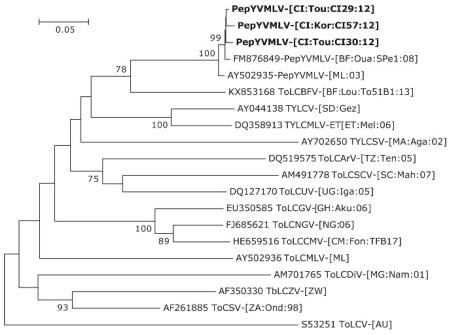


Figure 2

Table 1. Origin of pepper samples sampled in Côte d'Ivoire and results of PCR testing.

Sample	Locality[1]	Region	Date of sampling (dd/mm/yyyy)	GPS coordinates Latitude	GPS coordinates Longitude	PCR diagnosis[2]
CI7	Yamoussoukro	Centre	15/01/2012	06°51'21.30"N	05°23'50.80"W	+
CI29[a]	Toumodi	Centre	15/01/2012	06°51'07.30"N	05°00'20.20"W	+
CI30[b]	Toumodi	Centre	15/01/2012	06°51'07.30"N	05°00'20.20"W	+
CI31	Toumodi	Centre	15/01/2012	06°51'07.30"N	05°00'20.20"W	+
CI43	Abidjan	South	18/01/2012	05°15'30.20"N	03°56'08.70"W	+
CI49	Abidjan	South	18/01/2012	05°15'30.20"N	03°56'08.70"W	–
CI53	Abidjan	South	18/01/2012	05°15'30.20"N	03°56'08.70"W	–
CI56	Adzopé	South	24/01/2012	06°03'53.25"N	03°52'38.07"W	–
CI57[c]	Korhogo	North	28/01/2012	09°46'07.10"N	05°29'06.40"W	+
CI58	Bouaké	Centre	28/01/2012	07°32'35.90"N	05°00'27.60"W	–
CI59	Dabou	South	31/01/2012	05°20'15.80"N	04°27'40.40"W	+
CI62	Agboville	South	31/01/2012	05°40'02.60"N	04°07'29.00"W	–
CI63	Abidjan	South	31/01/2012	05°16'12.20"N	03°54'08.70"W	–
CI83	Yamoussoukro	Centre	24/08/2013	06°52'00.00"N	05°24'53.40"W	–
CI91	Toumodi	Centre	25/08/2013	06°24'13.70"N	05°05'27.20"W	–
CI92	Toumodi	Centre	25/08/2013	06°24'13.70"N	05°05'27.20"W	–
CI100	Toumodi	Centre	25/08/2013	06°24'13.70"N	05°05'27.20"W	–
CI108	Abidjan	South	26/08/2013	05°15'30.20"N	03°56'08.70"W	–
CI109	Abidjan	South	26/08/2013	05°15'30.20"N	03°56'08.70"W	–
CI115	Adzopé	South	28/08/2013	06°02'23.65"N	03°54'18.69"W	–

[1]Each sample corresponds to a distinct leaf from a different plant
[2]Degenerate primers designed to detect monopartite begomoviruses: Clust4CP-F342/Clust4CP-R1032
[a]PepYVMLV-[Cote d'Ivoire:Toumodi:CI29:2012]; GenBank Accession No. KY271077
[b]PepYVMLV-[Cote d'Ivoire:Toumodi:CI30:2012]; GenBank Accession No. KY271076
[c]PepYVMLV-[Cote d'Ivoire:Korhogo:CI57:2012]; GenBank Accession No. KY271075

Figure 3

To cite this report: Séka K, Ouattara A, Assiri KP, Kra KD, Hoareau M, Lefeuvre P, Atta Diallo H, Lett JM, 2017. First report of *Pepper yellow vein Mali virus* associated with pepper yellow vein disease in Cote d'Ivoire. *New Disease Reports* **35**, 11. http://dx.doi.org/10.5197/j.2044-0588.2017.035.011
©2017 The Authors
This report was published on-line at www.ndrs.org.uk where high quality versions of the figures can be found.

Article 5

First reports of Cotton leaf curl Gezira virus and Okra yellow crinkle virus associated with okra leaf curl disease in Côte d'Ivoire

New Disease Reports

First reports of *Cotton leaf curl Gezira virus* and *Okra yellow crinkle virus* associated with okra leaf curl disease in Côte d'Ivoire

K. Séka [1,2], A. Ouattara [3,4], K.P. Assiri [2], K.D. Kra [2], M. Hoareau [1], P. Lefeuvre [1], H. Atta Diallo [2] and J.M. Lett [1]*

[1] *CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 97410 Saint-Pierre, Ile de La Réunion, France*; [2] *Université Nangui Abrogoua, 02 BP 801 Abidjan, Côte d'Ivoire*; [3] *Université de Ouagadougou, 03 BP 7021, Ouagadougou 03, Burkina Faso*; [4] *INERA, LMI Patho-Bios, 01 BP 476, Ouagadougou 01, Burkina Faso*

*E-mail: lett@cirad.fr

Received: *02 Jun 2016.* ***Published:*** *23 Aug 2016.*

Okra leaf curl disease (OLCD) is commonly observed in okra (*Abelmoschus esculentus*) crops in several African countries (N'Guessan *et al*., 1992). Affected plants are severely stunted with apical leaf curl (upward or downward), distortion and thickening of the veins. In Africa, OLCD is associated with a complex of several strains of two begomovirus species: *Cotton leaf curl Gezira virus* (CLCuGV; Idris & Brown, 2002) and *Okra yellow crinkle virus* (OYCrV; Shih *et al*., 2007).

In 2012 and 2013, severe symptoms of leaf curling, deformation, and vein thickening (Fig. 1), resembling those of okra leaf curl disease were observed on okra in four localities in south-eastern Côte d'Ivoire (Table 1). Fourteen leaf samples with symptoms were collected and tested for the presence of begomoviruses using a polymerase chain reaction (PCR) assay with a set of degenerate primers designed to amplify the coat protein gene of Old World begomoviruses (Clust4CP-F342, 5′-TATMATCATTTCCACBCCVG-3′; Clust4CP-R1032, 5′-GCATGAGTACATGCCATATAC-3′). PCR products of the expected sizes were obtained for nine samples suggesting the presence of Old World monopartite begomoviruses in all four localities (Table 1).

PCR positive samples were further processed and full-length viral genomes were amplified from four samples (Table 1) by rolling-circle amplification, cloned using the *BamHI* restriction enzyme and sequenced. One complete genome sequence (GenBank Accession No. KX100570) showed the highest pairwise sequence identity of 99% (100% coverage) with isolates of the Niger strain of CLCuGeV from Niger (CLCuGeV-NE[NE:Sad:NG2FL:Ok:07], FJ469627) and Burkina Faso (CLCuGeV-NE[BF:Kap:Ok4:08], FN554524). The other three sequences (KX100571 to KX100573) showed the highest pairwise sequence identity of 98-99% (100% coverage) with isolates of the Mali strain of OYCrV from Mali (OYCrV-ML[ML:Mlo1:04], DQ902715; OYCrV-ML[ML:Bam4:06], EU024119). A maximum likelihood phylogenetic tree, produced from alignments of publicly available begomovirus sequences (MEGA6; Tamura *et al*., 2013), confirmed the genetic relationship of Ivorian CLCuGeV and OYCrV isolates with isolates from three neighbouring countries, Burkina Faso, Mali and Niger (Fig. 2).

To our knowledge, this is the first report of CLCuGeV and OYCrV implicated in OLCD in Côte d'Ivoire. Our results confirm that OLCD is associated with a complex of several strains of CLCuGeV and OYCrV in sub-Saharan Africa (Tiendrebeogo *et al*., 2010) and show that Ivorian isolates are genetically closely related to strains described in West Africa.

Acknowledgements

This study was funded by the *Région Réunion*, the European Union (FEDER) and the CIRAD.

References

Delatte H, Martin DP, Naze F, Goldbach R, Reynaud B, Peterschmitt M, Lett JM, 2005. South West Indian Ocean islands tomato begomovirus populations represent a new major monopartite begomovirus group. *Journal of General Virology* **86**, 1533-1542. http://dx.doi.org/10.1099/vir.0.80805-0

Idris AM, Brown JK, 2002. Molecular analysis of Cotton leaf curl virus-Sudan reveals an evolutionary history of recombination. *Virus Genes* **24**, 249-256. http://dx.doi.org/10.1023/A:1015380600089

N'Guessan KP, Fargette D, Fauquet C, Thouvenel JC, 1992. Aspects of the epidemiology of okra leaf curl virus in Côte d'Ivoire. Idris AM, Brown JK, 2002. *Tropical Pest Management* **38**, 122-126. http://dx.doi.org/10.1080/09670879209371668

Shih SL, Green SK, Tsai WS, Lee LM, Levasseur V, 2007. First report of a distinct begomovirus associated with okra yellow crinkle disease in Mali. *Plant Pathology* **56**, 718. http://dx.doi.org/10.1111/j.1365-3059.2007.01599.x

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

Tiendrébéogo F, Lefeuvre P, Hoareau M, Villemot J, Konaté G, Traoré AS, Barro N, Traoré VS, Reynaud B, Traoré 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. *Virology Journal* **7**, 48. http://dx.doi.org/10.1186/1743-422X-7-48

Figure 1

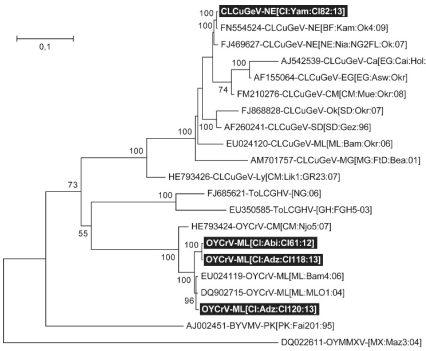


Figure 2

Table 1 Geographical origin of okra samples (each corresponding to a distinct leaf from a different plant) collected in Côte d'Ivoire and description of *Cotton leaf curl Gezira virus* (CLCuGV) and *Okra yellow crinkle virus* (OYCrV) isolates.

Sample	Locality[a]	Sampling date (dd/mm/yyyy)	Location (Latitude, Longitude)	PCR diagnosis (Clust4CP-F342 / R1032 primers)	Isolates	GenBank Accession No.
CI015	Toumodi	15/01/2012	06°5'107.30'N, 05°00'20.20'W	-	Not Available	-
CI032	Toumodi	15/01/2012	06°5'107.30'N, 05°00'20.20'W	+	Not Available	-
CI041	Toumodi	15/01/2012	06°5'107.30'N, 05°00'20.20'W	+	Not Available	-
CI042	Toumodi	15/01/2012	06°5'107.30'N, 05°00'20.20'W	+	Not Available	-
CI046	Abidjan	18/01/2012	05°15'30.20'N, 03°58'08.70'W	+	Not Available	-
CI061	Abidjan	18/01/2012	05°19'40.70'N, 04°16'41.30'W	+	OYCrV-Mali [Côte d'Ivoire: Abidjan CI61:2012]	KX100571
CI082	Yamoussoukro	24/08/2013	06°52'00.00'N, 05°24'53.40'W	+	CLCuGV-Niger [Côte d'Ivoire: Yamoussoukro CI82:2013]	KX100570
CI086	Yamoussoukro	24/08/2013	06°52'00.00'N, 05°24'53.40'W	-	Not Available	-
CI118	Adzopé	28/08/2013	06°02'23.65'N, 03°54'18.69'W	+	OYCrV-Mali [Côte d'Ivoire: Adzopé CI118:2013]	KX100572
CI119	Adzopé	28/08/2013	06°02'23.65'N, 03°54'18.69'W	-	Not Available	-
CI120	Adzopé	28/08/2013	06°02'23.65'N, 03°54'18.69'W	+	OYCrV-Mali [Côte d'Ivoire: Adzopé CI120:2013]	KX100573
CI128	Yamoussoukro	22/07/2013	06°52'36.80'N, 06°28'50.63'W	-	Not Available	-
CI129	Yamoussoukro	22/07/2013	06°52'36.80'N, 06°28'50.63'W	-	Not Available	-
CI132	Yamoussoukro	22/07/2013	06°52'36.80'N, 06°28'50.63'W	+	Not Available	-

Figure 3

To cite this report: Séka K, Ouattara A, Assiri KP, Kra KD, Hoareau M, Lefeuvre P, Atta Diallo H, Lett JM, 2016. First reports of *Cotton leaf curl Gezira virus* and *Okra yellow crinkle virus* associated with okra leaf curl disease in Côte d'Ivoire. *New Disease Reports* **34**, 8. http://dx.doi.org/10.5197/j.2044-0588.2016.034.008

©2016 The Authors *This report was published on-line at www.ndrs.org.uk where high quality versions of the figures can be found.*

CHAPITRE II

EVALUATION DES PRINCIPAUX PARAMETRES EPIDEMIOLOGIQUES ASSOCIES A L'EMERGENCE DU PepYVMLV AU BURKINA FASO

CHAPITRE II : EVALUATION DES PRINCIPAUX PARAMETRES EPIDEMIOLOGIQUES ASSOCIES A L'EMERGENCE DU PepYVMLV AU BURKINA FASO

Les maladies du jaunissement et/ou de l'enroulement foliaire de la tomate (*tomato [yellow] leaf curl disease*, TYLCD-ToLCD) figurent parmi les contraintes virales les plus importantes des cultures de tomate dans les régions tropicales et subtropicales du monde (Hanssen *et al.*, 2010; Navas-Castillo *et al.*, 2011). Ces maladies sont causées par une légion de bégomovirus (famille : *Geminiviridae*) transmis par l'aleurode *Bemisia tabaci*. Parmi ces virus figure un bégomovirus qui a été initialement caractérisé au Mali à partir d'échantillons foliaires de piment présentant des symptômes typiques de la maladie du jaunissement nervaire du piment (*pepper yellow vein disease*, PYVD ; Zhou *et al.*, 2008). Sur la base des symptômes de PYVD et de la plante hôte, l'espèce virale fut nommée *Pepper yellow vein Mali virus* (PepYVMLV, Chen *et al.*, 2009). Des études postérieures ont permis d'élargir l'aire de distribution de ce virus au Burkina Faso (Tiendrébéogo *et al.*, 2011) et plus récemment en Côte d'Ivoire (Séka *et al.*, 2017). Aucun ADN-B n'avait été détecté ou identifié en association avec l'ADN-A du PepYVMLV dans les échantillons foliaires de piment collectés au Mali, malgré l'usage d'amorces universelles ADN-B (Chen *et al.*, 2009). De plus, l'organisation génomique similaire à celle des bégomovirus monopartites de l'Ancien Monde, et le replacement phylogénétique au sein des bégomovirus monopartites africains ont conduit à définir le PepYVMLV comme un bégomovirus monopartite (Chen *et al.*, 2009; Tiendrébéogo *et al.*, 2011). Cependant, l'évaluation de la pathogénicité et de la gamme d'hôte du PepYVMLV par agroinoculation n'ont pas permis d'obtenir des symptômes de PYVD sur piment mais de sévères symptômes de froissement/déformation, nanisme et jaunisme foliaire sur tomate. L'absence des symptômes attendus de type PYVD sur piment en conditions contrôlées (agroinoculation de PepYVMLV seul) suggérait l'implication d'un ou de plusieurs composants additionnels dans la pathogénicité du PepYVMLV en plein champ.

Nos activités d'épidémio-surveillance au Burkina Faso ont montré que les maladies du ToLCD-TYLCD sont dues à la présence d'un complexe d'au moins cinq bégomovirus, dont le plus prévalent et le plus largement reparti est le PepYVMLV (Ouattara *et al.,* en préparation, Chapitre 1). De manière inattendue, nos efforts de clonage et de

séquençage ont permis la caractérisation d'un ADN-B en grande majorité en association avec le PepYVMLV. Même s'il possède une organisation génomique équivalente aux autres ADN-B décrits dans l'Ancien Monde, il est cependant assez distant génétiquement de ceux-ci (Ouattara *et al.,* en préparation, chapitre 1).

Au regard de la découverte d'un nouvel ADN-B en association avec l'ADN-A du PepYVMLV, et afin d'élucider l'énigme de leur association, ce chapitre est consacré à l'étude des principaux paramètres du pouvoir pathogène du PepYVMLV en association ou non avec le composant d'ADN-B. En conditions contrôlées, le pouvoir pathogène des deux molécules d'ADNs du PepYVMLV a été évalué en infection simple ou mixte. Des tests comparatifs de cinétiques de virulence (symptômes de maladie) ont été réalisés entre le PepYVMLV et la souche Israël du TYLCV (TYLCV-IL) considérée actuellement comme le bégomovirus le plus sévère sur les cultures de tomate dans le monde. Enfin, l'abondance virale (*fitness* relative) du PepYVMLV en infection simple ou mixte avec l'ADN-B et le taux de transmission individuel par l'aleurode *B. tabaci* ont été évalués.

L'ensemble de nos résultats ont permis de démontrer que même si le PepYVMLV n'est pas un bégomovirus bipartite strict, l'ADN-B représente à la fois un **activateur** fort de la **virulence**, de l'**accumulation** de l'ADN-A dans la plante et de la **transmission** par son insecte vecteur *B. tabaci*, qui semblent concourir à sa prédominance sur le terrain. Avec une sévérité comparable à celle du TYLCV-IL et des symptômes ultimes de ToLCD-TYLCD avec des arrêts de croissance pouvant aboutir à la mort de la plante en conditions contrôlées, on peut considérer que le PepYVMLV représente une menace importante pour la culture de la tomate au Burkina Faso, en Afrique et plus largement dans le monde.

Article 6

Stronger together: recruitment of a DNA-B component by an emerging monopartite begomovirus

Article 6

Stronger together: recruitment of a DNA-B component by an emerging monopartite begomovirus

Alassane Ouattara[1,2,3,4,5], Fidèle Tiendrébéogo[1,5], Pierre Lefeuvre[2], Nathalie Becker[6], Frédéric Chiroleu[2], Murielle Hoareau[2], Martial Grondin[2], Agathe Allibert[2], Edgar Valentin Traoré[1,5], Nicolas Barro[4], Oumar Traoré[1,5], Gaël Thébaud[7], Jean-Michel Lett[2]

[1] INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

[2] CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

[3] Université de La Réunion, UMR PVBMT, 97410, Saint-Pierre, Ile de La Réunion, France

[4] Université Ouaga I Professeur Joseph Ki-Zerbo, 03 BP 7021, Ouagadougou 03, Burkina Faso

[5] LMI Patho-Bios, IRD-INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

[6] MNHN, ISYEB, UMR 7205, 75005, Paris, France

[7] BGPI, INRA, SupAgro, Cirad, Univ. Montpellier, Montpellier, France

Abstract

While monopartite and segmented viruses infect all kinds of living organisms, multipartite or multicomponent viruses appear mostly restricted to plants. In recent decades, a legion of monopartite begomoviruses transmitted by the whitefly *Bemisia tabaci* has emerged as serious threats to vegetable crops in the tropical and subtropical regions of Africa. In Burkina Faso (West Africa), recent studies have reported the identification of at least five begomoviruses on tomato crops with the predominance of *Pepper yellow vein Mali virus* (PepYVMLV) which was surprisingly associated with a newly discovered DNA-B component. In order to improve our understanding of the role of this component in the emergence of PepYVMLV as the most prevalent and severe plant virus disease on tomato crops in Burkina Faso, biological traits related to virulence, virus accumulation in the plant and transmission rate of the virus were evaluated by *agrobacterium*- and whitefly-mediated inoculation. Even if the DNA-B component is not required for systemic movement and symptom development of PepYVMLV (non-strict association), their association induces more severe symptoms of leaf distortion and crumpling with growth arrest and death of 17% of inoculated plants. The higher virulence of PepYVMLV in infection with the DNA-B component was associated with a $2x10^5$-fold more viral DNA accumulation in plant tissues, a positive and linear relation between DNA-A and -B loads, and an 18% increase in the individual transmission rate by *B. tabaci*. Here, our results suggest that the recruitment of a DNA-B component by the monopartite PepYVMLV represents the key epidemiological factor of its emergence as the most prevalent and severe plant virus disease on tomato crops in Burkina Faso.

Introduction

Multipartite or multicomponent viruses have their genome divided into two or more nucleic acid segments which are each packaged into separate virus particles. While monopartite and segmented viruses (with both categories encapsidated in one single particle) infect all kinds of living organisms, multipartite viruses appear mostly restricted to plants (Sicard *et al.*, 2016). It has earlier been speculated that the multipartite architecture of the viral genome could be related to its RNA nature, with numerous examples of plant multipartite RNA viruses (Pressing & Reanney, 1984). However, it is now clear that DNA multipartite viruses are also frequent among plant viruses (Briddon & Stanley, 2006; Sicard *et al.*, 2016).

Geminivirus are plant viruses that have circular, single-stranded (ss) DNA genomes encapsidated in twinned icosahedral particles (Rojas *et al.*, 2005; Zerbini *et al.*, 2017). Based on host range, insect vector and genome organization, they are divided into nine genera, eight of which (*Becurtovirus, Capulavirus, Curtovirus, Eragrovirus, Grablovirus, Mastrevirus, Topocuvirus* and *Turncurtovirus*) consist of viruses with monopartite genomes only (Varsani *et al.*, 2014; Zerbini *et al.*, 2017). In contrast, the largest genus *Begomovirus* consists of viruses either with one or two genomic components (monopartite or bipartite). The component of begomoviruses known as DNA-A is homologous to the genomes of all geminiviruses and encodes proteins required for movement (MP/C4 for monopartite begomoviruses), replication, control of gene expression, overcoming host defenses, encapsidation and insect transmission. The second component, referred to as DNA-B, encodes two proteins (BC1 and BV1) with functions in intra- and intercellular movement in host plants, and are known to contribute to the increase of the viral load and the severity of the disease (Shafiq *et al.*, 2010; Fondong, 2013). The origin of the DNA-B component remains unclear. different segments can have distinct evolutionary histories and phylogeographies (Briddon & Stanley, 2006; Sicard *et al.*, 2016).

The proteins of DNA-A and -B are encoded by open reading frames (ORFs) on the virion and the complementary sense. The virion-sense ORFs and those in the complementary sense are separated by an intergenic region (IR), referred to as the

common region (CR), which contains sequences that are conserved between both DNA-A and -B components. The CR contains the origin of replication, which consists of a conserved hairpin structure with a nonanucleotide sequence in the loop and Rep binding iteron sequences located upstream of the hairpin (Hanley-Bowdoin *et al.*, 1999; Rojas *et al.*, 2005). The presence of iteron sequences in DNA-B similar to those in DNA-A ensures binding between the DNA A-encoded Rep and the cognate DNA-B.

Since the late 1980s, begomoviruses have emerged everywhere in the world and have become a major constraint in the production of vegetables and in particular tomato (*Solanum lycopersicum*; Navas-Castillo *et al.*, 2011). In Africa, a complex of at least twenty monopartite begomoviruses species is involved in tomato yellow leaf curl or tomato leaf curl diseases TYLCD-ToLCD including seven species described in West Africa (Rey *et al.*, 2012; Brown *et al.*, 2015). Among them, *Pepper yellow vein Mali virus* (PeYVMLV) has been identified as the most prevalent and severe tomato-infecting begomovirus on tomato and pepper in Burkina Faso (Ouattara *et al.*, unpublished, Chapter 1). Interestingly, whereas it was previously described as a West African monopartite begomovirus (Tiendrébéogo *et al.*, 2008; Zhou *et al.*, 2008), PepYVMLV was associated with a DNA-B component (Ouattara *et al.*, unpublished, Chapter 1). Although the two components of the majority of tomato-infecting bipartite begomoviruses have an obligate relationship (Hamilton *et al.*, 1983; Padidam *et al.*, 1995), this does not seem to be the case for all. Both *Tomato yellow leaf curl Thailand virus* (TYLCTHV) and *Tomato leaf curl Gujarat virus* (ToLCGV) DNA-A components are able to induce systemic and symptomatic infections on the experimental host *Nicotiana benthamiana* (Rochester *et al.*, 1994) and on tomato (Chakraborty *et al.*, 2003), respectively, in the absence of their cognate DNA-Bs, which suggests that these viruses likely represent evolutionary intermediates between monopartite and bipartite begomoviruses.

In order to understand the involvement of this DNA-B component in the emergence of PepYVMLV as the most prevalent and severe plant virus disease on tomato crops in Burkina Faso, biological traits related to virulence, virus accumulation in the plant and transmission rate of the virus by *B. tabaci* were evaluated by *agrobacterium*- and whitefly-mediated inoculation. We established that even if the relation between the two

components is not obligate, DNA-B increases the viral accumulation, the virulence and the transmission rate of PepYVMLV by *B. tabaci*. Finally, we demonstrate that the recruitment of a DNA-B component by the monopartite PepYVMLV represents a key epidemiological factor of its emergence.

Materials and Methods

Construction of infectious clones

Full-length DNA-A and DNA-B molecules of PepYVMLV [Burkina Faso:Sakabi:Tomato72:2013] ([BF:Sak:To72:13], EMBL: XXXX/XXXX) previously cloned into pGEM-3Zf (Promega,USA), were used for the construction of infectious clones in the binary vector pCambia0380 (Cambia, Australia). A 456-bp *Pst*I/*Bam*HI-digested fragment comprising the IR of the DNA-A was cloned to generate a 0.16-mer (pCambia0380-0.16). The full-length monomer was then cloned into *Bam*HI-digested pCambia0380-0.16 to generate a 1.16-mer of PepYVMLV DNA-A. For the DNA-B, a 2057-bp *Eco*RI/*Bam*HI-digested fragment comprising the IR was cloned to generate a 0.77-mer (pCambia0380-0.77). The full-length monomer was cloned into *Bam*HI-digested pCambia0380-0.77 to generate a 1.77-mer of PepYVMLV DNA-B. Recombinant plasmids were mobilized from *Escherichia coli* strain JM-109 cells into *Agrobacterium tumefaciens* (strain C58) by triparental mating using *E. coli* HMB101 harboring the helper plasmid pRK2013 (Ditta *et al.*, 1980). Along with PepYVMLV DNA-A and DNA-B, an agroinfectious clone of the Israel strain of tomato yellow leaf curl virus (TYLCV-IL, Péréfarres *et al.*, 2014) was used as a virulent control.

Agrobacterium-mediated inoculation experiments

Liquid culture of *Agrobacterium tumefaciens* harboring PepYVMLV DNA-A, PepYVMLV DNA-B or TYLCV-IL were grown for 48 hours and adjusted to an OD_{600nm} of 1.3 before inoculation. Tomato plants (*Farmer 209*, Known-You Seed) and *Nicotiana benthamiana* were mono-inoculated (PepYVMLV DNA-A, PepYVMLV DNA-B or TYLCV-IL DNA) or bi-inoculated (PepYVMLV DNA-A + PepYVMLV DNA-B and TYLCV-IL DNA + PepYVMLV DNA-B) at the three-leaf stage by injecting about 50 µL of *A. tumefaciens* culture. For mixed inoculations, *Agrobacterium* cultures were mixed in equal volumes. A total of 60 tomato and 40 *N. benthamiana*

plants were agroinoculated per condition. In addition, 20 plants of tomato and *N. benthamiana* were punctured with sterile needles and used as negative controls.

Inoculated plants were then arranged in a complete random block design and maintained 32 days in an insect-proof growth chamber at 25±4°C with 12 h photoperiod and 70±10% relative humidity. Symptoms were scored twice a week during 30 days post inoculation (dpi). The rating scale ranged from 1 (no symptoms) to 10 (dead plant), with grades 1-9 corresponding to the scale of Lapidot *et al* (2006). At 32 dpi, leaves were collected for the detection and the quantification of viral genomes using PCR and quantitative PCR, respectively, as described below. At the same time, the effect of PepYVMLV DNA-A and TYLCV-IL DNA in single or mixed infections with PepYVMLV DNA-B on the development of tomato plants was assessed by measuring the size of each plant.

Whitefly-mediated inoculation experiments

Non-viruliferous *B. tabaci* colonies of the cryptic species Middle East-Asia Minor 1 (MEAM1, formerly biotype B) were reared on cabbage plants (*Brassica oleracea*), known as a non-host plant for begomoviruses, in a growth chamber at 25 °C day and 20 °C night, 70% relative humidity and a 12 h photoperiod. Viruliferous whiteflies were obtained after a 72 h-acquisition access period (AAP) on tomato plants agroinoculated in single or mixed infections with PepYVMLV DNA-A and PepYVMLV DNA-B. After the AAP, adults were collected and deposited on healthy tomato seedlings (*Farmer 209*, Known-You Seed) at the one-leaf growth stage individually, and then placed under microcages for a 72h-inoculation access period (IAP). At the end of the IAP, insects were mechanically killed and the tomato seedlings were sprayed with insecticide (Confidor®, Bayer). In order to discard insects with an unknown time of IAP, only plants on which the insect had been found alive were maintained for the rest of the experiment (PepYVMLV DNA-A, n=80; PepYVMLV DNA-A + DNA-B, n=81). Negative control consisted in mock-inoculated plants (non viruliferous whiteflies, n=20). The plants were maintained in the same conditions as described above. After 30 days, symptoms were assessed and plants were tested for the presence of viral DNA by PCR.

Virus detection

Total DNA extraction was performed from 20 mg of dried plant material using the DNeasy Plant Miniprep Kit (Qiagen) according to the manufacturer's instructions with two successive 50 µL elutions with ultrapure water. Extracts were stored at -20 °C before use. Conventional PCR for TYLCV was carried out to detect viral DNA in samples collected at 32 dpi and 30 dpi in agro-inoculation and whitefly-mediated inoculation experiments, respectively, for TYLCV-IL DNA (Lefeuvre *et al.*, 2007) and PepYVMLV DNA-A and -B (Ouattara *et al.*, unpublished; Chapter 1).

Within-plant virus quantification

After the plant DNA extraction step, viral DNA was quantified using quantitative PCR, as in (Péréfarres *et al.*, 2011) and performed in triplicate, using *Solanum lycopersicum* 25S ribosomal RNA gene as internal control for each sample. Primers and TaqMan probes (Applied Biosystems, Foster City, USA; sequences and conditions of use in Supplementary table 1) were either designed using complete sequences of PepYVMLV DNA-A and -B characterized from Burkina Faso (Ouattara *et al.*, unpublished; Chapter 1), or designed previously (TYLCV and *Solanum lycopersicum* 25S ribosomal RNA gene; Becker et al., 2015; Péréfarres *et al.*, 2011). Standards were synthesized by Macrogen (Europe) (Supplementary table 1).

Quantitative PCR conditions

Two microliters of DNA samples, extracted as describe above, were tested in a 21 µl reaction mix comprising 1x TaqMan universal PCR master mix (Applied Biosystems, Foster City, USA), 0.64 µl of probe and 1.44 µl of primer. PCR reactions were carried out in the StepOnePlus real-time PCR system using 96-well optical reaction plates (Applied Biosystems, France). Cycling parameters were 50°C for 2 min and 95°C for 10 min followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Each sample was amplified in triplicate and a new aliquot of the standard was used in each run.

Statistical analysis

All statistical analyses were performed using the R statistical software (R Development Core Team, 2017). Nonlinear regression analyses were performed, testing different link

functions (Cauchy, cloglog, logistic, logit, loglog and probit), to fit the progression of disease severity with gnls function (Pinheiro *et al.*, 2016). Based on the likelihood and using Akaike's Information Criterion (AIC), the logit function appeared to be the most appropriate. In this model, written as Y ~ 1 + C / (1 + exp(- A * (X - B))), the disease severity (Y) is dependent on the dpi (X) and three biologically relevant parameters where A is the slope of the exponential phase, 1 + C is the disease severity at the plateau phase, and B is the time to reach 50% of the disease severity at the plateau phase. Once estimated, the parameters of the logistic disease progression model were compared between the different conditions by likelihood ratio tests in nested models. For these analyses, only plants for which single (DNA-A) or mixed (DNA-A and -B) infections were validated by PCR were used. Quantitative PCR data were expressed as the log of the ratio of the quantity of virus DNA to that of plant DNA. ANOVA *F*-test has been used to compare the amount of DNA-A component from single and mixed inoculation conditions. For the mixed infection case, linear regression was used to estimate the correlation between DNA-A and DNA-B components.

Results

Infectivity and virulence of agroinfectious clones

N. benthamiana and *S. lycopersicum* (tomato) plants agroinoculated with PepYVMLV DNA-A alone or associated with DNA-B developed differential symptoms of leaf crumpling with yellowing and stunting (Figure 1a and b). *N. benthamiana* plants agroinoculated with PepYVMLV DNA-A developed mild symptoms in 82.5% (33/40) of inoculated plants, with viral DNA-A PCR detection in 92.5% (37/40) of inoculated plants (Table 1, Figure 1a). In contrast, in mixed agroinfections with DNA-B very severe symptoms were observed on *N. benthamiana*, with plant growth arrest in 85% (34/40) of inoculated plants, with viral DNA-A and DNA-B PCR detection in 100% (40/40 and 40/40, respectively) of inoculated plants. When PepYVMLV DNA-A was agroinoculated to tomato plants, a low proportion (20%; Table 1, Figure 1b) developed very mild symptoms, although viral DNA-A was detected by conventional PCR in 80% (48/60) of inoculated plants. In contrast, in mixed agroinfection with DNA-B highly severe symptoms were observed on tomato, with plant growth arrest in 95% (57/60,

Table 1, Figure 1b) and plant death in 17% (10/60, Table 1) of inoculated plants, with viral DNA-A and DNA-B detected by conventional PCR analysis in 100% (60/60) and 87% (52/60) of inoculated plants, respectively. Tomato plants agroinoculated with TYLCV-IL confirmed the high infectivity of this monopartite begomovirus with 100% of symptomatic and PCR-positive plants in single or mixed infection with DNA-B (Table 1). Interestingly, DNA-B was only detected by PCR in 28% of mixed agroinfected tomato plants with TYLCV-IL. Mock plants and those agroinoculated with PepYVMLV DNA-B (alone) remained asymptomatic.

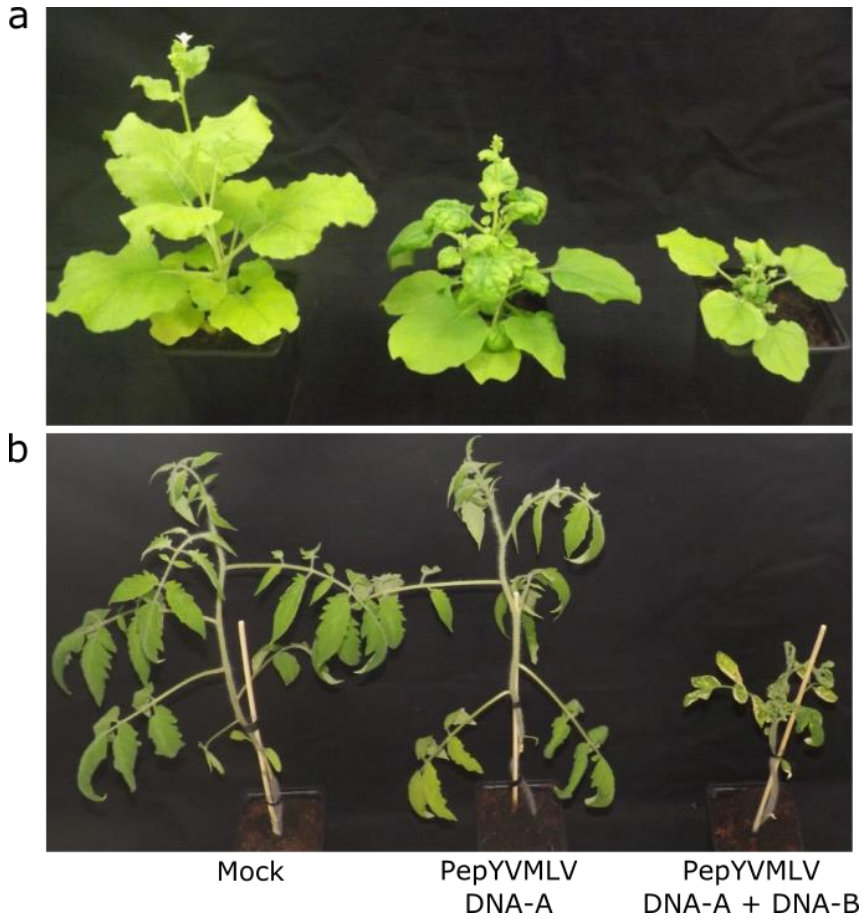


Figure 1: Disease symptoms on **(a)** *Nicotiana benthamiana* and **(b)** tomato (*Solanum lycopersicum*, cv. *Farmer 209*, Known-you seed) agroinoculated with PepYVMLV in single (DNA-A) or mixed (DNA-A and -B) infection at 30 days post inoculation. Control plants mock-inoculated (left).

Table 1: Infectivity of PepYVMLV and TYLCV-IL DNA-As in single or mixed association with PepYVMLV DNA-B after agroinoculation of *N. benthamiana* and tomato plants.

	Single infection		**Mixed infection**		
	Symptom	PCR-DNA-A	Symptom*	PCR-DNA-A	PCR-DNA-B
N. benthamiana/PepYVMLV	82.5% (33/40)	92.5% (37/40)	85% (34/40)	100% (40/40)	100% (40/40)
Tomato/PepYVMLV	20% (12/60)	80% (48/60)	95% (57/60) [10]	100% (60/60)	87% (52/60)
Tomato/TYLCV-IL	100% (60/60)	100% (60/60)	100% (60/60)	100% (60/60)	28% (17/60)

*Values in brackets indicate number of dead plants

Kinetics of symptom severity

The kinetics of symptom severity of PepYVMLV DNA-A and TYLCV-IL DNA-A were compared in single or mixed infection with DNA-B (Figure 2). Tomato plants agroinoculated with PepYVMLV DNA-A and -B started exhibiting symptoms of leaf crumpling, yellowing and stunting at 12 dpi (Figure 2, red curve). In contrast, for tomato plants agroinoculated with PepYVMLV DNA-A in single infection, the first mild symptoms of crumpling, yellowing and stunting were observed after 27 dpi (green curve). Typical symptoms of leaf curling, yellowing and dwarfism were observed on tomato plants agroinoculated with TYLCV-IL in single or mixed infection with DNA-B from 12 dpi (black and blue curves, respectively), similar to plants agroinoculated by PepYVMLV DNA-A and DNA-B. Symptom severity increased exponentially before reaching a plateau with highly severe disease symptoms (score = 9) at 22 and 32 dpi for PepYVMLV DNA-A and -B, and TYLCV-IL in single or mixed agroinfection, respectively. Whereas symptom severity did not significantly differ at the plateau phase ($p = 0.827$) and the slope of the exponential phase ($p = 0.062$) between the two viruses, the second parameter describing the speed of symptom progress (B; time to reach 50% of the disease severity) was very significantly different ($p < 10^{-4}$, Table 2). No significant difference was observed for the virulence kinetics of TYLCV-IL associated or not with PepYVMLV DNA-B (Table 2).

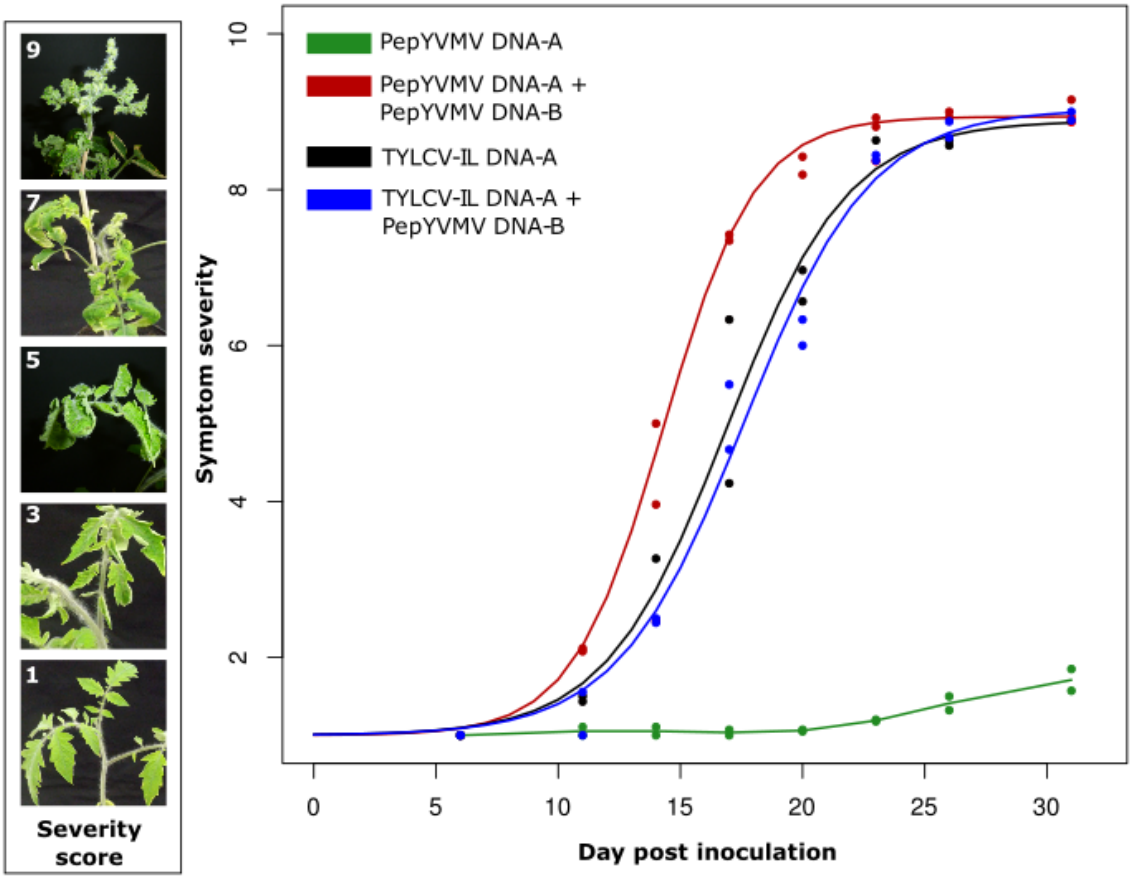


Figure 2: Kinetics of estimated symptom severity (solid line) of tomato (yellow) leaf curl disease following agroinoculation of tomato plants with PepYVMLV DNA-A and TYLCV-IL DNA-A in single or mixed infection with PepYVMLV DNA-B. The symptom severity scale (left scale bar) ranges from 1 (no symptom) to 10 (dead plant).

Table 2: Estimated parameters [95% confidence intervals] of the logit growth models governing symptom progression for tomato plants agroinoculated with PepYVMLV and TYLCV-IL in single or mixed inoculation with PepYVMLV DNA-B.

	Estimated logit model parameters		
Virus	**A**	**B**	**C**
PepYVMLV DNA-A + DNA-B	0.53 [0.46 - 0.60]	14.46 [14.17 - 14.74]	7.94 [7.74 - 8.13]
TYLCV-IL DNA-A	0.40 [0.30 - 0.51]	16.90 [16.13 - 17.67]	7.88 [7.35 - 8.42]
TYLCV-IL DNA-A + DNA-B	0.39 [0.30 - 0.47]	17.61 [16.95 - 18.27]	8.03 [7.56 - 8.51]

A: Slope during the exponential phase

B: Time to reach 50% of the viral load at the plateau phase

C: Viral load at the final plateau

Negative effect of DNA-B component on tomato growth

Surprisingly, our agroinoculation experiments showed no significant difference in size between the control and plants agroinoculated with PepYVMLV DNA-A alone (p = 0.428, Figure 3). As expected, a significant difference in size between the control and plants agroinoculated with TYLCV-IL alone was observed (p = 0.001). Interestingly, the agroinoculation of PepYVMV DNA-A and -B to tomato plants strongly impacted their growth, with a very highly significant reduction in size compared to the plants inoculated with PepYVMLV DNA-A alone ($p < 10^{-4}$, Figure 3). In contrast, no significant difference in size was observed between plants agroinoculated with TYLCV-IL DNA-A, in single or mixed inoculation with PepYVMLV DNA-B (p = 0.264).

