



HAL
open science

Evolution et écologie des géminivirus à l'échelle des écosystèmes réunionnais : impact du paysage végétal et de l'anthropisation

Sélim Ben Chehida

► **To cite this version:**

Sélim Ben Chehida. Evolution et écologie des géminivirus à l'échelle des écosystèmes réunionnais : impact du paysage végétal et de l'anthropisation. *Phytopathologie et phytopharmacie*. Université de la Réunion, 2023. Français. NNT : 2023LARE0037 . tel-04860568

HAL Id: tel-04860568

<https://theses.hal.science/tel-04860568v1>

Submitted on 1 Jan 2025

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.

UNIVERSITÉ DE LA RÉUNION
Faculté des Sciences et Technologies

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

École doctorale n° 542, Sciences, Technologies et Santé (EDSTS)

Thèse de doctorat pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ DE LA RÉUNION

Spécialité de doctorat : **Écologie et Évolution**

Evolution et écologie des phytovirus à l'échelle des écosystèmes réunionnais : impact du paysage végétal et de l'anthropisation

Thèse présentée et soutenue par

Sélim BEN CHEHIDA

Le 07 décembre 2023, devant le jury composé :

Bernard REYNAUD	Professeur, Université de La Réunion	Président
Charlotte TOLLENAERE	Chercheuse HDR, IRD, Montpellier	Rapporteuse
Benoit MOURY	Directeur de recherche, INRAE, Avignon	Rapporteur
Nathalie BECKER	Maîtresse de conférence, HDR, MNHN, Paris	Examinatrice
Jean-Michel LETT	Chercheur HDR, CIRAD, Réunion	Directeur de thèse
Pierre LEFEUVRE	Chercheur, CIRAD, Réunion	Co-encadrant de thèse

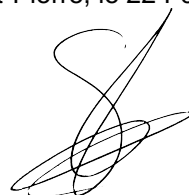
ATTESTATION SUR L'HONNEUR DE NON PLAGIAT

Je soussigné(e)Sélim BEN CHEHIDA..... déclare sur l'honneur que ce mémoire est le fruit d'un travail personnel et que je n'ai ni contrefait, ni falsifié, ni copié tout ou partie de l'œuvre d'autrui afin de la faire passer pour mienne.

Toutes les sources d'information utilisées (supports papiers, audiovisuels et numériques) et les citations d'auteur ont été mentionnées conformément aux usages en vigueur.

Je suis conscient(e) que le fait de ne pas citer une source ou de ne pas la citer clairement et complètement est constitutif de plagiat, que le plagiat est considéré comme une faute grave au sein de l'Université et qu'il peut être sévèrement sanctionné.

Date et signature de l'étudiant :
Saint-Pierre, le 22 Février 2024



Règlement intérieur de l'Université de la Réunion (tel que 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 »

« 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. »

REMERCIEMENTS

« Il faut s'entendre sur le mot "réussir", moi je crois qu'on réussit une seule chose, on réussit ses rêves. »

[J. Brel]

À Chedlia, Habiba et Chedly, vous aviez raison, je suis finalement bien devenu docteur.

Je remercie tout d'abord, l'ensemble des membres du jury de thèse **Bernard Reynaud, Charlotte Tollenaere, Benoît Moury et Nathalie Becker** d'avoir accepté de faire partie de ce jury et d'avoir consacré du temps pour évaluer ces travaux de thèse. Par ailleurs, je remercie aussi les membres de mon comité de thèse, **Philippe Roumagnac, Denis Filloux, Thierry Candresse, Christophe Lavergne et Frédéric Chiroleux**, pour le suivi et les conseils avisés qui m'ont permis de réaliser ce travail de thèse durant ces trois ans.