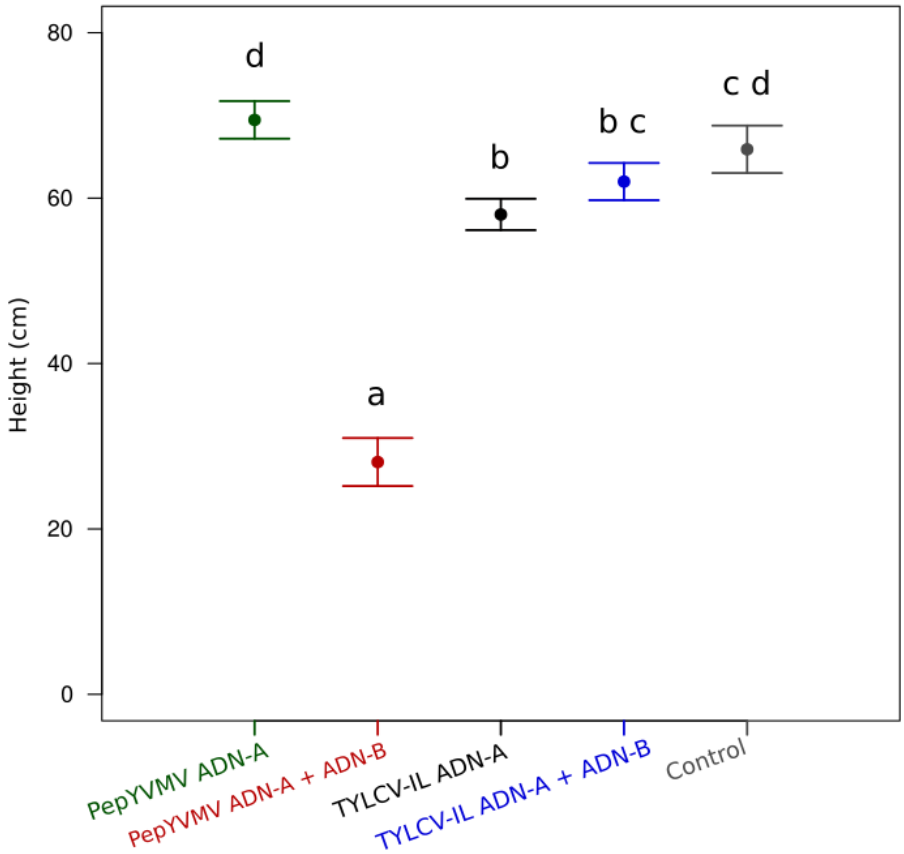


Figure 3: Mean size (in centimeters) of tomato plants after agroinfection with PepYVMLV DNA-A or TYLCV-IL in single or mixed infection with PepYVMLV DNA-B at 32 days post inoculation. Vertical bars around each dot represent the 95% confidence interval. Bars topped by the same letter are not significantly different.

Higher accumulation of PepYVMLV DNA-A in association with DNA-B

Within-plant accumulation of PepYVMLV DNA-A, associated or not with the DNA-B component, was evaluated in agroinoculated tomato plants at 32 dpi using quantitative PCR assays. In mixed infections (Figure 4), PepYVMLV DNA-A accumulated $2x10^5$-fold more than in the single infection ($p = 2.2x10^{-16}$, Figure 4a). The average number of PepYVMLV DNA-A copies per plant genomic DNA unit was about $1.28x10^{-1}$ in mixed infection against $5.9x10^{-7}$ in single infection. A positive linear relationship was found between the copy numbers of PepYVMLV DNA-A and DNA-B (Figure 4b), suggesting that there is a linear, DNA-B dose-dependent, increase of DNA-A viral load. Interestingly, considering the fact that DNA-A and DNA-B have been co-agroinoculated at equimolar quantities, their quantification at 32 dpi indicates a possible co-regulation of DNA-A with DNA-B.

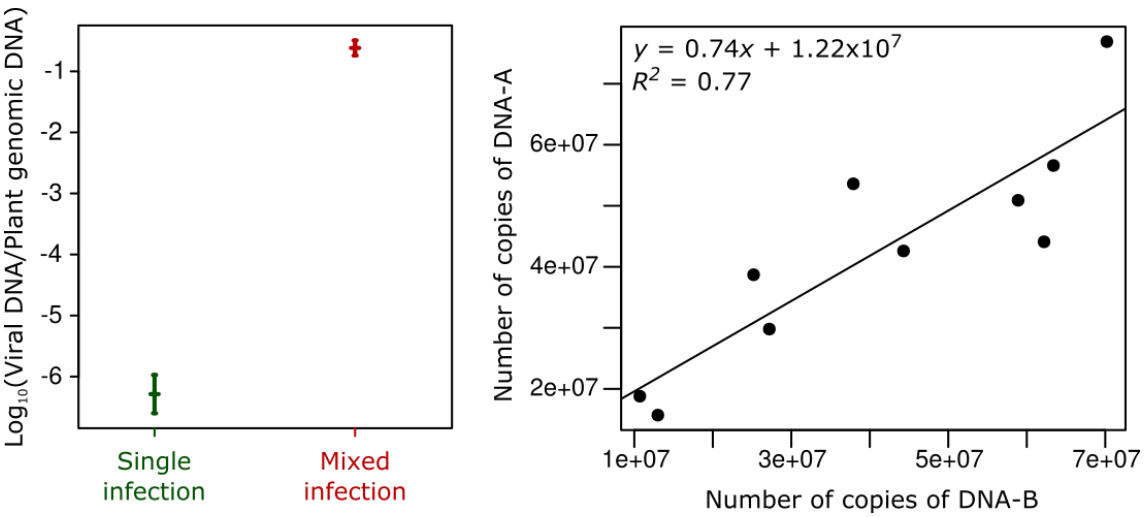


Figure 4: (a) Average of normalized viral DNA-A accumulation of PepYVMLV in single (DNA-A) or mixed (DNA-A and -B) infection of 10 tomato plants (*Solanum lycopersicum*, cv. *Farmer 209*, Known-You Seed). Vertical bars around each point represent the 95% confidence interval. **(b)** Graph representing the linear correlation between PepYVMLV DNA-A and -B loads. Each dot represents the quantity of DNA-A with respect to that of DNA-B.

Higher transmission rate of PepYVMLV DNA-A in association with DNA-B

Transmission experiments were performed twice independently with synchronous female adult individuals of the cryptic species MEAM1 of *B. tabaci*, fed on tomato plants agroinfected with PepYVMLV DNA-A alone or in association with DNA-B (Table 3). Based on disease symptom development, highly significant differences were found between single and mixed infections ($p < 10^{-4}$), with transmission rates of 52% (24/46) and 40% (14/35) for PepYVMLV DNA-A alone, and 71% (29/41) and 69% (27/39) for PepYVMLV DNA-A and DNA-B in the two experiments. Based on PepYVMLV DNA-A PCR detection, very highly significant differences were also found between single or mixed infections ($p = 4x10^{-5}$), with transmission rates of 80% (37/46) and 83% (29/35) for PepYVMLV DNA-A alone, and 100% for PepYVMLV DNA-A and DNA-B (41/41 and 39/39) in the two experiments. Here, we found that in association with the DNA-B component, PepYVMLV gained an average of 18% to reach 100% individual transmission efficiency by the *B. tabaci* MEAM1.

Table 3: Transmission rates of PepYVMLV by *Bemisia tabaci* MEAM1. Experiments were conducted with synchronous females after viral acquisition on tomato plants agroinoculated with PepYVMLV DNA-A in single or mixed infection with DNA-B. Percentage of transmission was evaluated based on disease symptom development and viral DNA detection by PCR.

	Single infection		**Mixed infection**		
	Symptoms	PCR DNA-A	Symptoms	PCR DNA-A	PCR DNA-B
Experiment 1	52% (24/46)	80% (37/46)	71% (29/41)	100% (41/41)	98% (40/41)
Experiment 2	40% (14/35)	83% (29/35)	69% (27/39)	100% (39/39)	100% (39/39)

Discussion

In Burkina Faso, recent studies have reported the identification of at least five begomoviruses on tomato crops (Tiendrébéogo *et al.*, 2008; Zhou *et al.*, 2008; Sattar *et al.*, 2015; Ouattara *et al.*, 2017), with the predominance of PepYVMLV which was surprisingly associated with a newly discovered DNA-B component (Ouattara et al, in preparation, Chapter 1). In order to uncover the role of the DNA-B component in the emergence of PepYVMLV as the most prevalent and severe plant virus disease on tomato crops in Burkina Faso, biological traits related to virulence, virus accumulation in the plant and individual transmission rate of the virus by *B. tabaci* were evaluated by *agrobacterium*- and whitefly-mediated inoculations.

Non-strict association between PepYVMLV DNA-A and -B

Agrobacterium-mediated inoculation experiments showed that PepYVMLV DNA-A alone induces systemic and symptomatic infections on *N. benthamiana* and tomato plants (Figure 1, Table 1). This result is compatible with the similar genome organization of PepYVMLV DNA-A with the Old World monopartite begomoviruses (Chen *et al.*, 2009; Tiendrébéogo *et al.*, 2011), and the evidence implicating V1 ORF in cell-to-cell movement (Rojas *et al.*, 2001). Such non-strict association between a monopartite begomovirus and a DNA-B component has been described for ToLCGV (Ranjan *et al.*, 2013), and TYLCTHV (Rochester *et al.*, 1994), which are able to induce systemic and symptomatic infections on the experimental host *N. benthamiana* (Rochester et al., 1994) and on tomato (Chakraborty et al, 2013), respectively, in the absence of their cognate DNA-Bs. In contrast, like for the New World bipartite begomoviruses, some Old Word tomato-infecting begomoviruses as *Tomato leaf curl New Dehli virus* (ToLCNDV) has an absolute requirement of their cognate DNA-B to induce a systemic movement and symptom development (Padidam *et al.*, 1995). Altogether, these experimental results suggest that these non-strict “mono-bipartite” begomoviruses likely represent evolutionary intermediates between monopartite and bipartite begomoviruses.

Higher virulence of PepYVMLV in association with DNA-B

Unlike single agroinfection of tomato plants with PepYVMLV, mixed agroinoculation with PepYVMLV DNA-A and -B resulted not only in stronger symptoms of leaf crumpling, yellowing and stunting, as observed in the field, but also led to plant death in 17% of inoculated plants. Several studies reported that DNA-B of begomoviruses contributes to symptom production (Jyothsna *et al.*, 2013; Ranjan *et al.*, 2013) and that BC1 protein is a pathogenicity determinant (Duan *et al.*, 1997; Hou *et al.*, 2000; Hussain *et al.*, 2005). However, to our knowledge, this is the first time that a DNA-B component is involved in plant mortality. Assuming that emerging viral diseases are mainly related to novel associations between introduced cultivated plants and native viruses or new viral associations (Jones, 2009), it could be assumed that the expression of disease symptoms as severe as the death of the plant is the consequence of a recent association between PepYVMLV DNA-A and DNA-B. This scenario is consistent with our findings suggesting that PepYVMLV components A and B originate from the Old World, while tomato is a New World plant and that ToLCD-TYLCD epidemics are described on tomato in West Africa since the 1990s (Konaté *et al.*, 1995).

TYLCV-IL has been reported as one of the most severe and devastating tomato virus worldwide (Péréfarres *et al.*, 2014). In order to compare PepYVMLV virulence in association with DNA-B to that of TYLCV-IL, we performed *Agrobacterium*-mediated inoculation experiments. Although the same levels of disease severity (score = 9) were observed at 32 dpi, PepYVMLV associated with DNA-B induced an earlier expression of disease symptoms and a more severe impact on plant growth (plant growth arrest and mortality). Taken together, these observations demonstrate that PepYVMLV associated with DNA-B is more virulent than the emblematic TYLCV-IL on tomato in controlled conditions, which highlights the new global risk for tomato culture if the PepYVMLV associated with DNA-B would rapidly spread worldwide under whitefly pressure, like TYLCV-IL.

Higher viral accumulation of PepYVMLV in association with DNA-B

Within-plant accumulation of PepYVMLV DNA-A in single or mixed infection with DNA-B in tomato plants using quantitative PCR assays showed that mixed-infected plants contained $2x10^5$-fold more viral DNA-A than in single infection. Even if BV1 and BC1 proteins encoded by DNA-B provide an analogous function to that performed by the V1/C4 proteins of the monopartite begomoviruses in viral movement (Rojas *et al.*, 2001), they appeared to facilitate the escape of some bipartite begomoviruses from the phloem and the infection of non-phloem tissues (Rojas *et al.*, 2005). It seems that DNA-B component allowed for the evolution of movement functions, independently of the constraints imposed by DNA-A. Based on this, we can speculate that DNA-B component is involved in the increase in PepYVMLV DNA-A load by enlarging cellular tropism and giving access to non-phloem tissues. Infection of these non-phloem tissues could enhance the opportunity for acquisition of virions by the insect vector (Rojas *et al.*, 2005). In addition, as the insect vectors are attracted to the color yellow, it may well be that by moving from the phloem into the surrounding tissues, the virus can induce larger regions of yellowing (Wege *et al.*, 2001) and thus create a stronger signal by which the vector is attracted.

Non-optimal association between TYLCV-IL and PepYVMLV DNA-B

As opposed to its major impact on PepYVMLV infection, PepYVMLV DNA-B had no significant impact on TYLCV-IL infection. Moreover, no increase in symptom severity was observed. Interestingly, DNA-B was only detected by PCR in 28% of mixed agroinfected tomato plants, indicating non-optimal association between TYLCV-IL and PepYVMLV DNA-B (Table 1). Geminivirus genomes replicate by a rolling circle mechanism which is initiated by the virus-encoded replication-associated protein (Rep) (Hanley-Bowdoin *et al.*, 1999, 2013; Fondong, 2013). Rep binds to specific sequences, known as iterons, in the intergenic region immediately upstream of a hairpin structure containing the nonanucleotide sequence (TAATATTAC). The DNA-A and DNA-B components of bipartite begomoviruses have the same iteron sequences, thereby ensuring that the DNA-A-encoded Rep may initiate replication of both components. However, mutational analyses suggested that begomoviruses may tolerate some

sequence variation in iteron sequences without deleterious effects on Rep recognition and it was showed that the first two or three bases may be more important in iteron recognition by Rep (Shafiq *et al.*, 2010). This does not seem to be the case for TYLCV-IL (iteron sequence: GGTGTCT) associated with PepYVMLV DNA-B (iteron sequence: GGGGTAC). Altogether, our results suggest that this association is not optimal and that DNA-B association with other tomato-infecting monopartite begomoviruses is not "universal".

Higher transmission efficiency of PepYVMLV in association with DNA-B

Transmission from host to host is a crucial step in the life cycle of pathogens, particularly of viruses, ensuring spread and maintenance in host populations. The immobile nature of plants and the strong pectin and cellulose barrier surrounding cells have constrained most plant virus to use vectors (mainly insects) for exit, transfer and entry from one host to another (Gutiérrez *et al.*, 2013). A link between within-plant viral load and transmission rate is generally accepted even if studies are scarce (Froissart *et al.*, 2010; Péréfarres *et al.*, 2014). Here, we demonstrated that PepYVMV accumulates much more within tomato plants ($2x10^5$-fold more) and is much better transmitted by whitefly (18% more) in mixed infection than in single infection (without DNA-B). This confirms the suggested link between within-plant viral load and transmission rate (Froissart *et al.*, 2010; Péréfarres *et al.*, 2014). Although the simplest explanation would be a mass-action mechanism (more virus genomes in the plant implying more virions in the insect implying a higher transmission rate), the increase in the transmission rate might also be explained by a better ability of the virus to initiate the infection of tomato plant cells thanks to the movement proteins encoded by the DNA-B.

Conclusion

Our study highlights the role of a DNA-B component in the emergence of a monopartite begomovirus. Our results suggest that the recruitment of a DNA-B component by a monopartite PepYVMLV represents likely the key epidemiological factor of its emergence as the most prevalent and severe plant virus disease on tomato crops in Burkina Faso. Although a DNA-B component is not required for systemic movement and symptom development of PepYVMLV (non-strict association), their association

induces a higher virulence with typical symptoms of crumpling, stunting and plant death, a much higher viral DNA accumulation in plant tissues and a much better whitefly transmission efficiency. At the agroecosystem level, this selective advantage to be more efficiently vectored is also associated with a large host range and the existence of weeds as potential reservoirs or alternative hosts widely present in the fields (Ouattara et al., in preparation, Chapter 1). Maintenance of the virus between epidemics in alternate hosts present in the cultivated area provides a means to survive through the seasonal cycle, and could also explain the predominance of PepYVMLV associated with DNA-B as the main causal agent of ToLCD-TYLCD in Burkina Faso. In the case of the non-strict association of PepYVMLV DNA-A and DNA-B, the latter may be regarded as an extra-viral component that can affect the pathogenicity, the accumulation, and the transmission of their cognate virus while themselves being dependent on the associated virus for their infection.

Acknowledgements

This work was supported by the European Union (ERDF), the Conseil Régional de La Réunion and CIRAD. AO was a recipient of a PhD fellowship from CIRAD.

Supplementary table 1: Primers, probes and standards developed and used in quantitative PCR

Name	Sequence 5'-3'	nM	Reporter dye
F-Sl-25S	CGCCCGGTCGTACTCATAA	100	none
R-Sl-25S	TCCATCGACCAGAGGCTGTT	100	none
P-Sl-25S-MGB	CGCATCAGGTCTCCA	250	VIC
Standard-25S	CGCCCGGTCGTACTCATAACCGCATCAGGTCTCCAAGGTGAACAGCCTCTGGTCGATGGA		
PepYVMLV-B-1606F	GCAGAACTAGGGTTGGGATTCA	900	none
PepYVMLV-B-1653R	TGTAGATTTCTGGTCCGTCGAA	900	none
PepYVMLV-B-1634P-MGB	TCCTTGATGTTAGACCTC	200	FAM
Standard-PepYVMLV-B	GCAGAACTAGGGTTGGGATTCAGTAATCTCCTTGATGTTAGACCTCGTTCGACGGACCAGAAATCTACA		none
PepYVMLV-A-2183F	CAGCGGAATCTGCATTTAAGG	50	none
PepYVMLV-A-2240R	GGAGGATGCCGAAATGCTAA	250	none
PepYVMLV-A-2205P-MGB	CTCTGCTGCCGCGTC	250	FAM
Standard-PepYVMLV-A	CAGCGGAATCTGCATTTAAGGCCTCTGCTGCCGCGTCGTTAGCATTTCGGCATCCTCC		

References

Briddon, R. & Stanley, J. (2006) Subviral agents associated with plant single-stranded DNA viruses. *Virology*, **344**, 198–210.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C.F., Fiallo-Olivé, E., Briddon, R.W., Hernández-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S. & Varsani, A. (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of Virology*, **160**, 1593–1619.

Chakraborty, S., Pandey, P.K., Banerjee, M.K., Kalloo, G. & Fauquet, C.M. (2003) Tomato leaf curl Gujarat virus a New Begomovirus Species Causing a Severe Leaf Curl Disease of Tomato in Varanasi, India. *Phytopathology*, **93**, 1485–1495.

Chen, L.-F., Rojas, M., Kon, T., Gamby, K., Xoconostle-ca, B. & Gilbertson, R.L. (2009) A severe symptom phenotype in tomato in Mali is caused by a reassortant between a novel recombinant begomovirus (Tomato yellow leaf curl Mali virus) and a betasatellite. *Molecular Plant Pathology*, **10**, 415–430.

Ditta, G., Stanfield, S. & Corbin, D. (1980) Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank of Rhizobium meliloti. *Proceedings of the National Academy of Sciences*, **77**, 7347–7351.

Duan, Y.P., Powell, C. a, Purcifull, D.E., Broglio, P. & Hiebert, E. (1997) Phenotypic variation in transgenic tobacco expressing mutated geminivirus movement/pathogenicity (BC1) proteins. *Molecular Plant- Microbe Interactions*, **10**, 1065–74.

Fondong, V.N. (2013) Geminivirus protein structure and function. *Molecular Plant Pathology*, **14**, 635–649.

Froissart, R., Doumayrou, J., Vuillaume, F., Alizon, S. & Michalakis, Y. (2010) The virulence-transmission trade-off in vector-borne plant viruses: a review of (non-) existing studies. *Philosophical Transactions of the Royal Society B*, **365**, 1907–1918.

Gutiérrez, S., Michalakis, Y., Van Munster, M. & Blanc, S. (2013) Plant feeding by insect vectors can affect life cycle , population genetics and evolution of plant viruses. *Functional Ecology*, **27**, 610–622.

Hamilton, W.D.O., Bisaro, D.M., Coutts, R.H.A. & Buck, K.W. (1983) Demonstration of the bipartite nature of the genome of a single-stranded DNA plant virus by infection with the doned DNA component. *Nucleic Acids Research*, **11**, 7387–7396.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D. & Mansoor, S. (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology*, **11**, 777–788.

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Robertson, D., Settlage, S.B., Orozco, B.M., Hanley-bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S. & Robertson, D. (1999) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Plant Sciences*, **18**, 71-106.

Hou, Y., Sanders, R., Ursin, V.M. & Gilbertson, R.L. (2000) Transgenic Plants Expressing Geminivirus Movement Proteins : Abnormal Phenotypes and Delayed Infection by Tomato mottle virus in Transgenic Tomatoes Expressing the Bean dwarf mosaic virus BV1 or BC1 Proteins. *Molecular Plant- Microbe Interactions*, **13**, 297–308.

Hussain, M., Mansoor, S., Iram, S., Fatima, A.N. & Zafar, Y. (2005) The Nuclear Shuttle Protein of Tomato Leaf Curl New Delhi Virus Is a Pathogenicity Determinant. *Journal of Virology*, **79**, 4434–4439.

Jones, R. (2009) Plant virus emergence and evoution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, **141**, 113–130.

Jyothsna, P., Haq, Q.M.I., Singh, P., Sumiya, K. V., Praveen, S., Rawat, R., Briddon, R.W. & Malathi, V.G. (2013) Infection of tomato leaf curl New Delhi virus

(ToLCNDV), a bipartite begomovirus with betasatellites, results in enhanced level of helper virus components and antagonistic interaction between DNA B and betasatellites. *Applied Microbiology and Biotechnology*, **97**, 5457–5471.

Konaté, G., Barro, N., Fargette, D. & Swanson, M.M. (1995) Occurrence of whitefly-transmitted geminiviruses in crops in Burkina Faso, and their serological detection and differentiation. *Annals of Applied Biology*, **126**, 121–130.

Lapidot, M., Cohen, L., Machbash, Z. & Levy, D. (2006) Development of a Scale for Evaluation of Tomato yellow leaf curl virus Resistance Level in Tomato Plants. *Phytopathology*, **96**, 1404–1408.

Lefeuvre, P., Hoareau, M., Delatte, H., Reynaud, B. & Lett, J.-M. (2007) A multiplex PCR method discriminating between the TYLCV and TYLCV-Mld clades of Tomato yellow leaf curl virus. *Journal of Virological Methods*, **144**, 165–168.

Navas-Castillo, J., Fiallo-Olivé, E. & Sanchez-Campos, S. (2011) Emerging Virus Diseases Transmitted by Whiteflies. *Annual Review of Phytopathology*, **49**, 219–248.

Ouattara, A., Tiendrébéogo, F., Lefeuvre, P., Claverie, S., Hoareau, M., Traoré, E.V., Barro, N., Traoré, O. & Lett, J.M. (2017) Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso. *Archives of Virology*, **162**, 1427–1429.

Padidam, R., Beachy, R.N. & Fauquet, C.M. (1995) Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *Journal of General Virology*, **76**, 25–35.

Péréfarres, F., Hoareau, M., Chiroleu, F., Reynaud, B., Dintinger, J. & Lett, J.-M. (2011) A novel synthetic quantification standard including virus and internal report targets: application for the detection and quantification of emerging begomoviruses on tomato. *Virology Journal*, **8**, 1–9.

Péréfarres, F., Thébaud, G., Lefeuvre, P., Chiroleu, F., Rimbaud, L., Hoareau, M., Reynaud, B. & Lett, J.-M. (2014) Frequency-dependent assistance as a way out of

competitive exclusion between two strains of an emerging virus. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 1–9.

Pinheiro, J.C., Bates, D.M., DebRoy, S., Sarkar, D. & Team, R.C. (2016) nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-128

Pressing, J. & Reanney, D.C. (1984) Divided Genomes and Intrinsic Noise. *Journal of Molecular Evolution*, **20**, 135–146.

Ranjan, P., Kumar, R.V. & Chakraborty, S. (2013) Differential pathogenicity among Tomato leaf curl Gujarat virus isolates from India. *Virus Genes*, **47**, 524–531.

Rey, M.E.C., Ndunguru, J., Berrie, L.C., Paximadis, M., Berry, S., Cossa, N., Nuaila, V.N., Mabasa, K.G., Abraham, N., Rybicki, E.P., Martin, D., Pietersen, G. & Esterhuizen, L.L. (2012) Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in southern Africa, including the South-west Indian ocean islands. *Viruses*, **4**, 1753–91.

Rochester, D.E., DePaulo, J.J., Fauquet, C.M. & Beachy, R.N. (1994) Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus, Thailand isolate. *Journal of General Virology*, **75**, 477–485.

Rojas, M.R., Hagen, C., Lucas, W.J. & Gilbertson, R.L. (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, **43**, 361–394.

Rojas, M.R., Jiang, H., Salati, R., Xoconostle-ca, B., Lucas, W.J. & Gilbertson, R.L. (2001) Functional Analysis of Proteins Involved in Movement of the Monopartite Begomovirus , Tomato Yellow Leaf Curl Virus. *Virology*, **291**, 110–125.

Sattar, M.N., Koutou, M., Husseini, S., Leke, W.N., Brown, J.K. & Kvarnheden, A. (2015) First Identification of Begomoviruses Infecting Tomato With Leaf Curl Disease in Burkina Faso. *Plant Disease*, **99**, 732–732.

Shafiq, M., Asad, S., Zafar, Y., Briddon, R.W. & Mansoor, S. (2010) Pepper leaf curl Lahore virus requires the DNA B component of Tomato leaf curl New Delhi virus

to cause leaf curl symptoms. *Virology Journal*, **7**, 367.

Sicard, A., Michalakis, Y., Gutiérrez, S. & Blanc, S. (2016) The Strange Lifestyle of Multipartite Viruses. *PLoS Pathogens*, **12**, 1–19.

Tiendrébéogo, F.., Traoré, E.V.S., Barro, N., Konaté, G., Traoré, A. & Traoré, O. (2008) Characterization of Pepper yellow vein Mali virus in Capsicum sp. in Burkina Faso. *Plant Pathology Journal*, **7**, 155–161.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Traoré, V.S.E., Barro, N., Péréfarres, F., Reynaud, B., Traoré, A.S., Konaté, G., Lett, J.-M. & Traoré, O. (2011) Molecular and biological characterization of Pepper yellow vein Mali virus (PepYVMV) isolates associated with pepper yellow vein disease in Burkina Faso. *Archives of Virology*, **156**, 483–487.

Varsani, A., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Brown, J.K., Murilo Zerbini, F. & Martin, D.P. (2014) Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. *Archives of Virology*, **59**, 2193-2203.

Wege, C., Saunders, K., Stanley, J. & Jeske, H. (2001) Comparative Analysis of Tissue Tropism of Bipartite Geminiviruses. *Journal of Phytopathology*, **149**, 359–368.

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

Zhou, Y.-C., Noussourou, M., Kon, T., Rojas, M.R., Jiang, H., Chen, L.-F., Gamby, K., Foster, R. & Gilbertson, R.L. (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. *Archives of Virology*, **153**, 693–706.

CHAPITRE III

DIVERSITE ET DYNAMIQUES DES POPULATIONS DE GEMINIVIRUS A L'ECHELLE DES SYSTEMES AGROECOLOGIQUES : APPROCHE METAGENOMIQUE CIBLEE

CHAPITRE III : DIVERSITE ET DYNAMIQUES DES POPULATIONS DE GEMINIVIRUS A L'ECHELLE DES SYSTEMES AGROECOLOGIQUES : APPROCHE METAGENOMIQUE CIBLEE

Jusqu'à présent, les efforts de caractérisation de la diversité virale se sont principalement focalisés sur les plantes d'intérêt agronomique, fortement impactées par les maladies émergentes. Toutefois, on considère que la majeure partie des maladies causées par des géminivirus sont liées au passage de virus indigènes, ayant co-évolué avec leur(s) hôte(s) naturel(s), sur des plantes introduites notamment pour la culture, non adaptés à ces virus et donc souvent plus sensibles à leurs effets (Fargette *et al.*, 2006; Nawaz-ul-Rehman & Fauquet, 2009). La découverte récente de nombreuses nouvelles espèces de géminivirus suite à l'étude de plante non cultivées (Varsani *et al.*, 2014; Bernardo *et al.*, 2017) suggère même qu'une grande partie de la diversité des géminivirus, sinon la vaste majorité, reste encore à être caractérisée à partir des plantes non-cultivées des écosystèmes naturels. En conséquence, outre les lacunes de nos connaissances de la diversité des populations virales naturelles, de leurs liens phylogénétiques avec les populations virales décrites sur les plantes cultivées et de leurs pathogénicités, on constate un manque de connaissances fondamentales sur l'écologie évolutive des virus en général (Alexander *et al.*, 2014; Roossinck *et al.*, 2015).

Ce biais dans la connaissance globale de la diversité des phytovirus présent dans les écosystèmes est particulièrement accru en Afrique, où la très grande majorité des espèces virales identifiées ont été caractérisées à partir d'un nombre restreint de plantes cultivées, comme le manioc, la tomate, le gombo, le tabac, le haricot et la patate douce.

Cependant, pour une compréhension complète de l'épidémiologie et pour l'élaboration de mesures de lutte appropriées, l'identification d'hôtes alternatifs est un aspect essentiel bien que souvent négligé (Leke *et al.*, 2015). À l'heure actuelle, parmi les géminivirus décrits sur les cultures maraîchères et vivrières en Afrique de l'Ouest et Centrale, seule une poignée d'espèces de plantes non cultivées potentiellement réservoirs de ces virus ont été décrits (Alabi *et al.*, 2008; Monde *et al.*, 2010; Leke *et al.*, 2012).

Ainsi, ce chapitre est consacré à l'étude de la diversité des populations de géminivirus à l'échelle des agroécosystèmes sur les plantes cultivées et non cultivées. Afin de détecter et d'identifier les espèces virales présentes dans nos échantillons sans a priori sur leurs

séquences nucléotidiques, une approche métagénomique ciblée sur les virus à ADN circulaire (RCA-RA-SGS) a été utilisée sur un large jeu d'échantillons (n=1068) collecté aléatoirement selon un système de grille dans deux localités du Burkina Faso. Sur ces échantillons et en globalité, 11 % des *reads* (courte séquence nucléotidique) ont été assignés à des virus. Parmi ceux-ci, la majorité (65%) a été assignée à des virus de la famille des *Geminiviridae*. Si le restant a surtout été assigné à des familles virales inféodées aux bactéries (25%), respectivement 3 % et 1 % des *reads* ont pu être assignés aux familles des *Caulimoviridae* (virus infectant les plantes) et des *Genomoviridae* (proche voisin des *Geminiviridae*).

Si on s'intéresse plus précisément aux géminivirus, en comparaison avec le diagnostic PCR, notre étude a démontré l'intérêt d'une approche métagénomique. Elle a non seulement permis de confirmer la présence d'une majorité des virus décrits préalablement avec les approches classiques, mais également permis la détection d'autres géminivirus appartenant potentiellement à de nouvelles espèces.

Ainsi 25% des échantillons se sont montrés positifs aux géminivirus, que ce soit par la présence de composant A ou B des bégomovirus, d'ADN satellites ou de mastrévirus. Les taux d'infection des plantes cultivés (48%) se sont aussi révélés plus élevés que ceux des plantes non cultivés (20%). Parmi les 19 genres de plantes collectés, les genres *Abelmoschus*, *Capsicum* et *Solanum* pour les plantes cultivées, et *Sida* et *Ipomoea* pour les plantes non cultivées ont montré des taux d'infection élevés variant de 30 à 90% suivant la localité d'échantillonnage. Au total, les géminivirus ont été détectés dans 14 genres de plantes, avec en plus de ceux cités précédemment, les genres des *Amaranthus*, *Ageratum*, *Boerhavia* et *Physalis* (décrits comme hôtes des géminivirus par ailleurs), et *Euphorbia*, *Heliotropium*, *Hibiscus*, *Melochia* et *Leucas* (hôtes nouvellement caractérisés).

Dans un second temps, les *reads* obtenus à partir de chacun des échantillons ont été regroupés suivant leurs profils d'assignation. Ces profils d'assignation sont révélateurs du ou des types de virus identifiés pour un échantillon ainsi que de leurs niveaux d'identité nucléotidique avec les espèces virales déjà connues. Au total, respectivement 16 et six types de profil d'ADN-A et d'ADN-B de bégomovirus, 13 types de profil d'ADN satellites et six types de profil de mastrévirus ont été identifiés lors de notre

étude. Si certaines espèces avaient déjà été caractérisés au Burkina Faso, comme le PepYVMLV, le ToLCMLV ou encore le CLCuGeV, d'autres représentent de nouvelles descriptions virales au Burkina Faso (*Hollyhock leaf crumple virus*, *Sweet potato leaf curl virus* et *Tobacco leaf curl Zimbabwe virus*) ou en Afrique en général (certains satellites), qui devront être néanmoins confirmés par le clonage et le séquençage de génomes complets. Par ailleurs, quatre profils d'assignation se sont montrés suffisamment distant avec des espèces déjà décrites, pour suspecter la présence de nouvelles espèces de géminivirus. Cette hypothèse a par ailleurs été confirmée par le clonage et le séquençage à partir d'un échantillon de *Melochia* sp. d'un nouveau begomovirus, très distant (74 % d'identité nucléotidique) des espèces déjà décrites et représentant une nouvelle espèce provisoirement appelée *Melochia associated virus*.

L'identification de ce cortège de virus, à partir d'un large nombre d'échantillons représentatifs des plantes de deux agroécosystèmes nous a aussi permis de décrire la structure des communautés de géminivirus. Celle-ci révèle la présence de degré variable de spécialisation. Aussi certaines plantes sont infectés par un nombre faible (comme celle du genre *Physalis*) ou important (comme celle du genre *Sida*) d'espèces virales. Dans leur globalité, l'analyse de l'association plante - virus permet de décrire des liens forts entre les cortèges de virus des plantes des genres *Sida* et *Abelmoschus* d'une part et des genres *Solanum*, *Capsicum* et *Ageratum* d'autre part. Ces deux groupes de virus ne présentent pas de liens directs mais certains genres d'hôtes (*Amaranthus* et *Euphorbia*) pourraient jouer le rôle d'intermédiaire pour la rencontre et l'interaction entre ces populations virales. Les cartographies des circuits d'infection virale et des échanges de matériel génétique par recombinaison entre groupes de virus devraient être révélatrices du fonctionnement de la communauté virale. Les travaux de clonage et de séquençage de divers virus représentatifs de chacun des clusters viraux découverts dans notre étude, devraient permettre d'évaluer l'importance à la fois des échanges de gènes entre les virus, de certaines plantes et de certains virus dans le fonctionnement des communautés virales à l'échelle des agroécosytèmes.

Article 7

Metagenomics approach to uncover geminiviruses diversity and community structure at the scale of agro-ecological system in two localities of Burkina Faso

Metagenomics approach to uncover geminiviruses diversity and community structure at the scale of agro-ecological system in two localities of Burkina Faso

Alassane Ouattara[1,2,3,4,5], Sohini Claverie[2], Bionimian Ezechiel Tibiri[1,5], Fidèle Tiendrébéogo[1,5], Murielle Hoareau[2], Edgar Valentin Traoré[1,5], Nicolas Barro[4], Oumar Traoré[1,5], Jean-Michel Lett[2], Pierre Lefeuvre[2]

[1] INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

[2] CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

[3] Université de La Réunion, UMR PVBMT, 97410, Saint-Pierre, Ile de La Réunion, France

[4] Université Ouaga I Professeur Joseph Ki-Zerbo, 03 BP 7021, Ouagadougou 03, Burkina Faso

[5] LMI Patho-Bios, IRD-INERA, 01 BP 476, Ouagadougou 01, Burkina Faso

Abstract

Geminiviruses have emerged to become one of the largest and most devastating groups of plants-infecting viruses. They constitute a major threat to vegetable production particularly in tropical and subtropical regions of the world. The vast majority of the geminiviruses known to date was characterized from crops. Nevertheless, it is thought that these crops infecting viruses represent only a small fraction of the existing diversity and that they emerged from viruses associated to uncultivated plants. However, the identification of alternative hosts or reservoirs was somewhat neglected in the past. Here, in order to describe the full diversity of geminiviruses associated to wild and cultivated plants in Burkina Faso, we extensively sampled two localities from distinct agroclimatic regions of the country before using a metagenomic protocol devised for small circular DNA viruses. A total of 1068 samples were collected from respectively 201 and 867 cultivated and uncultivated plants. A total of 270 plants (25%) were detected as positives with higher virus infection rates for cultivated (48%) than uncultivated plants (20%). At least 16, six, 13 and six distinct groups of begomovirus DNA-A, begomovirus DNA-B, satellites DNA and mastrevirus respectively were uncovered from our survey. Importantly, these groups comprised viruses highly similar or related to viruses already described along with putative new species which have yet to be characterized. Preliminary characterization assays result in the cloning of seven isolates of begomovirus and a new geminivirus species provisionally named *Melochia associated virus*. Seven alphasatellites and two betasatellites isolates were also obtained. Geminivirus community maps built from the analysis of virus-host and virus-virus association showed strong links between virus corteges found associated with groups of cultivated and uncultivated plant genus. This is the first effort to evaluate virus-host and virus-virus interactions at the scale of agro-ecological systems in West Africa. Further analysis of our networks using full genome sequences, selection and recombination analysis, would help understanding the functioning of this geminiviral community and in the end how viruses from wild plants can emerge as local and global threats to crops.

Introduction

Geminiviruses are plant-infecting viruses characterized by their unique twinned icosahedral virions and circular single-stranded DNA genomes (Rojas *et al.*, 2005; Zerbini *et al.*, 2017). These viruses have emerged to become one of the largest and most economically important groups of plant-infecting viruses (Fauquet *et al.*, 2008), and geminivirus-induced diseases are a major threat to worldwide vegetable production, particularly in tropical and subtropical regions of the world (Varma & Malathi, 2003; Mansoor *et al.*, 2006). Importantly, the accumulated body of work on some of the most important geminiviral associated diseases, such as the maize streak disease (for review see Shepherd *et al.*, 2010) and the cassava mosaic disease (Patil & Fauquet, 2009), clearly demonstrate the emergence of geminiviruses associated with wild plants onto imported crops. In fact, as on other pathosystems, it is now considered that most of these emerging viruses are the "offspring" of viral lineages present in "wild" plants, some of which would have adapted to cultivated plants (Jones, 2009). Moreover, recent metagenomic data suggested that the vast majority of viruses characterized from crops represent only a small fraction of the phytoviruses in general (Malmstrom *et al.*, 2011; Alexander *et al.*, 2014). The lack of comprehensiveness of current studies on the diversity of viral population results from the focus on crops of the majority of diversity studies and also from the difficulties in identifying possible symptoms associated with the infection of wild plants by viruses. Indeed, viruses that have long co-evolved with their host plants, very rarely induce harmful diseases in their natural ecosystems (Jones, 2009). In these ecological systems not disturbed by human activity, the interactions between vector, viruses and host plants are considered to be evolutionarily old and stable (Malmstrom *et al.*, 2011). However, due to environmental disturbances associated with modern agriculture, virus-host equilibriums are often severely disrupted, resulting in new virus vector-host-environment interactions, which can lead to the emergence of diseases (Fargette *et al.*, 2006; Lefeuvre & Moriones, 2015). In addition to the gaps in our knowledge of the diversity of natural viral populations, there is a lack of fundamental knowledge about the dynamics of viral communities at the scale of ecosystems and the evolutionary ecology of viruses (Alexander *et al.*, 2014; Roossinck *et al.*, 2015). Recent viral metagenomic studies have therefore started to look beyond

cropping systems to enable the study of plant viral diversity and evolution at the scale of entire ecosystems (Stobbe & Roossinck, 2014; Roossinck *et al.*, 2015). The development of a viral metagenomics approach that precisely linked sequence reads to information on the individual associated sampled host ('eco-genomic', Roossinck *et al.*, 2010), and its extension in a spatial context ('geo-metagenomic', Bernardo *et al.*, 2017) had paved the way towards quantitatively studying viral diversity, prevalence and spatial distributions at the agroecosystem scale.

In recent years, vegetables crops in Burkina Faso have been negatively impacted by diseases caused by geminiviruses. Early studies have reported *Cotton leaf curl Gezira virus* (CLCuGeV; (Tiendrébéogo *et al.*, 2010a,b) associated with okra leaf curl disease (OLCD), in combination or not with *Cotton leaf curl Gezira betasatellite* (CLCuGeB; Tiendrébéogo *et al.*, 2010a), *Cotton leaf curl Gezira alphasatellite* (CLCuGeA) and *Okra leaf curl Burkina Faso alphasatellite* (OLCBFA), (Tiendrébéogo *et al.*, 2010a). *Pepper yellow vein Mali virus* (PepYVMLV) has been associated with PepYVD (Tiendrébéogo *et al.*, 2011) and ToLCD-TYLCD (Ouattara *et al.*, unpublished, Chapter 1). *Tomato leaf curl Mali virus* (ToLCMLV; Sattar *et al.*, 2015), *Tomato yellow leaf curl Mali virus* (TYLCMLV; Sattar *et al.*, 2015) and, recently, *Tomato leaf curl Burkina Faso virus* (ToLCBFV; Ouattara *et al.*, 2017a), *Tomato leaf curl Ghana virus* (ToLCGHV; Ouattara *et al.*, unpublished, Chapter 1) and *Chickpea chlorotic dwarf virus* (CpCDV; Ouattara *et al.*, 2017b) have been characterized from tomato. However, the observation of varied virus-like symptoms, during field's surveys, suggested the involvement of several viruses possibly associated with satellites. Also, all of the above-mentioned geminiviruses have been characterized only based on the so-called dependent sequence methods. Given the pressing need for additional information on their diversity and distribution and the benefits of metagenomic methods, this study was initiated.

Here, we used a metagenomic approach based on DNA rolling circle amplification followed by random amplification and second generation sequencing (RCA-RA-SGS). After a large scale random sampling using a grid design, we used this metagenomic approach to assess the geminiviruses diversity associated with agro-ecological systems in two regions of Burkina Faso. Out of the 1068 collected plants, 270 were positive for the presence of geminiviruses. A large set of distinct geminivirus groups were detected,

with four groups of potentially new species. A geminivirus community map was built from the analysis of virus-host and virus-virus association. The map topology inspection allows establishing networks of viral sharing between host plants. The analysis of such community should give invaluable information on the process from which some virus may eventually emerge as local and global threats.

Materials and Methods

Samples collection

In 2016, a total of 1068 samples were collected from two different agro-ecosytems in two localities of Burkina Faso (Figure 1): Goué (Central-East near Ouagadougou; latitude: 12.567°, longitude: -1.471°) and Léguéma (South-West near Bobo-Diolasso; latitude: 11.226°, longitude: -4.164°), respectively in the Sudan and Sudan-Sahel areas. In order to sample randomly and without bias regarding the health status of plants, a grid sampling strategy was used. Each grid had dimensions of 500m X 1000m with meshes of 250m X 250m. In a radius of 20 meters of each point of the grid, up to 10 samples were collected per identified plant species. All samples were collected in envelopes and dried using an oven.

In addition to this random sampling scheme, from 2013 to 2016, 372 (98 cultivated and 274 uncultivated) plant leave samples were collected during surveys conducted in 29 localities from Burkina Faso. The plant species, the name of the locality and the geographical coordinates (latitude, longitude) of samples were recorded.

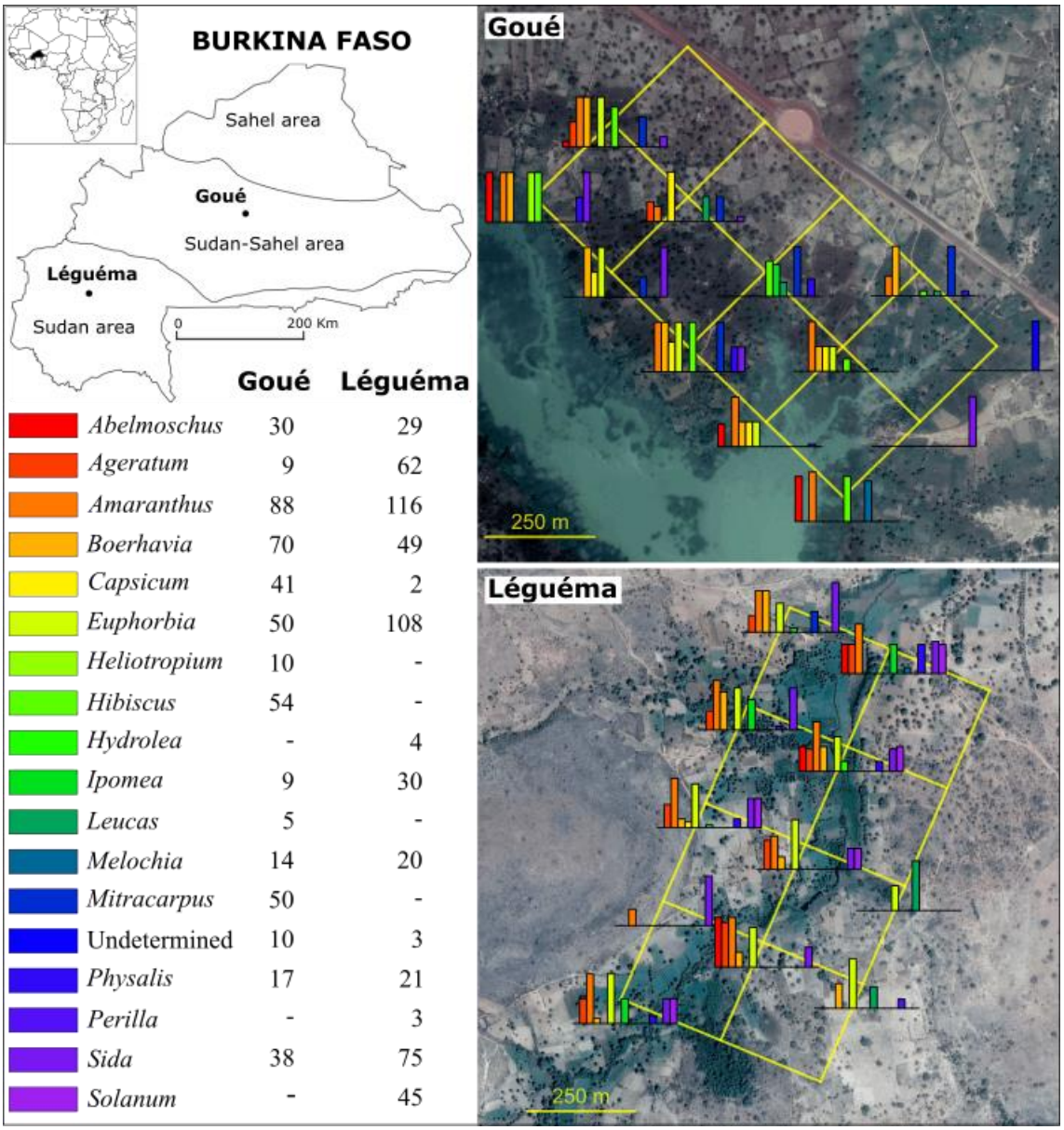
DNA extraction

Total DNA was extracted for all samples using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. DNA extracts were stored at -30°C before use.

Sample barcoding

Sequencing of the *matk* and *rbcl* genes were performed on samples for which no confident genus or species visual identification was obtained. PCR were conducted with the primers and conditions described in (Hollingsworth *et al.*, 2009) before direct

Sanger sequencing (Macrogen, South Korea). After quality control, sequences were classified using the RDP classifier (Wang *et al.*, 2007) against a database of *matk* and *rbcl* plant sequences obtained from GenBank.



	Goué	Léguéma
Abelmoschus	30	29
Ageratum	9	62
Amaranthus	88	116
Boerhavia	70	49
Capsicum	41	2
Euphorbia	50	108
Heliotropium	10	-
Hibiscus	54	-
Hydrolea	-	4
Ipomea	9	30
Leucas	5	-
Melochia	14	20
Mitracarpus	50	-
Undetermined	10	3
Physalis	17	21
Perilla	-	3
Sida	38	75
Solanum	-	45

Figure 1: Map of locations surveyed in Burkina Faso based on grid sampling strategy and diversity of collected plant samples. Each grid had dimensions of 500m X 1000m with meshes of 250m X 250m and covered a cultivated and uncultivated setting. Barplots correspond to the proportion of number of samples collected per plant genus per sampled grid node.

Begomovirus PCR detection

In order to compare the NGS virus detection and identification with a routinely used PCR diagnosis, PCR were performed on the 372 samples collected between 2013 and 2016 using the degenerate primer pair Clust4CP-F342 and Clust4CP-R1032 for the universal detection of the African begomoviruses as described in De Bruyn *et al* (2014). Amplicons were cleaned using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) before Sanger sequencing (Macrogen, South Korea).

Sample amplification for NGS

All the 1440 samples were treated with a metagenomic approach (independent sequence approach) devised for small circular DNA. First, a Rolling Circle Amplification (RCA, Inoue-Nagata *et al.*, 2004) of virus genomes was carried out using the Templiphi™ kit (Illustra GE Healthcare®). In a second step, a Random Amplification using Polymerase Chain Reaction (PCR-RA) was carried out on each of the RCA products. This PCR relied on the use of random primers each comprising on 5' extremity a single barcode of eight nucleotides, followed by 6 random nucleotides and the nucleotide motif TGGC (5'-BARCODE(8nt)-NNNNNN-TGGC-3'). A total of 160 distinct barcodes were defined according to Faircloth and Glenn (2012) with an edit distance of three. The PCR-RA screening of samples were carried out in 25 µL volumes with 5 µL of 5X buffer, 2.5 µL of dNTPs (2 mM), 1.5 µL of $MgCl_2$ (1.25mM), 10 µL of primers (4.8 µM) and 1 µL of GoTaq® Flexi DNA Polymerase (1.25 u). After an initial denaturation step at 95°C for 5 minutes, 35 cycles at 95°C for 1 minute, 50°C for 1 minute and 72°C for 1 minute were carried out before a final elongation of 5 minutes at 72°C. Importantly, each sample was treated twice with a distinct barcode. The PCR-RA products were diluted to 1/10 prior to a quantification using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) on a StepOnePlus Real-Time PCR system (Applied Biosystem). Based on the amplicons concentration, series of 160 amplicons were pooled with a maximum concentration ratio of 1.5 (*i.e.* no amplicons could present a concentration more than 1.5 higher than any another one in the same pool). Amplicon pools were then purified using the Illustra® GFX PCR DNA kit and gel purification (GE Healthcare®) according to the manufacturer's recommendations. Once purified, the

PCR-RA products were quantified using the Qubit dsDNA BR Assay Kit for Qubit fluorometer (Thermo Fisher Scientific) and check using the D5000 ScreenTape for 2200 TapeStation (Agilent Technologies). Pools were sent to GENEWIZ Next Generation Sequencing Services (USA). After Covaris shearing to fragment amplicons to the size range suitable for Illumina sequencing, libraries were constructed before sequencing on Illumina HiSeq 2500 and HiSeq X Ten with 2X150 bp paired-ends reads.

Metagenomic Data Analysis

Geminiviruses sequences were classified following the pipeline described in Claverie *et al.* (in preparation, Supplementary figure 1). Briefly, after quality control using Trimmomatic (Bolger *et al.*, 2014) sequences were demultiplexed using SABRE (https://github.com/najoshi/sabre) with a tolerance to one mutation in the barcode sequence. It is important to notice that due to the Covaris shearing step, not all the amplicons pairs present with an identifiable barcode on one or both reads. Pairs without any barcode were discarded from the analysis. After demultiplexing, sequence pairs were merged when possible using PANDAseq (Masella *et al.*, 2012) before dereplication using VSEARCH (Rognes *et al.*, 2016). Similarity search were performed on viral reference sequences obtained from Genbank on October 2017 using the “blastx“ algorithm of DIAMOND (Tang *et al.*, 2014). Another search was performed using the “usearch_global” algorithm of VSEARCH on a database restricted to the Geminiviruses and their satellites and obtained from Genbank on October 2017. Reads presenting with similarities to Geminivirus sequences were then clustered using SWARM v2 (Mahé *et al.*, 2015) with a distance parameter set to 3. Non-singleton clusters of sequences are then more precisely classified onto begomovirus, mastrevirus and satellites databases using pplacer (Matsen *et al.*, 2010) and the BoSSA R package (https://cran.r-project.org/package=BoSSA). It is important to notice here that reads were assigned without other assembly than simple pair merging. Whereas assembly would have results in larger contigs and probable better assignation, the presence of a TGGC clamp on the primer 3' side results in a partial viral genome coverage that prevents full component assembly in the vast majority of cases.

Cloning, sequencing and phylogenetic analysis of complete genome of geminiviruses

In order to reduce the cloning burden and to obtain global infection statistics, samples were clustered depending on their viral assignation profiles (*i.e.* virus taxonomic content contingency tables). These profiles (begomovirus DNA-A/A-like, begomovirus DNA-B, satellites and mastrevirus separetelly) were first converted in distance matrices (using the Morisita distance) before applying hierarchical clustering. Samples representative of each clusters were then subjected to RFLP-RCA procedure (Inoue-Nagata *et al.*, 2004) following by cloning and sequencing as described in Ouattara et al. (in preparation, see Chapter 1). Resulting sequences were edited and assembled using Geneious R8 (Biomatters Ltd, New Zealand) software. Nucleotide sequences were subjected to a BLAST search for preliminary species assignment. Alignment of the sequences along with representative geminivirus sequences available from GenBank were performed using the MUSCLE alignment tool (Edgar, 2004) available in MEGA 6 (Tamura *et al.*, 2013) and edited by eye. In order to assign sequences to virus species and strains, pairwise identity comparisons of nucleotide sequences were performed using SDT v1.2 with pairwise deletion of gaps (Muhire *et al.*, 2013). Maximum-likelihood (ML) phylogenetic trees were constructed using FastTree v2.1.7 (Price *et al.*, 2010) and edited using FigTree v1.3 (available at http://tree.bio.ed.ac.uk/software/figtree/).

Results

Sequence dependent approach (PCR) vs sequence independent approach (RCA-RA-SGS)

In order to compare the metagenomic approach to the more classical sequence dependent approach procedure, 372 samples collected in Burkina Faso from 2013 to 2016 were submitted to PCR using degenerated primers before amplicons sequencing and to the RCA-RA-SGS protocol. Based on PCR, 87/372 (23%) samples, comprising 43/98 (44%) cultivated and 44/274 (16%) uncultivated plants, were detected infected by a complex of six begomoviruses (Figure 2b). Detected begomovirus were CLCuGeV (4 cultivated and 26 uncultivated plants), PepYVMLV (34 cultivated and 14 uncultivated plants), *Tomato leaf curl Cameroon virus* (ToLCCMV, 2 uncultivated plants), ToLCGHV (2 cultivated plants), ToLCMLV (3 cultivated plants) and cases of mixed

infection with ToLCMLV and ToLCGHV (1 cultivated and 1 uncultivated plants) and with ToLCCMV and *Tomato leaf curl Nigeria virus* (ToLCNGV, 1 uncultivated plants).

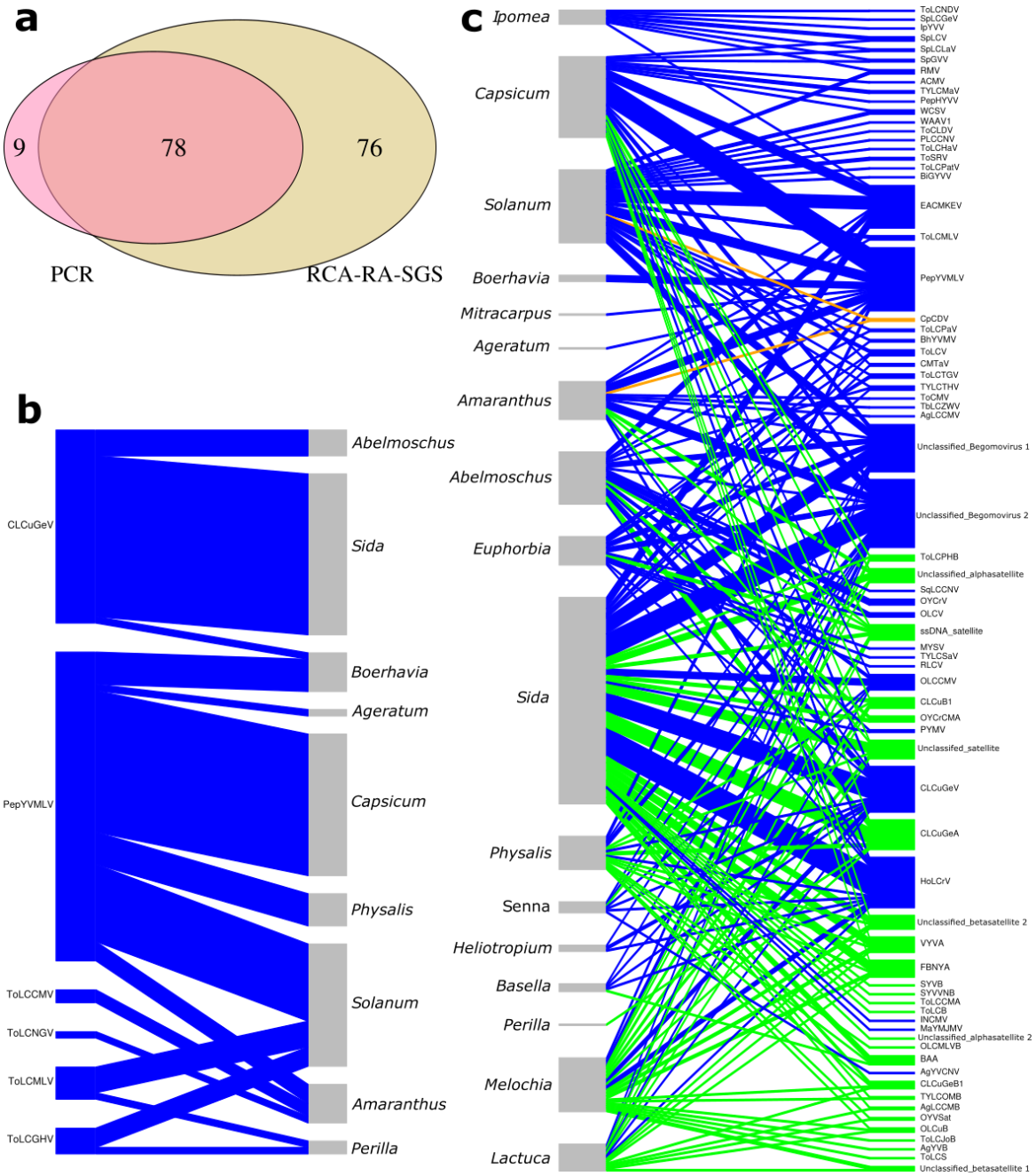


Figure 2: (a) Venn diagram illustrating the comparison of number of plants detected as host of geminivirus using classical PCR and RCA-RA-SGS method. Bipartite diagram showing a complex of begomoviruses detected from **(b)** the 87/372 positive plants by PCR and **(c)** the 154/372 positive plants by RCA-RA-SGS. Plant genera are represented on the left side of the diagram while viruses are represented on the right side of the diagram. Colored lines connect plant with the viruses detected with

blue for begomovirus DNA-A/A-like and DNA-B, green for satellites and orange for mastrevirus. Side bar height and line width are proportional to the number of infected plants.

Using the metagenomic procedure, geminiviruses were detected in 154/372 (41%) samples, comprising 58/98 (59%) cultivated and 96/274 (35%) uncultivated plants. Samples were considered as infected only if a minimum of 200 reads were assigned to one of the begomovirus DNA-A/A-like, begomovirus DNA-B, satellites or mastrevirus groups. Preliminary analysis of data confirms the presence of the begomovirus complex detected by PCR, and suggests the presence of other species of begomovirus and mastrevirus in single or mixed infections (Figure 2c). In addition, these preliminary data suggested the detection of alphasatellites and betasatellites.

Comparison of sequence dependent and independent approaches shows that 78/163 (48%) positive samples were detected with both approaches (Figure 2a). Nine (5%) positive samples were only detected by PCR while 76 (47%) were only detected by metagenomic method. Whereas the nine samples detected as positive by PCR were mostly from cultivated plants (6/9) and infected by PePYVMLV (8/9), the 76 samples detected as positive only with the metagenomic approach were plants from a larger set of genera (26 versus 5 genera) and positive to a larger set of viruses.

Detection of a wide diversity of geminivirus across the agro-ecological systems

The 1068 samples collected from Goué and Léguéma following a grid sampling strategy were processed using the RCA-RA-SGS procedure. Similarity search analysis results in 11% of the reads assigned to viruses. Among these, 65% were related to the *Geminiviridae* family. Interestingly, other plant infecting viral families were detected with *Caulimoviridae* (3%, 87 plants mostly from *Sida acuta*, *Ageratum conyzoides* and *Physalis ixocarpa* with more than 1000 reads assigned) and *Genomoviridae* (1%, 4 plants from two undetermined plants, one *Abelmoshus esculentus* and one *Sida rhombifolia* with more than 1000 reads assigned). Bacterial infecting viruses constituted a large part of the viral reads with *Myoviridae* (14%), *Siphoviridae* (8%), *Podoviridae* (2%) and *Phycodnaviridae* (1%). The remaining 6% of reads were shared with *Retroviridae* (2%), *Mimiviridae* (1%) and other families with low percentages.

Based on phylogenetic assignment and using a threshold of 200 reads of any of the begomovirus DNA-A/A-like, DNA-B, satellites or mastrevirus as a lower limit for detection, a total of 270 samples (25%) were detected as positive. A higher infection rate was obtained from samples collected in Goué (27%, 136/495) than in Léguéma (23%, 134/573). In both localities, begomoviruses (DNA-A/A-like), satellites and mastreviruses were detected (see Supplementary figure 2 and 3). Interestingly, one *Melochia* sample collected in Goué was also detected as positive for "unclassified geminivirus" and *Turncurtovirus*.

Identification of a large range of plant families, host of geminivirus

From Goué, 45% (57/125) and 21% (76/370) of cultivated and uncultivated plants were positive, respectively. The 57 positive cultivated plants belonged to two families comprising three genera, while the 76 positive uncultivated plants belonged to at least nine families comprising eleven genera (Table 1, Supplementary figure 2 and 4). From Léguéma, geminivirus-related reads were obtained from 51% (39/76) and 19% (95/497) of cultivated and uncultivated plants, respectively. Cultivated and uncultivated plants belonged to at least two families comprising two genera and six families comprising seven genera, respectively (Table 1, Supplementary figure 3 and 4). Notably, *Abelmoschus* (90%) and *Capsicum* (66%) from Goué and *Solanum* (77%) from Léguéma present with high infection rates. On both sampling grids, uncultivated plants from the genera *Ipomoea* and *Sida* were highly infected (30 to 89%).

Interestingly, having sampled a large set of plants from the same location, we were able to determine the effective host range of geminiviruses in natural conditions (Supplementary figure 2 and 3). Globally, it was apparent that satellites are mostly associated with the *Sida*, *Abelmoschus* and *Ageratum* genera. Few mastreviruses were detected and the majority was isolated from plant of the *Amaranthus* genus. A strong virus-host association was apparent from the association diagram, with most of the time, the majority of the viruses from a taxonomical group detected within a single plant genus. The specialization index was higher than expected by chance (p-value < 0.01).

Table 1: Percentage of infected plants collected from Goué and Léguéma after metagenomic virus detection

Grid location	Plant	Plant family	Plant genera	Positive rates
Goué	cultivated	*Malvaceae*	*Abelmoschus*	90% (27/30)
			Hibiscus	6% (3/54)
		Solanaceae	*Capsicum*	66% (27/41)
	uncultivated	*Amaranthaceae*	*Amaranthus*	20% (18/88)
		Asteraceae	*Ageratum*	33% (3/9)
		Boraginaceae	*Heliotropium*	10% (1/10)
		Convolvulaceae	*Ipomea*	89% (8/9)
		Euphorbiaceae	*Euphorbia*	22% (11/50)
		Lamiaceae	*Leucas*	40% (2/5)
		Malvaceae	*Melochia*	21% (3/14)
			Sida	61% (23/38)
		Nyctaginaceae	*Boerhavia*	7% (5/70)
		Rubiaceae	*Mitracarpus*	0% (0/50)
		Solanaceae	*Physalis*	12% (2/17)
		Undetermined	Undetermined1	0% (0/10)
Léguéma	cultivated	*Malvaceae*	*Abelmoschus*	14% (4/29)
		Solanaceae	*Capsicum*	0% (0/2)
			Solanum	78% (35/45)
	uncultivated	*Amaranthaceae*	*Amaranthus*	13% (15/116)
		Asteraceae	*Ageratum*	39% (24/62)
		Convolvulaceae	*Ipomea*	30% (9/30)
		Euphorbiaceae	*Euphorbia*	7% (8/108)
		Hydroleaceae	*Hydrolea*	0% (0/4)
		Lamiaceae	*Perilla*	0% (0/3)
		Malvaceae	*Abelmoschus*	0% (0/1)
			Melochia	20% (4/20)
			Sida	40% (30/75)
		Nyctaginaceae	*Boerhavia*	10% (5/49)
		Rubiaceae	*Mitracarpus*	0% (0/5)
		Solanaceae	*Physalis*	0% (0/21)
		Undetermined	Undetermined2	0% (0/3)

Viral clusters defined base on assignation profiles of samples

Assignation profiles (*i.e.* the putative geminivirus species composition) of each sample were clustered to define viral groups. On this basis, 16 begomovirus DNA-A/DNA-A like, six begomovirus DNA-B, six mastrevirus and 13 satellites clusters were defined. Among these clusters, varying degree of assignation confidence can be defined. For some clusters, the vast majority of the reads are assigned to a single species making these clear "assignation". Some other clusters present with assignation to multiple species, indicating either a co-infection of distinct viruses or the infection with viruses resembling to multiple known viruses (possibly due to recombination). Finally for some clusters, the assignation to "unknown" or "unclassified" geminiviruses, indicates the presence of a geminivirus without close relative in the database we used.

The 16 DNA-A/A-like component clusters were detected in 228 plants (Supplementary table 1). Four of these clusters were clearly associated to viruses already described in Burkina Faso with CLCuGeV (cluster A7) obtained from two cultivated plant genera (*Abelmoschus* and *Hibiscus*) and five genera of wild plants (*Amaranthus*, *Boerhavia*, *Euphorbia*, *Melochia* and *Sida*), PepYVMLV (cluster A3 and A4) was obtained from cultivated solanaceous (peppers and tomato) and six wild plant genera (*Ageratum*, *Amaranthus*, *Euphorbia*, *Ipomoea*, *Melochia* and *Physalis*) while ToLCMLV (cluster A13) was obtained from tomato and two uncultivated plant genera (*Ageratum* and *Amaranthus*). Among the other clusters, six had relatively clear assignation profiles and corresponds to new virus description in Burkina Faso, such as *Hollyhock leaf crumple virus* (HoLCrV) on *Amaranthus*, *Melochia* and *Sida* (cluster A2), *Tobacco leaf curl Zimbabwe virus* (TbLCZV) on *Amaranthus*, *Ageratum*, *Boerhavia Euphorbia* and *Melochia* (cluster A5), *Sweet potato leaf curl virus* (SPLCV) mainly on *Ipomoea* (cluster A14) and *Tomato leaf curl virus* on *Ageratum* (cluster A15). Some other clusters (n = 2) were defined by more complex profiles, possibly due to recombination. These clusters were mainly associated to CLCuGeV, HoLCrV and OYCrV (clusters A8 and A10). Finally four clusters (A1, A11, A12 and A16) have a large proportion of their reads, if not the majority presenting no clear relatedness with any of the already described begomovirus DNA-A/A-like components and most probably correspond to infections by new begomoviruses species.

The six clusters of begomovirus DNA-B were detected in 42 plants (Supplementary table 1). Cluster B1 obtained from cultivated solanaceous (peppers and tomato) and five wild plant genera (*Ageratum*, *Amaranthus*, *Boerhavia*, *Euphorbia*, and *Physalis*) was clearly associated to PepYVMLV DNA-B recently characterized in Burkina Faso (Ouattara *et al.*, in preparation, see Chapter 1). Cluster B5 obtained from okra and *Amaranthus* plants had relatively clear assignation profiles corresponding to a DNA-B of cassava mosaic viruses. Three clusters (cluster B3 to B5) had the majority of the reads unclassified and most probably correspond to infections by yet undescribed geminivirus-associated DNA-B. These clusters were obtained from tomato, and uncultivated plant (*Ageratum*, *Boerhavia* and *Sida*). The remaining cluster obtained on *Heliotropium* (cluster B6) had few reads assigned to *Kudzu mosaic virus* and *East African cassava mosaic Kenya virus*.

A total of 114 plants were found infected by 13 distinct satellites clusters (Supplementary table 1). Satellites were frequently detected without any other geminivirus (18%). Two of the 13 satellites clusters were associated to alphasatellites already described in Burkina Faso with CLCuGeA and OLCuMLB (clusters S7 and S9) obtained from cultivated plants (okra and *Hibiscus*) and from three genera of wild plants (*Amaranthus*, *Euphorbia*, and *Sida*). Two other clusters were associated to alphasatellites described in Cameroon. The first was associated with *Okra yellow crinkle Cameroon alphasatellite* (OYCrCMA, cluster S3) and obtained from cultivated plants (okra, pepper and *Hibiscus* sp.) and from four genera of wild plants (*Amaranthus*, *Boerhavia*, *Euphorbia* and *Sida*). The second was associated with *Ageratum leaf curl Cameroon alphasatellite* (cluster S2) and obtained from uncultivated plants belonging to *Ageratum* and *Amaranthus*. Among the other clusters, two had complex assignation profiles and may correspond to recombinant or co-infection. These clusters were mainly associated with CLCuGeA and OYCrCMA (clusters S5 and S8). Two clusters comprised assignation to satellites yet undescribed in West Africa with *Siegesbeckia yellow vein virus*-associated DNA beta (cluster S4) on *Capsicum* and the New World begomovirus associated satellite DNA (cluster S6) on okra and *Sida*. Finally, the remaining clusters (S1, S10, S11 and S12) had reads unclassified and most probably correspond to infections by new satellites or co-infection by new and known satellites.

Among the 17 plants detected as infected by mastrevirus, six clusters were defined (Supplementary table 1). Clusters M1, M2 and M6 were associated to *Chickpea chlorotic dwarf virus*, *Chickpea yellow dwarf virus* and related viruses. These clusters were described on okra, *Amaranthus*, *Ipomoea* and *Melochia*. Clusters M4 and M5 were surprisingly associated to *Maize streak virus* on tomato and *Sugarcane streak virus* on *Euphorbia* respectively. The remaining cluster (M3), described on okra and an uncharacterized plant was associated with a yet uncharacterized mastrevirus with some similarities to *Wheat dwarf virus*.

Characterization of known and unknown geminiviruses

In order to confirm the cluster viral composition and to later update the reference databases used for the assignations, we used the classical RCA-RFLP protocol to obtain clones of representative members of the distinct clusters. A total of eight DNA-A/A-like sequences comprising seven isolates of already described begomovirus and a probably new geminivirus were cloned (Supplementary figure 5a). The seven begomovirus isolates comprised five isolates of CLCuGeV (okra, n = 2; *Sida acuta*, n = 3) and two isolates of ToLCMLV (tomato, n = 1; *Sida acuta*, n = 1). Except ToLCMLV isolate obtained from a *Sida acuta* sample for which a yet unclassified virus was expected, all of these sequences confirmed the cluster composition inferred from the reads. Importantly, a new geminivirus genus/species sequence was obtained from a *Melochia* sp. sample. The genome of this new isolate was 2826 nucleotide long and possessed the typical genomic organization of a begomovirus with six ORFs (Figure 3a). Interestingly, whereas most of the genome shared high similarities to begomoviruses (73% and 74% with PepYVMLV and *Tomato curly stunt virus*, respectively; Figure 3b), the capsid ORF region displays its highest similarity with that of mastreviruses (33% to 34% nucleotide identities). SDT classification available within the Geminivirus Data Warehouse (Silva *et al.*, 2017) confirm that this sequence represents a new begomovirus species that we tentatively named *Melochia associated virus*.

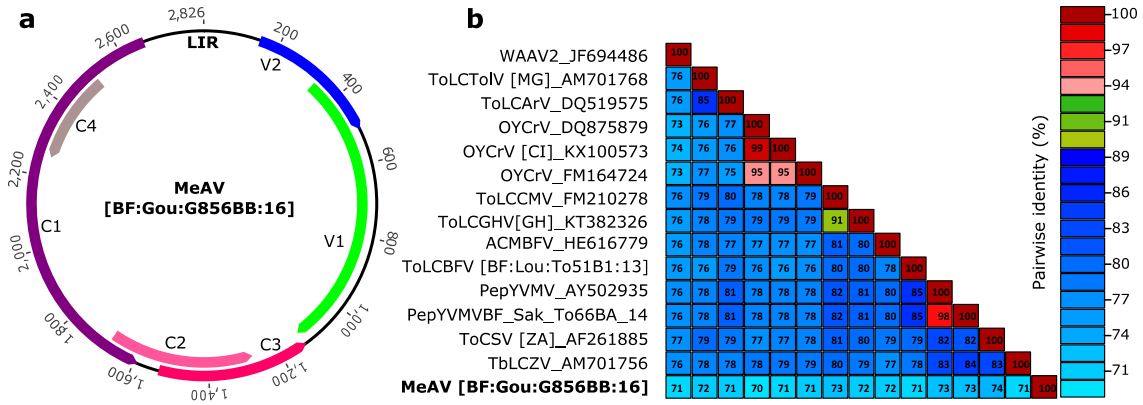


Figure 3: (a) Genome organization of MeAV comprising the six ORFs (V1, V2, C1, C2, C3, C4), the long intergenic region (LIR) including the origin of replication, the short intergenic region (SIR) and the common region (CR). **(b)** Colour-coded matrix of pairwise nucleotide identity inferred from alignments of complete DNA-A-like sequences of MeAV and selected Old World monopartite begomoviruses originating from Africa.

Nine complete sequences of satellites comprising seven alphasatellites (ranging from 1379 to 1468 nucleotides in length) and two betasatellites (1335 and 1343 nucleotides length) were obtained (Supplementary figure 5 b, c and d). Three of the seven alphasatellite sequences (89-96% identities between them) characterized from okra (n = 2) and an uncultivated plant were isolates of OLCuA. Another sequence shared 77-78% of nucleotide identities with OLCuA. The remaining sequences of alphasatellites (99-100% identities between them) shared 58-60% identities with CLCGeA. These two groups of sequences were all characterized from *Sida acuta* from which OLCuV and CLCuGeV were detected, respectively. According to the established alphasatellites species demarcation threshold of 83% (Mubin *et al.*, 2009), these alphasatellites could be classified as two distinct new species provisionally named *Okra leaf curl alphasatellite 2* (OLCuA2) and *Cotton leaf curl Gezira alphasatellite 2* (CLCuGeA2), respectively (Supplementary figure 5 b and c). Most of the characterized alphasatellites were congruent with the cluster assignation analysis, previously obtained from the reads. Both betasatellites sequences (97% identities between them) characterized from okra, were isolates of CLCuGeB (97-98% of similarity) (Supplementary figure 5 d). None were detected with the metagenomic approach.

Four partial satellite sequences (ranging between 913 and 1079) were also obtained from *Sida acuta*. One of them seems to be an isolate of OLCuA (87 % of similarity); two of them would be isolates of CLCuGeA (96-97% nucleotides identities) and the remaining one seems to be a new betasatellite related (75% nucleotide identity) to *Cotton leaf curl betasatellite* 1 (CLCuGeB1) already characterized in Burkina Faso, again from *Sida acuta* (Ouattara *et al.*, in preparation, chapter 1). Except the putative new betasatellite obtained from a *Sida acuta* sample for which CLCuGeV was detected, all of these partial sequences confirmed the cluster assignation obtained from reads

Networks of host-virus association

Based on cluster memberships, viral networks were constructed for each sampling grid. In these networks, each node represents a single geminivirus infected plant and each link represents the connection between samples sharing a similar cluster of virus (Figure 4 a and b). Due to the cluster nature of the data (samples are already grouped together depending on viral assignation), the global network appears as extremely modular. Thus, rather than uncovering co-infections with similar components (*e.g.* DNA-A with DNA-A), the examination of the network structure would help identifying possible hubs of begomovirus global diversity and the type of plant associated with.

The network from Goué was the more complex between the two sampling grids. It showed complex begomovirus-host, satellite-host and begomovirus DNA-B-host associations (Figure 4a). Beside some isolated groups and singletons, a large network interconnects most of the infected plants (124/136). Most of the clusters display low plant genus diversity but are connected to several other distinct clusters. As an exemple, cluster A7 (CLCuGeV, mostly on okra) is connected to seven other clusters (six satellites cluster and one begomovirus DNA-B cluster). As a notable exception to the other begomovirus DNA-B clusters, cluster B1 was connected to three distinct DNA-A, three satellites and one mastrevirus clusters. This cluster, associated to the newly described PepYVMLV DNA-B component was nevertheless mostly associated with plant from the *Capsicum* genus. Interestingly, three of the four mastreviruses clusters, detected in Goué, were found in co-infection with other geminiviruses in some plants.

Globally, plants from the same genus tend to be more connected to each other than expected by chance (Chi-square p-value < 2.2×10^{-16}).

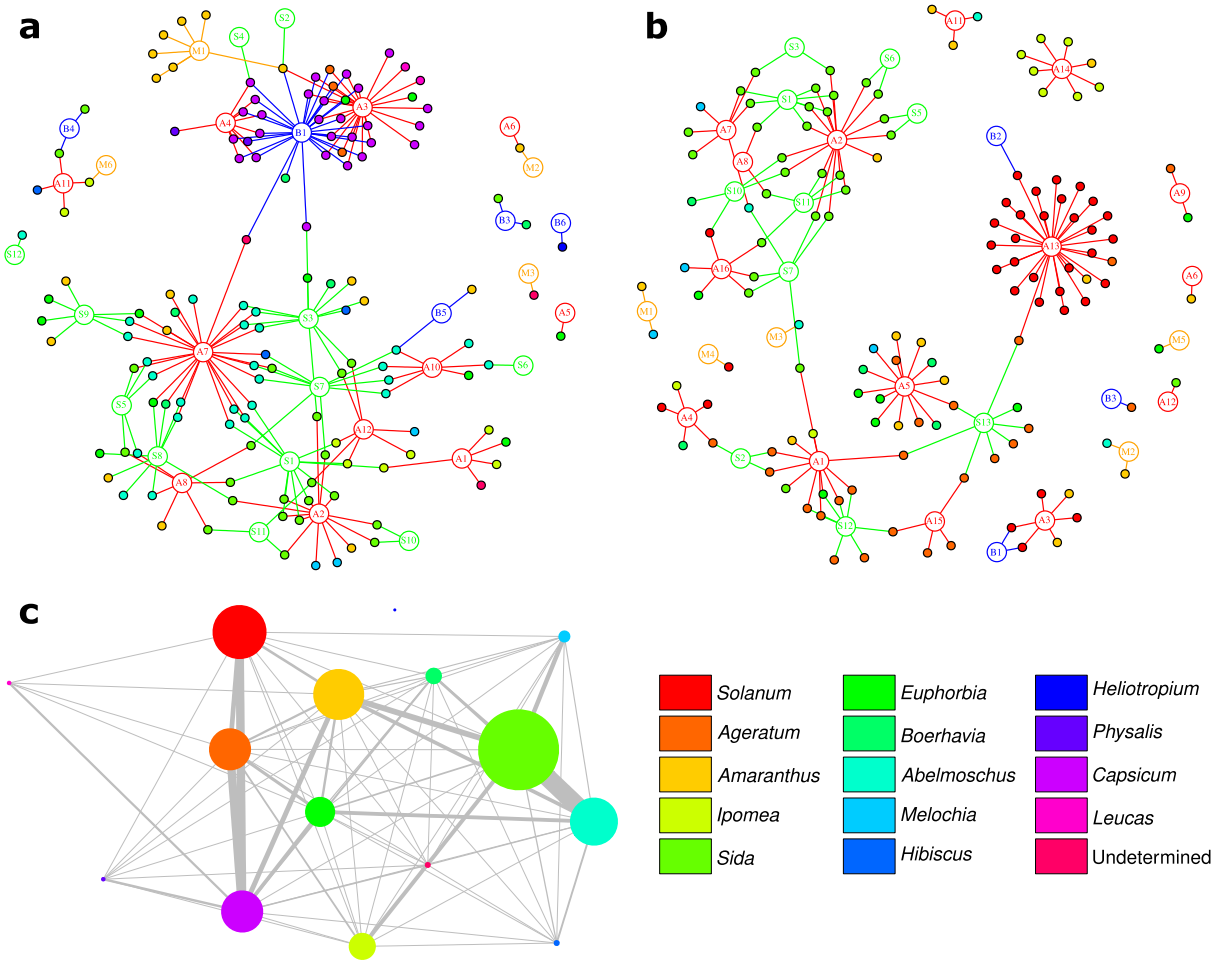


Figure 4: Network showing geminivirus community structure in Goué **(a)** and Léguéma **(b)** in Burkina Faso. Numbered nodes represent virus clusters. Letters A, B, S and M indicate cluster number for begomovirus DNA-A/A-like (red), begomovirus DNA-B (blue), satellites (green) and mastrevirus (yellow) respectively. Smaller nodes represent host plants and are colored according to their genus. **(c)** Interconnection between plants sharing similar geminiviruses. The size of dots is proportional to the number of infected plants and the line width is proportional to the number of shared virus clusters and the dots are colored according to their genus.

In Léguéma, a network presenting fewer connections was observed (Figure 4b). The main network connects 105 of the 134 plants together. Individuals sharing similar begomovirus constituted twelve isolated sub-clusters of small sizes. The most reticulate

structure involved a begomovirus DNA-A/A-like cluster (A2, *Hollyhock leaf crumple virus* mostly on *Sida*) that connects to seven distinct satellite clusters. Very few begomovirus DNA-B connection were present and the B clusters were always only associated to a single other virus. Mastrevirus clusters were isolated and never detected in more than two plants. Again, clusters and connected clusters tends to be composed of more similar plant genus than expect by chance (Chi-square p-value < 2.2×10^{-16}).

Taken together, the inspection of both networks showed that most of the begomovirus DNA-A and satellite clusters were interconnected to each other forming large structure associated with both cultivated and uncultivated plants. The analyses of the interconnection between plant genus (Figure 4c) highlights the interconnection of some of the cultivated and uncultivated plants with a strong structure between *Capsicum*, *Solanum* and *Ageratum* on one side and *Sida* and *Abelmoschus* on the other side. These two modules present very weak direct connections between each other but were strongly inter-connected through other uncultivated plant genera such as *Amaranthus* and *Euphorbia*.

Discussion

Need of combining both dependent and independent sequence approaches for virus detection

In order to detect and identify the geminivirus populations associated with plant samples collected in Burkina Faso without *a priori* on their nucleotide sequences, a targeted metagenomic (RCA-RA-SGS) approach-based on the amplification of viral nucleic acids was applied. This approach was compared to a procedure that includes PCR detection using degenerated primers designed to amplify African begomoviruses along with Sanger sequencing. Interestingly, our results showed that sequence independent approach permitted the detection of larger diversity of geminivirus than sequence dependent approach. This highlights the benefits of metagenomics approach for the study of the diversity of plant viruses. Whereas 47% of the positive samples were detected with the metagenomic approach, only 5% were detected by sequence

dependent approach. These nine samples were mainly *Solanum* and infected by PepYVMLV. Similar results were obtained when RT-PCR were compared to VANA-based metagenomic approach (Palanga *et al.*, 2016). Hypothesis of too low concentration of viral nucleic acid in samples was formulated (Palanga *et al.*, 2016). Whereas a less conservative criterion than the 200 reads threshold to consider a sample as positive could be used, it would also increase the false positive rate. Importantly, these results highlight that PCR detection remains a method of choice if one is interested in a single or few well characterized viral species.

Dynamics of geminivirus at the scale of agro-ecological systems

Studies suggested that wild plants may harbor a wide diversity of viruses, some of which are capable of moving into agricultural crops (Alexander *et al.*, 2014). Our results showed higher infection rates for cultivated plants (47%) than weeds (20%). Similar results have been reported in France (27% and 15% of cultivated and uncultivated plants, respectively) and South Africa (19% and 8% of cultivated and uncultivated plants, respectively) in an earlier study (Bernardo *et al.*, 2017). In Bernardo *et al.* (2017), cultivated and uncultivated settings were spatially separated from each other. The differences in infection rates appeared to be partly related to the decrease of plants diversity in cultivated areas due to human activities. In fact, density variations induce changes in the number of hosts available in a given space and at a given time and the distance between potential hosts (Gilbert, 2002; Keesing *et al.*, 2006, 2010). The decrease of plants diversity in cultivated areas leads to the increase of the number of cultivated plants (potential host) which may facilitates the feeding of insect vector and then the transmission of viruses. In our sampling grids, where uncultivated and cultivated plants are more spatially mixed (see Figure 1 and supplementary figure 4), the higher infection for cultivated plants may also be explained by a higher plant susceptibility due to a possible lesser adaptation to their infecting viruses.

Large diversity of geminiviruses in Burkina Faso

In Burkina Faso, previous studies have reported the presence of at least six geminivirus including five begomovirus with CLCuGeV (Tiendrébéogo *et al.*, 2010a), PepYVMLV (Tiendrébéogo *et al.*, 2011), ToLCBFV (Ouattara *et al.*, 2017a), ToLCMLV and

TYLCMLV (Sattar *et al.*, 2015) and a mastrevirus with CpCDV (Ouattara *et al.*, 2017b). Here, metagenomic data suggested, based on read assignment, the presence of more large diversity of these virus in this country with eleven identified begomovirus species, a flurry of satellites and mastrevirus species. The actual number of species would require further cloning, sequencing and sequence comparison to be properly established. Nevertheless, as we were able to clone the detected DNA-A/A like components from some clusters, we are confident in the estimation obtained from the metagenomic study. In fact, based on the assignation profiles, we expect at least three additional new geminivirus species to be uncovered from our samples. Conversely, the estimation of the number of satellites circulating in Burkina Faso must be taken with caution. The database used for sequences assignment seems far from exhaustive. Many clusters present with assignation to more distant relative satellites and only cloning and sequencing will help uncover the actual diversity of satellites. It must also be noticed here that we were not able to detect any helper virus for some satellites. It is possible that competition during the amplification steps cause the absence of detection. Smaller circular matrices, such as satellites compared to DNA-A component, may be more efficiently amplified by RCA and end up being largely preponderant in the amplification product.

Interestingly, we were able to detect the PepYVMV DNA-B component, previously described in Burkina Faso (Ouattara *et al.*, in preparation, chapters 1 and 2), in 33 samples. It was detected in association with PepYVMV DNA-A for 30 samples, with CLCuGV for one sample and without any DNA-A component for two samples. Again, the absence of DNA-A detection for the remaining two samples could be associated to template competition during the RCA and the RA amplification steps. A total of six distinct mastrevirus groups were detected in the samples. The detection of sequences related to CpCDV on *Amaranthus*, *Abelmoschus* and *Euphorbia* plants confirms its presence in Burkina Faso, as described previously (Ouattara *et al.*, 2017b). The other viral detection should be confirmed with cloning of full genomes.