À **Jean-Michel**, mon très cher mentor, mein General, merci de m'avoir fait rentrer dans le monde de la virologie végétale, et de tout ce qui l'entoure, d'avoir investi tant de temps et d'énergie à former et corriger le jeune padawan que je suis. En somme, beaucoup de rouge, du boulguibouga, mais le manuscrit de Jean-Sélim lé la !
Hopla

À **Pierre**, pour tout ton savoir, ta patience que j'ai mis à rude épreuve, ton envie de transmettre et d'expliquer (un peu plus qu'une fois), ta franchise, tes schémas légendaires, ton aisance bio-informatique, ton œil esthétique et surtout ton soutiens inébranlable. Merci pour énormément de choses que je n'aurai pas la place de lister ici. Je n'oublierai pas tous les débats gastronomiques que nous avons pu avoir. Je pense que mon seul regret est que tu n'aies pas pu (encore) goûter à mon couscous, mais ça, ça se rattrape (je parle encore de bouffe c'est fou). Entre nous, le plus humain de nous deux, je me demande si, avec une p-value inférieure à 0.05, ce n'est pas significativement toi Sidi. Merci

Bon ben voilà, **Maman et Papa** le quatrième docteur Bench' est là. Je vous remercierai jamais assez pour tout l'immense soutiens et présence que vous m'apportez depuis le tout début, il s'en est passé des choses depuis mon départ de Tunis et quelles épreuves, du coyote mauve à ce manuscrit... Vous avez tous les deux étaient là pour me permettre d'aller le plus loin possible malgré les différentes péripéties et me donner l'envie de faire ce que je fais aujourd'hui. Merci infiniment !!

A **Yacine**, bouffa, sans qui jamais je n'aurai mis un orteil dans ce monde là. Merci d'avoir toujours été inspirant, à fond dans tout ce que tu entreprends, et surtout motivant ! L'aventure est loin d'être fini, mais je suis sûr que ça va valoir le coup !
Salon de coiffure

Mourad & Maud, ce n'est pas un vocal de 30 minutes qui me permettra d'exprimer toute la gratitude que j'ai envers vous deux. Moi aussi je fête les 11 ans avec vous finalement haha ! Vous êtes la rencontre la plus importante depuis mon départ de Tunis. Merci de m'avoir écouté, d'avoir été là quand c'était pas forcément le plus facile, de m'avoir toujours si bien reçu, à Toulouse ou à Paris, de m'avoir vu dans tous les états possibles et imaginables et de me permettre de systématiquement passer de bons moments avec vous. Aller Maud, aller aller ! Lâches ces sirènes maintenant !

Hassilou, je suis content que ça se soit bien passé

Khouya **Sofiane**, merci d'avoir été ma source principale de rire et de bleada depuis si (trop ?) longtemps maintenant. Un jour nous remonterons sur les planches ensemble, tous les deux en tant que docteurs en plus, ceci est non négociable !! Dire que tu prépares Montreux, pendant ce temps là... Ama 9olli, tekeli fik ?

Un infini merci à tous ceux qui ont été là au début, à la fin, au milieu, de temps en temps, quasiment jamais, tout le temps... pour tous ces messages drôles, rassurants, motivants et qui m'ont donné l'énergie de tout donner et de terminer cette **dernière ligne droite**, qui dure depuis bien trop longtemps maintenant

À ma famille montpelliéraine qui m'a offert la plus belle découverte de l'île et d'avoir été là depuis le début des 015, et de ne jamais avoir perdu sa bonne humeur !! Merci d'être venus me voir, de m'avoir écrit si souvent et d'avoir été là tout du long.

Baptiste c'est bon tu peux revenir pour retenter le volcan ! **TheDupe** je t'attends pour le niveau 3 ? **Maiwenn**, le champion nuisette est là ! **Ophélie**, ton fan numéro 2 rentre bientôt ! **Orion**, j'ai trouvé un tacos avec 500g de viande, c'est pas écolo, mais faut trop que tu viennes le tester ! **Hugo**, ton ti-punch est prêt, t'arrives ? **Ada**, je t'ai peut-être sauvé la vie une fois, mais t'as sauvé mon mental des centaines d'autres. **Perle**, on en parlait ensemble comme d'un lointain objectif, bientôt la thèse ne sera qu'un lointain souvenir ! **Thomas**, tes vélos traversent les océans ? **Romane**, j'espère que t'es prête à faire la visite guidée de ton île ! **Pégaze**, on l'écrira notre histoire de jazz des bas fonds du Harlem ! **Gautrelle**, la queen, tu penses répondre au vocaux avant ou après ta démission ? **Chloé Susu**, j'espère que mon assistance téléphonique par décalage horaire t'aura été d'un bon soutien. **laptouff**, on se fait quand Piton des Neiges – Mada en parapente ? **Thib** aller sors de ton antre et vient prendre une dodo ! **Totor**, merci de m'avoir transmis ta passion des mots fléchés, j'ai pas vu la fin de la thèse passée.. **Jehad** et **Loicia** préparez vos estomacs, y a du buffet à volonté à rattraper ! Sans oublier **Boulette**, **Huguette**, **Chlochlo**, **Zoé**, **Zoé**, **Debby**, **Gisèle**, **Melou**, **Elo**, **Simo**, **Marie**, **Cécile**, **DJAAVAAN**

Hahaha t'as cru que j'allais t'omettre **Anton**, mais j'ai pas PERDU mon sens du spectacle et de la surprise !!! macdo ? Et tu sais ce qu'on dit : « *Moskau, Moskau, Wirf die Gläser an den Wand, Russland ist ein schönes Land, Ho-ho-ho-ho-ho !* »

Bon et sinon, Raaaaaaaaaaaaage Caaaaaaaaaaaaaaage ! *bruits métallique de poêle*

À tous les incroyables copains qui m'ont adopté à la Réunion et avec qui j'ai découvert ce caillou magique. Merci pour ce soutien quotidien, et ces instants que j'aurai aimé vivre et revivre 974 autres fois encore avec vous tous

Emma, présente depuis mon (quasi) premier jour jusqu'à mon (possible) dernier jour sur l'île, merci pour tout, ta présence, ton soutien, la 0-3000, la confiance, le moral, la confiance... et bien plus de dieu ! Aaaaah mais ça vient de là !! **Margot**, par où commencer ? Il me faudrait une thèse entière pour décrire toutes les choses pour lesquelles j'aimerais te remercier, mais je dirai que d'avoir accepté tant et tant de ma part a été l'un des plus gros soutien que j'ai reçu. Merci infiniment. **Mathilde**, ta présence perpétuelle depuis 3 ans, toujours motivante motivée et dans la joie m'est très précieuse ! Et encore plus quand il s'agit de créer des jeux d'équipes ! Aller, p'tite pause café ? **Stan**, le constructeur de l'extrême, le grimpeur intense ! Vas-y on dit « Je ne l'ai pas » comme code ! Ca sera sûrement une bonne idée ! Haha ! A ça de gagner Run Express 2. **Filou**, mon modèle de thésard à toute épreuve ! Je suis content d'avoir partagé un bout de cette aventure avec toi, et qui sait, ptet' qu'on se retrouvera en post-doc dans le pays des kangourous ? En attendant, vient on va plonger ! **Lysa**, je crois que si quelqu'un devait représenter la notion de positivité, ça devrait être toi ! Toujours partante, souriante et dans le partage. Tellement heureux d'avoir profiter de tous ces moments avec toi ! **Thomas**, un ami ? Nan, bien plus ! Merci de m'avoir fait retrouver le chemin de ma vraie maison tant de fois ! **July**, la main verte, le bricolo, le touche à tout, l'homme au mille et une passions, qui ont très souvent donné leur fruits (ouah un jeu de mots!). Change vraiment rien, sauf les encombrants ! Arrêtes de ramasser les encombrants ! Bee Happy ! **Corette**, je t'avoue que ton départ, ton café tes clopes ta georgette et ta passion pour les moulins de vols et l'escalade me manque beaucoup ! La Némobile te salue. **Paulette**, la MJ par excellence ! Tant de créativité et d'enthousiasme dans une seule personne. J'espère qu'on se retrouvera vite autour d'un love letter au milieu d'un cirque en bivouac. **Maria**, si l'Italie devait avoir une représentante internationale, ça devrait être la pizza aux pâtes ? Haha ! Nan ça serait toi ! Oui oui oui !! Ta joie, ton tiramisu et ton authenticité me manquent ! **Nina**, la plus grande artiste bien trop en avance sur son temps. Merci d'être la formidable personne que tu es. **Jérémy**, l'homme à la moustache inébranlable, vainqueur des Alpes et conquérant de Mafate, on se refait la fin de la MTT quand tu veux. **Kosh**, à tes aventures de pêche, ta sudisterie inégalable et nos début de journées bien trop matinaux. **Simon**, on se les fait quand ces savons ?? **Chloé**, à ta joie, ta bonne humeur, ta positivité et ta motivation toujours au tip top ! **Lélé**, à la douce et agréable personne que tu es, toujours motivante et motivée. **Juliette**, et tous ces petits moments, souvent autour d'une bonne assiette, en ton agréable compagnie. **Hugo & Olga**, a votre authenticité et bienveillance, votre folie et drôlerie. **Gégé**, maintenant que la thèse m'a tué, on se refait une murder ? A **Margaux**, **Antoine**, **Emma**, **Micka**, **Migui**, **Jules**, **Charlène**, **Julie**, **Juliette**, **Charlie**, **Paulin**, **Ambre**, **Clairette**.