Wide host range of geminivirus

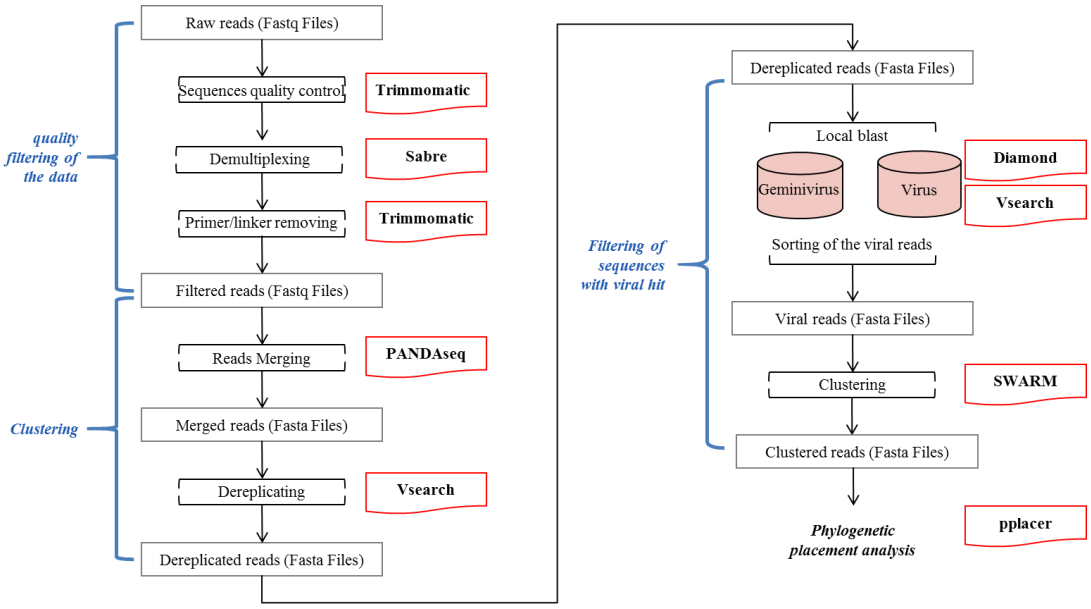
Historically, plant virus epidemiology has mainly focused on virus spread within crop fields and effects on crop yield (Alexander *et al.*, 2014). For complete understanding of the epidemiology and for developing appropriate control strategy, the identification of alternative hosts is an aspect of the epidemiology that has been somewhat neglected in the past (Leke *et al.*, 2015). Here, we have identified eleven genera of uncultivated plants as host of geminiviruses. More importantly, ten of these plants genus were potential alternative hosts to crops infecting geminiviruses. The genera *Sida* and *Ageratum* were particularly connected to cultivated crops with specificity to viruses also infecting *Capsicum* and *Solanum* for *Ageratum* and *Abelmoschus* for *Sida*. Interestingly, these plant modules were connected by *Euphorbia* and *Amaranthus*. Whether these later genera could act as hubs for virus exchange between *Abelmoschus* on one side and *Capsicum* and *Solanum* on the other side has to be investigated. Future analysis of gene flow through recombination along this network should help to answer this essential question.

Concluding remarks

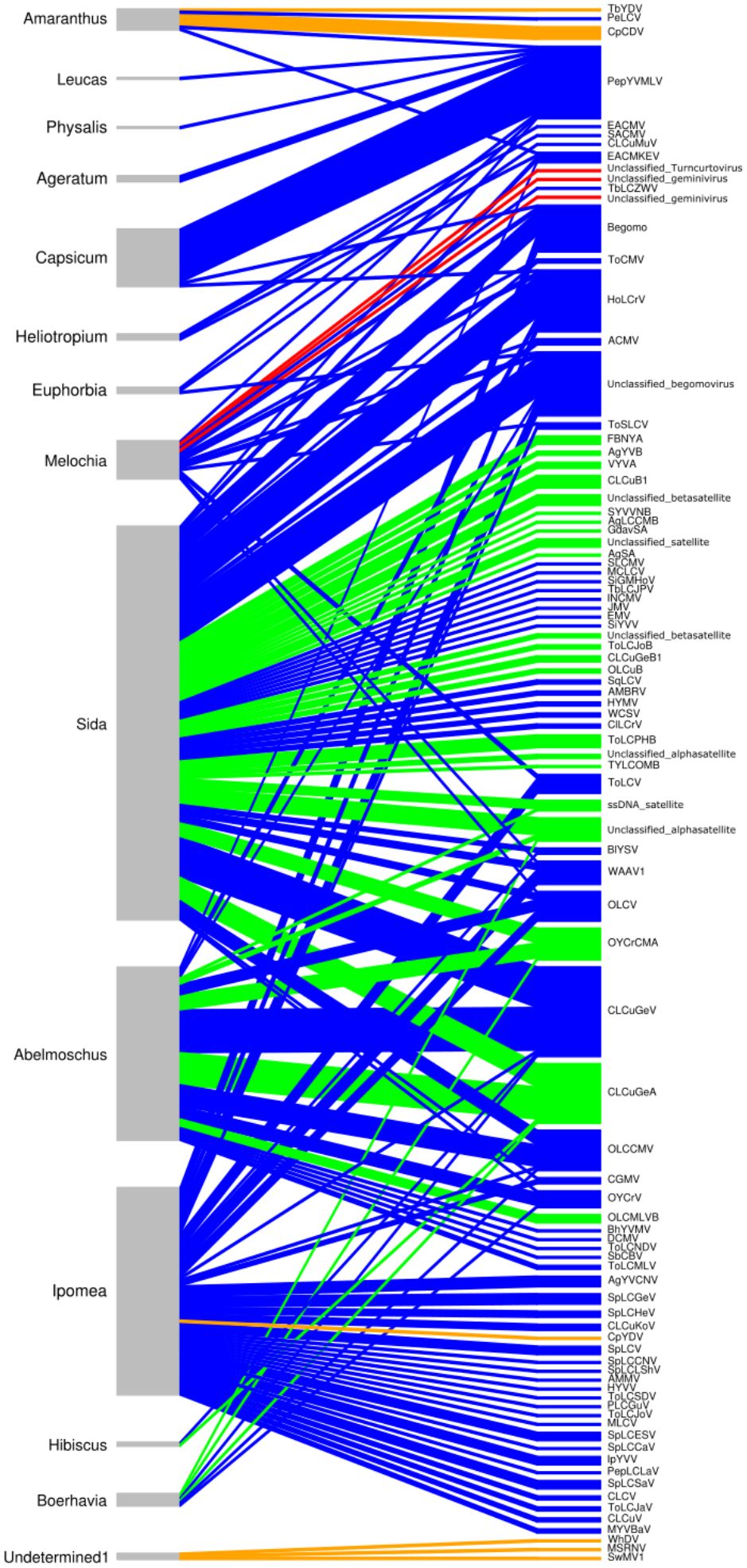
Our findings highlight the benefits of the metagenomic approach to study diversity of plant viruses. They also highlight the extent and abundance of plant-infecting geminiviruses at the scale of agro-ecosystems. In our study and previous work, crops were more infected than wild plants. Inversely, a largest diversity of geminivirus was found in wild plants than crops. Eleven genera of uncultivated plants were identified as host of geminiviruses. More importantly, ten of these plant genera were potential alternative hosts to crops infecting geminiviruses. The analysis of the virus sharing map allowed establishing viral sharing networks between host plants. Besides confirming the high prevalence of geminivirus as the whole agro-system scale, our targeted metagenomic-based approach has provided a first complete description of the geminiviral community structure. A more refined characterization of the virus groups of this community would help understand the functioning of natural viral community and how the whole community participates to the emergence of new viral variants.

Acknowledgements

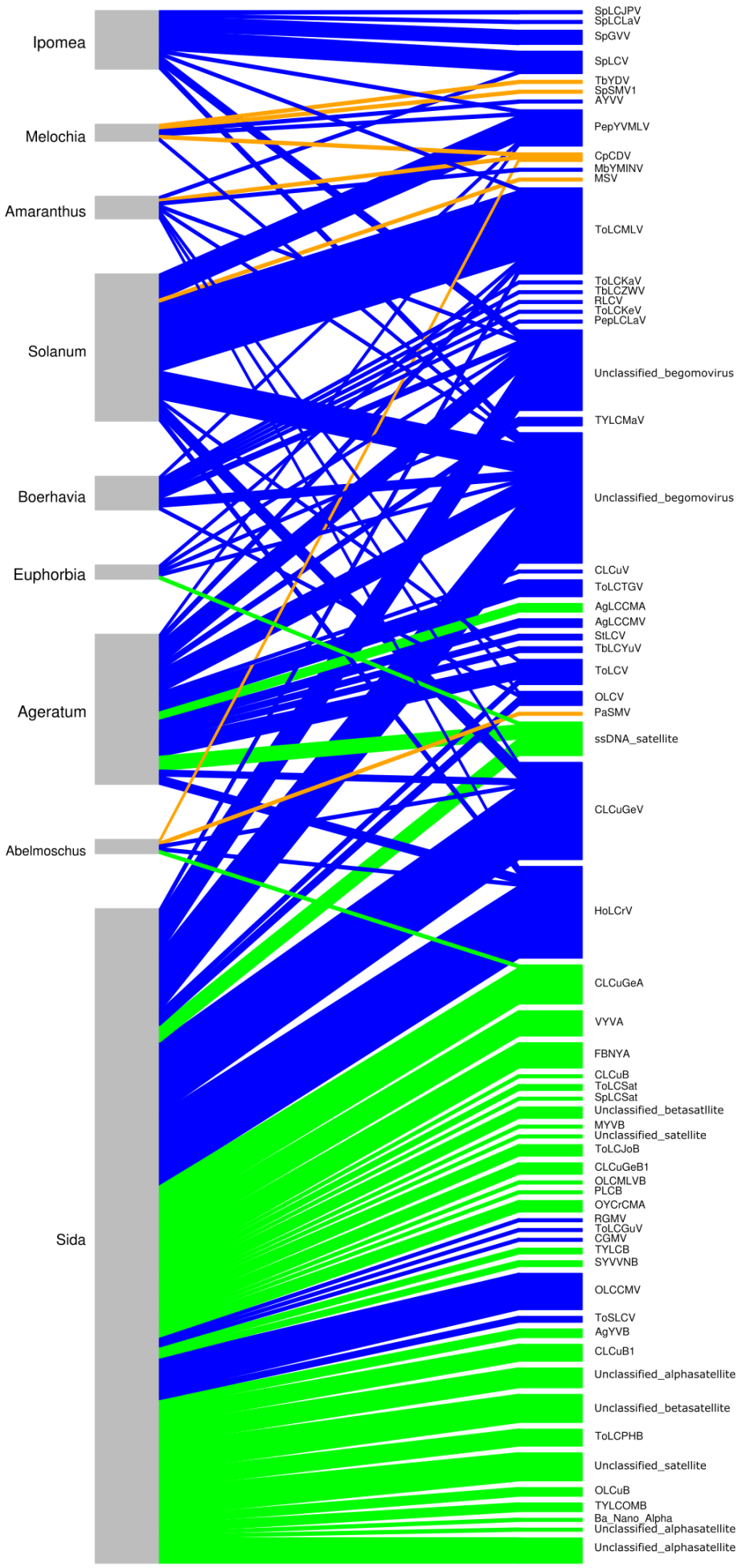
This work was supported by the European Union (ERDF), the Conseil Régional de La Réunion and CIRAD. AO was a recipient of a PhD fellowship from CIRAD.



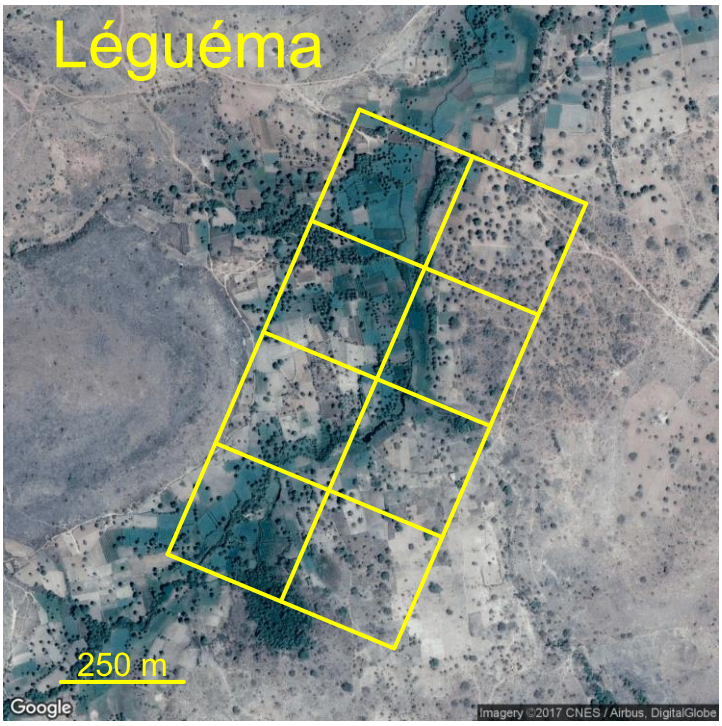
Supplementary figure 1: Metagenomic data analysis pipeline.

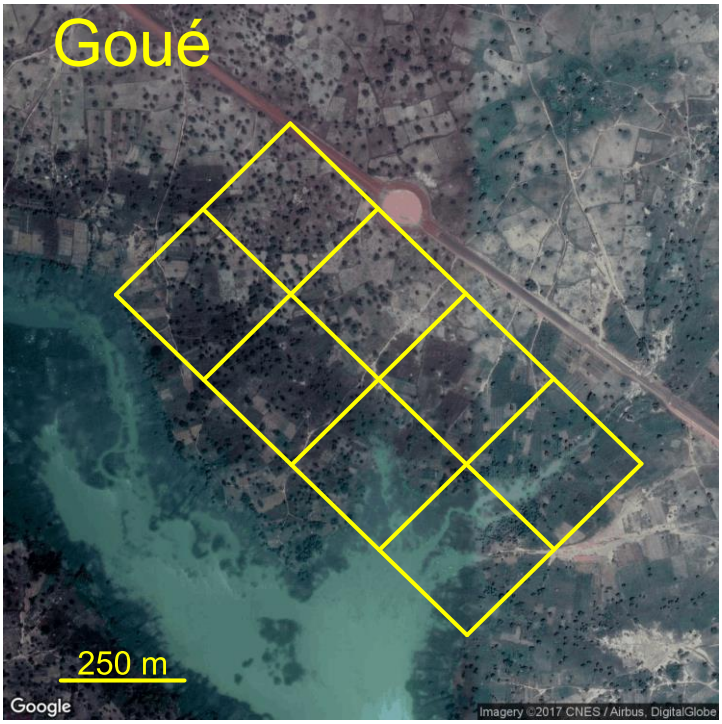

Amaranthus
Leucas
Physalis
Ageratum
Capsicum
Heliotropium
Euphorbia
Melochia
Sida
Abelmoschus
Ipomea
Hibiscus
Boerhavia
Undetermined1
TbYDV
PeLCV
CpCDV
PepYVMLV
EACMV
SACMV
CLCuMuV
EACMKEV
Unclassified_Turncurtovirus
Unclassified_geminivirus
TbLCZWV
Unclassified_geminivirus
Begomo
ToCMV
HoLCrV
ACMV
Unclassified_begomovirus
ToSLCV
FBNYA
AgYVB
VYVA
CLCuB1
Unclassified_betasatellite
SYVVNB
AgLCCMB
GdavSA
Unclassified_satellite
AgSA
SiCMV
MCLCV
SiGMHoV
TbLCJPV
INCMV
JMV
EMV
SiYVV
Unclassified_betasatellite
ToLCJoB
CLCuGeB1
OLCuB
SqLCV
AMBRV
HYMV
WCSV
CILCrV
ToLCPHB
Unclassified_alphasatellite
TYLCOMB
ToLCV
ssDNA_satellite
Unclassified_alphasatellite
BIYSV
WAAV1
OLCV
OYCrCMA
CLCuGeV
CLCuGeA
OLCCMV
CGMV
OYCrV
OLCMLVB
BhYVMV
DCMV
ToLCNDV
SbCBV
ToLCMLV
AgYVCNV
SpLCGeV
SpLCHeV
CLCuKoV
CpYDV
SpLCV
SpLCCNV
SpLCLShV
AMMV
HYVV
ToLCSDV
PLCGuV
ToLCJoV
MLCV
SpLCESV
SpLCCaV
IpYVV
PepLCLaV
SpLCSaV
CLCV
ToLCJaV
CLCuV
MYVBaV
WhDV
MSRNV
SwMV1

Supplementary figure 2: Bipartite graph showing the diversity of geminivirus from Goué and the corresponding plant genera from which viral hits were obtained. Plant genera are represented on the left side of the diagram while viruses are represented on the right side of the diagram. Colored lines connect plant with the viruses detected with blue for begomovirus DNA-A/A-like and DNA-B, green for satellites, orange for mastrevirus and red for other geminiviruses. Side bar height and line width are proportional to the number of plant infected.


Ipomea
Melochia
Amaranthus
Solanum
Boerhavia
Euphorbia
Ageratum
Abelmoschus
Sida
SpLCJPV
SpLCLaV
SpGVV
SpLCV
TbYDV
SpSMV1
AYVV
PepYVMLV
CpCDV
MbYMINV
MSV
ToLCMLV
ToLCKaV
TbLCZWV
RLCV
ToLCKeV
PepLCLaV
Unclassified_begomovirus
TYLCMaV
Unclassified_begomovirus
CLCuV
ToLCTGV
AgLCCMA
AgLCCMV
StLCV
TbLCYuV
ToLCV
OLCV
PaSMV
ssDNA_satellite
CLCuGeV
HoLCrV
CLCuGeA
VYVA
FBNYA
CLCuB
ToLCSat
SpLCSat
Unclassified_betasatllite
MYVB
Unclassified_satellite
ToLCJoB
CLCuGeB1
OLCMLVB
PLCB
OYCrCMA
RGMV
ToLCGuV
CGMV
TYLCB
SYVVNB
OLCCMV
ToSLCV
AgYVB
CLCuB1
Unclassified_alphasatellite
Unclassified_betasatellite
ToLCPHB
Unclassified_satellite
OLCuB
TYLCOMB
Ba_Nano_Alpha
Unclassified_alphasatellite
Unclassified_alphasatellite

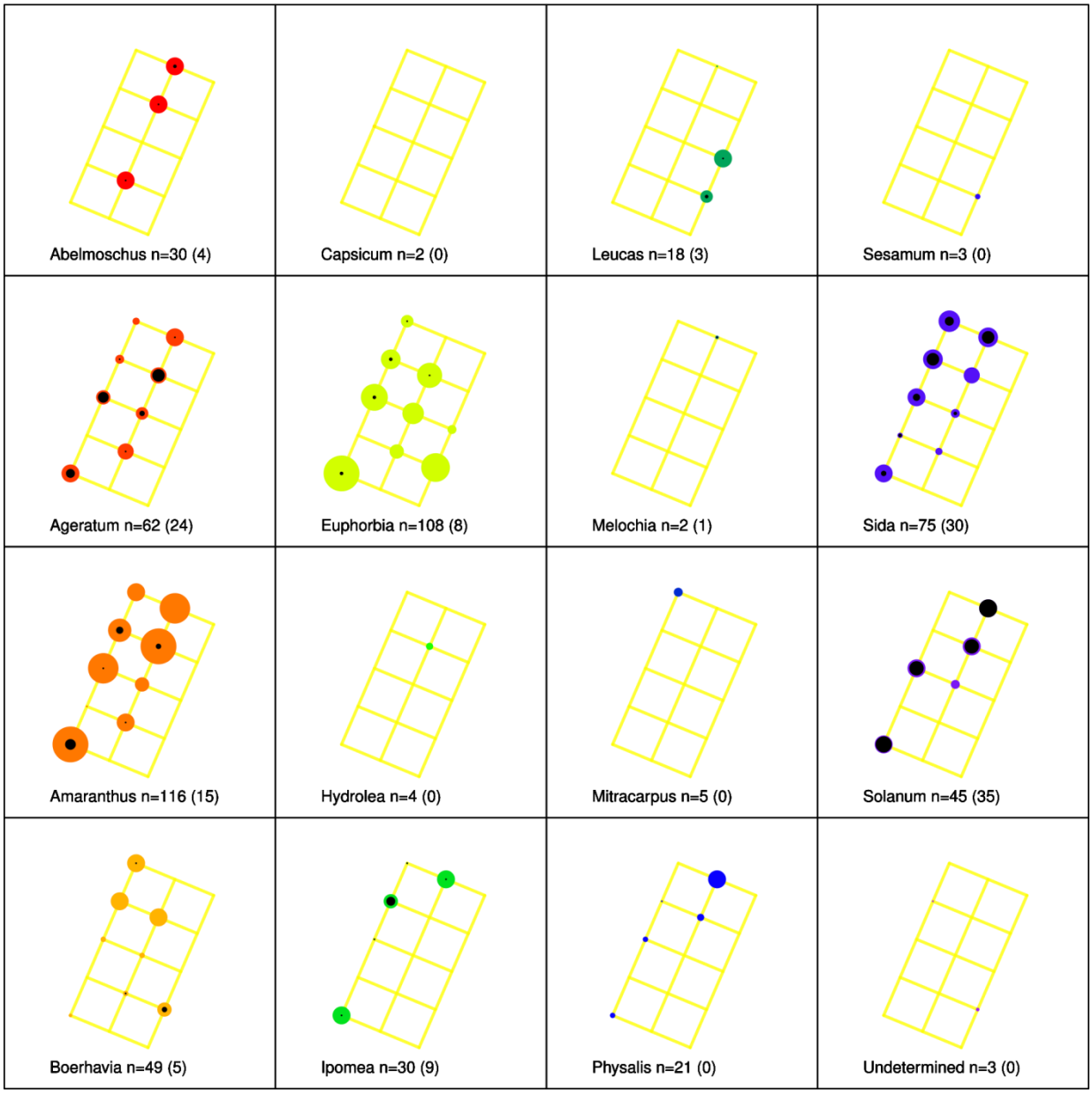
Supplementary figure 3: Bipartite graph showing the diversity of geminivirus from Léguéma and the corresponding plant genera from which viral hits were obtained. Plant genera are represented on the left side of the diagram while viruses are represented on the right side of the diagram. Colored lines connect plant with the viruses detected with blue for begomovirus DNA-A/A-like and DNA-B, green for satellites and orange for mastrevirus. Side bar height and line width are proportional to the number of plant infected.





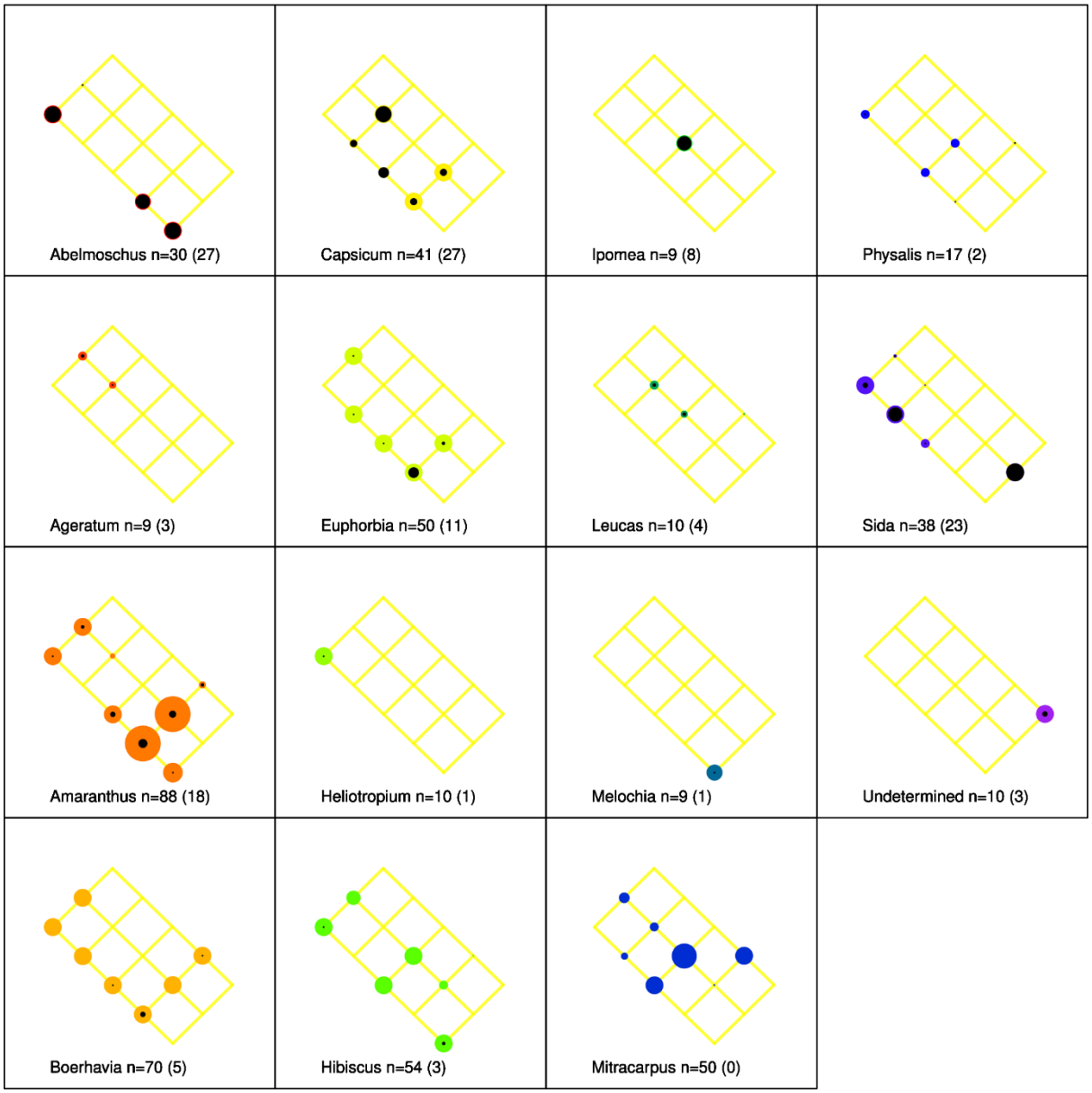
Supplementary figure 4 (continued below)

Léguéma



Supplementary figure 4 (continued below)

Goué



Supplementary figure 4: Geographical distribution of distinct genera of plants collected from Goué and Léguéma sampling grids. Dots are colored according to plant genus. The black dots sizes indicate the proportion of positive plants according to each genus.

a
ToLCToIV-[MG:Men:Ly]_AM701768
ToLCArV_DQ519575
ToLCArV [AFTT23]_EF194760
ToLCSV [ZA]_AF261885
MeLCV [BF:Gou:MeG856:16]
WAAV2_JF694486
TYLCSV_AY02650
ToLCSDVs
TYLCVs/ TYLCMLVs
CMGs
EAToLCVs
ToLCBFV-[BF:Lou:To51B1:13]
PepYVMLVs
WAAV1_JF694484
CLCuGV-Si[BF:Krs:Si1002BE:15]
CLCuGV-Si[BF:Leg:SiG192BF:16]
CLCuGV-Si[BF:Leg:SiG299BI:16]
CLCuGV-Si[BF:Dje:Si690BE:15]
CLCuGV-Si[BF:Sak:Si70PD:14]
CLCuGV-Si[BF:Bnk:Si1132BE:15]
CLCuGV-Si[BF:Bnk:Si1135BA:15]
CLCuGV-Si[BF:Kay:Si29P:14]
CLCuGV-Si[BF:Si965BB:16]
CLCuGV-NE[BF:Gou:OkG1163B12j:16]
CLCuGV-NE[BF:Bsr:Am1124B7g:16]
CLCuGV-NE[BF:Kam:Ok:08]_FN554523
CLCuGV-NE[BF:Kam:Ok94PC:14]
CLCuGV-NE[BF:Po:Ok:08]_FN554538
CLCuGV-NE[BF:Kon:Ok99B5:13]
CLCuGV-NE[BF:Kam:Ok100B2:13]
CLCuGV-NE[BF:Kam:Ok103B1:13]
CLCuGV-NE[BF:Kon:Ok87B1:13]
CLCuGV-NE[BF:Kam:Ok:08]_FN554521
CLCuGV-NE[BF:Leg:SiG301P10b:16]
CLCuGV-NE[BF:Gou:OkG1158B10g:16]
CLCuGV-NE[BF:Lou:Ok100PD:14]
CLCuGV-NE[CI:Yam:CI82:13]_KX100570
CLCuGV-NE[BF:Kam:Ok:08]_FN554527
CLCuGV-NE[BF:Kam:Ok:08]_FN554529
CLCuGV-NE[NE:Sad:Ok:07]_EU432374
CLCuGV-NE[NE:Sad:OkNG2FL:07]_FJ469627
CLCuGV-NE[NE:Sad:Ok:07]_EU432373
CLCuGV-NE[NE:Sad:OkNG1FL:07]_FJ469626
CLCuGV-NE[ML:Bam]_EU024120
CLCuGV-NE[BF:Lou:Ok57B3:13]
CLCuGV-NE[BF:Kam:Ok:08]_FN554526
CLCuGV-SD
CLCuGV-CM
CLCuGV-EG
HoLCrVs
OLCuCMV-[CM:Lik:07]_HE793426
OLCuCMV
OYCrV
ToLCMLV-[ML]_AY502936
ToLCMLV-BF[BF:Dje:To656BA:15]
ToLCMLV-BF[BF:Dje:To671BA:15]
ToLCMLV-BF[BF:Dje:To675BA:15]
ToLCMLV-BF[BF:Dje:To685BA:15]
ToLCMLV-BF[BF:Dje:To672BE:15]
ToLCMLV-BF[BF:Dje:To668BE:15]
ToLCMLV-BF[BF:Dje:To665BA:15]
ToLCMLV-BF[BF:Dje:To679BB:15]
ToLCMLV-BF[BF:Dje:To680BB:15]
ToLCMLV-BF[BF:Krs:To940BC:15]
ToLCMLV-BF[BF:Dje:To683BD:15]
ToLCMLV-BF[BF:Dje:To684BC:15]
ToLCMLV-BF[BF:Leg:To579E7c:16]
ToLCMLV-BF[BF:Sak:To81B1:13]
ToLCMLV-BF[BF:Tou:To529B14c:16]
ToLCMLV-BF[BF:Leg:ToG5B14c:16]
ToLCMLV-BF[BF:Leg:To590BA:16]
ToLCMLV-BF[BF:Gou:Si868B15c:16]
ToLCMLV-BF[BF:Leg:To574BA:16]
ToLCMLV-BF[BF:Krs:To960BD:15]
ToLCMLV-BF[BF:Sak:To83B4:13]
ToLCMLV-BF[BF:Tom-145]_LM651404
ToLCGHVs
ToLCCMVs/ALCCMVs
0.06
b
LeYCrA/PaLCA
FbNSA/MVDA
OELCA
HoLYVA [PK]_FR772086
MeYVMA
OLCuA [OBKG]_FM164739
CLCuGeA [BF:Kon:Ok97B1:13]
CLCuGEA [BF:Gue:Ok1119B3a:16]
CLCuGeA [BF]_FN554582
CLCuGeA [BF]_FN554583
CLCuGeA [BF]_FN554581
CLCuGeA [IsSq-C171]_KT099176
CLCuGeA [SD]_KC763632
CLCuGeA [SD]_KC763634
CLCuGeA [SD]_KC763633
CLCuGeA [SD]_KC763631
OYCrCMA [CM]_FN675288
OYCrCMA [Cm]_FN675287
CLuGeA [BF:Gou:SiG593P10a:16]
OYCrCMA
CLCuGeA[BF:Gou:SiG861PC2a:16]
CLCuGeA [BF:Kam:Ok100B1:13]
GMSA/GDSA
FbNYA
DrAA_NC_019498
DrAA_JX458742
Alphasatellite_2_KT099172
VYVA
OLCuA_FM164740
CLCuGeA2 [BF:Gou:UnG1199BL:16]
CLCuGeA2 [BF:Leg:SiG301BI:16]
CLCuGeA2 [BF: Gou:SiG1201BI:16]
DLCA_KC206077
0.3
c
ToLCCMA/ALCCMA
OLCuA_FJ868830
OLCuA_NC_010620
OLCuA_EU589450
OLCuA [BF:Lou:Ok100bBB:14]
OLCuA [BF:Lou:Ok57B1:13]
OLCuA [BF:OkG1126B07a:16]
OLCuA [BF:Am1123B5b:16]
CLCuGeA [KSA27]_HG530546
CLCuGeA [KSA27]_HG530545
CLCuGeA [KSA27]_HG530544
CLCuGeA [KSA27]_HG530547
CLCuGeA [SH7]_HM446369
CLCuGeA_NC_013593
CLCuGeA [BF]_FN554580
CLCuGeA [Njo5]_HE858192
OLCuMLA [CM]_FN675290
OLCuMLA [CM]_FN675289
OLCuA[BF:Gou:OkG727P6c:16]
OLCuA[BF:Gou:Un983B19e:16]
OLCuA [BF:Sos:Ok87EF:14]
OLCuA [BF:Sos:Ok86PG:14]
OLCuA [BF:Leg:Un:G143P14g:16]
OLCuA2 [BF:Leg:SiG192BJ:16]
CMMGA
MaYMCMA
0.3
d
ChYMBJB
ALCCMB
ToLCTGB
CLCuGeB2 [BF:SiG242B14h:16]
CLCuGeB1 [BF:Kay:Si28BB:14]
CLCuGeB1 [BF:Kay:Si30BA:14]
CLCuGeB1 [BF:Kay:Si29BA:14]
OLCuB_FJ187796
CLCuGeB [IsSq1]_KU095847
CLCuGeB [IsSqC112]_KT099177
CLCuGeB
OLCuMLB
CLCuGeB
OLCuMLB [BF:Sos:Ok86PE:14]
OLCuMLB
OLCuMLB [BF]_FN554578
CLCuGeB [BF:Kon:Ok87P1:13]
CLCuGeB [BF:Lou:Ok97PF:14]
CLCuGeB_EU727088
CLCuGeB_FJ469629
CLCuGeB [BF:Sos:Ok87PA:14]
CLCuGeB [BF]_FN554577
CLCuGeB [BF]_FN554573
CLCuGeB_EU727087
CLCuGeB_FJ469628
CLCuGeB [BF:OkG117P7d:16]
CLCuGeB [BF:OkG727P6a:16]
CLCuGeB [BF:Kon:Ok97P4:13]
CLCuGeB_KM279620
CLCuGeB [KSA2]_HG530542
OLCuB [OKB1]_KF267444
OLCuB [Ama]_JX649952
OLCuB_AF397217
OLCuB_FJ187798
CLCuGeB [IsSqC9]_KT099178
OLCuB [Hom]_KJ396939
OkLCB
OLCuBF_J455515
PepGMB_FJ436005
ToLCSDB/ToLCYMB
0.3

Supplementary figure 5: Maximum-likelihood phylogenetic tree inferred from alignments of **(a)** begomovirus DNA-A/DNA-A like sequences, **(b)** alphasatellites and **(c)** betasatellites associated with begomovirus groups. The sequences obtained in this study are shown in bold.

Supplementary table 1**:** Table presenting main virus(es) associated with each assignation profiles-based cluster obtained from samples collected from Goué and Léguéma using grid sampling strategy.

Cluster ID	Type	Main Virus(es)	Nb plant
A1	Begomovirus DNA-A/A-like	Unclassified begomovirus	17
A2	Begomovirus DNA-A/A-like	Hollyhock leaf crumple virus	32
A3	Begomovirus DNA-A/A-like	Pepper yellow vein Mali virus	32
A4	Begomovirus DNA-A/A-like	Pepper yellow vein Mali virus / Tomato leaf curl mali virus	15
A5	Begomovirus DNA-A/A-like	Tobacco leaf curl Zimbabwe virus	15
A6	Begomovirus DNA-A/A-like	Malvastrum yellow vein virus	2
A7	Begomovirus DNA-A/A-like	Cotton leaf curl Gezira virus	34
A8	Begomovirus DNA-A/A-like	Cotton leaf curl Gezira virus / Hollyhock leaf crumple virus	9
A9	Begomovirus DNA-A/A-like	Sweet potato leaf curl Henan virus	2
A10	Begomovirus DNA-A/A-like	Cotton leaf curl Gezira virus / Okra yellow crinkle virus	7
A11	Begomovirus DNA-A/A-like	West African Asystasia virus 1 / Unclassified begomovirus	7
A12	Begomovirus DNA-A/A-like	Unclassified begomovirus / Hollyhock leaf crumple virus	8
A13	Begomovirus DNA-A/A-like	Tomato leaf curl mali virus	30
A14	Begomovirus DNA-A/A-like	Sweet potato leaf curl virus	8
A15	Begomovirus DNA-A/A-like	Tomato leaf curl virus	4
A16	Begomovirus DNA-A/A-like	Unclassified begomovirus	6
B1	Begomovirus DNA-B	Pepper yellow vein Mali virus	33
B2	Begomovirus DNA-B	Unclassified begomovirus	1
B3	Begomovirus DNA-B	Unclassified begomovirus / East African cassava mosaic virus	3
B4	Begomovirus DNA-B	Unclassified begomovirus / Abutilon mosaic Brazil virus	2
B5	Begomovirus DNA-B	East African cassava mosaic Kenya virus	2
B6	Begomovirus DNA-B	Kudzu mosaic virus	1
S1	ssDNA Satellites	Unclassified ssDNA satellite	21
S2	ssDNA Satellites	Ageratum leaf curl Cameroon alphasatellite	4
S3	ssDNA Satellites	Okra yellow crinkle Cameroon alphasatellite	15
S4	ssDNA Satellites	Siegesbeckia yellow vein virus-associated DNA beta	1
S5	ssDNA Satellites	Cotton leaf curl Gezira alphasatellite / Unclassified ssDNA satellite	7
S6	ssDNA Satellites	New World begomovirus associated satellite DNA	3
S7	ssDNA Satellites	Cotton leaf curl Gezira alphasatellite	16
S8	ssDNA Satellites	Cotton leaf curl Gezira alphasatellite / Okra yellow crinkle Cameroon alphasatellite	9
S9	ssDNA Satellites	Okra leaf curl Mali virus satellite DNA beta	7
S10	ssDNA Satellites	Unclassified ssDNA satellite	6
S11	ssDNA Satellites	Unclassified ssDNA satellite	8
S12	ssDNA Satellites	Unclassified ssDNA satellite	9
S13	ssDNA Satellites	Ageratum leaf curl Cameroon betasatellite	8
M1	Mastrevirus	Chickpea chlorotic dwarf virus	9
M2	Mastrevirus	Unclassfied mastrevirus / Chickpea chlorotic dwarf virus	3
M3	Mastrevirus	Unclassfied mastrevirus	2
M4	Mastrevirus	Maize streak virus	1
M5	Mastrevirus	Sugarcane streak virus	1
M6	Mastrevirus	Chickpea yellow dwarf virus	1

References

Alexander, H.M., Mauck, K.E., Whitfield, a. E., Garrett, K. a. & Malmstrom, C.M. (2014) Plant-virus interactions and the agro-ecological interface. *European Journal of Plant Pathology*, **138**, 529–547.

Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Filloux, D., Hartnady, P., Rebelo, T.A., Cousins, S.R., Mesleard, F., Cohez, D., Yavercovski, N., Varsani, A., Harkins, G.W., Peterschmitt, M., Malmstrom, C.M., Martin, D.P. & Roumagnac, P. (2017) Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME Journal*, 1–12.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

De Bruyn, A. (2014) Diversité génétique , Evolution et Dissémination des bégomovirus africains dans les îles du sud-ouest de l'océan Indien.Thèse de Doctorat; Université de la Réunion, 240 p

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Fargette, D., Konaté, G., Fauquet, C., Muller, E., Peterschmitt, M. & Thresh, J.M. (2006) Molecular ecology and emergence of tropical plant viruses. *Annual Review of Phytopathology*, **44**, 235–60.

Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M. & Zhou, X. (2008) Geminivirus strain demarcation and nomenclature. *Archives of Virology*, **153**, 783–821.

Gilbert, G.S. (2002) Evolutionary ecology of plant diseases in natural ecosystems. *Annual Review of Phytopathology*, **40**, 13–43.

Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J., Graham,

S.W., James, K.E., Kim, K.-J., John Kress, W., Schneider, H., Van Alphen Stahl, J., Barrett, S.C.H., Van Der Berg, C., Bogarin, D., Burgess, K.S., Cameron, K.M., Carine, M., Chacón, J., Clark, A., Clarkson, J.L., Conrad, F., Devey, D.S., Ford, C.S., Hedderson, T.A.J., Hollingsworth, M.L., Husband, B.C., Kelly, L.J., Kesanakurti, P.R., Kim, J.S., Kim, Y.-D., Lahaye, R., Kim, H.-L., Long, D.G., Madriñán, S., Maurin, O., Meusnier, I., Newmaster, S.G., Park, C.-W., Percy, D.M., Petersen, G., Logo, O., Richardson, J.E., Salazar, G.A., Savolainen, V., Seberg, O., Logo, O., Wilkinson, M.J., Yi, D.-K. & Little, D.P. (2009) A DNA Barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 12794–12797.

Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B. & Nagata, T. (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage ?29 DNA polymerase. *Journal of Virological Methods*, **116**, 209–211.

Jones, R. (2009) Plant virus emergence and evoution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, **141**, 113–130.

Keesing, F., Belden, L.K., Daszak, P., Dobson, A., Harvell, C.D., Holt, R.D., Hudson, P., Jolles, A., Jones, K.E., Mitchell, C.E., Myers, S.S., Bogich, T. & Ostfeld, R.S. (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, **468**, 647–652.

Keesing, F., Holt, R.D. & Ostfeld, R.S. (2006) Effects of species diversity on disease risk. *Ecology Letters*, **9**, 485–498.

Lefeuvre, P. & Moriones, E. (2015) Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. *Current Opinion in Virology*, **10**, 14–19.

Leke, W.N., Mignouna, D.B., Brown, J.K. & Kvarnheden, A. (2015) Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. *Agriculture & Food Security*, **4**, 1.

Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, **3**, e1420.

Malmstrom, C.M., Melcher, U. & Bosque-Pérez, N. a. (2011) The expanding field of plant virus ecology: Historical foundations, knowledge gaps, and research directions. *Virus Research*, **159**, 84–94.

Mansoor, S., Zafar, Y. & Briddon, R.W. (2006) Geminivirus disease complexes: the threat is spreading. *Trends in Plant Science*, **11**.

Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. & Neufeld, J.D. (2012) PANDAseq: PAired-eND Assembler for Illumina sequences. *BMC Bioinformatics*, **13**, 31.

Matsen, F.A., Kodner, R.B. & Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, **11**, 538.

Mubin, M., Briddon, R.W. & Mansoor, S. (2009) Complete nucleotide sequence of chili leaf curl virus and its associated satellites naturally infecting potato in Pakistan. *Archives of Virology*, **154**, 365–368.

Muhire, B., Darren, P.M., Brown, J.K., Navas-Castillo, J., Moriones, E., Zerbini, F.M., Rivera-Bustamante, R., Malathi, V.G., Briddon, R.W. & Varsani, A. (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of Virology*, **158**, 1411-1424.

Ouattara, A., Tiendrébéogo, F., Lefeuvre, P., Claverie, S., Hoareau, M., Traoré, E.V., Barro, N., Traoré, O. & Lett, J.M. (2017a) Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso. *Archives of Virology*, **162**, 1427–1429.

Ouattara, A., Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Claverie, S., Traoré, E.V., Barro, N., Traoré, O., Varsani, A. & Lett, J.M. (2017b) New strains of chickpea chlorotic dwarf virus discovered on diseased papaya and tomato plants in Burkina

Faso. *Archives of Virology*, **162**, 1791–1794.

Palanga, E., Filloux, D., Martin, D.P., Fernandez, E., Gargani, D., Ferdinand, R., Zabre, J., Bouda, Z., Neya, B., Sawadogo, M., Traore, O., Peterschmitt, M. & Roumagnac, P. (2016) Metagenomic-Based Screening and Molecular Characterization of Cowpea- Infecting Viruses in Burkina Faso. *PloS one*, **11**, 1–21.

Patil, B.L. & Fauquet, C.M. (2009) Cassava mosaic geminiviruses : actual knowledge and perspectives. *Plant Molecular Pathology*, **10**, 685–701.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2010) Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, **5**.

Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, **4**, 1–22.

Rojas, M.R., Hagen, C., Lucas, W.J. & Gilbertson, R.L. (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, **43**, 361–394.