A **Nathanaelleuh**, t'es sûr ? Tu veux toujours pas te lancer dans l'épopée doctorale ? Mais à qui je vais bizuther les sabots moi maintenant ? Faut que j'aille au fin fond des Alpes c'est ça ? Su-Per Nils, bravo pour l'appareil photo.. A **Lucie**, promis je serai à toutes représentation en fanfare ! A **Flo' & Aurélie**, le duo choc du Cirad, ça part en Maison Glace sunset V&B? Moi suis chaud ! A **Juliette**, du coup si je trouve un post-doc à Mtp, c'est moi qui te suis ? A **Paul**, la surprise du trail réunionnais, tu sais supporter Gaab, c'est la plus belle preuve d'endurance ! **Pascal**, je t'attends

toujours sur MK8 !! A **Camille** et **Floriane** pour la joie de vous retrouver à chaque fois et votre inspiration artistique unique en son genre !

A la meilleure équipe du 3P, la GeminiDream Team du Cirad ! A la meilleure dynamique possible, permise par super **Mumu** la reine du labo, **Babbitha** qui a fait de la magie et sans qui le NV4 ne serait qu'une utopie, et biensûr ma chère stagiaire unique et préférée, bientôt tout autant docteur, **Heemee** ! A madame docteur **Delphine Massé** ! On l'a fait !! A base de longues discussions et de litres de thé, mais le manuscrit est là ! Haha !!

A l'ensemble de l'African Phytovirus Way of Life

« Mmmhmm... BON ! » « Ca va aller, ça va aller ! » comment oublier ces phrases que tu répétais sans cesse pour m'encourager et me permettre de continuer malgré que « certaines fois ça te donne envie de casser R » mon très très cher voisin, M. le président **Alassane** !! Quelle belle année en ta présence, quelle fantastique rencontre ! Hâte de venir te voir, toi et tous tes projets !! « Ah ces gars là qui regarde leur télévision »

N'Chooooooooooooo !!!!!!! Ah sacré **Jocelin**! Sache que monsieur Chichon a été heureux de partir avec toi à Aussois ! A quand le prochain rhum ?

Daniel, mon très très cher Daniel, chef de l'Attieke, j'espère que t'as préparé le stock de vitamines pour cette dernière année ! Parce que tu sais ce qu'on dit ? « La thèse, c'est pas de la foutaise ! » Bon courage frérot !! Prépare Abidjan, parce que je vais débarquer sans prévenir, j'te le dis.

M-V is in the place ! Mais oui, on t'adopte dans notre collectif très officiel, et puis c'est connu, Mada, c'est l'Afrique ! Ca va me manquer de venir, te piquer des sucreries et m'asseoir et regarder par la fenêtre en racontant ma vie à ton bureau. Bon courage pour cette 3ème/4ème année de thèse !

Sans oublié **Olivier** !! Notre entomo !! Beaucoup de belles choses dans cette thèse qui t'attend !!

J'ajoute à ce petit beau monde le maître des cicadelles, roi du maïs, **Martial** à l'ingéniosité sans limite et à l'aide unique, on part à la retraite ensemble M. Grondin ! Et bien sûr, la fine équipe **Jimbo** & **Sergio** à la bonne humeur continue. **Stéphanie** et ta présence précieuse en micro-bio mais surtout au bureau. **Véro&Véro**, **Claudine**, **Jean-Jacques**, **Jérémy**.

Mais aussi à notre Ironman à tous, sidi **PYT** et ton comté qui m'a servi de potion magique. **Nath'** et tous tes souvenirs tunisiens, ta mlokhia, ton soutien infini, tes ploufs et ta joie incommensurable. **Fred**, j'espère qu'assister à la fin des thèses de deux frères n'a pas été trop dur! **Fred** pour ton humour, les malheurs de R, et ton gilet jaune ! **Olive** et les cafés/thés aux aurores du cirad. **Gilles** et ton esprit... sportif même à l'épreuve de la thèse ! **Yann** pour tes heures aux cirad quasiment supérieures au miennes. Je me demande si tu faisais ça pour ne pas que je sois seul ou si le doctorat c'est pas si pire en fait ? Merci pour tout ton aide avec mes gestions TRÈS organisées de plantes. **Nico**, et tes rêves dignes des meilleurs scénarios... Tu pourras m'aider pour mon générique de fin de soutenance ?

A l'ensemble des post-docs, docteurs et doctorants à qui je ne souhaite que du courage, de la force, de l'endurance et de trouver du bonheur dans cette aventure pleine de... de... d'embûches. A **Nico**, quelle idée d'être pap... euh docteur ! A **Paola**, merci de m'avoir donné le goût de rester tard, d'avoir un appart de luxe professionnel et une porte de bureau memesthique. A **Ismael**, à ton humour délirants, tes références dithyrambiques (THYRAMBIQUES). A **Gaëlle&Gaëlle** vous qui avez réussi à trouver une suite si vite... A **Benoît** mon troisième jumeau, promis, on terminera plus jamais à 20h ! C'est quand qu'on monte notre food truck Chez Benoit&Benji, je suis sûr que **Juliette** (que je remercie entre ces parenthèses, parce que chacun des moments avec toi était de douces parenthèses durant cette thèse) sera heureuse de nous soutenir et de nous faire la plus belle fresque du monde. A **Marine**, dont la thèse n'est qu'un détail de la vie. Tu devais pas soutenir en même temps que moi d'ailleurs ? Finalement, je passerai bien plus de 20 minutes avec toi. A **L.A.**, du politiquement pas correct, du tarot, des stats du cinéma et des bouses de vaches... qu'est ce qu'on peut partager d'autres entre petits hommes trapus myopes barbus à la chevelure éparse ? A **Emma** qui nous a quitté bien trop tôt pour... Versailles ?! Nan mais t'as cru qu'on allait venir te voir en plus ?? Bon okay, j'arrive.. A **Roxane**, c'est quand que tu deviens formatrice officielle du cirad ? A **Quentinou**, tu sais, je pense qu'il suffit que tu appliques ta passion pour les animes à la thèse, le temps d'une saison des feux de l'amour, et hop c'est dans la poche ! **Ze Q**, la tour operator du 3P espère que vous êtes partis assez tôt du cirad et vous souhaitez d'atteindre le même classement au beach tennis qu'à l'ED. A **Eva** et nos points matinaux pour le courage de chaque journée. A **Terriane**, ça y est, cette fois-ci tu peux me demander si j'ai un enfant, elle s'appelle Thèse, c'est original je sais, mais Terriane aussi c'est original et alors ? A **Océane** et **Thomas** pour votre agréabilité, ça se dit pas, mais vous m'êtes agréables, qu'est ce que j'y peux moi ? A **Margaux**, t'inquiètes, on va le saigner le Wingspan !!

A la communautés de stagiaires beaucoup trop nombreuse et je sais que je vais pas arriver à tous vous nommer, mais je sais que j'ai passé de formidables moments avec vous, la première ascension du Piton des Neiges (Rox in the coeur), les kaz, les soirées aux kaz, les jeux-dis, ce bon vieux «insérer un prénom », les tarots, les couscous pour 30, 1234, les poules mâles et du coup **Roxane**, **Ludo**, **Laeti**, **Charline**, **Melina**, **Lyna**, **Nato**, **Maelle**, **Alain**, **Nolwenn**, **Nico** (la crevette curieuse... *miaulement de chat sur les toits de Paris*), **Robin**, **Paul**, **Clément**, **Alexis**, **Justine**, **Justine**, **Céline**, **Lucas**, **Julien**, **Manon**, **Bastien**, **Nino**, **Vincent**, **Orane**, **Leila**, **Pauline**, **Pierre**, **Sam**, **Quentin**, **Axelle**, **Blandine**, **Jeanne**, **Blanche**, **Lavena**, **Romane**, **Coline**, **Hugo**, **Pablo**, **Moïra**, **Corey**, **Anne**, **Clémence**, **Luna**...