Roossinck, M.J., Martin, D.P. & Roumagnac, P. (2015) Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology*, **105**, 716–727.

Roossinck, M.J., Saha, P., Wiley, G.B., Quan, J., White, J.D., Lai, H., Chavarría, F., Shen, G. & Roe, B.A. (2010) Ecogenomics: Using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*, **19**, 81–88.

Sattar, M.N., Koutou, M., Husseini, S., Leke, W.N., Brown, J.K. & Kvarnheden, A. (2015) First Identification of Begomoviruses Infecting Tomato With Leaf Curl Disease in Burkina Faso. *Plant Disease*, **99**, 732–732.

Shepherd, D.N., Martin, D.P., Walt, E.V.A.N.D.E.R., Dent, K., Varsani, A. & Rybicki, E.P. (2010) Pathogen profile Maize streak virus : an old and complex " emerging " pathogen. *Molecular Plant Pathology*, **11**, 1–12.

Silva, J.C.F., Carvalho, T.F.M., Basso, M.F., Deguchi, M., Pereira, W.A., Sobrinho,

R.R., Vidigal, P.M.P., Brustolini, O.J.B., Silva, F.F., Dal-Bianco, M., Fontes, R.L.F., Santos, A.A., Zerbini, F.M., Cerqueira, F.R. & Fontes, E.P.B. (2017) Geminivirus data warehouse: a database enriched with machine learning approaches. *BMC Bioinformatics*, **18**, 240.

Stobbe, A.H. & Roossinck, M.J. (2014) Plant virus metagenomics: what we know and why we need to know more. *Frontiers in Plant Science*, **5**, 1–4.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.

Tang, W., Yu, L., He, W., Yang, G., Ke, F., Baxter, S.W., You, S., Douglas, C.J. & You, M. (2014) DBM-DB: The diamondback moth genome database. *Database*, **2014**, 1–8.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Traoré, V.S.E., Barro, N., Péréfarres, F., Reynaud, B., Traoré, A.S., Konaté, G., Lett, J.-M. & Traoré, O. (2011) Molecular and biological characterization of Pepper yellow vein Mali virus (PepYVMV) isolates associated with pepper yellow vein disease in Burkina Faso. *Archives of Virology*, **156**, 483–487.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Villemot, J., Konaté, G., Traoré, A.S., Barro, N., Traoré, V.S., Reynaud, B., Traoré, O. & Lett, J.-M. (2010a) Molecular diversity of cotton leaf curl Gezira virus isolates and their satellite DNAs associated with okra leaf curl disease in Burkina Faso. *Virology Journal*, **7**, 48.

Tiendrébéogo, F., Traoré, V.S.E., Lett, J., Barro, N., Konaté, G., Traoré, A.S. & Traoré, O. (2010b) Impact of okra leaf curl disease on morphology and yield of okra. *Crop Protection*, **29**, 712-716..

Varma, A. & Malathi, V.G. (2003) Emerging geminivirus problems: a serious threat to crop production. *Annals of Applied Biology*, **142**, 145–164.

Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and*

Environmental Microbiology, **73**, 5261–5267.

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

Discussion générale

Discussion générale

Depuis plusieurs décennies, les maladies virales émergentes sur les cultures causent des dégâts considérables et constituent une menace importante pour la sécurité alimentaire notamment dans les régions où les populations dépendant étroitement de la production agricole pour leur subsistance (Anderson *et al.*, 2004; Jones, 2009). Parmi les facteurs impliqués dans ces phénomènes d'émergence, le développement des activités humaines semble avoir profondément perturbé les équilibres écologiques et favorisé l'apparition de nouvelles interactions entre plantes, virus et insectes vecteurs (Buddenhagen, 1977; Thresh, 1980; Fargette *et al.*, 2006; Jones, 2009). D'une part, les virus introduits par l'intermédiaire des plantes cultivées peuvent infecter les plantes indigènes sauvages et menacer la biodiversité. A l'inverse, des virus indigènes peuvent devenir émergents sur des plantes cultivées introduites. Enfin, ces deux scénarios peuvent être facilités par la dissémination d'espèces vectrices comme les insectes polyphages, et par la multiplication des zones de contact et d'interface entre les milieux naturels et cultivés (Webster *et al.*, 2007; Jones, 2009; Alexander *et al.*, 2014).

En raison de leur impact majeur sur le rendement de nombreuses cultures, l'intérêt pour les virus à ADN de la famille des *Geminiviridae* n'a cessé d'augmenter. A ce titre, la communauté des virologues s'est rapidement concentrée sur la caractérisation de la diversité de ces virus afin de comprendre leurs origines et les mécanismes associés à leurs émergences. Toutefois, si de par le monde les travaux récents d'épidémiologie moléculaire ont ainsi permis de décrire une multitude de nouvelles espèces de bégomovirus sur les cultures fortement impactées par les maladies émergentes (voir entre autres pour revues Shepherd *et al.*, 2010; Sattar *et al.*, 2013; Leke *et al.*, 2015; Sudarshana *et al.*, 2015), le manque de travaux portant sur les virus associés aux plantes non-cultivés et aux écosytèmes sauvages en général biaise profondément la vision de la diversité connue en géminivirus. En effet, les données récentes de la métagénomique suggèrent que la grande majorité des virus caractérisés sur les cultures ne représentent qu'une faible fraction des virus infectant les plantes en général (Malmstrom *et al.*, 2011; Alexander *et al.*, 2014). Ce biais dans la connaissance globale de la diversité des phytovirus présent dans les écosystèmes est particulièrement accru en Afrique de

l'Ouest et Centrale, où la très grande majorité des espèces virales identifiées ont été caractérisées à partir d'un nombre restreint de plantes cultivées, comme le gombo, le haricot, le manioc, la patate douce et la tomate (pour revue Leke *et al.*, 2015).

Dans le cadre de ces travaux de thèse, nous avons contribué à la sécurité des productions maraîchères en Afrique de l'Ouest, notamment au Burkina Faso et en Côte d'Ivoire, en générant des informations scientifiques sur les aspects épidémiologiques majeurs des maladies à géminivirus. Dans un premier temps, nous avons étudié la diversité génétique des géminivirus associés aux maladies sur les cultures maraîchères au Burkina Faso et en Côte d'Ivoire et évalué en plein champ les principaux paramètres épidémiologiques associés dont notamment la prévalence des géminivirus sur les cultures de tomate et leur gamme de plantes hôtes. Dans un second temps, nous avons évalué quelques principaux paramètres épidémiologiques (virulence, abondance virale, taux de transmission) du bégomovirus le plus sévère et prévalent sur les cultures de tomate au Burkina Faso, à savoir le PepYVMLV en association avec un composant ADN-B que nous avons nouvellement décrit. Enfin, nous avons étudié la diversité et la structure des populations de géminivirus à l'échelle des systèmes agroécologiques sur les plantes cultivées et non cultivées par une approche de métagénomique ciblée.

Diversité de géminivirus en Afrique de l'Ouest : importance des réseaux d'épidémiosurveillance

L'utilisation conjointe d'une approche de biologie moléculaire classique pour le clonage de génomes de géminivirus (Inoue-Nagata *et al.*, 2004; Chapitre 1) et d'une approche de métagénomique spécifique aux molécules d'ADN circulaire de petite taille (Chapitre 3) a permis à la fois la détection d'espèces de géminivirus déjà décrites au Burkina Faso ou en Afrique mais aussi de plusieurs autres espèces de géminivirus qui était encore inconnues. S'il faudra confirmer certaines de ces découvertes par le clonage et le séquençage de génomes complets, ces résultats confirment l'existence d'une grande diversité de géminivirus sur plantes cultivées et non cultivées.

Les séquences virales obtenues à partir des prospections réalisées de 2013 à 2016 au Burkina Faso et en Côte d'Ivoire ont permis la description sur plantes maraîchères (tomate, poivron, piment, et gombo) de **cinq espèces de bégomovirus** préalablement décrites en Afrique de l'Ouest (CLCuGeV, ToLCMLV, ToLCGHV, OYCrV, PepYVMLV ; Chapitres 1 et 3; Articles 1, 4 et 5), de **deux nouvelles espèces** de géminivirus dont **un bégomovirus** (ToLCBF ; Chapitre 1 ; Article 2) et **un mastrévirus** (CpCDV ; Chapitre 1 ; Article 3) et sur plantes non cultivées **d'une nouvelle souche de CLCuGeV** (Chapitre 1, Article 1) et d'**une nouvelle espèce de bégomovirus MeAV** (Chapitre 3, Article 7) infectant respectivement *Sida acuta* et *Melochia* sp.

Les premières descriptions du CLCuGeV et de l'OYCrV associées à la maladie de l'OLCD sur gombo (Chapitre 1, Article 5), et du PepYVMLV associé à la maladie du PYVD sur piment en Côte d'Ivoire (Chapitre 1, Article 4) et du ToLCGHV associée à la maladie du ToLCD-TYLCD sur tomate au Burkina Faso (Chapitre 1, Article 1), confirment les suspicions d'implication de bégomovirus dans ces maladies des cultures maraichères en Afrique de l'Ouest et démontrent la large répartition géographique et le caractère invasif de certains de ces bégomovirus. De plus, la première description d'un mastrévirus inféodé aux dicotylédones sur tomate en Afrique, le CpCDV, représente une nouveauté de grande importance pour la gestion de cette culture en Afrique de l'Ouest et plus largement dans l'Ancien Monde. En effet, ce groupe de mastrévirus inféodés aux dicotylédones a été essentiellement caractérisé sur légumineuses au Moyen-Orient et dans le Sud du continent Indien, et n'aurait été détecté jusqu'à lors qu'une seule fois sur tomate au Pakistan en 2012 (Données non publiées). Par ailleurs, en plus de sa détection sur tomate, le clonage d'un génome complet à partir d'un échantillon symptomatique de papayer du Burkina Faso ouvre de nombreuses questions sur sa gamme d'hôte réelle et sa répartition géographique dans l'Ancien Monde.

Ces données soulignent l'importance de la surveillance épidémiologique, des descriptions sentinelles et de la caractérisation moléculaire des populations virales présentes sur les plantes cultivées à l'échelle d'un réseau régional. Autours du réseau de protection des végétaux en Afrique sub-saharienne (ProVeg) et du laboratoire mixte interdisciplinaire PathoBios, une initiative pourrait inclure tous les pays d'Afrique de l'Ouest et Central, notamment ceux pour lesquels il existe peu ou pas d'information sur

la présence d'espèces ou souches virales impliqués dans certaines formes sévères de maladies. A titre d'exemple, ce type d'initiative permet aujourd'hui de retracer précisément l'historique d'invasion, encore en cours, de la souche Uganda de l'*East African cassava mosaic virus* (EACMV-UG), impliquée dans la forme sévère de la maladie de la mosaïque du manioc (Legg *et al.*, 2014). En effet, cette souche a été détectée récemment dans de nombreux pays d'Afrique de l'Ouest et Centrale dont le Burkina Faso (Tiendrébéogo et al., 2009), le Cameroun (Akinbade *et al.*, 2010), la Centrafrique et le Tchad (Zinga *et al.*, 2012), et la Guinée Equatoriale (Valam-Zango *et al.*, 2015). En conjonction de descriptions plus anciennes en République Démocratique du Congo (Neuenschwander *et al.*, 2002)et au Gabon (Legg *et al.*, 2004), ces descriptions sentinelles démontrent une propagation rapide d'Est en Ouest de l'EACMV-UG en Afrique qui font craindre une aggravation de la situation sanitaire du manioc en Afrique de l'Ouest et Centrale avec des répercussions sur la situation alimentaire des populations locales. Au-delà de permettre la mise en place d'un plan de lutte et la diffusion de matériel résistant adapté, ces initiatives permettent d'analyser de manière rétrospective les facteurs associés à une émergence panafricaine et ainsi de mieux prendre en compte les menaces futures du même type.

Bégomovirus et multipartisme, une nouvelle menace pour la production des solanées

Que ce soit par des méthodes classiques de biologie moléculaire ou par métagénomique, le PepYVMLV a été retrouvé majoritairement en association avec un composant d'ADN-B jusqu'alors inconnu (Chapitres 1 et 3). Afin de comprendre l'implication de ce composant dans l'étiologie du ToLCD-TYLCD au Burkina Faso, une étude des principaux paramètres épidémiologiques (virulence, accumulation virale et le taux de transmission par l'insecte vecteur) a été conduite en conditions contrôlées. Nos résultats démontrent que même si l'ADN-B n'est pas nécessaire pour le mouvement systémique du PepYVMLV et le développement des symptômes (association non-stricte), leur association induit une plus forte virulence, une charge virale supérieure de l'ADN-A ainsi qu'une meilleure efficacité de transmission par l'aleurode *B. tabaci.* Ainsi, le

recrutement d'une molécule d'ADN-B par le bégomovirus monopartite PepYVMLV représente vraisemblablement un des éléments clés de son émergence en tant que maladie virale la plus prévalente et sévère sur les cultures de tomate au Burkina Faso (Chapitre 2). Bien que les deux composants de la majorité des bégomovirus bipartites décrits comme associés à la maladie du ToLCV-TYLCD de par le monde présentent une relation obligatoire comme c'est le cas notamment de tous les bégomovirus bipartites du Nouveau Monde (Padidam *et al.*, 1995), d'autres cas d'associations non-strictes ont été décrites dans la littérature dans l'Ancien Monde comme celles du *Tomato yellow leaf curl Thailand virus* (TYLCTHV ; Rochester *et al.*, 1994) et du *Tomato leaf curl Gujarat virus* (ToLCGV; Chakraborty *et al.*, 2003), suggérant que ces virus représentent des intermédiaires évolutifs entre les bégomovirus monopartites et bipartites strictes.

Face à l'expression, en conditions contrôlées, de symptômes supérieurs à ceux du TYLCV-IL (Péréfarres *et al.*, 2014), avec notamment des arrêts de croissance pouvant aboutir à la mort de la plante, on peut considérer que le PepYVMLV en association avec le composant d'ADN-B représente une menace importante pour la culture de la tomate non seulement pour le Burkina Faso, mais également pour l'Afrique et plus largement le monde. Outre les effets de ces associations dans les traits viraux comme la virulence, l'abondance virale ou la capacité de transmission, elles pourraient également jouer un rôle clef dans l'élargissement de la gamme d'hôtes des bégomovirus et *in fine* dans leur capacité d'émergence sur de nouveaux hôtes. En effet, dans d'autres cas d'associations de certains bégomovirus monopartites de l'Ancien Monde avec des ADN satellites (Briddon & Stanley, 2006), leurs rôles dans l'expression de symptômes plus sévères de la maladie (Saunders *et al.*, 2000; Briddon *et al.*, 2001) mais aussi dans l'élargissement de la gamme d'hôte (cas du *Sri Lankan cassava mosaic virus* (Saunders *et al.*, 2002), ont été montrés.

Importance de la recombinaison dans l'évolution des bégomovirus (recombinaison inter-genres MeAV)

La capacité d'un virus à s'adapter à un hôte est conditionnée par sa variabilité génétique, qui est liée à la fois à son taux de mutation et de recombinaison. L'obtention d'un nombre croissant de séquences génomiques complètes et le développement des méthodes de détection de la recombinaison ont permis de montrer que ce processus jouait un rôle essentiel dans la genèse de la diversité de nombreux virus en général et des géminivirus en particulier (Padidam *et al.*, 1999). La caractérisation de la diversité des populations géminivirales à l'échelle des agro-systèmes, nous a amené à décrire une nouvelle espèce de bégomovirus présentant une structure génomique classique au bégomovirus monopartite et nommée provisoirement *Melonchia associated virus* (MeAV) (Chapitre 3). Les premières analyses du génomes permettent de suspecter très fortement que ce virus est issue d'une recombinaison inter-genre entre un bégomovirus apparenté aux PepYVMLV et au *Tomato curly stunt virus* et un mastrévirus ne présentant une parenté que très lointaine avec le *Sweetpotato symptomless mastrevirus 1*, nouvellement caractérisé en Chine sur patate douce (Cao *et al.*, 2017). Un nombre réduit de cas de recombinaison inter-genres entre géminivirus ont été décrites dans la littérature. Cette rareté semble s'expliquer à la fois par les contraintes liées au partage d'une gamme d'hôte commune, à la coinfection de la même cellule hôte par au moins deux espèces virales (Morilla *et al.*, 2004) et aux maintiens des interactions épistatiques encodés par le génome dans le cas de recombinaison entre parents très distants (Martin *et al.*, 2005, 2011).

Association entre les géminivirus et leurs plantes hôtes, des relations complexes au sein des agro-écosystèmes

Le caractère non pérenne de la plupart des plantes cultivées concernées par les maladies virales suggère l'existence d'hôtes alternatifs ou réservoirs permettant aux populations virales de se maintenir hors des périodes de culture et de ré-émerger en fonction de la disponibilité de leur hôte(s) cultivé(s) (Funayama-Noguchi, 2001). Les plantes

adventices présentes aux abords des cultures ont très vite été suspectées dans ce rôle de plantes réservoirs, et certaines sont considérées comme des « générateurs » de variants recombinants en raison des infections multiples qui les caractérisent (Ooi *et al.*, 1997; Jovel *et al.*, 2004; García-Andrés *et al.*, 2006). Cependant, les connaissances de la diversité des populations virales naturelles, de leurs répartitions géographiques, de leurs liens phylogénétiques avec les populations virales décrites sur les plantes cultivées et de leurs pathogénicités restent insuffisantes. Ainsi, on constate un manque de connaissances fondamentales sur la dynamique des communautés virales à l'échelle des écosystèmes, l'écologie évolutive des virus en général et leur capacité à changer d'hôte en particulier (Alexander *et al.*, 2014 ; Roossinck *et al.*, 2015). Dans cette thèse, la combinaison des approches RCA-RFLP et des techniques de séquençage à haut débit nous a apporté des informations précieuses sur la prévalence des géminivirus et sur leurs gammes de plantes hôtes.

Dans un premier temps, nous avons montré à l'échelle des agroécosystèmes que la prévalence des géminivirus sur plante cultivée est généralement plus élevée que sur plante non-cultivée (Chapitre 3). Ces données confirment les résultats obtenus par des études récentes ayant portées sur la diversité et la prévalence des populations virales présentes dans deux agroécosystèmes méditerranéens (Bernardo *et al.*, 2017). Plusieurs éléments pourraient expliquer cette différence de prévalence virale entre les plantes cultivées et non cultivées. Ainsi, les prévalences virales élevées dans les milieux cultivés pourraient s'expliquer en premier lieu par la sensibilité des plantes cultivées aux infections virales et la présence d'une diversité de plante réduite. En effet, selon la théorie de l'allocation des ressources (Herms & Mattson, 1992), une plante qui investit ses ressources dans la croissance diminuerait son investissement dans la résistance aux agents pathogènes. De plus, lorsqu'une variété est réputée pour ses rendements exceptionnels, elle est généralement choisie par la majorité des producteurs et se retrouve ainsi cultivée sur de grandes superficies. Cette augmentation du nombre de plantes cultivées (hôtes potentiels) aura tendance à favoriser le développement et la pullulation de certains insectes vecteurs, la transmission virale et ainsi augmenter les risques épidémiques (Roossinck & García-Arenal, 2015). A l'inverse, les faibles prévalences des populations virales dans les milieux naturels serait en partie liées à la

diversité des plantes non cultivées et au phénomène de dilution de l'hôte (Ostfeld & Keesing, 2012).

Nos travaux ont permis d'identifier plus d'une dizaine de genres de plantes non cultivées comme hôtes potentiels des géminivirus (Chapitres 1 et 3). La plupart de ces plantes représentent potentiellement des hôtes alternatifs ou réservoirs des géminivirus infectant les cultures. Néanmoins, même si ces données nécessitent d'être vérifiées par la caractérisation moléculaire de génomes complets des espèces virales détectées, elles confirment d'ores et déjà l'implication des plantes non cultivées dans le maintien des populations de géminivirus durant les périodes d'inter-cultures. De plus, certaines des plantes non-cultivées que nous avons caractérisées présentent de remarquables capacités de résistance à la sécheresse, permettant aux populations virales de se maintenir pendant les longues périodes de sécheresse caractéristiques du climat sahélien. C'est notamment le cas du *Sida acuta* (*Malvaceae*), qui du fait de sa large distribution dans les trois zones agroclimatiques du Burkina Faso, de son caractère de plante pérenne et son omniprésence dans les cultures comme une des principales plante adventice, semble jouer un rôle fondamental dans le maintien de certaines populations virales dans les agro-écosystèmes. Premièrement, nous avons pu confirmer que *S. acuta* héberge le CLCuGeV (Chapitre 1, article 1), un bégomovirus monopartite inféodé aux malvacées. Ce bégomovirus est actuellement le principal agent causal de la maladie émergente de l'enroulement foliaire du gombo, qui représente la principale contrainte de cette culture dans tous les pays sahéliens producteurs (Tiendrébéogo *et al.*, 2010). Deuxièmement, nous avons pu montrer que parmi les cinq géminivirus associés à la maladie du ToLCD-TYLCD au Burkina Faso (Chapitres 1 et 3), seul le PepYVMLV est hébergé par *S. acuta*. Outre ses traits biologiques remarquables en association avec le composant d'ADN-B avec une augmentation notable de sa capacité d'accumulation virale, de sa virulence et de son taux de transmission (Chapitre 2), le PepYVMLV semble également bénéficier d'un autre avantage adaptatif supplémentaire avec l'utilisation du *S. acuta* comme plante hôte réservoir ou alternative, particulièrement bien adaptée aux périodes d'intercultures.

Existence de géminivirus ubiquistes : généraliste *versus* spécialiste ?

Nous avons montré à l'échelle des agro-écosystèmes (Chapitres 1 et 3) que des lignés virales apparentées à des bégomovirus connus ou non encore caractérisées infectaient des plantes appartenant à différentes familles notamment celle des *Amaranthaceae*, *Asteraceae*, *Convolvulaceae*, *Euphorbiaceae*, *Malvaceae*, *Nyctaginaceae* et *Solanaceae*. Le potentiel adaptatif visiblement impressionnant de ces lignées de bégomovirus suggère qu'elles pourraient représenter des populations virales généralistes telles que l'emblématique TYLCV, responsable de la forme la plus sévère de la maladie du TYLCD sur tomate au niveau mondial, dont les données actuelles suggèrent que sa gamme d'hôte couvre cinq familles différentes de végétaux : *Compositae, Leguminosae, Malvaceae, Plantaginaceae et Solanaceae* (Ioannou *et al.*, 1987). On peut également citer en exemple le CpCDV, auparavant décrit sur des légumineuses en Moyen-Orient et en Inde (Kraberger *et al.*, 2015), et pour lequel nous avons élargi sa gamme d'hôte à deux autres familles botaniques les *Solanaceae* et *Caricacea*. Ce mastrévirus pourrait également être considéré comme un généraliste telle que la souche A sévère et invasive du *Maize streak virus* (MSV-A) qui présente une gamme d'hôte plus large que les autres souches de ce mastrévirus (Varsani *et al.*, 2008; Shepherd *et al.*, 2010). Le potentiel d'émergence des virus de type généraliste pourrait s'avérer particulièrement important et nécessitera des expérimentations en laboratoire pour étudier leur gamme d'hôte après inoculation à l'aide de clones agro-infectieux ou de leurs insectes vecteurs.

Par ailleurs, de nombreux cas de coinfection vraisemblablement lié à l'ubiquité des géminivirus dits généralistes ont été observés. En effet, le partage d'une partie de leur gamme d'hôte favorise les phénomènes non seulement de recombinaison, qui pourraient aboutir à l'émergence de nouvelles souches ou espèces virales, mais aussi de synergie pouvant être à l'origine d'épidémies plus sévères, comme observé en Afrique de l'Est avec la forme sévère de la maladie de la mosaïque du manioc (Legg & Fauquet, 2004).

En parfait contraste avec ces groupes viraux qualifiés de généralistes, plusieurs des groupes de géminivirus de type spécialiste ont été détectés, sur des plantes appartenant au genre *Amaranthus*, *Ipomoea* et *Sida*. En particulier, un virus proche des ToLCVs sur *Amaranthus*, le *Sweet potato leaf curl virus* sur *Ipomoea* et le HoLCrV sur *Sida* montre

des gammes de plantes hôtes très restreintes alors même que leurs prévalences restent relativement élevées sur les hôtes cités.

Vers une description complète du fonctionnement des communautés de géminivirus ?

Les travaux entrepris durant cette thèse confirment l'existence d'une importante diversité de géminivirus en Afrique de l'Ouest et Centrale, à la fois sur plantes cultivées et sauvages (Leke *et al.*, 2015). Au-delà, ils permettent de cartographier la structure de l'agencement des populations virales à l'échelle de deux agro-écosytèmes aux caractéristiques climatiques contrastées du Burkina Faso. La multitude d'espèces de geminivirus décrites, le large panel de plantes hôtes associées, les gammes d'hôtes individuelles variées et la présence fréquente de co-infection, décrivent une **communauté complexe de géminivirus** (*i.e.* un assemblage de populations d'espèces distinctes). Si celle-ci se caractérise par la présence de virus avec des degrés divers de spécialisation, l'analyse des réseaux de co-occurence de virus suivant le type d'hôte permet aussi de faire émerger deux modules principaux et distincts de plantes, avec d'un côté les plantes des genres *Ageratum*, *Capsicum* et *Solanum* et de l'autre les plantes des genres *Abelmoschus* et *Sida*. La présence de deux modules comportant à la fois des plantes cultivées et non-cultivées souligne une nouvelle fois, l'importance des plantes dites « alternatives » ou « réservoirs » dans l'épidémiologie des maladies des cultures.

De manière intéressante, si ces deux modules principaux ne présentent pas ou peu de connexion directe (*i.e.* ils ne partagent que peu de similitude dans les cortèges de virus les infectant), d'autres genres de plantes, avec en particulier les genres *Amaranthus* et *Euphorbia*, forment une jonction entre ces modules et partagent en partie les groupes viraux propres à chacun de ceux-ci. L'existence de tels *hubs* ou carrefours de diversité au sein de la communauté pourrait permettre à des virus circulants au sein de chacun des modules, d'avoir des effets sur la structure globale de la communauté. Ainsi, la mise en relation de l'historique des échanges de matériel génétique par recombinaison entre virus et de la structure de ces réseaux pourrait nous informer sur le rôle potentiel de ces

modules et *hubs* dans l'évolution des individus et de la structure globale de la communauté. L'inter-connexion des gammes d'hôtes de chacun des virus détectés indique aussi que c'est bien la communauté virale complète (concept de virome) voir la communauté microbienne dans son ensemble (concept de pathobiome; Vayssier-Taussat *et al.*, 2014), qu'il faut considérer si l'on souhaite comprendre l'émergence de maladies et en particulier si l'on souhaite déterminer si il existe des compositions ou des structures de communautés plus propices à ces émergences. C'est aussi en tentant de comprendre comment de telles communautés évoluent et répondent aux changements, comme ceux imposés par l'agriculture moderne, que nous pourrons mesurer l'impact des pratiques humaines sur l'apparition de maladie virale.

Bien sûr, même pour des réseaux d'interactions relativement réduit, comme ceux décrits ici, le nombre de facteurs pouvant intervenir (tels que la diversité des vecteurs et leurs préférences alimentaires, la densité et la répartition des plantes, les interactions virus-virus ou plante virus…), associé à la difficulté à les mesurer précisément rendent encore difficile la prédiction qualitative ou quantitative des évolutions à venir de ces communautés (Geoghegan & Holmes, 2017). C’est toutefois en poursuivant l'inventaire systématique des populations virales associées à des écosystèmes que nous pourrons émettre des hypothèses sur l'évolution des communautés virales et l'émergence de maladies, et les évaluer dans le cadre d’études comparatives.

Références bibliographiques

Références bibliographiques

Abhary, M., Patil, B.L. & Fauquet, C.M. (2007) *Molecular Biodiversity, Taxonomy, and Nomenclature of Tomato Yellow Leaf Curl-Like Viruses*. *Tomato Yellow Leaf Curl Virus Disease* (ed. by H. Czosnek), pp. 85–118. Springer Netherlands.

Akhtar, K.P., Ahmad, M., Shah, T.M. & Atta, B.M. (2011) Transmission of chickpea chlorotic dwarf virus in chickpea by the leafhopper Orosius albicinctus (Distant) in Pakistan -short communication. *Plant Protection Science*, **47**, 1–4.

Akhtar, S., Khan, A.J. & Briddon, R.W. (2014) A distinct strain of Chickpea chlorotic dwarf virus infecting pepper in Oman. *Plant Disease*, **98**, 286–287.

Akinbade, S.A., Hanna, R., Nguenkam, A., Njukwe, E., Fotso, A., Doumtsop, A., Ngeve, J., Tenku, S.T.N. & Kumar, P.L. (2010) First report of the East African cassava mosaic virus-Uganda (EACMV-UG) infecting cassava (Manihot esculenta) in Cameroon. *Plant Pathology*, **21**, 22.

Alabi, O.J., Ogbe, F.O., Bandyopadhyay, R., Lava Kumar, P., Dixon, A.G.O., D'A. Hughes, J. & Naidu, R.A. (2008) Alternate hosts of African cassava mosaic virus and East African cassava mosaic Cameroon virus in Nigeria. *Archives of Virology*, **153**, 1743–1747.

Alexander, H.M., Mauck, K.E., Whitfield, a. E., Garrett, K. a. & Malmstrom, C.M. (2014) Plant-virus interactions and the agro-ecological interface. *European Journal of Plant Pathology*, **138**, 529–547.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R. & Daszak, P. (2004) Emerging infectious diseases of plants: pathogen pollution , climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, **19**, 535–544.

Antignus, Y. (2007) *The management of Tomato yellow Leaf Curl Virus in Greenhouses and the Open Field, a Strategy of Manipulation BT - Tomato Yellow Leaf Curl*

Virus Disease: Management, Molecular Biology, Breeding for Resistance. Tomato Yellow Leaf Curl Virus Disease (ed. by H. Czosnek), pp. 263–278. Springer Netherlands, Dordrecht.

Antignus, Y., Nestel, D., Cohen, S. & Lapidot, M. (2001) Ultraviolet-deficient greenhouse environment affects whitefly attraction and flight-behavior. *Environmental Entomology*, **30**, 394–399.

Arguello-Astorga, G., Ascencio-Ibáñez, J.T., Dallas, M.B., Orozco, B.M. & Hanley-Bowdoin, L. (2007) High-frequency reversion of geminivirus replication protein mutants during infection. *Journal of Virology*, **81**, 11005–15.

Ascencio-Ibáñez, J.T., Diaz-Plaza, R., Méndez-Lozano, J., Monsalve-Fonnegra, Z.I., Argüello-Astorga, G.R. & Rivera-Bustamante, R.F. (1999) First Report of Tomato Yellow Leaf Curl Geminivirus in Yucatán, México. *Plant Disease*, **83**, 1178.

Barro, N. (1994) Caractérisation sérologique, biologique et aspects écologiques de quelques virus infectant les plantes maraîchères au Burkina Faso. Thèse de doctorat. Université de Ouagadougou, 171p.

De Barro, P. & Ahmed, M. (2011) Genetic networking of the Bemisia tabaci cryptic species complex reveals pattern of biological invasions. *PLoS ONE*, **6**.

De Barro, P.J., Liu, S.-S., Boykin, L.M. & Dinsdale, A.B. (2011) *Bemisia tabaci* : A Statement of Species Status. *Annual Review of Entomology*, **56**, 1–19.

Becker, N., Rimbaud, L., Chiroleu, F., Reynaud, B., Thébaud, G. & Lett, J.-M. (2015) Rapid accumulation and low degradation: key parameters of Tomato yellow leaf curl virus persistence in its insect vector Bemisia tabaci. *Nature Publishing Group*, 1–10.

Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Filloux, D., Hartnady, P., Rebelo, T.A., Cousins, S.R., Mesleard, F., Cohez, D., Yavercovski, N., Varsani, A., Harkins, G.W., Peterschmitt, M., Malmstrom, C.M., Martin, D.P. & Roumagnac, P. (2017) Geometagenomics illuminates the impact of agriculture

on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME Journal*, 1–12.

Bethke, J. a, Paine, T.D. & Nuessly, G.S. (1991) Comparative biology, morphometrics, and development of two populations of Bemisia tabaci (Homoptera: Aleyrodidae) on cotton and poinsettia. *Annals of the Entomological Society of America*, **84**, 407–411 ST–Comparative biology, morphometrics,.

Bigarré, L., Chazly, M., Salah, M., Ibrahim, M., Padidam, M., Nicole, M., Peterschmitt, M., Fauquet, C. & Thouvenel, J.C. (2001) Characterization of a new begomovirus from Egypt infecting hollyhock (Althea rosea). *European Journal of Plant Pathology*, **107**, 701–711.

Al Bitar, L. & Luisoni, E. (1995) Tomato yellow leaf curl geminivirus: Serological evaluation of an improved purification method. *Bulletin OEPP*, **25**, 269–276.

Blackmer, J.L., Lindley, V.A. & Byrne, D.N. (1995) Histological examination of flight muscle development and breakdown in Bemisia tabaci (Homoptera: Aleyrodidae): Relationship to age and flight behvior. *Journal of Morphology*, **226**, 213–221.

Blanc, S., Drucker, M. & Uzest, M. (2014) Localizing Viruses in Their Insect Vectors. *Annual Review of Phytopathology*, **52**, 403–425.

Blanc, S., Uzest, M. & Drucker, M. (2011) New research horizons in vector-transmission of plant viruses. *Current Opinion in Microbiology*, **14**, 483–491.

Blauth, S.L., Churchill, G.A. & Mutschler, M.A. (1998) Identification of quantitative trait loci associated with acylsugar accumulation using intraspeci c populations of the wild tomato, Lycopersicon pennellii. *Theoretical and Applied Genetics*, **96**, 458–467.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F. & Rohwer, F. (2002) Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences*, **99**, 14250–14255.

Briddon, R. & Stanley, J. (2006) Subviral agents associated with plant single-stranded DNA viruses. *Virology*, **344**, 198–210.

Briddon, R.W. (2003) Cotton leaf curl disease, a multicomponent begomovirus complex. *Molecular Plant Pathology*, **4**, 427–434.

Briddon, R.W., Bedford, I.D., Tsai, J.H. & Markham, P.G. (1996) Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology*, **219**, 387–394.

Briddon, R.W., Bull, S.E., Amin, I., Mansoor, S., Bedford, I.D., Rishi, N., Siwatch, S.S., Zafar, Y., Abdel-Salam, A.M. & Markham, P.G. (2004) Diversity of DNA 1: A satellite-like molecule associated with monopartite begomovirus-DNA complexes. *Virology*, **324**, 462–474.

Briddon, R.W., Mansoor, S., Bedford, I.D., Pinner, M.S., Saunders, K., Stanley, J., Zafar, Y., Malik, K.A. & Markham, P.G. (2001) Identification of DNA components required for induction of cotton leaf curl disease. *Virology*, **285**, 234–243.

Briddon, R.W., Patil, B.L., Bagewadi, B., Nawaz-ul-Rehman, M.S. & Fauquet, C.M. (2010) Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evolutionary Biology*, **10**, 97.

Brown, J. & Bird, J. (1995) Variability within the Bemisia tabaci species complex and its relation to new epidemics caused by geminiviruses. *Ceiba*, **36**, 73–80.

Brown, J.K., Fauquet, C.M., Briddon, R.W., Zerbini, M., Moriones, E. & Navas-Castillo, J. (2011) *Family Geminiviridae*. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses* (ed. by A.M.Q. King), M.J. Adams), E.B. Carstens), and E.J. Lefkowitz), pp. 351–373. Elsevier-Academic Press, Amsterdam, The Netherlands.

Brown, J.K., Idris, A.M., Banks, G.K. & Wyatt, S.D. (2001) The core region of the coat protein gene is highly useful for establishing the provisional identification and

classification of begomoviruses. *Archives of Virology*, 1581–1598.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C.F., Fiallo-Olivé, E., Briddon, R.W., Hernández-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S. & Varsani, A. (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of Virology*, **160**, 1593–1619.

Van Brunschot, S.L., Persley, D.M., Geering, A.D.W., Campbell, P.R. & Thomas, J.E. (2010) Tomato yellow leaf curl virus in Australia: Distribution, detection and discovery of naturally occurring defective DNA molecules. *Australasian Plant Pathology*, **39**, 412–423.

De Bruyn, A. (2014) Diversité génétique , Evolution et Dissémination des bégomovirus africains dans les îles du sud-ouest de l'océan Indien. Thèse de Doctorat; Université de la Réunion, 240 p.

Buddenhagen, I.W. (1977) Resistance and Vulnerability of Tropical Crops in Relation To Their Evolution and Breeding. *Annals of the New York Academy of Sciences*, **287**, 309–326.

Bull, S.E., Briddon, R.W., Sserubombwe, W.S., Ngugi, K., Markham, P.G. & Stanley, J. (2007) Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. *Journal of General Virology*, **88**, 1624–1633.

Byrne, D.N. & Bellows, T.S. (1991) Whitefly Biology. *Annual Review of Entomology*, **36**, 431–457.

Byrne, F.J., Castle, S., Prabhaker, N. & Toscano, N.C. (2003) Biochemical study of resistance to imidacloprid in B biotype Bemisia tabaci from Guatemala. *Pest Management Science*, **59**, 347–352.

Candresse, T., Filloux, D., Muhire, B., Julian, C., Galzi, S., Fort, G., Bernardo, P., Daugrois, J.H., Fernandez, E., Martin, D.P., Varsani, A. & Roumagnac, P. (2014) Appearances can be deceptive: Revealing a hidden viral infection with deep

sequencing in a plant quarantine context. *PLoS ONE*, **9**.

Cao, M., Lan, P., Li, F., Abad, J., Zhou, C. & Li, R. (2017) Genome characterization of sweet potato symptomless virus 1: a mastrevirus with an unusual nonanucleotide sequence. *Archives of Virology*, **162**, 2881–2884.

Caro, M., Verlaan, M.G., Julián, O., Finkers, R., Wolters, A.M.A., Hutton, S.F., Scott, J.W., Kormelink, R., Visser, R.G.F., Díez, M.J., Pérez-de-Castro, A. & Bai, Y. (2015) Assessing the genetic variation of Ty-1 and Ty-3 alleles conferring resistance to tomato yellow leaf curl virus in a broad tomato germplasm. *Molecular Breeding*, **35**, 1–13.

Chakraborty, S., Pandey, P.K., Banerjee, M.K., Kalloo, G. & Fauquet, C.M. (2003) Tomato leaf curl Gujarat virus a New Begomovirus Species Causing a Severe Leaf Curl Disease of Tomato in Varanasi, India. *Phytopathology*, **93**, 1485–1495.

Chen, L.-F., Rojas, M., Kon, T., Gamby, K., Xoconostle-ca, B. & Gilbertson, R.L. (2009) A severe symptom phenotype in tomato in Mali is caused by a reassortant between a novel recombinant begomovirus (Tomato yellow leaf curl Mali virus) and a betasatellite. *Molecular Plant Pathology*, **10**, 415–430.

Cohen, S. & Antignus, Y. (1994) Tomato yellow leaf curl virus, a whitefly-borne geminivirus of tomatoes. *Advances in Disease Vector Research*, **10**, 259–288.

Coltart, C.E.M., Lindsey, B., Ghinai, I., Johnson, A.M. & Heymann, D.L. (2017) The Ebola outbreak, 2013–2016: old lessons for new epidemics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160297.

Cooper, I. & Jones, R.A.C. (2006) Wild Plants and Viruses: Under-Investigated Ecosystems. *Advances in Virus Research*, **67**, 1–47.

Cooper, J.I., Jones, A.T., Road, M., Section, V. & Crop, S. (1983) Responses of Plants to Viruses: Proposals for the Use of Terms. *Phytopathology*, **73**, 127–128.

Crameri, A., Raillard, S.-A., Bermudez, E. & Stemmer, W.P.C. (1998) DNA shuffling

of a family of genes from diverse species accelerates directed evolution. *Nature*, **391**, 288–91.

Cui, X., Li, G., Wang, D., Hu, D. & Zhou, X. (2005) A begomovirus DNAb-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. *Journal of Virology*, **79**, 10764–10775.

Czosnek, H. (2007) *Interactions of Tomato Yellow Leaf Curl Virus with its Whitefly Vector*, *Tomato Yellow Leaf Curl Virus Disease*, 157-170.

Czosnek, H., Ghanim, M., Morin, S., Rubinstein, G., Fridman, V. & Zeidan, M. (2001) Whiteflies : Vectors and victims Geminiviruses. *Advances in Virus Research*, **57**, 291–322.

Dayaram, A., Galatowitsch, M.L., Argüello-Astorga, G.R., van Bysterveldt, K., Kraberger, S., Stainton, D., Harding, J.S., Roumagnac, P., Martin, D.P., Lefeuvre, P. & Varsani, A. (2016) Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infection, Genetics and Evolution*, **39**, 304–316.

Dekker, E.L., Woolston, C.J., Xue, Y., Cox, B. & Mullineaux, P.M. (1991) Transcript mapping reveals different expression strategies for the bicistronic RNAs of the geminivirus wheat dwarf virus. *Nucleic Acids Research*, **19**, 4075–4081.

Delatte, H., Duyck, P.F., Triboire, A., David, P., Becker, N., Bonato, O. & Reynaud, B. (2008) Differential invasion success among biotypes: Case of Bemisia tabaci. *Biological Invasions*, **11**, 1059–1070.

Delatte, H., Holota, H., Reynaud, B. & Dintinger, J. (2006) Characterisation of a quantitative resistance to vector transmission of Tomato yellow leaf curl virus in Lycopersicon pimpinellifolium. *European Journal of Plant Pathology*, **114**, 245–253.

Dinsdale, A., Cook, L., Riginos, C., Buckley, Y.M. & De Barro, P. (2010) Refined Global Analysis of <I>Bemisia tabaci</I> (Hemiptera: Sternorrhyncha:

Aleyrodoidea: Aleyrodidae) Mitochondrial Cytochrome Oxidase 1 to Identify Species Level Genetic Boundaries. *Annals of the Entomological Society of America*, **103**, 196–208.

Dry, I.B., Krake, L.R., Rigden, J.E. & Rezaian, M.A. (1997) A novel subviral agent associated with a geminivirus: The first report of a DNA satellite. *Proceedings of the National Academy of Sciences*, **94**, 7088–7093.

Duffy, S. & Holmes, E.C. (2007) Multiple introductions of the old world begomovirus Tomato yellow leaf curl virus into the new world. *Applied and Environmental Microbiology*, **73**, 7114–7117.

Duffy, S. & Holmes, E.C. (2008) Phylogenetic Evidence for Rapid Rates of Molecular Evolution in the Single-Stranded DNA Begomovirus Tomato Yellow Leaf Curl Virus. *Journal of Virology*, **82**, 957–965.

Duffy, S. & Holmes, E.C. (2009) Validation of high rates of nucleotide substitution in geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses. *Journal of General Virology*, **90**, 1539–1547.

Duffy, S., Shackelton, L.A. & Holmes, E.C. (2008) Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, **9**, 267–276.

Duncan, J.M. & Torrrance, L. (1992) Techniques for the Rapid Detection of Plant Pathogens. Blackwell Scientific Publications. *FAO, 1989, Production Year Book*, **43**.

Elbert, A. & Nauen, R. (2000) Resistance of Bemisia tabaci (Homoptera: Aleyrodidae) to insecticides in southern Spain with special reference to neonicotinoids. *Pest Management Science*, **56**, 60–64.

Esterhuizen, L.L., van Heerden, S.W., Rey, M.E.C. & van Heerden, H. (2012) Genetic identification of two sweet-potato-infecting begomoviruses in South Africa. *Archives of Virology*, **157**, 2241–2245.

Fahmy, I.F., Taha, O. & El-Ashry, A.N. (2015) First genome analysis and molecular characterization of Chickpea chlorotic dwarf virus Egyptian isolate infecting squash. *Virus Disease*, **26**, 33–41.

Fargette, D., Konaté, G., Fauquet, C., Muller, E., Peterschmitt, M. & Thresh, J.M. (2006) Molecular ecology and emergence of tropical plant viruses. *Annual Review of Phytopathology*, **44**, 235–60.

Farrar, J., Focks, D., Gubler, D., Barrera, R., Guzman, M.G., Simmons, C., Kalayanarooj, S., Lum, L., McCall, P.J., Lloyd, L., Horstick, O., Dayal-Drager, R., Nathan, M.B. & Kroeger, A. (2007) Editorial: Towards a global dengue research agenda. *Tropical Medicine and International Health*, **12**, 695–699.

Farzadfar, S., Pourrahim, R., Golnaraghi, A.R. & Ahoonmanesh, A. (2008) PCR detection and partial molecular characterization of Chickpea chlorotic dwarf virus in naturally infected sugar beet plants in Iran. *Journal of Plant Pathology*, **90**, 247–251.

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.

Fondong, V.N. (2013) Geminivirus protein structure and function. *Molecular Plant Pathology*, **14**, 635–649.

Fondong, V.N. (2017) The Search for Resistance to Cassava Mosaic Geminiviruses: How Much We Have Accomplished, and What Lies Ahead. *Frontiers in Plant Science*, **8**, 1–19.

Fontenele, R.S., Lamas, N.S., Lacorte, C., Lacerda, A.L.M., Varsani, A. & Ribeiro, S.G. (2017) A novel geminivirus identified in tomato and cleome plants sampled in Brazil. *Virus Research*.

Frohlich, D.H., Torres-Jerez, I., Bedford, I.D., Markham, P.G. & Brown, J.K. (1999) A phylogeographical analysis of the Bemisia tabaci species complex based on

mitochondrial DNA markers. *Molecular Ecology*, **8**, 1683–1691.

Fuentes, A., Ramos, P.L., Fiallo, E., Callard, D., Sánchez, Y., Peral, R., Rodríguez, R. & Pujol, M. (2006) Intron-hairpin RNA derived from replication associated protein C1 gene confers immunity to tomato yellow leaf curl virus infection in transgenic tomato plants. *Transgenic Research*, **15**, 291–304.

Funayama-Noguchi, S. (2001) Ecophysiology of virus-infected plants: a case study of Eupatorium makinoi infected by Geminivirus. *Plant Biology*, **3**, 251–262.

Gafni, Y. & Epel, B.L. (2002) The role of host and viral proteins in intra- and inter-cellular tra ¤ cking of geminiviruses. *Physiological and Molecular Plant Pathology*, **60**, 231–241.

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.

Garcia-Diaz, M. & Bebenek, K. (2007) Multiple functions of DNA polymerases. *Critical reviews in plant sciences*, **26**, 105–122.

Ge, L., Zhang, J., Zhou, X. & Li, H. (2007) Genetic structure and population variability of tomato yellow leaf curl China virus. *Journal of Virology*, **81**, 5902–5907.

Geoghegan, J.L. & Holmes, E.C. (2017) Predicting virus emergence amid evolutionary noise. *Open Biology*, **7**, 170189.

Ghanim, M. (2014) A review of the mechanisms and components that determine the transmission efficiency of Tomato yellow leaf curl virus (Geminiviridae; Begomovirus) by its whitefly vector. *Virus Research*, **186**, 47–54.

Gnankiné, O., Mouton, L., Savadogo, A., Martin, T., Sanon, A., Dabire, R.K., Vavre, F. & Fleury, F. (2013) Biotype status and resistance to neonicotinoids and carbosulfan in Bemisia tabaci (Hemiptera: Aleyrodidae) in Burkina Faso, West Africa. *International Journal of Pest Management*, 1–8.

Gray, S., Cilia, M. & Ghanim, M. (2014) *Circulative, "Nonpropagative" virus transmission: An orchestra of virus-, insect-, and plant-derived instruments*, *Advances in Virus Research*, 2014, vol. 89, p. 141-199.

Gray, S.M. & Banerjee, N. (1999) Mechanisms of Arthropod Transmission of Plant and Animal Viruses. *Microbiology and Molecular Biology Reviews*, **63**, 128–148.

Grubman, M. & Baxt, B. (2004) Foot-and-mouth disease. *Clinical Microbiology Reviews*, **17**, 465–493.

Gu, Y.H., Tao, X., Lai, X.J., Wang, H.Y. & Zhang, Y.Z. (2014) Exploring the polyadenylated RNA virome of sweet potato through high-throughput sequencing. *PLoS ONE*, **9**, 1–10.

Gueguen, N., Vavre, F., Gnankine, O., Charif, D., Chiel, E. & Gottlieb, Y. (2010) Endosymbiont metacommunities , mtDNA diversity and the evolution of the Bemisia tabaci (Hemiptera : Aleyrodidae) species complex. *Molecular Ecology*, **19**, 4365–4378.

Guo, J., Lai, X.P., Li, J.X., Yue, J.Q., Zhang, S.Y., Li, Y.Y., Gao, J.Y., Wang, Z.R., Duan, H.F. & Yang, J.D. (2015) First Report on Citrus Chlorotic Dwarf Associated Virus on Lemon in Dehong Prefecture, Yunnan, China. *Plant Disease*, **99**, 1287.

Gutierrez, C. (1999) Geminivirus DNA replication. *Cellular and Molecular Life Sciences*, **56**, 313–329.

Hadfield, J., Martin, D.P., Stainton, D., Kraberger, S., Owor, B.E., Shepherd, D.N., Lakay, F., Markham, P.G., Greber, R.S., Briddon, R.W. & Varsani, A. (2011) Bromus catharticus striate mosaic virus : a new mastrevirus infecting Bromus catharticus from Australia. *Archives of Virology*, **156**, 335–341.

Hadfield, J., Thomas, J.E., Schwinghamer, M.W., Kraberger, S., Stainton, D., Dayaram, A., Parry, J.N., Pande, D., Martin, D.P. & Varsani, A. (2012) Molecular characterisation of dicot-infecting mastreviruses from Australia. *Virus Research*, **166**, 13–22.

Haible, D., Kober, S. & Jeske, H. (2006) Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *Journal of Virological Methods*, **135**, 9–16.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D. & Mansoor, S. (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology*, **11**, 777–788.

Hanley-Bowdoin, L., Elmer, J.S. & Rogers, S.G. (1989) Functional Expression of the Leftward Open Reading Frames of the A Component of Tomato Golden Mosaic Virus in Transgenic Tobacco Plants. *The Plant Cell*, **1**, 1057–1067.

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Robertson, D., Settlage, S.B., Orozco, B.M., Hanley-bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S. & Robertson, D. (1999) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Plant Sciences*, **18**, 71-106.

Hanssen, I.M., Lapidot, M. & Thomma, B.P.H.J. (2010) Emerging Viral Diseases of Tomato Crops. *Molecular Plant- Microbe Interactions*, **23**, 539–548.

Harimalala, M., De Bruyn, A., Hoareau, M., Andrianjaka, A., Ranomenjanahary, S., Reynaud, B., Lefeuvre, P. & Lett, J.-M. (2013) Molecular characterization of a new alphasatellite associated with a cassava mosaic geminivirus in Madagascar. *Archives of Virology*, **158**, 1829–32.

Harrison, B.D. & Robinson, D.J. (1999) Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). *Annual Review of Phytopathology*, **37**, 369–398.

Herms, D.A. & Mattson, W.J. (1992) The Dilemma of Plants: To Grow or Defend. *The Quarterly Review of Biology*, **67**, 283–335.

Holmes, E.C. & Rambaut, A. (2004) Viral evolution and the emergence of SARS coronavirus. *Philosophical Transactions of the Royal Society B: Biological*

Sciences, **359**, 1059–1065.

Horn, N.M., Reddy, S. V., Roberts, I.M. & Reddy, D.V.R. (1993) Chickpea chlorotic dwarf virus, a new leafhopper-transmitted geminivirus of chickpea in India. *Annals of Applied Biology*, **122**, 467–479.

Horowitz, A.R., Kontsedalov, S., Khasdan, V. & Ishaaya, I. (2005) Biotypes B and Q of Bemisia tabaci and their relevance to neonicotinoid and pyriproxyfen resistance. *Archives of Insect Biochemistry and Physiology*, **58**, 216–225.

Hou, Y.-M., Paplomatas, E.J. & Gilbertson, R.L. (1998) Host Adaptation and Replication Properties of Two Bipartite Geminiviruses and Their Pseudorecombinants. *Molecular Plant-Microbe Interactions*, **11**, 208–217.

Hull, R. (2014) *Plant Virology*, Fifth Edit. (ed. by Academic Press).

Idris, A., Al-Saleh, M., Piatek, M.J., Al-Shahwan, I., Ali, S. & Brown, J.K. (2014) Viral metagenomics: Analysis of begomoviruses by illumina high-throughput sequencing. *Viruses*, **6**, 1219–1236.

Idris, A.L.I.M. & Brown, J.K. (2002) Molecular Analysis of Cotton Leaf Curl Virus-Sudan Reveals an Evolutionary History of Recombination. *Virus Genes*, **24**, 249–256.

Idris, A.M. & Brown, J.K. (2005) Evidence for interspecific-recombination for three monopartite begomoviral genomes associated with the tomato leaf curl disease from central Sudan Brief Report. *Archives of Virology*, **150**, 1003-1012.

Idris, A.M., Shahid, M.S., Briddon, R.W., Khan, a J., Zhu, J.-K. & Brown, J.K. (2011) An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. *Journal of General Virology*, **92**, 706–17.

Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B. & Nagata, T. (2004) A simple method for cloning the complete begomovirus genome using the

bacteriophage phi29 DNA polymerase. *Journal of Virological Methods*, **116**, 209–211.

Ioannou, N., Kyriakou, A. & Hadjinicolis, A. (1987) Host range and natural reservoirs of tomato yellow leaf curl virus. *Cyprus Agricultural Research Institute Technical Bulletin.*, **85**, 1–8.

Isnard, M., Granier, M., Frutos, R., Reynaud, B. & Peterschmitt, M. (1998) Quasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars. *Journal of General Virology*, **79**, 3091–3099.

Jenkins, G.M., Rambaut, A., Pybus, O.G. & Holmes, E.C. (2002) Rates of molecular evolution in RNA viruses: A quantitative phylogenetic analysis. *Journal of Molecular Evolution*, **54**, 156–165.

Jeske, H. (2009) *Geminivirus. TT Viruses - The Still Elusive Human Pathogens* (ed. by E.-M. De Villiers) and H. Zur Hausen), pp. 185–226. Springer-Verlag Berlin Heidelberg.

Jeske, H., Lu, M. & Preiû, W. (2001) DNA forms indicate rolling circle and recombination- dependent replication of Abutilon mosaic virus. *The EMBO Journal*, **20**, 6158–6167.

Jones, D.R. (2003) Plant viruses transmitted by white ies. *European Journal of Plant Pathology*, **109**, 195–219.

Jones, R. (2009) Plant virus emergence and evoution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, **141**, 113–130.

Jovel, J., Reski, G., Rothenstein, D., Ringel, M., Frischmuth, T. & Jeske, H. (2004) Sida micrantha mosaic is associated with a complex infection of begomoviruses different from Abutilon mosaic virus. *Archives of Virology*, **149**, 829–841.

Kato, K., Onuki, M., Fuji, S. & Hanada, K. (1998) The first occurrence of tomato yellow leaf curl virus in tomato (Lycopersicon esculentum Mill.) in Japan. *Annals of the Phytopathological Society of Japan*, **64**, 552–559.

Klute, K.A., Nadler, S.A. & Stenger, D.C. (1996) Horseradish curly top virus is a distinct subgroup II geminivirus species with rep and C4 genes derived from a subgroup III ancestor. *Journal of General Virology*, **77**, 1369–1378.

Kon, T. & Gilbertson, R.L. (2011) 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. *Archives of Virology*, **157**, 107–20.

Kon, T., Rojas, M.R., Abdourhamane, I.K. & Gilbertson, R.L. (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *Journal of General Virology*, **90**, 1001–1013.

Konaté, G., Barro, N., Fargette, D. & Swanson, M.M. (1995) Occurrence of whitefly-transmitted geminiviruses in crops in Burkina Faso, and their serological detection and differentiation. *Annals of Applied Biology*, **126**, 121–130.

Koonin, E. V. & Ilyina, T. V. (1992) Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *Journal of General Virology*, **73**, 2763–2766.

Kraberger, S. (2014) Towards understanding mastrevirus dynamics and the use of viral metagenomic approaches to identify novel gemini-like circular DNA viruses. PhD, University of Canterbury, 317p.

Kraberger, S., Argüello-Astorga, G.R., Greenfield, L.G., Galilee, C., Law, D., Martin, D.P. & Varsani, A. (2015a) Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. *Infection, Genetics and Evolution*, **31**, 73–86.

Kraberger, S., Harkins, G.W., Kumari, S.G., Thomas, J.E., Schwinghamer, M.W.,

Sharman, M., Collings, D.A., Briddon, R.W., Martin, D.P. & Varsani, A. (2013) Evidence that dicot-infecting mastreviruses are particularly prone to inter-species recombination and have likely been circulating in Australia for longer than in Africa and the Middle East. *Virology*, **444**, 282–291.

Kraberger, S., Kumari, S.G., Hamed, A.A., Gronenborn, B., Thomas, J.E., Sharman, M., Harkins, G.W., Muhire, B.M., Martin, D.P. & Varsani, A. (2015b) Molecular diversity of Chickpea chlorotic dwarf virus in Sudan: High rates of intra-species recombination - a driving force in the emergence of new strains. *Infection, Genetics and Evolution*, **29**, 203–215.

Kraberger, S., Saumtally, S., Pande, D., Khoodoo, M.H.R., Dhayan, S., Dookun-Saumtally, A., Shepherd, D.N., Hartnady, P., Atkinson, R., Lakay, F., Hanson, B., Redhi, D., Monjane, A.L., Windram, O., Walters, M., Oluwafemi, S., Michel-Lett, J., Lefeuvre, P., Martin, D.P. & Varsani, A. (2017) Molecular diversity, geographic distribution and host range of monocot-infecting mastreviruses in Africa and surrounding islands. *Virus Research*, **238**, 171–178.

Krake, L.R., Rezaian, M.A. & Dry, I.B. (1998) Expression of the Tomato Leaf Curl Geminivirus C4 Gene Produces Viruslike Symptoms in Transgenic Plants. *Molecular Plant- Microbe Interactions*, **11**, 413–417.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I. & Simon, R. (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology*, **388**, 1–7.

Krupovic, M. (2013) Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Current Opinion in Virology*, **3**, 578–586.

Krupovic, M., Ravantti, J.J. & Bamford, D.H. (2009) Geminiviruses: a tale of a plasmid becoming a virus. *BMC Evolutionary Biology*, **9**, 112.

Kumar, J., Kumar, J., Singh, S.P. & Tuli, R. (2014) Association of Satellites with a Mastrevirus in Natural Infection: complexity of Wheat dwarf India virus disease.

Journal of Virology, **88**, 7093–7104.

Kumari, S.G., Makkouk, K.M., Attar, N., Ghulam, W. & Lesemann, D.E. (2004) First report of chickpea chlorotic dwarf virus infecting spring chickpea in Syria. *Plant Disease*, **88**, 424.

Labonté, J.M. & Suttle, C.A. (2013) Previously unknown and highly divergent ssDNA viruses populate the oceans. *The ISME Journal*, **7**, 2169–2177.

Lefeuvre, P. (2008) Recombinaison et Émergence Virale : Le modèle des Begomovirus. Thèse de doctorat, Université de La Réunion, 88p.

Lefeuvre, P., Lett, J.-M., Varsani, a & Martin, D.P. (2009) Widely conserved recombination patterns among single-stranded DNA viruses. *Journal of Virology*, **83**, 2697–2707.

Lefeuvre, P., Martin, D.P., Harkins, G., Lemey, P., Gray, A.J. a, Meredith, S., Lakay, F., Monjane, A., Lett, J.-M., Varsani, A. & Heydarnejad, J. (2010) The spread of tomato yellow leaf curl virus from the Middle East to the world. *PLoS Pathogens*, **6**, e1001164.

Lefeuvre, P. & Moriones, E. (2015) Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. *Current Opinion in Virology*, **10**, 14–19.

Legg, J., Kumar, P.L., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., Ntawuruhunga, P. & Cuellar, W. (2014a) Cassava Virus Diseases: Biology, Epidemiology, and Management. *Advances in Virus Research*, 85–142.

Legg, J., Somado, E.A., Barker, I., Beach, L., Ceballos, H., Cuellar, W., Elkhoury, W., Gerling, D., Helsen, J., Hershey, C., Jarvis, A., Kulakow, P., Kumar, L., Lorenzen, J., Lynam, J., McMahon, M., Maruthi, G., Miano, D., Mtunda, K., Natwuruhunga, P., Okogbenin, E., Pezo, P., Terry, E., Thiele, G., Thresh, M., Wadsworth, J., Walsh, S., Winter, S., Tohme, J. & Fauquet, C. (2014b) A global alliance declaring war on cassava viruses in Africa. *Food Security*, **6**, 231–248.

Legg, J.P. & Fauquet, C.M. (2004) Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology*, **56**, 585–599.

Legg, J.P., Ndjelassili, F. & Okao-Okuja, G. (2004) First report of cassava mosaic disease and cassava mosaic geminiviruses in Gabon. *Plant Pathology*, **53**, 232.

Legg, J.P. & Ogwal, S. (1998) Changes in the incidence of African cassava mosaic virus disease and the abundance of its whitefly vector along south–north transects in Uganda. *Journal of Applied Entomology*, **122**, 169–178.

Leke, W.N., Brown, J.K., Ligthart, M.E., Sattar, N., Njualem, D.K. & Kvarnheden, A. (2012) Ageratum conyzoides: a host to a unique begomovirus disease complex in Cameroon. *Virus Research*, **163**, 229–37.

Leke, W.N., Kvarnheden, A., Ngane, E.B., Titanji, V.P.K. & Brown, J.K. (2011) Molecular characterization of a new begomovirus and divergent alphasatellite from tomato in Cameroon. *Archives of Virology*, **156**, 925–8.

Leke, W.N., Mignouna, D.B., Brown, J.K. & Kvarnheden, A. (2015) Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. *Agriculture & Food Security*, **4**, 1.

Leke, W.N., Sattar, M.N., Ngane, E.B., Ngeve, J.M., Kvarnheden, A. & Brown, J.K. (2013) Molecular characterization of begomoviruses and DNA satellites associated with okra leaf curl disease in Cameroon. *Virus Research*, **174**, 116–125.

Lett, J.M., Granier, M., Hippolyte, I., Grondin, M., Royer, M., Blanc, S., Reynaud, B. & Peterschmitt, M. (2002) Spatial and temporal distribution of geminiviruses in leafhoppers of the genus Cicadulina monitored by conventional and quantitative polymerase chain reaction. *Phytopathology*, **92**, 65–74.

Lett, J.M., Lefeuvre, P., Couston, L., Hoareau, M., Thierry, M., Reynaud, B., Martin, D.P. & Varsani, A. (2009) Complete genomic sequences of Tomato yellow leaf curl Mali virus isolates infecting tomato and pepper from the North Province of Cameroon. *Archives of Virology*, **154**, 535–540.

Lett, J.M., Reynaud, B. & Peterschmitt, M. (2001) Les mécanismes de la transmission circulante des virus de plante par insecte piqueurs - suceurs. *Virologie*, **5**, 169–181.

Li, Y., Qin, L., Zhao, J., Muhammad, T., Cao, H., Li, H., Zhang, Y. & Liang, Y. (2017) SlMAPK3 enhances tolerance to tomato yellow leaf curl virus (TYLCV) by regulating salicylic acid and jasmonic acid signaling in tomato (Solanum lycopersicum). *PLoS ONE*, **12**, 1–21.

Liang, P., Navarro, B., Zhang, Z., Wang, H., Lu, M., Xiao, H., Wu, Q., Zhou, X., Di Serio, F. & Li, S. (2015) Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. *Journal of General Virology*, **96**, 2411–2420.

Liu, L., Van Tonder, T., Pietersen, G., Davies, J.W. & Stanley, J. (1997) Molecular characterization of a subgroup I geminivirus from a legume in South Africa. *Journal of General Virology*, **78**, 2113–2117.

Lobin, K., Druffel, K.L., Pappu, H.R. & Benimadhu, S.P. (2010) First report of Tomato yellow leaf curl virus in tomato in Mauritius. *Plant Disease*, **94**, 1261.

Loconsole, G., Saldarelli, P., Doddapaneni, H., Savino, V., Martelli, G.P. & Saponari, M. (2012) Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. *Virology*, **432**, 162–172.

Lozano, G., Trenado, H.P., Fiallo-Olivé, E., Chirinos, D., Geraud-Pouey, Francis Briddon, R.W. & 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. *Frontiers in Microbiology*, **7**, 1–13.

Lu, Q.-Y., Wu, Z.-J., Xia, Z.-S. & Xie, L.-H. (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (Morus alba L.). *Archives of Virology*, **160**, 1–4.

Ma, Y., Navarro, B., Zhang, Z., Lu, M., Zhou, X., Chi, S., Serio, F. Di & Li, S. (2015) Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease. *Journal of General Virology*, **96**, 2421–2434.

Mabvakure, B., Martin, D.P., Kraberger, S., Cloete, L., van Brunschot, S., Geering, A.D.W., Thomas, J.E., Bananej, K., Lett, J.M., Lefeuvre, P., Varsani, A. & Harkins, G.W. (2016) Ongoing geographical spread of Tomato yellow leaf curl virus. *Virology*, **498**, 257–264.

Mahuku, G., Lockhart, B.E., Wanjala, B., Jones, M.W., Kimunye, J.N., Stewart, L.R., Cassone, B.J., Sevgan, S., Nyasani, J.O., Kusia, E., Kumar, P.L., Niblett, C.L., Kiggundu, A., Asea, G., Pappu, H.R., Wangai, A., Prasanna, B.M. & Redinbaugh, M.G. (2015) Maize Lethal Necrosis (MLN), an Emerging Threat to Maize-Based Food Security in Sub-Saharan Africa. *Phytopathology*, **105**, 956–965.

Malmstrom, C.M., Melcher, U. & Bosque-Pérez, N. a. (2011) The expanding field of plant virus ecology: Historical foundations, knowledge gaps, and research directions. *Virus Research*, **159**, 84–94.

Mansoor, S., Khan, S.H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. a., Briddon, R., Stanley, J. & Markham, P.G. (1999) Identification of a Novel Circular Single-Stranded DNA Associated with Cotton Leaf Curl Disease in Pakistan. *Virology*, **259**, 190–199.

Manzoor, M.T., Ilyas, M., Shafiq, M., Haider, M.S., Shahid, A.A. & Briddon, R.W. (2014) A distinct strain of chickpea chlorotic dwarf virus (genus Mastrevirus, family Geminiviridae) identified in cotton plants affected by leaf curl disease. *Archives of Virology*, **159**, 1217–1221.

Martin, D.P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P. & Varsani, A. (2011a) Recombination in eukaryotic single stranded DNA viruses. *Viruses*, **3**, 1699–1738.

Martin, D.P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.Y., Dijoux, B.,

Vincent, C., Reynaud, B. & Lett, J.M. (2011b) Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. *PLoS Pathogens*, **7**.

Martin, D.P. & Shepherd, D.N. (2009) The epidemiology , economic impact and control of maize streak disease. *Food Security*, **1**, 305–315.

Martin, D.P., Van Walt, E. Der, Posada, D. & Rybicki, E.P. (2005) The evolutionary value of recombination is constrained by genome modularity. *PLoS Genetics*, **1**, 0475–0479.

Martin, R.R., Constable, F. & Tzanetakis, I.E. (2016) Quarantine Regulations and the Impact of Modern Detection Methods. *Annual Review of Phytopathology*, **54**, 189–205.

Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M., Reynard, J.S., Rumbou, A., Saldarelli, P., Škoric, D., Vainio, E.J., Valkonen, J.P.T., Vanderschuren, H., Varveri, C. & Wetzel, T. (2017) A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Frontiers in Microbiology*, **8**, 1–7.

Mbanzibwa, D.R., Tugume, A.K., Chiunga, E., Mark, D. & Tairo, F.D. (2014) Small RNA deep sequencing-based detection and further evidence of DNA viruses infecting sweetpotato plants in Tanzania. *Annals of Applied Biology*, **165**, 329–339.

Melgarejo, T.A., Kon, T., Rojas, M.R., Paz-carrasco, L., Zerbini, F.M. & Gilbertson, R.L. (2013) Characterization of a New World Monopartite Begomovirus Causing Leaf Curl Disease of Tomato in Ecuador and Peru Reveals a New Direction in Geminivirus Evolution. *Journal of Virology*, **87**, 5397–5413.

Miano, D.W., LaBonte, D.R., Clark, C.A., Valverde, R.A., Hoy, M.W., Hurtt, S. & Li, R. (2006) First report of a begomovirus infecting sweetpotato in Kenya. *Plant Disease*, **90**, 832.

Mokili, J.L., Rohwer, F. & Dutilh, B.E. (2012) Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology*, **2**, 63–77.

Momotaz, A., Scott, J.W. & Schuster, D.J. (2010) Identification of Quantitative Trait Loci Conferring Resistance to Bemisia tabaci in an F2 Population of Solanum lycopersicum × Solanum habrochaites Accession LA1777. *Journal of the American Society for Horticultural Science*, **135**, 134–142.

Monci, F., Sanchez-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.

Monde, G., Walangululu, J., Winter, S. & Bragard, C. (2010) Dual infection by cassava begomoviruses in two leguminous species (Fabaceae) in Yangambi, Northeastern Democratic Republic of Congo. *Archives of Virology*, **155**, 1865–1869.

Morilla, G., Janssen, D., García-Andrés, S., Moriones, E., Cuadrado, I.M. & Bejarano, E.R. (2005) Pepper (Capsicum annuum) is a dead-end host for Tomato yellow leaf curl virus. *Phytopathology*, **95**, 1089–1097.

Morilla, G., Jeske, H., Bejarano, E.R. & Wege, C. (2004) Tête à Tête of Tomato Yellow Leaf Curl Virus and Tomato Yellow Leaf Curl Sardinia Virus in Single Nuclei. *Journal of Virology*, **78**, 10715–10723.

Moriones, E. & Navas-Castillo, J. (2000) Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. *Virus Research*, **71**, 123–134.

Morris, B.A.M., Richardson, K.I.M.A., Haley, A.N.N., Zhan, X. & Thomas, J.E. (1992) The Nucleotide Sequence of the Infectious Cloned DNA Component of Tobacco Yellow Dwarf Virus Reveals Features of Geminiviruses Infecting Monocotyledonous Plants. *Virology*, **187**, 633–642.

Mosig, G. (1998) Recombination and recombination-dependent DNA replication in bacteriophage T4. *Annual Review of Genetics*, **32**, 379–413.

Mosig, G., Gewin, J., Luder, A., Colowick, N. & Vo, D. (2001) Two recombination-dependent DNA replication pathways of bacteriophage T4, and their roles in mutagenesis and horizontal gene transfer. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 8306–8311.

Mound, L.A. (1978) *Whitefly of the world; a systematic catalogue of the Aleyrodidae (Homoptera) with host plant and natural enemy data*, Chichester, Sussex (UK) Wiley, Wiley, London, UK.

Moury, B., Fereres, A., Garcia-Arenal, F. & Lecoq, H. (2011) *Sustainable management of plant resistance to viruses. Recent advances in plant virology* (ed. by C. Caranta), M. Aranda), M. Tepfer), and J. Lopez Moya), pp. 219–236. Caister Academic Press, Wymondham.

Muhire, B., Darren, P.M., Brown, J.K., Navas-Castillo, J., Moriones, E., Zerbini, F.M., Rivera-Bustamante, R., Malathi, V.G., Briddon, R.W. & Varsani, A. (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of Virology*, **158**, 1411-1424..

Muigai, S.G., Schuster, D.J., Snyder, J.C., Scott, J.W., Bassett, M.J. & McAuslane, H.J. (2002) Mechanisms of resistance in Lycopersicon germplasm to the whitefly Bemisia argentifolii. *Phytoparasitica*, **30**, 347–360.

Mullineaux, P.M., Guerineau, F. & Accotto, G.P. (1990) Processing of complementary sense RNAs of Digitaria streak virus in its host and in transgenic tobacco. *Nucleic Acids Research*, **18**, 7259–7265.

N'Guessan, K.P., Fargette, D., Fauquet, C. & Thouvenel, J. (1992) Aspects off the epidemiology of okra leaf curl virus in Côte d'lvoire. *Tropical Pest Management*, **38**, 122–126.

Nahid, N., Amin, I., Mansoor, S., Rybicki, E.P., Van Der Walt, E. & Briddon, R.W. (2008) Two dicot-infecting mastreviruses (family Geminiviridae) occur in Pakistan. *Archives of Virology*, **153**, 1441–1451.

Nair, S.K., Babu, R., Magorokosho, C., Mahuku, G., Semagn, K., Beyene, Y., Das, B., Makumbi, D., Lava Kumar, P., Olsen, M. & Boddupalli, P.M. (2015) Fine mapping of Msv1, a major QTL for resistance to Maize Streak Virus leads to development of production markers for breeding pipelines. *Theoretical and Applied Genetics*, **128**, 1839–1854.

Navas-Castillo, J., Fiallo-Olivé, E. & Sanchez-Campos, S. (2011) Emerging Virus Diseases Transmitted by Whiteflies. *Annual Review of Phytopathology*, **49**, 219–248.

Nawaz-ul-Rehman, M.S. & Fauquet, C.M. (2009) Evolution of geminiviruses and their satellites. *FEBS Letters*, **583**, 1825–1832.

Neuenschwander, P., Hughes, J. d'A., Ogbe, F., Ngatse, J.M. & Legg, J.P. (2002) Occurrence of the Uganda variant of East African Cassava Mosaic Virus (EACMV-Ug) in Western Democratic Republic of Congo. *Plant Pathology*, **51**, 385.

Ng, T.F.F., Duffy, S., Polston, J.E., Bixby, E., Vallad, G.E. & Breitbart, M. (2011) Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies. *PLoS ONE*, **6**.

Novick, R.P. (1998) Contrasting lifestyles of rolling-circle phages and plasmids. *Trends in Biochemical Sciences*, **23**, 434–438.

Ohnishi, J., Kitamura, T., Terami, F. & Honda, K.I. (2009) A selective barrier in the midgut epithelial cell membrane of the nonvector whitefly Trialeurodes vaporariorum to Tomato yellow leaf curl virus uptake. *Journal of General Plant Pathology*, **75**, 131–139.

Oliveira, M.R. V, Henneberry, T.J. & Anderson, P. (2001) History , current status , and collaborative research projects for Bemisia tabaci. *Crop Protection*, **20**, 709–723.

Ooi, K., Ohshita, S., Ishii, I. & Yahara, T. (1997) Molecular phylogeny of geminivirus infecting wild plants in Japan. *Journal of Plant Research*, **110**, 247–257.

Osei, M., Akromah, R., Shih, S., Lee, L. & Green, S. (2008) First report and molecular characterization of three distinct Begomoviruses associated with tomato leaf curl disease in Ghana. *Plant Disease*, **92**, 1585–1585.

Ostfeld, R.S. & Keesing, F. (2012) Effects of Host Diversity on Infectious Disease. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 157–182.

Otoidobiga, L.C., Vincent, C. & Stewart, R.K. (2003) Susceptibility of field populations of adult Bemisia tabaci Gennadius (Homoptera: Aleyrodidae) and Eretmocerus sp (Hymenoptera: Aphelinidae) to cotton insecticides in Burkina Faso (West Africa). *Pest Management Science*, **59**, 97–106.

Owor, B.E., Martin, D.P., Shepherd, D.N., Edema, R., Monjane, A.L., Rybicki, E.P., Thomson, J.A. & Varsani, A. (2007) Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. *Journal of General Virology*, **88**, 3154–3165.

Padidam, M., Sawyer, S. & Fauquet, C.M. (1999) Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology*, **265**, 218–225.

Padidam, R., Beachy, R.N. & Fauquet, C.M. (1995) Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *Journal of General Virology*, **76**, 25–35.

Papayiannis, L.C., Katis, N.I., Idris, A.M. & Brown, J.K. (2011) Identification of Weed Hosts of *Tomato yellow leaf curl virus* in Cyprus. *Plant Disease*, **95**, 120–125.

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

Patel, V.P., Rojas, M.R., Paplomatas, E.J. & Gilbertson, R.L. (1993) Cloning biologically active geminivirus DNA using PCR and overlapping Primers. *Nucleic Acids Research*, **21**, 1325–1326.

Patil, B.L. & Fauquet, C.M. (2009) Cassava mosaic geminiviruses : actual knowledge and perspectives. *Plant Molecular Pathology*, **10**, 685–701.

Patil, B.L. & Fauquet, C.M. (2010) Differential interaction between cassava mosaic geminiviruses and geminivirus satellites. *Journal of General Virology*, **91**, 1871–82.

Patil, B.L., Legg, J.P., Kanju, E. & Fauquet, C.M. (2015) Cassava brown streak disease: A threat to food security in Africa. *Journal of General Virology*, **96**, 956–968.

Péréfarres, F., De Bruyn, A., Kraberger, S., Hoareau, M., Barjon, F., Lefeuvre, P., Pellegrin, F., Caplong, P., Varsani, A. & Lett, J.-M. (2012) Occurrence of the Israel strain of Tomato yellow leaf curl virus in New Caledonia and Loyalty Islands. *New Disease Reports*, **25**, 5197.

Péréfarres, F., Thébaud, G., Lefeuvre, P., Chiroleu, F., Rimbaud, L., Hoareau, M., Reynaud, B. & Lett, J.-M. (2014) Frequency-dependent assistance as a way out of competitive exclusion between two strains of an emerging virus. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 1–9.

Perring, T.M. (2001) The Bemisia tabaci species complex. *Crop Protection*, **20**, 725–737.

Peterschmitt, M., Granier, M., Mekdoud, R., Dalmon, A., Gambin, O., Vayssieres, J.F. & Reynaud, B. (1999) First report of tomato yellow leaf curl virus in Réunion Island. *Plant Disease*, **83**, 303.

Pialoux, G., Gaüzère, B.-A., Jauréguiberry, S. & Strobel, M. (2007) Review Chikungunya, an epidemic arbovirosis. *The Lancet Infectious Diseases*, **7**, 319–327.

Picó, B., Díez, M.J. & Nuez, F. (1996) Viral diseases causing the greatest economic losses to the tomato crop. II. The Tomato yellow leaf curl virus — a review. *Scientia Horticulturae*, **67**, 151–196.

Pilartz, M. & Jeske, H. (1992) Abutilon mosaic geminivirus double-stranded DNA is packed into minichromosomes. *Virology*, **189**, 800–802.

Pita, J.S., Fondong, V.N., Sangar, A., Kokora, R.N.N. & Fauquet, C.M. (2001a) Genomic and biological diversity of the African cassava geminiviruses. *Euphytica*, **120**, 115–125.

Pita, J.S., Fondong, V.N., Sangare, A., Ogwal, S. & Fauquet, C.M. (2001b) Recombination , pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology*, **82**, 655–665.

Polston, J.E., Bois, D., Keinath, A.P. & Chellemi, D.O. (1995) Occurrence of tomato mottle geminivirus in South Carolina, Tennessee, and Virginia. *Plant Disease*, **79**, 539.

Polston, J.E., Bois, D., Serra, C.A. & Concepción, S. (1994) First report of a tomato yellow leaf curl-like geminivirus in the Western Hemisphere. *Plant Disease*, **78**, 831.

Polston, J.E. & Lapidot, M. (2007) *Management of Tomato yellow leaf curl virus: US and Israel Perspectives BT - Tomato Yellow Leaf Curl Virus Disease: Management, Molecular Biology, Breeding for Resistance*. *Tomato Yellow Leaf Curl Virus Disease* (ed. by H. Czosnek), pp. 251–262. Springer Netherlands, Dordrecht.

Preiss, W. & Jeske, H. (2003) Multitasking in replication is common among geminiviruses. *Journal of Virology*, **77**, 2972–2980.

Prins, M., Laimer, M., Noris, E., Schubert, J., Wassenegger, M. & Tepfer, M. (2008) Strategies for antiviral resistance in transgenic plants. *Molecular Plant Pathology*, **9**, 73–83.

Rector, A., Tachezy, R. & Ranst, M. Van (2004) A Sequence-Independent Strategy for Detection and Cloning of Circular DNA Virus Genomes by Using Multiply Primed

Rolling-Circle Amplification. *Journal of Virology*, **78**, 4993–4998.

Rey, M.E.C., Ndunguru, J., Berrie, L.C., Paximadis, M., Berry, S., Cossa, N., Nuaila, V.N., Mabasa, K.G., Abraham, N., Rybicki, E.P., Martin, D., Pietersen, G. & Esterhuizen, L.L. (2012) Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in southern Africa, including the South-west Indian ocean islands. *Viruses*, **4**, 1753–91.

Reynaud, B. (1988) Transmission des virus de la striure, du stripe et de la mosaïque du maïs par leurs vecteurs Cicadulina mbila (Naude, 1924) et Peregrinus maidis (Ashmead, 1890)(Homoptera). Thèse de Doctorat. Université de la Réunion, 130p.

Rochester, D.E., DePaulo, J.J., Fauquet, C.M. & Beachy, R.N. (1994) Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus, Thailand isolate. *Journal of General Virology*, **75**, 477–485.

Rodríguez-López, M.J., Garzo, E., Bonani, J.P., Fereres, A., Fernández-Muñoz, R. & Moriones, E. (2011) Whitefly Resistance Traits Derived from the Wild Tomato Solanum pimpinellifolium Affect the Preference and Feeding Behavior of Bemisia tabaci and Reduce the Spread of Tomato yellow leaf curl virus. *Phytopathology*, **101**, 1191–1201.

Rojas, M.R., Hagen, C., Lucas, W.J. & Gilbertson, R.L. (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, **43**, 361–394.

Rojas, M.R., Jiang, H., Salati, R., Xoconostle-ca, B., Lucas, W.J. & Gilbertson, R.L. (2001) Functional Analysis of Proteins Involved in Movement of the Monopartite Begomovirus , Tomato Yellow Leaf Curl Virus. *Virology*, **291**, 110–125.

Romay, G., Chirinos, D., Geraud-Pouey, F. & Desbiez, C. (2010) Association of an atypical alphasatellite with a bipartite New World begomovirus. *Archives of Virology*, **155**, 1843–1847.

Roossinck, M.J. (2015) Metagenomics of plant and fungal viruses reveals an abundance

of persistent lifestyles. *Frontiers in Microbiology*, **5**, 1–3.

Roossinck, M.J. (2013) Plant Virus Ecology. *PLoS Pathogens*, **9**, 9–11.

Roossinck, M.J. & García-Arenal, F. (2015) Ecosystem simplification, biodiversity loss and plant virus emergence. *Current Opinion in Virology*, **10**, 56–62.

Roossinck, M.J., Martin, D.P. & Roumagnac, P. (2015) Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology*, **105**, 716–727.

Roossinck, M.J., Saha, P., Wiley, G.B., Quan, J., White, J.D., Lai, H., Chavarría, F., Shen, G. & Roe, B.A. (2010) Ecogenomics: Using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*, **19**, 81–88.

Rosario, K., Marr, C., Varsani, A., Kraberger, S., Stainton, D., Moriones, E., Polston, J.E. & Breitbart, M. (2016) Begomovirus-associated satellite DNA diversity captured through vector-enabled metagenomic (VEM) surveys using whiteflies (Aleyrodidae). *Viruses*, **8**, 1–16.

Rosario, K., Seah, Y.M., Marr, C., Varsani, A., Kraberger, S., Stainton, D., Moriones, E., Polston, J.E., Duffy, S. & Breitbart, M. (2015) Vector-enabled metagenomic (VEM) surveys using whiteflies (Aleyrodidae) reveal novel begomovirus species in the new and old worlds. *Viruses*, **7**, 5553–5570.

Roux, S., Enault, F., Bronner, G., Vaulot, D., Forterre, P. & Krupovic, M. (2013) Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nature Communications*, **4**.

Al Rwahnih, M., Alabi, O.J., Westrick, N.M., Golino, D. & Rowhani, A. (2017) Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. *Phytopathology*, **107**, 240–251.

Saiz, J.-C., Martín-Acebes, M.A., Bueno-Marí, R., Salomón, O.D., Villamil-Jiménez, L.C., Heukelbach, J., Alencar, C.H., Armstrong, P.K., Ortiga-Carvalho, T.M., Mendez-Otero, R., Rosado-de-Castro, P.H. & Pimentel-Coelho, P.M. (2017) Zika

Virus: What Have We Learnt Since the Start of the Recent Epidemic? *Frontiers in Microbiology*, **8**, 1–25.

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 Research*, **1**, 1–8.

Sánchez-Campos, S., Navas-Castillo, J., Camero, R., Soria, C., Díaz, J.A. & Moriones, E. (1999) Displacement of Tomato Yellow Leaf Curl Virus (TYLCV) -Sr by TYLCV-Is in Tomato Epidemics in Spain. *Virology*, **89**, 1039–1043.

Sanderfoot, A.A., Ingham, D.J. & Lazarowitz, S.C. (1996) Vira1 Movement Protein as a Nuclear Shuttle: The Geminivirus BR1 Movement Protein Contains Domains Essential for lnteraction with BL1 and Nuclear Localization. *Plant Physiology*, **110**, 23–33.

Sattar, M.N., Koutou, M., Husseini, S., Leke, W.N., Brown, J.K. & Kvarnheden, A. (2015) First Identification of Begomoviruses Infecting Tomato With Leaf Curl Disease in Burkina Faso. *Plant Disease*, **99**, 732–732.

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

Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M. & Stanley, J. (2000) A unique virus complex causes Ageratum yellow vein disease. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6890–6895.

Saunders, K., Salim, N., Mali, V.R., Malathi, V.G., Briddon, R., Markham, P.G. & Stanley, J. (2002) Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology*, **293**, 63–74.

Saunders, K. & Stanley, J. (1999) A nanovirus-like DNA component associated with yellow vein disease of Ageratum conyzoides: evidence for interfamilial recombination between plant DNA viruses. *Virology*, **264**.

Seal, S.E., VandenBosch, F. & Jeger, M.J. (2006) Factors Influencing Begomovirus Evolution and Their Increasing Global Significance: Implications for Sustainable Control. *Critical Reviews in Plant Sciences*, **25**, 23–46.

Séka, K., Ouattara, A., Assiri, K.P., Kra, K.D., Hoareau, M., Lefeuvre, P., Diallo, H.A. & Lett, J.M. (2017) First report of Pepper yellow vein Mali virus associated with pepper yellow vein disease in Cote d'Ivoire. *New Disease Reports*, **35**, 5197.

Servin, V.R., Martinez-Carrillo, J.L. & Hirales, L. (1999) Weeds and cultivated hosts of the silverleaf whitefly Bemisia argentifolii Bellows and Perring in Baja California Sur, Mexico. *Southwestern Entomologist*, **v. 24**.

Shepherd, D.N., Martin, D.P., Lefeuvre, P., Monjane, A.L., Owor, B.E., Rybicki, E.P. & Varsani, A. (2008) A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. *Journal of Virological Methods*, **149**, 97–102.

Shepherd, D.N., Martin, D.P. & Thomson, J.A. (2009) Transgenic strategies for developing crops resistant to geminiviruses. *Plant Science*, **176**, 1–11.

Shepherd, D.N., Martin, D.P., Van Der Walt, E., Dent, K., Varsani, A. & Rybicki, E.P. (2010a) Maize streak virus: an old and complex "emerging" pathogen. *Molecular Plant Pathology*, **11**, 1–12.

Shepherd, D.N., Martin, D.P., Walt, E.V.A.N.D.E.R., Dent, K., Varsani, A. & Rybicki, E.P. (2010b) Pathogen profile Maize streak virus : an old and complex " emerging " pathogen. *Molecular Plant Pathology*, **11**, 1–12.

Shih, S.L., Green, S.K., Tsai, W.S., Lee, L.M. & Levasseur, V. (2007) First report of a distinct begomovirus associated with okra yellow crinkle disease in Mali. *Plant Pathology*, **56**, 718.

Shih, S.L., Kumar, S., Tsai, W.S., Lee, L.M. & Green, S.K. (2009) Complete nucleotide sequences of okra isolates of Cotton leaf curl Gezira virus and their associated DNA- b from Niger. *Archives of Virology*, **154**, 369–372.