A **Oumi**, wiliwiliwiliwiliwili ! Tu crois qu'on pourra mettre Ya Rayah à ta soutenance ? Ca y est c'est quasiment fait, we are douctoouurrr & douctoouurrra ! A **Charlotte & Laurence** le meilleur duo de l'univers de tous les temps ! Merci pour tout tout tooouuut votre soutien, votre bonne humeur, mon sens de l'orientation, le fromage et le saucisson ! A quand le prochain échantillonnage ? Mamma !! A **Manu**, tes histoires, les manières de les raconter et tes millions de reads en sortie de MinION... J'espère rencontrer un rhinocéros avec toi un jour ! A **Denis**, roi de la bio-info et du baby-foot... au sol ? A **Mathilde**, **Noun**, **Thomas**, **Melia**, **Babil**, en espérant avoir

l'occasion de vous retrouver un jour, pour parler phytovirus ou savourer du rhum arrangé A **Nils**, mon co-roi de la piste de danse !! A Madame **Claire Neema**, parce qu'être arabe c'est bien, mais visiblement libanais c'est encore mieux ! Merci pour ta gentillesse, ta sincérité et ta joie légendaire ! Au cheikh qui fait bouger toute ce beau monde, notre chef à tous, **Sidi Abdel Roumagnac** ! Merci pour ta contribution et ton soutien dans ma thèse, mais surtout ta bonne humeur et enthousiasme perpétuel.

Au Dr. **Alex** OUTIIIIIIIIIIIIIIIIIIII !! C'est bon, je peux me remettre au trail, on traverse les Alpes ?

Au Dr. **Manon**, c'est sympa ça comme couleur corail ?

Au Dr. **Tristan**, j'te prends au Catan quand tu veux ! Un ti kebab ? Bêêêêêh

Au Dr. **Nico**, el Sharknado Cabron, un ti broc et c'est reparti !

Au Dr. **Qwentine**, j'espère que les phosphatières du Quercy ne sont pas trop hautes

Au Dr. **Narla**, prepara o Brasil e a cachaça, estou chegando !!!

Ayéééé Je suis docteur aussi les coupaings !!! Gras-gredi autour d'un Darwin's Choice ?

A Dr **Eluette**, Dooouuuuce Eluette !

A **Alex, Kris** et doctour **Griita**

Au meilleur club de plongée et ses formidables plongeurs !!

Aquabulle

A **Johanna**, j'espère que t'es prête à faire de moi un plongeur autonome, parce que c'est pas gagné gagné. A **Damien**, seigneur du bateau, roi du royaume sous la mer, merci de m'avoir mis la tête dans l'eau ! A **Sophie**, chokran beaucoup pour ta gentillesse, et Mich zum Meeresgrundreiniger gemacht zu haben. Et tombes pas dans les virus hein ! A **David**, c'est notre luxembourgeois à tous, mais c'est surtout notre ami, il marquera l'histoire. A **Julie**, pour ta gentillesse, ton soutien et ta persévérance à toute épreuve ! A **Marion**, ayéééé ! Premier grand raid terminé, je peux me mettre à celui qui traverse l'île avec toi ! **Serena & Arnaud** les parents modèles du XXI^{ème} siècle, quelle belle mais courte rencontre, préparez Lilles, je vais leur montrer ce que c'est qu'être tunisien ! To'tuuuu

Au TER avec qui j'ai passé une formidable année sur les planches et continue à passer de beaux moments autour de kaskrout tunisiens

Audrey, BeneBene, Michel, ah pardon c'est pas la bonne oreille, **Oliv, Nolwenn, Elsa, Alex**

A **Julien** merci pour tes messages drôles et de souvent penser à moi

A mon quatuor inséparable, **Lorena, Carole & Joëlle**, de l'internat à aujourd'hui (enfin sauf géographiquement). Merci pour ces quelques appels, c'est remontage de moral, ces mise à jour sur nos vies et ces encouragement ! Vous me manquez tant !

A **Thibault, Mélanie, Anais, Besbess, Manon, Nils & Celia**, les meilleurs marraines et parrains qui se sont occupés comme il faut de leur petit et ce jusqu'à aujourd'hui.

A **Ornella**, ton soutien précieux été un bel atout dans cette épopée !

A **Robin**, véto aux TAAF, le gars cumule les rêves en trois mots... ça c'est mon frerot ! Hâte de te retrouver, et encore mieux sur scène !!

A **Ondine** kikiiii ! Non, je ne me lancerai pas dans un de ces messages beaucoup tr... quoi, c'est déjà trop long ? Nan ça ne le sera jamais assez pour te dire à quel point je te remercie pour tes appels, tes lettres et tes messages. A **Estelle** qui apparaît aux moments les plus importants ! Merci de ta présence

Aux pitis bizuths, trop attachés à leur vieux spé, venus le rejoindre à la Réunion. Heureux d'avoir passé ces moments avec vous la **Vieille championne**, le **Ninja**, **Cavex**, **Loana**, **Thibaud** et **Mathilde** ! J'espère que vous êtes prêts à vous occuper du vieux sénile à la retraite

A **Julie**, pour avoir co-habité quelques jours avec l'ermite que j'étais et m'avoir nourri au coup de tarte et de chouchou!

A **Sandra**, **Féryel**, **Amine**, le Némoto est enfin docteur !! Maintenant, j'ai hâte de vous retrouver !!!

A **Kenza**, à toutes ces années, où rien à changer et à toutes celles qui nous attendent

A mes anciens petits colocs, **Camille**, **Antoine**, **Sophie**, **Julie**, **Clément**, **Thibaut**, **Elise** et cette folle année en votre compagnie

Dites moi que c'est la dernière étape **Emilie**, **Dave**, et maintenant on peut monter notre labo pour bosser sur le Hihi ! On est docteuuuuur

A ceux qui m'ont fait mettre les mains dans le cambouis de la génétique et de l'écologie : **Erwan**, **Nico**, **Bruno**, **Jeff**, **Samuel**

A minha **Moon** de Pira ! Obrigado pra a sua motivação e seus messagens !! Vou voltar com certeza

« Toutes les choses ont une fin, sauf la saucisse qui en a deux. »

[Néti]

Parce que la thèse, c'est comme la 0-3000, chaque année est une épreuve, et sans bon copain, ça se termine pas ! La dernière ligne droite a été franchie !!!!

Sommaire

Liste des acronymes viraux selon le comité international de taxonomie virale	i
Liste des abréviations	ii
Liste des figures	vi
Liste des tableaux.....	x
Introduction.....	1
Chapitre 1 : Mise au point et optimisation d'un protocole de séquençage haut débit pour la détection des virus à ADN simple brin circulaire.....	49
Article 1	52
<i>Nanopore sequencing is a credible alternative to recover complete genomes of geminiviruses</i>	
Chapitre 2 : Exploration de la diversité génomique et de la structure des communautés phytovirales à ADN circulaire à La Réunion : Association étroite des mastrevirus avec les milieux agricoles.....	67
Article 2	71
<i>A novel chimeric circular Rep-encoding single stranded DNA virus with interfamilial gene exchange illustrates the considerable evolutionary capacity of ssDNA viruses</i>	
Article 3	99
<i>Reconstructing the Mastrevirus communities structure in Reunion: The tale of agricultural associated pathogens</i>	

Chapitre 3 Diversité et structure des communautés virales des plantes herbacées indigènes et exotiques de différents écosystèmes réunionnais	156
Article 4	160
<i>Unraveling the enigmatic influences of phytoviruses in plant invasions on Reunion Island</i>	
Discussion générale	207
Références	222