Shivaprasad, P. V, Akbergenov, R., Trinks, D., Veluthambi, K., Hohn, T., Mikhail, M., Rajeswaran, R. & Pooggin, M.M. (2005) Promoters , Transcripts , and Regulatory Proteins of Mungbean Yellow Mosaic Geminivirus Promoters , Transcripts , and Regulatory Proteins of Mungbean Yellow Mosaic Geminivirus. *Journal of Virology*, **79**, 8149–8163.

Sivalingam, P.N. & Varma, A. (2012) Role of betasatellite in the pathogenesis of a bipartite begomovirus affecting tomato in India. *Archives of Virology*, **157**, 1081–1092.

Stemmer, W.P.C. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proceedings of the National Academy of Sciences*, **91**, 10747–10751.

Storey, H.H. (1938) Investigations of the Mechanism of the Transmission of Plant Viruses by Insect Vectors. II. The Part Played by Puncture in Transmission. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **125**, 455–477.

Sudarshana, M.R., Perry, K.L. & Fuchs, M.F. (2015) Grapevine Red Blotch-Associated Virus, an Emerging Threat to the Grapevine Industry. *Phytopathology*, **105**, 1026–1032.

Sunter, G. & Bisaro, D.M. (1997) Regulation of a Geminivirus Coat Protein Promoter by AL2 Protein (TrAP): Evidence for Activation and Derepression Mechanisms. *Virology*, **232**, 269–280.

Sunter, G. & Bisaro, D.M. (1991) Transactivation in a Geminivirus: AL2 Gene Product Is Needed for Coat Protein Expression. *Virology*, **180**, 416–419.

Sunter, G. & Bisaro, D.M. (1989) Transcription map of the B genome component of

tomato golden mosaic virus and comparison with A component transcripts. *Virology*, **173**, 647–655.

Suttle, C.A. (2005) Viruses in the sea. *Nature*, **437**, 356–361.

Swanson, M.M. & Harrison, B.D. (1994) Properties, relationships and distribution of cassava mosaic geminiviruses. *Tropical Science*, **34**, 15–25.

Swanson, M.M. & Harrison, B.D. (1993) Serological relationships and epitope profile of an isolate of okra leaves curl geminivirus from Africa and Middle East. *Biochimie*, **75**, 707–711.

Thomas, J.E., Parry, J.N., Schwinghamer, M.W. & Dann, E.K. (2010) Two novel mastreviruses from chickpea (Cicer arietinum) in Australia. *Archives of Virology*, **155**, 1777–1788.

Thresh, J.M. (1980) The origins and epidemiology of some important plant virus diseases. *Applied Biology*, **5**, 2–65.

Tiendrébéogo, F.., Traoré, E.V.S., Barro, N., Konaté, G., Traoré, A. & Traoré, O. (2008) Characterization of Pepper yellow vein Mali virus in Capsicum sp. in Burkina Faso. *Plant Pathology Journal*, **7**, 155–161.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Harimalala, M. a, De Bruyn, A., Villemot, J., Traoré, V.S.E., Konaté, G., Traoré, A.S., Barro, N., Reynaud, B., Traoré, O. & Lett, J.-M. (2012) Evolution of African cassava mosaic virus by recombination between bipartite and monopartite begomoviruses. *Virology Journal*, **9**, 67.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Traoré, V.S.E., Barro, N., Péréfarres, F., Reynaud, B., Traoré, A.S., Konaté, G., Lett, J.-M. & Traoré, O. (2011) Molecular and biological characterization of Pepper yellow vein Mali virus (PepYVMV) isolates associated with pepper yellow vein disease in Burkina Faso. *Archives of Virology*, **156**, 483–487.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Traoré, V.S.E., Barro, N., Reynaud, B.,

Traoré, A.S., Konaté, G., Traoré, O. & Lett, J.M. (2009) Occurrence of East African cassava mosaic virus-Uganda (EACMV-UG) in Burkina Faso. *Plant Pathology*, **58**, 783.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Villemot, J., Konaté, G., Traoré, A.S., Barro, N., Traoré, V.S., Reynaud, B., Traoré, O. & Lett, J.-M. (2010) Molecular diversity of cotton leaf curl Gezira virus isolates and their satellite DNAs associated with okra leaf curl disease in Burkina Faso. *Virology Journal*, **7**, 48.

Traoré, N.S., Dakouo, D., Ouédraogo, I. & Traoré, D. (2003) Evolution de la densité de population de cicadulina spp dans différents biotopes de deux zones écologiques et de leur capacité de survie en saison sèche au Burkina Faso. *Journal des Sciences*, **3**, 35–43.

Uchibori, M., Hirata, A., Suzuki, M. & Ugaki, M. (2013) Tomato yellow leaf curl virus accumulates in vesicle-like structures in descending and ascending midgut epithelial cells of the vector whitefly, Bemisia tabaci, but not in those of nonvector whitefly Trialeurodes vaporariorum. *Journal of General Plant Pathology*, **79**, 115–122.

Valam-Zango, A., Zinga, I., Hoareau, M., Tocko-Marabena, B.K., Mvila, A.C., Semballa, S. & Lett, J.M. (2015) First report of cassava mosaic geminiviruses and the Uganda strain of East African cassava mosaic virus (EACMV-UG) associated with cassava mosaic disease in Equatorial Guinea. *Plant Pathology*, **32**, 29.

Vanitharani, R., Chellappan, P. & Fauquet, C.M. (2005) Geminiviruses and RNA silencing. *Trends in Plant Science*, **10**, 144–151.

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

Varsani, A., Roumagnac, P., Fuchs, M., Moriones, E., Idris, A., Briddon, R.W., Zerbini, F.M. & Martin, D.P. (2017) Capulavirus and Grablovirus : two new genera in the

family Geminiviridae. *Archives of Virology*, **162**, 1819-1831.

Varsani, A., Shepherd, D.N., Monjane, A.L., Owor, B.E., Erdmann, J.B., Rybicki, E.P., Peterschmitt, M., Briddon, R.W., Markham, P.G., Oluwafemi, S., Windram, O.P., Lefeuvre, P., Lett, J.M. & Martin, D.P. (2008) Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *Journal of General Virology*, **89**, 2063–2074.

Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.-F., Jacques, M.-A., Lebrun, M.-H., Le Loir, Y., Ogliastro, M., Petit, M.-A., Roumagnac, P. & Candresse, T. (2014) *Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics*. *Frontiers in Cellular and Infection Microbiology*, **4**.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.-H. & Smith, H.O. (2004) Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science*, **304**, 66–74.

Voinnet, O., Pinto, Y.M. & Baulcombe, D.C. (1999) Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 14147–52.

van der Walt, E., Martin, D.P., Varsani, A., Polston, J.E. & Rybicki, E.P. (2008) Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. *Virology Journal*, **5**, 104.

Wasswa, P., Otto, B., Maruthi, M.N., Mukasa, S.B., Monger, W. & Gibson, R.W. (2011) First identification of a sweet potato begomovirus (sweepovirus) in Uganda: Characterization, detection and distribution. *Plant Pathology*, **60**, 1030–1039.

Webb, M.D. (1987) Species recognition in Cicadulina leafhoppers (Hemiptera: Cicadellidae), vectors of pathogens of Gramineae. *Bulletin of Entomological*

Research, **77**, 683–712.

Webster, C.G., Coutts, B.A., Jones, R.A.C., Jones, M.G.K. & Wylie, S.J. (2007) Virus impact at the interface of an ancient ecosystem and a recent agroecosystem: studies on three legume-infecting potyviruses in the southwest Australian floristic region. *Plant Pathology*, **56**, 729–742.

Wezel, R. Van, Liu, H., Tien, P., Stanley, J. & Hong, Y. (2001) Gene C2 of the Monopartite Geminivirus Tomato yellow leaf curl virus -China Encodes a Pathogenicity Determinant That Is Localized in the Nucleus. *Molecular Plant-Microbe Interactions*, **14**, 1125–1128.

Wong, M., Grandgirard, J., Hascoat, J., Hoareau, M., Simiand, C., Reynaud, B., Lefeuvre, P., Delatte, H. & Lett, J.M. (2017) Occurrence of the Israel strain of Tomato yellow leaf curl virus and the whitefly Bemisia tabaci MEAM1 species in French Polynesia. *New Disease Reports*, **35**, 29-29.

Wright, E.A., Heckel, T., Groenendijk, J., Davies, J.W. & Boulton, M.I. (1997) Wright_et_al_1997_Splicing features in Maize streak virus virion- and complementary-sense gene expression.pdf. *The Plant Journal*, **12**, 1285–1297.

Wu, J.-X., Shang, H.-L., Xie, Y., Shen, Q.-T. & Zhou, X.-P. (2012) Monoclonal Antibodies Against the Whitefly-Transmitted Tomato Yellow Leaf Curl Virus and Their Application in Virus Detection. *Journal of Integrative Agriculture*, **11**, 263–268.

Wyant, P.S., Strohmeier, S., Schäfer, B., Krenz, B., Assunção, I.P., Lima, G.S. de A. & Jeske, H. (2012) Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. *Virology*, **427**, 151–157.

Xie, Q., Sanz-Burgos, A.P., Guo, H., García, J.A. & Gutiérrez, C. (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant Molecular Biology*, **39**, 647–656.

Yongping, Z., Weimin, Z., Huimei, C., Yang, Q., Kun, S., Yanhui, W., Longying, Z., Li, Y. & Zhang, H. (2008) Molecular identification and the complete nucleotide sequence of TYLCV isolate from Shanghai of China. *Virus Genes*, **36**, 547–551.

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

Zhang, W., Olson, N.H., Baker, T.S., Faulkner, L., Agbandje-McKenna, M., Boulton, M.I., Davies, J.W. & McKenna, R. (2001) Structure of the Maize streak virus geminate particle. *Virology*, **279**, 471–7.

Zhou, X. (2013) Advances in understanding begomovirus satellites. *Annual review of phytopathology*, **51**, 357–81.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G.W., Robinson, D.J. & Harrison, B.D. (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, **78**, 2101–2111.

Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z. & Fauquet, C.M. (2003) Characterization of DNAβ associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. *Journal of General Virology*, **84**, 237–247.

Zhou, Y.-C., Noussourou, M., Kon, T., Rojas, M.R., Jiang, H., Chen, L.-F., Gamby, K., Foster, R. & Gilbertson, R.L. (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. *Archives of Virology*, **153**, 693–706.

Zinga, I., Harimalala, M., Bruyn, A. De, Hoareau, M., Mandakombo, N., Semballa, S., Reynaud, B. & Lefeuvre, P. (2012) East African cassava mosaic virus-Uganda (EACMV-UG) and African cassava mosaic virus (ACMV) reported for the first time in Central African Republic and Chad. *New Disease Reports*, **26**, 17.

Annexes

Supp. table 1: Summary of the 2063 collected samples from 2013 to 2016 in Burkina Faso

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2013	Sahel	Kongoussi	13,309	-1,519	Yes	Malvaceae	Abelmoschus esculentus	2
2013	Sahel	Kongoussi	13,309	-1,519	Yes	Malvaceae	Hibiscus sabdariffa	1
2013	Sahel	Kongoussi	13,309	-1,519	Yes	Solanaceae	Capsicum frutescens	1
2013	Sahel	Kongoussi	13,309	-1,519	Yes	Solanaceae	Solanum lycopersicum	4
2013	Sahel	Kongoussi	13,309	-1,519	No	UD	UD	1
2013	Sahel	Yako	12,948	-2,253	Yes	Solanaceae	Solanum lycopersicum	1
2013	Soudan	Sakabi	11,224	-4,278	No	Amaranthaceae	Amaranthus hybridus	1
2013	Soudan	Sakabi	11,224	-4,278	No	Asteraceae	Ageratum sp	1
2013	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Capsicum annuum	3
2013	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Capsicum frutescens	1
2013	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Solanum lycopersicum	8
2013	Soudan_sahel	Boulmiougou	12,330	-1,587	Yes	Brassicaceae	Brassica rapa	1
2013	Soudan_sahel	Boulmiougou	12,330	-1,587	Yes	Solanaceae	Capsicum frutescens	6
2013	Soudan_sahel	Boulmiougou	12,330	-1,587	Yes	Solanaceae	Solanum lycopersicum	10
2013	Soudan_sahel	Kamboinse	12,492	-1,432	No	Convolvulaceae	Ipomoea dichroa	1
2013	Soudan_sahel	Kamboinse	12,492	-1,432	Yes	Malvaceae	Abelmoschus esculentus	5
2013	Soudan_sahel	Kamboinse	12,492	-1,432	No	Malvaceae	Corchorus sp	2
2013	Soudan_sahel	Kamboinse	12,492	-1,432	Yes	Solanaceae	Solanum lycopersicum	3
2013	Soudan_sahel	Kamboinse	12,492	-1,432	No	UD	UD	3
2013	Soudan_sahel	Loumbila	12,485	-1,417	No	Amaranthaceae	Amaranthus viridis	1
2013	Soudan_sahel	Loumbila	12,485	-1,417	No	Boraginaceae	Heliotropium indicum	1
2013	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Malvaceae	Abelmoschus esculentus	2
2013	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Capsicum annuum	4
2013	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Solanum lycopersicum	11
2013	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Solanum melongena	3
2013	Soudan_sahel	Ouagadougou	12,388	-1,507	Yes	Solanaceae	Capsicum frutescens	3
2013	Soudan_sahel	Ouagadougou	12,388	-1,507	Yes	Solanaceae	Solanum lycopersicum	10
2013	Soudan_sahel	Pabré	12,511	-1,580	Yes	Solanaceae	Capsicum frutescens	5
2013	Soudan_sahel	Pabré	12,511	-1,580	Yes	Solanaceae	Solanum lycopersicum	1
2013	Soudan_sahel	Yamtenga	12,344	-1,452	Yes	Solanaceae	Capsicum annuum	2
2013	Soudan_sahel	Yamtenga	12,344	-1,452	Yes	Solanaceae	Solanum lycopersicum	3
2013	Soudan_sahel	Ziniaré	12,568	-1,280	No	Amaranthaceae	Amaranthus viridis	2
2013	Soudan_sahel	Ziniaré	12,568	-1,280	No	Malvaceae	Sida acuta	2
2013	Soudan_sahel	Ziniaré	12,568	-1,280	Yes	Solanaceae	Solanum melongena	4
2014	Sahel	Kaya	13,180	-1,165	No	Amaranthaceae	Amaranthus sp	1
2014	Sahel	Kaya	13,180	-1,165	No	Malvaceae	Sida acuta	3
2014	Sahel	Kaya	13,180	-1,165	Yes	Solanaceae	Capsicum frutescens	10
2014	Sahel	Kaya	13,180	-1,165	No	UD	UD	2
2014	Soudan	Banfora	10,643	-4,782	No	Amaranthaceae	Amaranthus sp	1
2014	Soudan	Banfora	10,643	-4,782	Yes	Solanaceae	Capsicum frutescens	2

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2014	Soudan	Banfora	10,643	-4,782	No	UD	UD	1
2014	Soudan	Bobo-dioulasso	11,130	-4,267	No	Lamiaceae	Perilla frutescens	1
2014	Soudan	Bobo-dioulasso	11,130	-4,267	Yes	Solanaceae	Capsicum annuum	3
2014	Soudan	Bobo-dioulasso	11,130	-4,267	Yes	Solanaceae	Solanum lycopersicum	1
2014	Soudan	Sakabi	11,224	-4,278	No	Amaranthaceae	Amaranthus sp	1
2014	Soudan	Sakabi	11,224	-4,278	No	Asteraceae	Ageratum sp	1
2014	Soudan	Sakabi	11,224	-4,278	No	Fabaceae	Crotalaria spectabilis	1
2014	Soudan	Sakabi	11,224	-4,278	No	Malvaceae	Sida acuta	1
2014	Soudan	Sakabi	11,224	-4,278	No	Nyctaginaceae	Boerhavia erecta	3
2014	Soudan	Sakabi	11,224	-4,278	No	Rubiaceae	Leucaena glota	2
2014	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Capsicum frutescens	2
2014	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Solanum lycopersicum	13
2014	Soudan	Sakabi	11,224	-4,278	No	UD	UD	4
2014	Soudan	Sossogona	11,251	-4,457	Yes	Malvaceae	Abelmoschus esculentus	2
2014	Soudan	Sossogona	11,251	-4,457	No	Malvaceae	Sida acuta	1
2014	Soudan	Sossogona	11,251	-4,457	Yes	Solanaceae	Capsicum frutescens	4
2014	Soudan	Sossogona	11,251	-4,457	Yes	Solanaceae	Solanum lycopersicum	4
2014	Soudan	Toussiana	10,856	-4,640	Yes	Solanaceae	Capsicum frutescens	6
2014	Soudan	Toussiana	10,856	-4,640	Yes	Solanaceae	Solanum lycopersicum	3
2014	Soudan_sahel	Boulmiougou	12,330	-1,587	No	Lamiaceae	Ocimum basilicum	1
2014	Soudan_sahel	Boulmiougou	12,330	-1,587	Yes	Solanaceae	Capsicum annuum	2
2014	Soudan_sahel	Boulmiougou	12,330	-1,587	Yes	Solanaceae	Capsicum frutescens	1
2014	Soudan_sahel	Fada	12,064	-0,357	No	Malvaceae	Sida acuta	1
2014	Soudan_sahel	Fada	12,064	-0,357	Yes	Solanaceae	Capsicum annuum	2
2014	Soudan_sahel	Fada	12,064	-0,357	Yes	Solanaceae	Capsicum frutescens	1
2014	Soudan_sahel	Fada	12,064	-0,357	Yes	Solanaceae	Solanum lycopersicum	1
2014	Soudan_sahel	Kamboinse	12,492	-1,432	Yes	Malvaceae	Abelmoschus esculentus	1
2014	Soudan_sahel	Kamboinse	12,492	-1,432	No	Malvaceae	Sida acuta	1
2014	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Malvaceae	Abelmoschus esculentus	2
2014	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Capsicum frutescens	2
2014	Soudan_sahel	Mogtédo	12,299	-0,835	No	Amaranthaceae	Amaranthus sp	1
2014	Soudan_sahel	Mogtédo	12,299	-0,835	No	Euphorbiaceae	Euphorbia hirta	1
2014	Soudan_sahel	Mogtédo	12,299	-0,835	No	Lamiaceae	Ocimum basilicum	1
2014	Soudan_sahel	Mogtédo	12,299	-0,835	No	Malvaceae	Sida acuta	1
2014	Soudan_sahel	Mogtédo	12,299	-0,835	Yes	Solanaceae	Capsicum frutescens	5
2015	Sahel	Dakiri	13,311	-0,278	No	Nyctaginaceae	Boerhavia erecta	1
2015	Sahel	Dakiri	13,311	-0,278	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Dakiri	13,311	-0,278	No	UD	UD	4
2015	Sahel	Dem	13,178	-1,168	No	Amaranthaceae	Amaranthus sp	2
2015	Sahel	Dem	13,178	-1,168	No	Convolvulaceae	Ipomoea sp	1
2015	Sahel	Dem	13,178	-1,168	No	Euphorbiaceae	Euphorbia hirta	1
2015	Sahel	Dem	13,178	-1,168	No	Fabaceae	Crotalaria spectabilis	1
2015	Sahel	Dem	13,178	-1,168	Yes	Malvaceae	Hibiscus sabdariffa	1
2015	Sahel	Dem	13,178	-1,168	No	Solanaceae	Physalis angulata	2
2015	Sahel	Dem	13,178	-1,168	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Dem	13,178	-1,168	No	UD	UD	2
2015	Sahel	Kongoussi	13,309	-1,519	No	Amaranthaceae	Amaranthus sp	1
2015	Sahel	Kongoussi	13,309	-1,519	No	Solanaceae	Physalis angulata	2
2015	Sahel	Kongoussi	13,309	-1,519	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Kongoussi	13,309	-1,519	No	UD	UD	2
2015	Sahel	Ouahigouya	13,556	-2,449	No	Amaranthaceae	Amaranthus sp	1

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2015	Sahel	Ouahigouya	13,556	-2,449	No	Amaranthaceae	Amaranthus spinosus	2
2015	Sahel	Ouahigouya	13,556	-2,449	No	Euphorbiaceae	Euphorbia hirta	2
2015	Sahel	Ouahigouya	13,556	-2,449	No	Nyctaginaceae	Boerhavia erecta	1
2015	Sahel	Ouahigouya	13,556	-2,449	No	Solanaceae	Physalis angulata	1
2015	Sahel	Ouahigouya	13,556	-2,449	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Ouahigouya	13,556	-2,449	No	UD	UD	3
2015	Sahel	Pibaoré	12,906	-0,823	No	Amaranthaceae	Amaranthus sp	1
2015	Sahel	Pibaoré	12,906	-0,823	No	Euphorbiaceae	Euphorbia hirta	1
2015	Sahel	Pibaoré	12,906	-0,823	No	Nyctaginaceae	Boerhavia erecta	2
2015	Sahel	Pibaoré	12,906	-0,823	No	UD	UD	1
2015	Sahel	Tougouri	13,325	-0,512	No	Amaranthaceae	Amaranthus sp	1
2015	Sahel	Tougouri	13,325	-0,512	No	Solanaceae	Physalis angulata	1
2015	Sahel	Tougouri	13,325	-0,512	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Tougouri	13,325	-0,512	No	UD	UD	3
2015	Sahel	Yalgo	13,602	-0,277	No	Solanaceae	Physalis angulata	1
2015	Sahel	Yalgo	13,602	-0,277	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Yalgo	13,602	-0,277	No	UD	UD	4
2015	Sahel	Zéguédéguin	12,957	-0,453	No	Amaranthaceae	Amaranthus sp	1
2015	Sahel	Zéguédéguin	12,957	-0,453	No	Amaranthaceae	Amaranthus spinosus	1
2015	Sahel	Zéguédéguin	12,957	-0,453	No	Solanaceae	Physalis angulata	2
2015	Sahel	Zéguédéguin	12,957	-0,453	Yes	Solanaceae	Solanum lycopersicum	30
2015	Sahel	Zéguédéguin	12,957	-0,453	No	UD	UD	1
2015	Soudan	Banakeledara	11,361	-4,308	Yes	Malvaceae	Abelmoschus esculentus	1
2015	Soudan	Banakeledara	11,361	-4,308	No	Malvaceae	Sida acuta	6
2015	Soudan	Banakeledara	11,361	-4,308	Yes	Solanaceae	Capsicum frutescens	2
2015	Soudan	Banakeledara	11,361	-4,308	Yes	Solanaceae	Solanum lycopersicum	1
2015	Soudan	Banfora	10,643	-4,782	No	Amaranthaceae	Amaranthus sp	2
2015	Soudan	Banfora	10,643	-4,782	No	Convolvulaceae	Ipomoea sp	2
2015	Soudan	Banfora	10,643	-4,782	Yes	Cucurbitaceae	Cucumis sativus	1
2015	Soudan	Banfora	10,643	-4,782	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan	Banfora	10,643	-4,782	Yes	Malvaceae	Abelmoschus esculentus	3
2015	Soudan	Banfora	10,643	-4,782	Yes	Malvaceae	Hibiscus sabdariffa	2
2015	Soudan	Banfora	10,643	-4,782	No	Malvaceae	Sida acuta	2
2015	Soudan	Banfora	10,643	-4,782	No	Solanaceae	Physalis angulata	4
2015	Soudan	Banfora	10,643	-4,782	Yes	Solanaceae	Solanum lycopersicum	1
2015	Soudan	Banfora	10,643	-4,782	No	UD	UD	2
2015	Soudan	Diarradougou	11,281	-4,444	No	Fabaceae	Crotalaria spectabilis	1
2015	Soudan	Diarradougou	11,281	-4,444	No	Malvaceae	Sida acuta	1
2015	Soudan	Diarradougou	11,281	-4,444	No	UD	UD	3
2015	Soudan	Djéri	11,006	-5,023	No	Amaranthaceae	Amaranthus sp	1
2015	Soudan	Djéri	11,006	-5,023	Yes	Malvaceae	Hibiscus sabdariffa	1
2015	Soudan	Djéri	11,006	-5,023	No	Malvaceae	Sida acuta	1
2015	Soudan	Djéri	11,006	-5,023	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan	Djéri	11,006	-5,023	No	UD	UD	2
2015	Soudan	Karangasso-sambla	11,281	-4,444	No	Amaranthaceae	Amaranthus spinosus	1
2015	Soudan	Karangasso-sambla	11,281	-4,444	Yes	Cucurbitaceae	Cucumis sativus	2
2015	Soudan	Karangasso-sambla	11,281	-4,444	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan	Karangasso-sambla	11,281	-4,444	No	Malvaceae	Sida acuta	3
2015	Soudan	Karangasso-sambla	11,281	-4,444	No	Solanaceae	Physalis angulata	1
2015	Soudan	Karangasso-sambla	11,281	-4,444	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan	Karangasso-sambla	11,281	-4,444	No	UD	UD	2

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2015	Soudan	Karfiguela	10,689	-4,808	No	Amaranthaceae	Amaranthus sp	1
2015	Soudan	Karfiguela	10,689	-4,808	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan	Karfiguela	10,689	-4,808	No	Malvaceae	Sida acuta	2
2015	Soudan	Karfiguela	10,689	-4,808	No	Nyctaginaceae	Boerhavia erecta	2
2015	Soudan	Karfiguela	10,689	-4,808	No	Solanaceae	Physalis angulata	1
2015	Soudan	Karfiguela	10,689	-4,808	No	UD	UD	3
2015	Soudan	Kompienbiga	11,268	0,587	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan	Kompienbiga	11,268	0,587	No	Fabaceae	Crotalaria spectabilis	1
2015	Soudan	Kompienbiga	11,268	0,587	No	Malvaceae	Sida acuta	1
2015	Soudan	Kompienbiga	11,268	0,587	No	Solanaceae	Physalis angulata	2
2015	Soudan	Leguéma	11,229	-4,164	No	Amaranthaceae	Amaranthus sp	1
2015	Soudan	Leguéma	11,229	-4,164	No	Amaranthaceae	Amaranthus spinosus	1
2015	Soudan	Leguéma	11,229	-4,164	Yes	Malvaceae	Hibiscus sabdariffa	1
2015	Soudan	Leguéma	11,229	-4,164	No	Solanaceae	Physalis angulata	1
2015	Soudan	Leguéma	11,229	-4,164	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan	Leguéma	11,229	-4,164	No	UD	UD	1
2015	Soudan	Sakabi	11,224	-4,278	No	Amaranthaceae	Amaranthus sp	3
2015	Soudan	Sakabi	11,224	-4,278	No	Amaranthaceae	Amaranthus spinosus	2
2015	Soudan	Sakabi	11,224	-4,278	Yes	Convolvulaceae	Ipomoea batatas	1
2015	Soudan	Sakabi	11,224	-4,278	No	Solanaceae	Physalis angulata	2
2015	Soudan	Sakabi	11,224	-4,278	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan	Sakabi	11,224	-4,278	No	UD	UD	2
2015	Soudan	Tingréla	10,563	-4,815	Yes	Malvaceae	Hibiscus sabdariffa	1
2015	Soudan	Tingréla	10,563	-4,815	No	Solanaceae	Physalis angulata	1
2015	Soudan	Tingréla	10,563	-4,815	No	UD	UD	3
2015	Soudan	Toussiana	10,856	-4,640	No	Amaranthaceae	Amaranthus sp	5
2015	Soudan	Toussiana	10,856	-4,640	No	Amaranthaceae	Amaranthus spinosus	1
2015	Soudan	Toussiana	10,856	-4,640	No	Solanaceae	Physalis angulata	1
2015	Soudan	Toussiana	10,856	-4,640	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan	Toussiana	10,856	-4,640	Yes	Solanaceae	Solanum melongena	1
2015	Soudan	Toussiana	10,856	-4,640	No	UD	UD	2
2015	Soudan_sahel	Di	13,185	-3,391	Yes	Solanaceae	Solanum melongena	3
2015	Soudan_sahel	Guiédougou	12,995	-3,433	No	Amaranthaceae	Amaranthus sp	2
2015	Soudan_sahel	Guiédougou	12,995	-3,433	No	Amaranthaceae	Amaranthus viridis	2
2015	Soudan_sahel	Guiédougou	12,995	-3,433	No	Euphorbiaceae	Euphorbia hirta	2
2015	Soudan_sahel	Guiédougou	12,995	-3,433	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan_sahel	Guiédougou	12,995	-3,433	No	UD	UD	4
2015	Soudan_sahel	Korsimoro	12,824	-1,052	No	Amaranthaceae	Amaranthus sp	3
2015	Soudan_sahel	Korsimoro	12,824	-1,052	No	Convolvulaceae	Ipomoea alba	1
2015	Soudan_sahel	Korsimoro	12,824	-1,052	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan_sahel	Korsimoro	12,824	-1,052	No	Solanaceae	Physalis angulata	1
2015	Soudan_sahel	Korsimoro	12,824	-1,052	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan_sahel	Korsimoro	12,824	-1,052	No	UD	UD	4
2015	Soudan_sahel	Koupela	12,197	-0,390	Yes	Solanaceae	Solanum melongena	3
2015	Soudan_sahel	Loumbila	12,485	-1,417	No	Amaranthaceae	Amaranthus sp	1
2015	Soudan_sahel	Loumbila	12,485	-1,417	No	Solanaceae	Physalis angulata	3
2015	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan_sahel	Loumbila	12,485	-1,417	No	UD	UD	1
2015	Soudan_sahel	Sao	12,258	-3,003	No	Amaranthaceae	Amaranthus sp	3
2015	Soudan_sahel	Sao	12,258	-3,003	No	Malvaceae	Sida acuta	1
2015	Soudan_sahel	Sao	12,258	-3,003	Yes	Solanaceae	Solanum lycopersicum	30

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2015	Soudan_sahel	Sao	12,258	-3,003	No	UD	UD	1
2015	Soudan_sahel	Tandjari	12,190	0,328	No	Euphorbiaceae	Euphorbia hirta	1
2015	Soudan_sahel	Tandjari	12,190	0,328	No	Solanaceae	Physalis angulata	2
2015	Soudan_sahel	Tandjari	12,190	0,328	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan_sahel	Tandjari	12,190	0,328	No	UD	UD	2
2015	Soudan_sahel	Ziniaré	12,568	-1,280	No	Solanaceae	Physalis angulata	1
2015	Soudan_sahel	Ziniaré	12,568	-1,280	Yes	Solanaceae	Solanum lycopersicum	30
2015	Soudan_sahel	Ziniaré	12,568	-1,280	No	UD	UD	4
2016	Sahel	Dakiri	13,311	-0,278	No	Amaranthaceae	Amaranthus viridis	2
2016	Sahel	Dakiri	13,311	-0,278	Yes	Solanaceae	Solanum lycopersicum	90
2016	Sahel	Dem	13,178	-1,168	No	Amaranthaceae	Amaranthus viridis	5
2016	Sahel	Dem	13,178	-1,168	No	Euphorbiaceae	Euphorbia hirta	1
2016	Sahel	Dem	13,178	-1,168	No	Rubiaceae	Mitracarpus villosus	6
2016	Sahel	Dem	13,178	-1,168	No	Solanaceae	Physalis angulata	2
2016	Sahel	Dem	13,178	-1,168	Yes	Solanaceae	Solanum lycopersicum	90
2016	Sahel	Ouahigouya	13,556	-2,449	No	Amaranthaceae	Amaranthus hybridus	8
2016	Sahel	Ouahigouya	13,556	-2,449	No	Amaranthaceae	Amaranthus spinosus	4
2016	Sahel	Ouahigouya	13,556	-2,449	No	Amaranthaceae	Amaranthus viridis	3
2016	Sahel	Ouahigouya	13,556	-2,449	No	Euphorbiaceae	Euphorbia heterophylla	5
2016	Sahel	Ouahigouya	13,556	-2,449	No	Euphorbiaceae	Euphorbia hirta	5
2016	Sahel	Ouahigouya	13,556	-2,449	No	Malvaceae	Sida acuta	1
2016	Sahel	Ouahigouya	13,556	-2,449	No	Nyctaginaceae	Boerhavia erecta	5
2016	Sahel	Ouahigouya	13,556	-2,449	Yes	Solanaceae	Capsicum frutescens	1
2016	Sahel	Ouahigouya	13,556	-2,449	Yes	Solanaceae	Solanum lycopersicum	90
2016	Sahel	Ouahigouya	13,556	-2,449	No	UD	UD	2
2016	Soudan	Guénon	11,080	-0,902	Yes	Malvaceae	Abelmoschus esculentus	2
2016	Soudan	Karangasso-sambla	11,281	-4,444	No	Amaranthaceae	Amaranthus viridis	6
2016	Soudan	Karangasso-sambla	11,281	-4,444	No	Asteraceae	Ageratum conyzoides	3
2016	Soudan	Karangasso-sambla	11,281	-4,444	No	Solanaceae	Physalis angulata	3
2016	Soudan	Karangasso-sambla	11,281	-4,444	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan	Leguéma	11,229	-4,164	No	Amaranthaceae	Amaranthus spinosus	8
2016	Soudan	Leguéma	11,229	-4,164	No	Amaranthaceae	Amaranthus viridis	4
2016	Soudan	Leguéma	11,229	-4,164	No	Asteraceae	Ageratum conyzoides	1
2016	Soudan	Leguéma	11,229	-4,164	No	Convolvulaceae	Ipomoea alba	1
2016	Soudan	Leguéma	11,229	-4,164	No	Euphorbiaceae	Euphorbia heterophylla	1
2016	Soudan	Leguéma	11,229	-4,164	No	Euphorbiaceae	Euphorbia hirta	1
2016	Soudan	Leguéma	11,229	-4,164	No	Malvaceae	Sida acuta	4
2016	Soudan	Leguéma	11,229	-4,164	No	Nyctaginaceae	Boerhavia erecta	1
2016	Soudan	Leguéma	11,229	-4,164	No	Rubiaceae	Mitracarpus villosus	2
2016	Soudan	Leguéma	11,229	-4,164	No	Solanaceae	Physalis angulata	2
2016	Soudan	Leguéma	11,229	-4,164	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan	Tiébélé	11,120	-0,964	No	Amaranthaceae	Amaranthus hybridus	1
2016	Soudan	Tiébélé	11,120	-0,964	No	Euphorbiaceae	Euphorbia heterophylla	3
2016	Soudan	Tiébélé	11,120	-0,964	No	Euphorbiaceae	Euphorbia hirta	6
2016	Soudan	Tiébélé	11,120	-0,964	No	Malvaceae	Sida acuta	4
2016	Soudan	Tiébélé	11,120	-0,964	No	Nyctaginaceae	Boerhavia erecta	4
2016	Soudan	Tiébélé	11,120	-0,964	Yes	Solanaceae	Capsicum frutescens	6
2016	Soudan	Tiébélé	11,120	-0,964	No	Solanaceae	Physalis angulata	8
2016	Soudan	Toussiana	10,856	-4,640	No	Amaranthaceae	Amaranthus hybridus	1
2016	Soudan	Toussiana	10,856	-4,640	No	Amaranthaceae	Amaranthus spinosus	2
2016	Soudan	Toussiana	10,856	-4,640	No	Amaranthaceae	Amaranthus viridis	2

Sampling year	Agroclimatic areas	Locality	Latitude	Longitude	Cultivated	Host family	Host species	Number of plants sampled
2016	Soudan	Toussiana	10,856	-4,640	No	Asteraceae	Ageratum conyzoides	2
2016	Soudan	Toussiana	10,856	-4,640	No	Euphorbiaceae	Euphorbia heterophylla	1
2016	Soudan	Toussiana	10,856	-4,640	No	Euphorbiaceae	Euphorbia hirta	1
2016	Soudan	Toussiana	10,856	-4,640	No	Malvaceae	Sida acuta	4
2016	Soudan	Toussiana	10,856	-4,640	No	Nyctaginaceae	Boerhavia erecta	5
2016	Soudan	Toussiana	10,856	-4,640	No	Rubiaceae	Mitracarpus villosus	1
2016	Soudan	Toussiana	10,856	-4,640	No	Solanaceae	Physalis angulata	1
2016	Soudan	Toussiana	10,856	-4,640	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan_sahel	Bon_Srima	11,928	-1,218	No	Amaranthaceae	Amaranthus hybridus	2
2016	Soudan_sahel	Bon_Srima	11,928	-1,218	No	Malvaceae	Sida acuta	2
2016	Soudan_sahel	Bon_Srima	11,928	-1,218	Yes	Solanaceae	Capsicum annuum	3
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Amaranthaceae	Amaranthus hybridus	1
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Amaranthaceae	Amaranthus viridis	6
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Euphorbiaceae	Euphorbia heterophylla	1
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Euphorbiaceae	Euphorbia hirta	3
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Solanaceae	Physalis angulata	8
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	Solanaceae	Physalis sp	3
2016	Soudan_sahel	Guiédougou	12,995	-3,433	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan_sahel	Guiédougou	12,995	-3,433	No	UD	UD	3
2016	Soudan_sahel	Korsimoro	12,824	-1,052	No	Amaranthaceae	Amaranthus viridis	2
2016	Soudan_sahel	Korsimoro	12,824	-1,052	No	Malvaceae	Sida acuta	5
2016	Soudan_sahel	Korsimoro	12,824	-1,052	No	Solanaceae	Physalis angulata	2
2016	Soudan_sahel	Korsimoro	12,824	-1,052	No	Solanaceae	Physalis sp	1
2016	Soudan_sahel	Korsimoro	12,824	-1,052	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan_sahel	Loumbila	12,485	-1,417	No	Boraginaceae	Heliotropium indicum	5
2016	Soudan_sahel	Loumbila	12,485	-1,417	No	Malvaceae	Sida rhombifolia	4
2016	Soudan_sahel	Loumbila	12,485	-1,417	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan_sahel	Loumbila	12,485	-1,417	No	UD	UD	7
2016	Soudan_sahel	Tandjari	12,190	0,328	Yes	Solanaceae	Solanum lycopersicum	90
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Amaranthaceae	Amaranthus hybridus	3
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Amaranthaceae	Amaranthus viridis	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Amaranthaceae	Spinacia oleracea	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Asteraceae	Ageratum conyzoides	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Asteraceae	Lactuca sativa	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Brassicaceae	Brassica oleracea	2
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Euphorbiaceae	Euphorbia hirta	2
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Nyctaginaceae	Boerhavia erecta	3
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	Rubiaceae	Mitracarpus villosus	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Solanaceae	Capsicum annuum	2
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Solanaceae	Capsicum frutescens	1
2016	Soudan_sahel	Tionkui	12,508	-3,457	Yes	Solanaceae	Solanum lycopersicum	5
2016	Soudan_sahel	Tionkui	12,508	-3,457	No	UD	UD	4

Supp. table 2: General information about the 1068 samples collected in Goué and Léguéma using sampling grids in 2016.

Locality	Stitch	Latitudes	Longitudes	Plants	Sample number
Goué	A1	-1,474	12,571	*Abelmoschus esculentus*	10
Goué	A1	-1,474	12,571	*Amaranthus viridis*	10
Goué	A1	-1,474	12,571	*Boerhavia erecta*	10
Goué	A1	-1,474	12,571	*Hibiscus* sp	10
Goué	A1	-1,474	12,571	*Heliotropium indicum*	10
Goué	A1	-1,474	12,571	*Physalis ixocarpa*	5
Goué	A1	-1,474	12,571	*Sida* sp	10
Goué	A2	-1,472	12,572	*Abelmoschus esculentus*	1
Goué	A2	-1,472	12,572	*Ageratum conyzoides*	5
Goué	A2	-1,472	12,572	*Amaranthus viridis*	10
Goué	A2	-1,472	12,572	*Boerhavia erecta*	10
Goué	A2	-1,472	12,572	*Euphorbia hirta*	10
Goué	A2	-1,472	12,572	*Hibiscus* sp	8
Goué	A2	-1,472	12,572	*Mitracarpus villosus*	6
Goué	A2	-1,472	12,572	*Sida acuta*	2
Goué	B1	-1,472	12,569	*Boerhavia erecta*	10
Goué	B1	-1,472	12,569	*Capsicum frutescens*	5
Goué	B1	-1,472	12,569	*Euphorbia hirta*	10
Goué	B1	-1,472	12,569	*Mitracarpus villosus*	4
Goué	B1	-1,472	12,569	*Sida acuta*	10
Goué	B2	-1,471	12,571	*Ageratum conyzoides*	4
Goué	B2	-1,471	12,571	*Amaranthus viridis*	3
Goué	B2	-1,471	12,571	*Capsicum frutescens*	10
Goué	B2	-1,471	12,571	*Mitracarpus villosus*	1
Goué	B2	-1,471	12,571	UD	4
Goué	B2	-1,471	12,571	UD1	5
Goué	B2	-1,471	12,571	*Sida* sp	1
Goué	C1	-1,471	12,567	*Amaranthus viridis*	10
Goué	C1	-1,471	12,567	*Boerhavia erecta*	10
Goué	C1	-1,471	12,567	*Capsicum frutescens*	6
Goué	C1	-1,471	12,567	*Euphorbia hirta*	10
Goué	C1	-1,471	12,567	*Hibiscus* sp	10
Goué	C1	-1,471	12,567	*Mitracarpus villosus*	10
Goué	C1	-1,471	12,567	*Physalis ixocarpa*	5
Goué	C1	-1,471	12,567	*Sida acuta*	3
Goué	C1	-1,471	12,567	*Sida* sp	2
Goué	C2	-1,469	12,569	*Hibiscus* sp	10
Goué	C2	-1,469	12,569	*Ipomea alba*	9
Goué	C2	-1,469	12,569	*Mitracarpus villosus*	9
Goué	C2	-1,469	12,569	UD2	4
Goué	C2	-1,469	12,569	UD3	5
Goué	C2	-1,469	12,569	*Physalis ixocarpa*	5
Goué	D1	-1,469	12,566	*Abelmoschus esculentus*	9
Goué	D1	-1,469	12,566	*Amaranthus spinosus*	10
Goué	D1	-1,469	12,566	*Amaranthus viridis*	10
Goué	D1	-1,469	12,566	*Boerhavia erecta*	10
Goué	D1	-1,469	12,566	*Capsicum frutescens*	10
Goué	D1	-1,469	12,566	*Euphorbia hirta*	10
Goué	D1	-1,469	12,566	*Physalis ixocarpa*	1
Goué	D2	-1,467	12,567	*Amaranthus spinosus*	10
Goué	D2	-1,467	12,567	*Amaranthus viridis*	10
Goué	D2	-1,467	12,567	*Boerhavia erecta*	10

Locality	Stitch	Latitudes	Longitudes	Plants	Sample number
Goué	D2	-1,467	12,567	*Capsicum annuum*	10
Goué	D2	-1,467	12,567	*Euphorbia hirta*	10
Goué	D2	-1,467	12,567	*Hibiscus* sp	5
Goué	D2	-1,467	12,567	UD	1
Goué	D3	-1,466	12,569	*Amaranthus spinosus*	1
Goué	D3	-1,466	12,569	*Amaranthus viridis*	3
Goué	D3	-1,466	12,569	*Boerhavia erecta*	10
Goué	D3	-1,466	12,569	*Hibiscus* sp	1
Goué	D3	-1,466	12,569	*Mitracarpus villosus*	10
Goué	D3	-1,466	12,569	UD1	1
Goué	D3	-1,466	12,569	*Physalis ixocarpa*	1
Goué	E1	-1,467	12,564	*Abelmoschus esculentus*	10
Goué	E1	-1,467	12,564	*Amaranthus spinosus*	1
Goué	E1	-1,467	12,564	*Amaranthus viridis*	10
Goué	E1	-1,467	12,564	*Hibiscus* sp	10
Goué	E1	-1,467	12,564	UD4	9
Goué	E2	-1,466	12,566	*Sida acuta*	10
Goué	E3	-1,464	12,567	UD *(OCIMUM)*	10
Léguéma	A1	-4,168	11,223	*Ageratum conyzoides*	10
Léguéma	A1	-4,168	11,223	*Amaranthus spinosus*	10
Léguéma	A1	-4,168	11,223	*Amaranthus viridis*	10
Léguéma	A1	-4,168	11,223	*Boerhavia erecta*	2
Léguéma	A1	-4,168	11,223	*Euphorbia heterophylla*	10
Léguéma	A1	-4,168	11,223	*Euphorbia hirta*	10
Léguéma	A1	-4,168	11,223	*Ipomea alba*	10
Léguéma	A1	-4,168	11,223	*Physalis ixocarpa*	3
Léguéma	A1	-4,168	11,223	*Sida acuta*	10
Léguéma	A1	-4,168	11,223	*Solanum lycopersicum*	10
Léguéma	A2	-4,167	11,225	*Amaranthus viridis*	1
Léguéma	A2	-4,167	11,225	*Sida acuta*	3
Léguéma	A3	-4,166	11,227	*Ageratum conyzoides*	8
Léguéma	A3	-4,166	11,227	*Amaranthus spinosus*	7
Léguéma	A3	-4,166	11,227	*Amaranthus viridis*	10
Léguéma	A3	-4,166	11,227	*Boerhavia erecta*	3
Léguéma	A3	-4,166	11,227	*Capsicum frutescens*	2
Léguéma	A3	-4,166	11,227	*Euphorbia heterophylla*	4
Léguéma	A3	-4,166	11,227	*Euphorbia hirta*	11
Léguéma	A3	-4,166	11,227	*Ipomea alba*	1
Léguéma	A3	-4,166	11,227	*Physalis ixocarpa*	3
Léguéma	A3	-4,166	11,227	*Sida acuta*	7
Léguéma	A3	-4,166	11,227	*Sida rhombifolia*	3
Léguéma	A3	-4,166	11,227	*Solanum lycopersicum*	10
Léguéma	A4	-4,165	11,229	*Ageratum conyzoides*	5
Léguéma	A4	-4,165	11,229	*Amaranthus spinosus*	10
Léguéma	A4	-4,165	11,229	*Amaranthus viridis*	3
Léguéma	A4	-4,165	11,229	*Boerhavia erecta*	10
Léguéma	A4	-4,165	11,229	*Euphorbia heterophylla*	5
Léguéma	A4	-4,165	11,229	*Euphorbia hirta*	6
Léguéma	A4	-4,165	11,229	*Ipomea alba*	8
Léguéma	A4	-4,165	11,229	UD	1
Léguéma	A4	-4,165	11,229	*Physalis ixocarpa*	1
Léguéma	A4	-4,165	11,229	*Sida acuta*	10
Léguéma	A4	-4,165	11,229	*Sida rhombifolia*	1
Léguéma	A5	-4,164	11,231	*Ageratum conyzoides*	4
Léguéma	A5	-4,164	11,231	*Amaranthus viridis*	10

Locality	Stitch	Latitudes	Longitudes	Plants	Sample number
Léguéma	A5	-4,164	11,231	*Boerhavia erecta*	10
Léguéma	A5	-4,164	11,231	*Euphorbia heterophylla*	2
Léguéma	A5	-4,164	11,231	*Euphorbia hirta*	5
Léguéma	A5	-4,164	11,231	*Ipomea alba*	1
Léguéma	A5	-4,164	11,231	*Mitracarpus villosus*	5
Léguéma	A5	-4,164	11,231	*Sida acuta*	12
Léguéma	B2	-4,165	11,224	*Abelmoschus esculentus*	10
Léguéma	B2	-4,165	11,224	*Ageratum conyzoides*	9
Léguéma	B2	-4,165	11,224	*Amaranthus spinosus*	1
Léguéma	B2	-4,165	11,224	*Amaranthus viridis*	9
Léguéma	B2	-4,165	11,224	*Boerhavia erecta*	3
Léguéma	B2	-4,165	11,224	*Euphorbia heterophylla*	2
Léguéma	B2	-4,165	11,224	*Euphorbia hirta*	6
Léguéma	B2	-4,165	11,224	*Sida acuta*	2
Léguéma	B2	-4,165	11,224	*Sida rhombifolia*	2
Léguéma	B3	-4,164	11,226	*Ageratum conyzoides*	7
Léguéma	B3	-4,164	11,226	*Amaranthus spinosus*	5
Léguéma	B3	-4,164	11,226	*Amaranthus viridis*	3
Léguéma	B3	-4,164	11,226	*Boerhavia erecta*	3
Léguéma	B3	-4,164	11,226	*Euphorbia heterophylla*	6
Léguéma	B3	-4,164	11,226	*Euphorbia hirta*	6
Léguéma	B3	-4,164	11,226	*Sida acuta*	5
Léguéma	B3	-4,164	11,226	*Solanum lycopersicum*	5
Léguéma	B4	-4,163	11,228	*Abelmoschus esculentus*	9
Léguéma	B4	-4,163	11,228	*Ageratum conyzoides*	9
Léguéma	B4	-4,163	11,228	*Amaranthus spinosus*	10
Léguéma	B4	-4,163	11,228	*Amaranthus viridis*	10
Léguéma	B4	-4,163	11,228	*Boerhavia erecta*	10
Léguéma	B4	-4,163	11,228	*Euphorbia heterophylla*	5
Léguéma	B4	-4,163	11,228	*Euphorbia hirta*	9
Léguéma	B4	-4,163	11,228	UD	5
Léguéma	B4	-4,163	11,228	*Physalis ixocarpa*	4
Léguéma	B4	-4,163	11,228	*Sida acuta*	9
Léguéma	B4	-4,163	11,228	*Solanum lycopersicum*	10
Léguéma	B5	-4,162	11,23	*Abelmoschus esculentus*	10
Léguéma	B5	-4,162	11,23	*Ageratum conyzoides*	10
Léguéma	B5	-4,162	11,23	*Amaranthus hybridis*	1
Léguéma	B5	-4,162	11,23	*Amaranthus spinosus*	9
Léguéma	B5	-4,162	11,23	*Amaranthus viridis*	7
Léguéma	B5	-4,162	11,23	*Ipomea alba*	10
Léguéma	B5	-4,162	11,23	UD*1*	3
Léguéma	B5	-4,162	11,23	*Physalis ixocarpa*	10
Léguéma	B5	-4,162	11,23	*Sida acuta*	10
Léguéma	B5	-4,162	11,23	*Sida rhombifolia*	1
Léguéma	B5	-4,162	11,23	*Solanum lycopersicum*	10
Léguéma	C2	-4,163	11,223	*Boerhavia erecta*	8
Léguéma	C2	-4,163	11,223	*Euphorbia heterophylla*	10
Léguéma	C2	-4,163	11,223	*Euphorbia hirta*	6
Léguéma	C2	-4,163	11,223	UD0	2
Léguéma	C2	-4,163	11,223	UD1	7
Léguéma	C2	-4,163	11,223	*sesame s*	3
Léguéma	C3	-4,162	11,225	*Euphorbia hirta*	5
Léguéma	C3	-4,162	11,225	UD1	10

Supp. table 3: General information about the 372 samples randomly collected in Burkina Faso from 2013 to 2015.

Area	Locality	Plants family	Plants species	Samples number
Sahel	Dem	Amaranthaceae	*Amaranthus viridus*	5
Sahel	Dem	Euphorbiaceae	*Euphorbia hirta*	1
Sahel	Dem	Rubiaceae	*Mitracarpus villosus*	3
Sahel	Dem	Solanaceae	*Physalis ixocarpa*	2
Sahel	Dem	Solanaceae	*Solanum lycopersicum*	1
Sahel	Kaya	Solanaceae	*Capsicum annuum*	2
Sahel	Kaya	Solanaceae	*Capsicum frutescens*	2
Sahel	Ouahigouya	Amaranthaceae	*Amaranthus hybridus*	6
Sahel	Ouahigouya	Amaranthaceae	*Amaranthus sp*	1
Sahel	Ouahigouya	Amaranthaceae	*Amaranthus spinosus*	4
Sahel	Ouahigouya	Amaranthaceae	*Amaranthus viridus*	3
Sahel	Ouahigouya	Euphorbiaceae	*Euphorbia heterophylla*	4
Sahel	Ouahigouya	Euphorbiaceae	*Euphorbia hirta*	7
Sahel	Ouahigouya	Lamiaceae	*Leucas martinicens*	2
Sahel	Ouahigouya	Malvaceae	*Sida acuta*	1
Sahel	Ouahigouya	Nyctaginaceae	*Boerhavia erecta*	6
Sahel	Ouahigouya	Solanaceae	*Capsicum frutescens*	1
Sahel	Ouahigouya	Solanaceae	*Physalis angulata*	1
Sahel	Ouahigouya	Solanaceae	*Solanum lycopersicum*	1
Sahel	Ouahigouya	Undetermined	*Undetermined*	3
Sahel	Zéguédéguin	Solanaceae	*Solanum lycopersicum*	3
Soudan_sahel	Bon_Srima	Amaranthaceae	*Amaranthus hybridus*	2
Soudan_sahel	Bon_Srima	Malvaceae	*Sida acuta*	2
Soudan_sahel	Bon_Srima	Solanaceae	*Capsicum annuum*	3
Soudan_sahel	Dakiri	Amaranthaceae	*Amaranthus viridus*	2
Soudan_sahel	Di	Solanaceae	*Solanum melongena*	2
Soudan_sahel	Guiédougou	Amaranthaceae	*Amaranthus hybridus*	1
Soudan_sahel	Guiédougou	Amaranthaceae	*Amaranthus sp*	2
Soudan_sahel	Guiédougou	Amaranthaceae	*Amaranthus viridis*	2
Soudan_sahel	Guiédougou	Amaranthaceae	*Amaranthus viridus*	4
Soudan_sahel	Guiédougou	Euphorbiaceae	*Euphorbia heterophylla*	1
Soudan_sahel	Guiédougou	Euphorbiaceae	*Euphorbia hirta*	5
Soudan_sahel	Guiédougou	Lamiaceae	*Leucas martinicens*	3
Soudan_sahel	Guiédougou	Solanaceae	*Physalis ixocarpa*	7
Soudan_sahel	Guiédougou	Solanaceae	*Physalis sp*	3
Soudan_sahel	Guiédougou	Solanaceae	Solanum lycopersicum	1
Soudan_sahel	Guiédougou	Undetermined	*Undetermined*	4
Soudan_sahel	Korsimoro	Amaranthaceae	*Amaranthus viridus*	2
Soudan_sahel	Korsimoro	Malvaceae	*Sida acuta*	3
Soudan_sahel	Korsimoro	Solanaceae	*Physalis ixocarpa*	2
Soudan_sahel	Korsimoro	Solanaceae	*Physalis sp*	1
Soudan_sahel	Korsimoro	Solanaceae	*Solanum lycopersicum*	1
Soudan_sahel	Koupela	Solanaceae	*Solanum melongena*	2
Soudan_sahel	Loumbila	Boraginaceae	*Heliotropium indicum*	5
Soudan_sahel	Loumbila	Malvaceae	*Melochia sp.*	7
Soudan_sahel	Loumbila	Malvaceae	*Sida rhombifolia*	4
Soudan_sahel	Loumbila	Solanaceae	*Capsicum annuum*	4
Soudan_sahel	Loumbila	Solanaceae	*Capsicum frutescens*	2
Soudan_sahel	Loumbila	Solanaceae	*Solanum lycopersicum*	9
Soudan_sahel	Loumbila	Solanaceae	*Solanum melongena*	1

Area	Locality	Plants family	Plants species	Samples number
Soudan_sahel	Mogtédo	Amaranthaceae	*Amaranthus sp*	1
Soudan_sahel	Mogtédo	Euphorbiaceae	*Euphorbia hirta*	1
Soudan_sahel	Mogtédo	Lamiaceae	*Ocimum basilicum*	1
Soudan_sahel	Mogtédo	Malvaceae	*Sida acuta*	1
Soudan_sahel	Mogtédo	Solanaceae	*Capsicum frutescens*	5
Soudan_sahel	Sao	Solanaceae	*Solanum lycopersicum*	3
Soudan_sahel	Tandjari	Euphorbiaceae	*Euphorbia hirta*	1
Soudan_sahel	Tandjari	Solanaceae	*Physalis angulata*	2
Soudan_sahel	Tandjari	Undetermined	*Undetermined*	2
Soudan_sahel	Tionkui	Amaranthaceae	*Amaranthus hybridus*	3
Soudan_sahel	Tionkui	Amaranthaceae	*Amaranthus sp.*	1
Soudan_sahel	Tionkui	Amaranthaceae	*Amaranthus viridus*	1
Soudan_sahel	Tionkui	Amaranthaceae	*Spinacia oleracea_epinard*	1
Soudan_sahel	Tionkui	Asteraceae	*Ageratum conyzoides*	1
Soudan_sahel	Tionkui	Asteraceae	*Lactuca sativa*	1
Soudan_sahel	Tionkui	Basellaceae	*Basella alba*	1
Soudan_sahel	Tionkui	Brassicaceae	*Brassica oleracea*	2
Soudan_sahel	Tionkui	Euphorbiaceae	*Euphorbia hirta*	2
Soudan_sahel	Tionkui	Fabaceae	*Senna occidentalis*	2
Soudan_sahel	Tionkui	Nyctaginaceae	*Boerhavia erecta*	3
Soudan_sahel	Tionkui	Rubiaceae	*Mitracarpus villosus*	1
Soudan_sahel	Tionkui	Solanaceae	*Capsicum annuum*	2
Soudan_sahel	Tionkui	Solanaceae	*Capsicum frutescens*	1
Soudan_sahel	Tionkui	Solanaceae	*Solanum lycopersicum*	5
Soudan_sahel	Ziniaré	Solanaceae	*Solanum lycopersicum*	3
Soudan	Banakeledara	Amaranthaceae	*Amaranthus sp*	2
Soudan	Banakeledara	Convolvulaceae	*Ipomea sp*	2
Soudan	Banakeledara	Cucurbitaceae	*Cucumis sativus*	1
Soudan	Banakeledara	Euphorbiaceae	*Euphorbia hirta*	1
Soudan	Banakeledara	Malvaceae	*Abelmoschus esculentus*	4
Soudan	Banakeledara	Malvaceae	*Hibiscus sabdariffa*	2
Soudan	Banakeledara	Malvaceae	*Sida acuta*	8
Soudan	Banakeledara	Solanaceae	*Capsicum frutescens*	2
Soudan	Banakeledara	Solanaceae	*Physalis angulata*	4
Soudan	Banakeledara	Solanaceae	*Solanum lycopersicum*	2
Soudan	Banakeledara	Undetermined	*Undetermined*	2
Soudan	Bobo	Lamiaceae	*Perilla frutescens*	1
Soudan	Diarradougou	Fabaceae	*Crotalaria spectabilis*	1
Soudan	Diarradougou	Malvaceae	*Sida acuta*	1
Soudan	Diarradougou	Undetermined	*Undetermined*	3
Soudan	Djéri	Solanaceae	*Solanum lycopersicum*	1
Soudan	Guénon	Malvaceae	*Abelmoschus esculentus*	2
Soudan	Karangasso-sambla	Amaranthaceae	*Amaranthus spinosus*	1
Soudan	Karangasso-sambla	Amaranthaceae	*Amaranthus viridus*	6
Soudan	Karangasso-sambla	Asteraceae	*Ageratum conyzoides*	3
Soudan	Karangasso-sambla	Cucurbitaceae	*Cucumis sativus*	2
Soudan	Karangasso-sambla	Euphorbiaceae	*Euphorbia hirta*	1
Soudan	Karangasso-sambla	Malvaceae	*Sida acuta*	3
Soudan	Karangasso-sambla	Solanaceae	*Physalis angulata*	1
Soudan	Karangasso-sambla	Solanaceae	*Physalis ixocarpa*	3
Soudan	Karangasso-sambla	Undetermined	*Undetermined*	2
Soudan	Kompienbiga	Euphorbiaceae	*Euphorbia hirta*	1
Soudan	Kompienbiga	Fabaceae	*Crotalaria spectabilis*	1
Soudan	Kompienbiga	Malvaceae	*Sida acuta*	1

Area	Locality	Plants family	Plants species	Samples number
Soudan	Kompienbiga	Solanaceae	*Physalis angulata*	2
Soudan	Léguéma	Amaranthaceae	*Amaranthus sp*	1
Soudan	Léguéma	Amaranthaceae	*Amaranthus spinosus*	9
Soudan	Léguéma	Amaranthaceae	*Amaranthus viridus*	4
Soudan	Léguéma	Asteraceae	*Ageratum conyzoides*	1
Soudan	Léguéma	Convolvulaceae	*Ipomea alba*	1
Soudan	Léguéma	Euphorbiaceae	*Euphorbia heterophylla*	1
Soudan	Léguéma	Euphorbiaceae	*Euphorbia hirta*	1
Soudan	Léguéma	Malvaceae	*Hibiscus sabdariffa*	1
Soudan	Léguéma	Malvaceae	*Sida acuta*	4
Soudan	Léguéma	Nyctaginaceae	*Boerhavia erecta*	1
Soudan	Léguéma	Rubiaceae	*Mitracarpus villosus*	2
Soudan	Léguéma	Solanaceae	*Physalis angulata*	1
Soudan	Léguéma	Solanaceae	*Physalis ixocarpa*	2
Soudan	Léguéma	Solanaceae	*Solanum lycopersicum*	2
Soudan	Léguéma	Undetermined	*Undetermined*	1
Soudan	Sakabi	Asteraceae	*Ageratum sp*	1
Soudan	Sakabi	Nyctaginaceae	*Boerhavia erecta*	2
Soudan	Sakabi	Solanaceae	*Capsicum annuum*	1
Soudan	Sakabi	Solanaceae	*Solanum lycopersicum*	5
Soudan	Sakabiabi	Amaranthaceae	*Amaranthus sp*	3
Soudan	Sakabiabi	Amaranthaceae	*Amaranthus spinosus*	2
Soudan	Sakabiabi	Convolvulaceae	*Ipomea batatas*	1
Soudan	Sakabiabi	Solanaceae	*Physalis angulata*	2
Soudan	Sakabiabi	Undetermined	*Undetermined*	2
Soudan	Sossogona	Malvaceae	*Abelmoschus esculentus*	2
Soudan	Sossogona	Malvaceae	*Sida acuta*	1
Soudan	Sossogona	Solanaceae	*Capsicum frutescens*	4
Soudan	Sossogona	Solanaceae	*Solanum lycopersicum*	4
Soudan	Tiébélé	Amaranthaceae	*Amaranthus hybridus*	1
Soudan	Tiébélé	Euphorbiaceae	*Euphorbia heterophylla*	3
Soudan	Tiébélé	Euphorbiaceae	*Euphorbia hirta*	6
Soudan	Tiébélé	Malvaceae	*Sida acuta*	4
Soudan	Tiébélé	Nyctaginaceae	*Boerhavia erecta*	4
Soudan	Tiébélé	Solanaceae	*Capsicum frutescens*	6
Soudan	Tiébélé	Solanaceae	*Physalis ixocarpa*	8
Soudan	Toussiana	Amaranthaceae	*Amaranthus hybridus*	1
Soudan	Toussiana	Amaranthaceae	*Amaranthus spinosus*	2
Soudan	Toussiana	Amaranthaceae	*Amaranthus viridus*	2
Soudan	Toussiana	Asteraceae	*Ageratum conyzoides*	2
Soudan	Toussiana	Euphorbiaceae	*Euphorbia heterophylla*	1
Soudan	Toussiana	Euphorbiaceae	*Euphorbia hirta*	1
Soudan	Toussiana	Malvaceae	*Sida acuta*	4
Soudan	Toussiana	Nyctaginaceae	*Boerhavia erecta*	3
Soudan	Toussiana	Rubiaceae	*Mitracarpus villosus*	1
Soudan	Toussiana	Solanaceae	*Physalis ixocarpa*	1
Soudan	Toussiana	Solanaceae	*Solanum lycopersicum*	1

Supp. table 4: List of the 270 positive samples collected from Goué and Léguéma using grid sampling strategy. Columns named Cluster A, B, S and M indicate assignation profiles-based cluster memberships for begomovirus DNA-A/A-like, begomovirus DNA-B, satellites and mastrevirus, respectively

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G1_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G2_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G3_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	4	-	-	-
G5_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G6_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G7_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	3	-	-	-
G9_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	3	1	-	-
G11_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	9	-	-	-
G12_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	3	1	-	-
G14_BF_200316	Amaranthus	Amaranthus viridis	Léguéma	1	13	-	-	-
G17_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	4	-	-	-
G30_BF_200316	Amaranthus	Amaranthus viridis	Léguéma	1	3	-	-	-
G33_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	5	-	-	-
G34_BF_200316	Amaranthus	Amaranthus viridis	Léguéma	1	3	-	-	-
G41_BF_200316	Ipomea	Ipomea alba	Léguéma	1	4	-	-	-
G51_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	5	-	13	-
G54_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	-	-	13	-
G55_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	-	-	13	-
G59_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	13	-	13	-
G79_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	10	-
G86_BF_200316	Sida	Sida acuta	Léguéma	1	12	-	-	-
G96_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	11	-	-	-
G98_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	5	-	-	-
G101_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	1	-
G111_BF_200316	Euphorbia	Euphorbia hirta	Léguéma	1	9	-	-	-
G122_BF_200316	Euphorbia	Euphorbia heterophylla	Léguéma	1	1	-	12	-
G143_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	11	-
G144_BF_200316	Sida	Sida acuta	Léguéma	1	1	-	7	-
G147_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	-	-	-	4
G148_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G149_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G150_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G151_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G153_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G154_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G155_BF_200316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G166_BF_200316	Sida	Sida rhombifolia	Léguéma	1	7	-	-	-

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G167_BF_200316	Sida	Sida acuta	Léguéma	1	7	-	3	-
G168_BF_200316	Sida	Sida acuta	Léguéma	1	1	-	-	-
G171_BF_200316	Euphorbia	Euphorbia hirta	Léguéma	1	5	-	-	-
G173_BF_200316	Euphorbia	Euphorbia hirta	Léguéma	1	-	-	13	-
G191_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	-	-	13	-
G192_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	1	-
G193_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G194_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	2	-
G212_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	12	-
G213_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	15	-	-	-
G219_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	12	-
G222_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	-	-	12	-
G227_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	5	-	-	-
G229_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	6	-
G230_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	1	-
G231_BF_200316	Sida	Sida acuta	Léguéma	1	8	-	1	-
G235_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	11	-	-	-
G238_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G242_BF_200316	Sida	Sida acuta	Léguéma	1	7	-	1	-
G251_BF_200316	Euphorbia	Euphorbia heterophylla	Léguéma	1	5	-	-	-
G263_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	2	-	-	-
G264_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	6	-
G273_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	11	-
G274_BF_200316	Amaranthus	Amaranthus spinosus	Léguéma	1	6	-	-	-
G279_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	5	-
G280_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G282_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G285_BF_200316	Euphorbia	Euphorbia heterophylla	Léguéma	1	-	-	-	5
G290_BF_200316	Ageratum	Ageratum conyzoides	Léguéma	1	5	-	-	-
G292_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G294_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G295_BF_200316	Amaranthus	Amaranthus viridis	Léguéma	1	14	-	-	-
G298_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	7	-
G299_BF_200316	Sida	Sida acuta	Léguéma	1	8	-	11	-
G300_BF_200316	Ipomea	Ipomea alba	Léguéma	1	14	-	-	-
G301_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	1	-
G303_BF_200316	Boerhavia	Boerhavia erecta	Léguéma	1	-	-	10	-
G307_BF_200316	Euphorbia	Euphorbia hirta	Léguéma	1	5	-	-	-
G331_BF_200316	Sida	Sida acuta	Léguéma	1	16	-	11	-
G333_BF_200316	Sida	Sida acuta	Léguéma	1	2	-	7	-
G363_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	13	-
G372_BF_210316	Amaranthus	Amaranthus viridis	Léguéma	1	-	-	-	2
G386_BF_210316	Boerhavia	Boerhavia erecta	Léguéma	1	4	-	-	-
G401_BF_210316	Abelmoschus	Abelmoschus esculentus	Léguéma	1	-	-	-	3

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G444_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	5	-	-	-
G445_BF_210316	Sida	Sida acuta	Léguéma	1	7	-	1	-
G446_BF_210316	Sida	Sida acuta	Léguéma	1	7	-	-	-
G447_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	15	-	-	-
G448_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	2	-
G449_BF_210316	Amaranthus	Amaranthus spinosus	Léguéma	1	-	-	-	1
G454_BF_210316	Amaranthus	Amaranthus spinosus	Léguéma	1	5	-	-	-
G456_BF_210316	Euphorbia	Euphorbia hirta	Léguéma	1	16	-	-	-
G478_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	12	-
G480_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	15	-	13	-
G484_BF_210316	Amaranthus	Amaranthus spinosus	Léguéma	1	1	-	-	-
G508_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	15	-	12	-
G512_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	1	-	12	-
G522_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G523_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G524_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	16	-	10	-
G525_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G526_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G527_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G528_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G529_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G533_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	13	-	-	-
G534_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	4	-	2	-
G535_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	-	-	12	-
G555_BF_210316	Abelmoschus	Abelmoschus esculentus	Léguéma	1	7	-	7	-
G566_BF_210316	Sida	Sida acuta	Léguéma	1	2	-	10	-
G567_BF_210316	Sida	Sida acuta	Léguéma	1	2	-	1	-
G571_BF_210316	Abelmoschus	Abelmoschus esculentus	Léguéma	1	11	-	-	-
G579_BF_210316	Abelmoschus	Abelmoschus esculentus	Léguéma	1	-	-	-	2
G581_BF_210316	Ipomea	Ipomea alba	Léguéma	1	1	-	-	-
G590_BF_210316	Sida	Sida acuta	Léguéma	1	16	-	7	-
G593_BF_210316	Sida	Sida acuta	Léguéma	1	2	-	3	-
G594_BF_210316	Sida	Sida acuta	Léguéma	1	2	-	11	-
G595_BF_210316	Sida	Sida acuta	Léguéma	1	2	-	5	-
G597_BF_210316	Sida	Sida acuta	Léguéma	1	16	-	7	-
G621_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G622_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G623_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G624_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G625_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	3	-	-	-
G626_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G627_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G628_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G629_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	2	-	-

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G630_BF_210316	Solanum	Solanum lycopersicum	Léguéma	1	13	-	-	-
G640_BF_210316	Ageratum	Ageratum conyzoides	Léguéma	1	-	3	-	-
G651_BF_210316	Melochia	Melochia sp.	Léguéma	1	5	-	-	-
G662_BF_210316	Melochia	Melochia sp.	Léguéma	1	-	-	-	1
G670_BF_210316	Boerhavia	Boerhavia erecta	Léguéma	1	5	-	-	-
G673_BF_210316	Boerhavia	Boerhavia erecta	Léguéma	1	5	-	-	-
G674_BF_210316	Boerhavia	Boerhavia erecta	Léguéma	1	5	-	-	-
G698_BF_210316	Melochia	Melochia sp.	Léguéma	1	16	-	-	-
G701_BF_210316	Melochia	Melochia sp.	Léguéma	1	7	-	-	-
G710_BF_230416	Sida	Sida sp	Goué	1	-	4	-	-
G711_BF_230416	Sida	Sida sp	Goué	1	11	4	-	-
G714_BF_230416	Sida	Sida sp	Goué	1	-	3	-	-
G725_BF_230416	Hibiscus	Hibiscus sp	Goué	1	7	-	7	-
G726_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	5	-
G727_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	5	7	-
G728_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	-	7	-
G729_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	7	-
G730_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	-	7	-
G731_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	1	-
G732_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	-	-	12	-
G733_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	1	-
G734_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	1	-
G736_BF_230416	Heliotropium	Heliotropium indicum	Goué	1	-	6	-	-
G768_BF_230416	Amaranthus	Amaranthus viridis	Goué	1	7	-	-	-
G780_BF_230416	Physalis	Physalis ixocarpa	Goué	1	4	1	-	-
G792_BF_230416	Amaranthus	Amaranthus viridis	Goué	1	7	-	-	-
G795_BF_230416	Amaranthus	Amaranthus viridis	Goué	1	-	5	-	-
G802_BF_230416	Ageratum	Ageratum conyzoides	Goué	1	3	1	-	-
G803_BF_230416	Ageratum	Ageratum conyzoides	Goué	1	3	1	-	-
G808_BF_230416	Euphorbia	Euphorbia hirta	Goué	1	3	-	-	-
G821_BF_230416	Sida	Sida acuta	Goué	1	7	-	5	-
G831_BF_230416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	7	-
G833_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G834_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G835_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G836_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G837_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G838_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G839_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G840_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G842_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G847_BF_230416	Ageratum	Ageratum conyzoides	Goué	1	3	1	-	-
G856_BF_230416	Melochia	Melochia sp.	Goué	1	2	-	-	-
G859_BF_230416	Melochia	Melochia sp.	Goué	1	2	-	-	-

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G861_BF_230416	Sida	Sida acuta	Goué	1	2	-	3	-
G862_BF_230416	Sida	Sida acuta	Goué	1	12	-	3	-
G863_BF_230416	Sida	Sida acuta	Goué	1	2	-	1	-
G864_BF_230416	Sida	Sida acuta	Goué	1	2	-	1	-
G865_BF_230416	Sida	Sida acuta	Goué	1	2	-	8	-
G866_BF_230416	Sida	Sida acuta	Goué	1	2	-	11	-
G867_BF_230416	Sida	Sida acuta	Goué	1	2	-	1	-
G869_BF_230416	Sida	Sida acuta	Goué	1	8	-	11	-
G881_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G883_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G884_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G885_BF_230416	Capsicum	Capsicum frutescens	Goué	1	3	-	-	-
G888_BF_230416	Euphorbia	Euphorbia hirta	Goué	1	3	1	-	-
G902_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	2	-	-	-
G904_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	-	1
G905_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	-	1
G919_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	10	-	-	-
G943_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G944_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G945_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G946_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G947_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G949_BF_240416	Boerhavia	Boerhavia erecta	Goué	1	-	3	-	-
G964_BF_240416	Sida	Sida acuta	Goué	1	7	-	7	-
G966_BF_240416	Capsicum	Capsicum frutescens	Goué	1	3	1	-	-
G983_BF_240416	Leucas	Leucas martinicens	Goué	1	3	-	-	-
G985_BF_240416	Leucas	Leucas martinicens	Goué	1	3	-	-	-
G992_BF_240416	Ipomea	Ipomea alba	Goué	1	12	-	-	-
G993_BF_240416	Ipomea	Ipomea alba	Goué	1	1	-	-	-
G994_BF_240416	Ipomea	Ipomea alba	Goué	1	12	-	1	-
G995_BF_240416	Ipomea	Ipomea alba	Goué	1	11	-	-	-
G997_BF_240416	Ipomea	Ipomea alba	Goué	1	11	-	-	6
G998_BF_240416	Ipomea	Ipomea alba	Goué	1	1	-	1	-
G999_BF_240416	Ipomea	Ipomea alba	Goué	1	12	-	1	-
G1000_BF_240416	Ipomea	Ipomea alba	Goué	1	1	-	-	-
G1021_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	3	1	2	1
G1028_BF_240416	Boerhavia	Boerhavia erecta	Goué	1	-	1	-	-
G1033_BF_240416	Physalis	Physalis ixocarpa	Goué	1	4	-	-	-
G1035_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	-	1
G1037_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	1	-	-	-
G1040_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	5	-	-	-
G1050_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	-	1
G1057_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	-	1
G1079_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	-	-	-	1

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G1084_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	6	-	-	2
G1088_BF_240416	Capsicum	Capsicum annuum	Goué	1	3	-	-	-
G1092_BF_240416	Capsicum	Capsicum annuum	Goué	1	3	-	-	-
G1093_BF_240416	Capsicum	Capsicum annuum	Goué	1	4	1	-	-
G1094_BF_240416	Capsicum	Capsicum annuum	Goué	1	3	1	-	-
G1107_BF_240416	Capsicum	Capsicum frutescens	Goué	1	3	-	-	-
G1108_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	-	-
G1109_BF_240416	Capsicum	Capsicum frutescens	Goué	1	-	1	3	-
G1113_BF_240416	Capsicum	Capsicum frutescens	Goué	1	4	1	4	-
G1117_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	3	-
G1118_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	1	-
G1119_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	8	-	5	-
G1120_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	-	6	-
G1121_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	-	7	-
G1123_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	10	-	-	-
G1124_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	-	-
G1126_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	8	-
G1131_BF_240416	Boerhavia	Boerhavia erecta	Goué	1	7	-	8	-
G1135_BF_240416	Boerhavia	Boerhavia erecta	Goué	1	-	-	3	-
G1136_BF_240416	Boerhavia	Boerhavia erecta	Goué	1	7	-	-	-
G1137_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	-	-	3	-
G1138_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	8	-	-	-
G1139_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	-	-	9	-
G1140_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	-	-	9	-
G1142_BF_240416	Amaranthus	Amaranthus spinosus	Goué	1	-	-	3	-
G1148_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	-	-	8	-
G1149_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	-	-	9	-
G1150_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	-	-	3	-
G1151_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	7	-	9	-
G1152_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	-	-	9	-
G1155_BF_240416	Euphorbia	Euphorbia hirta	Goué	1	7	-	8	-
G1158_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	3	-
G1159_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	9	-
G1160_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	3	-
G1161_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	-	-	8	-
G1163_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	3	-
G1164_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	8	-	8	-
G1165_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	5	-
G1166_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	8	-
G1167_BF_240416	Abelmoschus	Abelmoschus esculentus	Goué	1	7	-	9	-
G1171_BF_240416	Hibiscus	Hibiscus sp	Goué	1	-	-	3	-
G1172_BF_240416	Hibiscus	Hibiscus sp	Goué	1	11	-	-	-
G1186_BF_240416	Amaranthus	Amaranthus viridis	Goué	1	-	-	8	-
G1189_BF_240416	Melochia	Melochia sp.	Goué	1	12	-	-	-

Ech ID	Genre	Host	Locality	Positive	Cluster A	Cluster B	Cluster S	Cluster M
G1199_BF_240416	Sida	Sida acuta	Goué	1	12	-	3	-
G1200_BF_240416	Sida	Sida acuta	Goué	1	2	-	7	-
G1201_BF_240416	Sida	Sida acuta	Goué	1	2	-	1	-
G1202_BF_240416	Sida	Sida acuta	Goué	1	8	-	5	-
G1203_BF_240416	Sida	Sida acuta	Goué	1	2	-	10	-
G1204_BF_240416	Sida	Sida acuta	Goué	1	2	-	1	-
G1205_BF_240416	Sida	Sida acuta	Goué	1	2	-	10	-
G1206_BF_240416	Sida	Sida acuta	Goué	1	8	-	7	-
G1207_BF_240416	Sida	Sida acuta	Goué	1	12	-	11	-
G1208_BF_240416	Sida	Sida acuta	Goué	1	8	-	1	-
G1213_BF_240416	Undetermined	Undetermined	Goué	1	7	1	-	-
G1215_BF_240416	Undetermined	Undetermined	Goué	1	1	-	-	-
G1218_BF_240416	Undetermined	Undetermined	Goué	1	-	-	-	3

Molecular epidemiology of geminiviruses responsible for emerging diseases on vegetable crops in Burkina Faso

Geminiviruses have emerged to become one of the largest and most economically important groups of plant-infecting viruses, and geminivirus-induced diseases are a major threat to worldwide vegetable production, particularly in tropical and subtropical regions of the world. Importantly, the accumulated body of work on some of the most important geminiviral associated diseases clearly demonstrate the role of geminiviruses associated with wild plants on the emergence of disease on imported crops. Moreover, recent metagenomic data suggested that the vast majority of viruses characterized from crops represent only a small fraction of the phytoviruses in general. It is therefore of prime interest to obtain a better knowledge of viral diversity infecting crops and wild plants, the main epidemiological parameters involved in their emergence and their dynamic at the scale of agro-ecological systems.

In this work, a survey of solanaceous crop fields and their surrounding uncultivated plants from 48 localities in Burkina Faso and Ivory Coast was performed. The sample analysis using classical molecular biology tools both confirm the incompleteness of our knowledge of the geminivial diversity and the existence of numerous alternative wild host plants. At least five species of begomovirus and mastrevirus were found in association with the ToLCD-TYLCD. A North to South increasing gradient of complexity of viruses populations was uncover with PepYVMLV being the most prevalent on cultivated solanaceous plants.

The discovery of the association of a newly described DNA-B component with the PepYVMLV also lead to the study of the epidemiological parameters of this co-infection. Despite this association being relaxed, it was demonstrated that the virulence of the disease, the viral accumulation and the transmission by *Bemisia tabaci* were increased with the presence of the co-infection with the DNA-B component. All these factors are probably associated with the success of this association on the field. Because of the extreme severity of the resulting disease, the diffusion of this new DNA-B component at a larger scale would represent a major threat to tomato culture in Burkina Faso, Africa and the world in general.

The use of a metagenomic approach, allow the generalization of our findings to full agro-ecological settings. Besides confirming previous species discovery, species yet undescribed in Burkina Faso along with completely new begomovirus species were described. The inspection of the virus-plant and virus-virus associations networks allow to uncover strong links existing between the viral corteges associated to groups of cultivated and uncultivated plants. These findings emphasized the necessity to consider full agro-ecological settings plant diversity rather than only crops in order to understand and prevent geminiviruses associated diseases.

Globally, our results highlight the necessity to carry on the ongoing plant disease monitoring work and the inventory of viral populations associated with cultivated and uncultivated plants in order to understand the functioning of natural geminiviral community and the impact of human practices on the emergence of viral disease.

Epidémiologie moléculaire des géminivirus responsables de maladies émergentes sur les cultures maraîchères au Burkina Faso

Les épidémies virales constituent une menace majeure pour les cultures dans de nombreuses régions tropicales et subtropicales du monde. En Afrique de l'Ouest et Centrale, malgré la caractérisation d'une multitude de nouvelles espèces de géminivirus, les connaissances sur l'épidémiologie des maladies virales émergentes restent incomplètes, notamment concernant la diversité et l'identification précise des espèces virales incriminées. Paradoxalement, si l'on considère que la plupart des maladies émergentes causées par les géminivirus sont liées à des virus infectant naturellement les plantes sauvages, et qui se seraient adaptés aux plantes cultivées introduites, la diversité des populations virales dans leurs écosystèmes naturels reste très largement méconnue. De plus, le caractère non pérenne de la majorité des plantes maraîchères suggère l'existence de plantes hôtes alternatives ou réservoirs, qui permettent aux populations virales de se maintenir lors des périodes d'intercultures. Par conséquent, pour lutter contre ces nouvelles maladies virales dévastatrices, il est primordial d'acquérir de meilleures connaissances de la diversité des populations virales des plantes cultivées et sauvages, de leurs répartitions géographiques, de leurs liens phylogénétiques, de leurs pathogénicités et des principaux facteurs épidémiologiques à l'échelle des agro-écosystèmes.

Ainsi, les travaux entrepris dans cette thèse ont porté sur l'analyse d'échantillons de plantes maraichères et non cultivées provenant de 48 localités réparties sur le territoire du Burkina Faso et de Côte d'Ivoire. Nos principaux résultats confirment à la fois la non-exhaustivité de notre connaissance de la diversité des géminivirus présent sur les cultures maraîchères, avec l'existence d'au moins cinq espèces de géminivirus impliqués dans la maladie du ToLCD-TYLCD, d'un gradient Nord-Sud de complexité des espèces de géminivirus au Burkina Faso, la prédominance du PepYVMLV sur les solanées cultivées et la présence de nombreuses plantes alternatives ou réservoirs.

Au regard de la découverte d'un nouveau composant d'ADN-B en association avec le PepYVMLV, les principaux paramètres du pouvoir pathogène du PepYVMLV en infection simple ou mixte avec le composant d'ADN-B ont été évalués en conditions contrôlées. Nos résultats démontrent que même si le PepYVMLV n'est pas un bégomovirus bipartite strict, l'ADN-B représente à la fois un activateur fort de la virulence, de l'accumulation du virus dans la plante et de la transmission par son insecte vecteur *Bemisia tabaci*, qui semblent concourir à sa prédominance sur le terrain. Une diffusion plus large de ce composant d'ADN-B représenterait une menace majeure pour la culture de la tomate au Burkina Faso, en Afrique et plus largement dans le monde.

L'élargissement des analyses par l'utilisation d'une approche métagénomique à non seulement permis de confirmer la présence d'une majorité des virus décrits préalablement avec les approches classiques, mais également la caractérisation de nouvelles espèces virales. L'analyse des réseaux d'associations virus-plantes hôtes et virus-virus ont montré des liens forts entre les populations virales associés à des groupes de genres de plantes cultivées et non cultivées soulignant la nécessité de considérer les agro-écosystèmes dans leur totalité pour lutter contre les maladies à géminivirus.

L'ensemble de nos résultats soulignent l'importance des travaux d'épidémiosurveillance et d'inventaire des populations virales sur les plantes cultivées et sauvages afin de comprendre le fonctionnement des communautés virales à l'échelle des agroécosystèmes et l'impact des perturbations anthropiques sur les processus d'émergence de nouvelles maladies.

POLE RECHERCHE
Ecoles Doctorales

LETTRE D'ENGAGEMENT DE NON-PLAGIAT

Je, soussigné(e) Alassane OUATTARA, en ma qualité de doctorant(e) de l'Université de La Réunion, déclare être conscient(e) que le plagiat est un acte délictueux passible de sanctions disciplinaires. Aussi, dans le respect de la propriété intellectuelle et du droit d'auteur, je m'engage à systématiquement citer mes sources, quelle qu'en soit la forme (textes, images, audiovisuel, internet), dans le cadre de la rédaction de ma thèse et de toute autre production scientifique, sachant que l'établissement est susceptible de soumettre le texte de ma thèse à un logiciel anti-plagiat.

Fait à Saint-Pierre le : 09/11/2017

Signature :

Extrait du Règlement intérieur de l'Université de La Réunion
(validé par le Conseil d'Administration en date du 11 décembre 2014)

Article 9. Protection de la propriété intellectuelle – Faux et usage de faux, contrefaçon, plagiat

L'utilisation des ressources informatiques de l'Université implique le respect de ses droits de propriété intellectuelle ainsi que ceux de ses partenaires et plus généralement, de tous tiers titulaires de tels droits.
En conséquence, chaque utilisateur doit :
- utiliser les logiciels dans les conditions de licences souscrites ;
- ne pas reproduire, copier, diffuser, modifier ou utiliser des logiciels, bases de données, pages Web, textes, images, photographies ou autres créations protégées par le droit d'auteur ou un droit privatif, sans avoir obtenu préalablement l'autorisation des titulaires de ces droits.

La contrefaçon et le faux
Conformément aux dispositions du code de la propriété intellectuelle, toute représentation ou reproduction intégrale ou partielle d'une œuvre de l'esprit faite sans le consentement de son auteur est illicite et constitue un délit pénal.
L'article 444-1 du code pénal dispose : « Constitue un faux toute altération frauduleuse de la vérité, de nature à causer un préjudice et accomplie par quelque moyen que ce soit, dans un écrit ou tout autre support d'expression de la pensée qui a pour objet ou qui peut avoir pour effet d'établir la preuve d'un droit ou d'un fait ayant des conséquences juridiques ».
L'article L335_3 du code de la propriété intellectuelle précise que : « Est également un délit de contrefaçon toute reproduction, représentation ou diffusion, par quelque moyen que ce soit, d'une œuvre de l'esprit en violation des droits de l'auteur, tels qu'ils sont définis et réglementés par la loi. Est également un délit de contrefaçon la violation de l'un des droits de l'auteur d'un logiciel (…) ».

Le plagiat est constitué par la copie, totale ou partielle d'un travail réalisé par autrui, lorsque la source empruntée n'est pas citée, quel que soit le moyen utilisé. Le plagiat constitue une violation du droit d'auteur (au sens des articles L 335-2 et L 335-3 du code de la propriété intellectuelle). Il peut être assimilé à un délit de contrefaçon. C'est aussi une faute disciplinaire, susceptible d'entraîner une sanction.
Les sources et les références utilisées dans le cadre des travaux (préparations, devoirs, mémoires, thèses, rapports de stage…) doivent être clairement citées. Des citations intégrales peuvent figurer dans les documents rendus, si elles sont assorties de leur référence (nom d'auteur, publication, date, éditeur…) et identifiées comme telles par des guillemets ou des italiques.

Les délits de contrefaçon, de plagiat et d'usage de faux peuvent donner lieu à une sanction disciplinaire indépendante de la mise en œuvre de poursuites pénales.