Liste des acronymes viraux selon le comité international de taxonomie virale

Acronyme Virus		Espèce virale	
		Ancienne nomenclature	Nouvelle nomenclature binomiale
ACMV :	African cassava mosaic virus	<i>African cassava mosaic virus</i>	<i>Begomovirus manihotis</i>
ACSV :	Axonopus compressus streak virus	<i>Axonopus compressus streak virus</i>	<i>Mastrevirus axonopi</i>
BSMV :	barley stripe mosaic virus	<i>Barley stripe mosaic virus</i>	En cours d'actualisation
BYDV :	barley yellow dwarf virus	<i>Barley yellow dwarf virus</i>	<i>Luteovirus pavhordei</i>
CEAV :	Cenchrus echinatus associated virus	<i>Cenchrus echinatus associated virus</i>	<i>Mastrevirus cenchri</i>
CLRDV :	cotton leafroll dwarf virus	<i>Cotton leafroll dwarf virus</i>	En cours d'actualisation
CMV :	cucumber mosaic virus	<i>Cucumber mosaic virus</i>	En cours d'actualisation
CPAV :	Cenchrus purpureus associated virus	<i>Cenchrus purpureus associated virus</i>	<i>Cenpurvirus bourbonense</i>
CPMSV :	Cenchrus purpureus mild streak virus	<i>Cenchrus purpureus mild streak virus</i>	<i>Mastrevirus purpurei</i>
EACMV :	East African cassava mosaic virus	<i>East African cassava mosaic virus</i>	<i>Begomovirus manihotisafricaense</i>
EcmLV :	Euphorbia caput-medusae latent virus	<i>Euphorbia caput-medusae latent virus</i>	<i>Capulavirus euphorbiae</i>
EIAV :	Eleusine indica associated virus	<i>Eleusine indica associated virus</i>	<i>Mastrevirus eleusinae</i>
EMSV :	Eragrostis minor streak virus	<i>Eragrostis minor streak virus</i>	<i>Mastrevirus eragrostisminoris</i>
ESV :	Eragrostis streak virus	<i>Eragrostis streak virus</i>	<i>Mastrevirus eragrostis</i>
LCV :	lettuce chlorosis virus	<i>Lettuce chlorosis virus</i>	En cours d'actualisation
MeRAV :	Melinis repens associated virus	<i>Melinis repens associated virus</i>	<i>Mastrevirus melinis</i>
MSDV :	maize streak dwarfing virus	<i>Maize streak dwarfing virus</i>	<i>Mastrevirus fulleri</i>
MSRV :	maize streak Reunion virus	<i>Maize streak Reunion virus</i>	<i>Mastrevirus bourbonense</i>
MSV :	maize streak virus	<i>Maize streak virus</i>	<i>Mastrevirus storeyi</i>
NAMaMV :	North American maize streak-associated mastrevirus	<i>North American maize streak-associated mastrevirus</i>	<i>Mastrevirus americaense</i>
SAAV :	Sorghum arundinaceum associated virus	<i>Sorghum arundinaceum associated virus</i>	<i>Mastrevirus sorghi</i>
SacSV :	Saccharum streak virus	<i>Saccharum streak virus</i>	<i>Mastrevirus sachari</i>
SCSV :	sugarcane chlorotic streak virus	<i>Sugarcane chlorotic streak virus</i>	<i>Mastrevirus saccharumpallidi</i>
SCYLV :	sugarcane yellow leaf virus	<i>Sugarcane yellow leaf virus</i>	En cours d'actualisation
SSEV :	sugarcane streak Egypt virus	<i>Sugarcane streak Egypt virus</i>	<i>Mastrevirus saccharumegyptense</i>
SSRV :	sugarcane streak Reunion virus	<i>Sugarcane streak Reunion virus</i>	<i>Mastrevirus saccharumbourbonense</i>
SSV :	sugarcane streak virus	<i>Sugarcane streak virus</i>	<i>Mastrevirus saccharofficinari</i>
SWSV :	sugarcane white streak	<i>Sugarcane white streak</i>	<i>Mastrevirus saccharumalbusvenae</i>
PanSV :	Panicum streak virus	<i>Panicum streak virus</i>	<i>Mastrevirus panici</i>
PVY :	potyvirus Y	<i>Potyvirus Y</i>	En cours d'actualisation
TMV :	tobacco mosaic virus	<i>Tobacco mosaic virus</i>	En cours d'actualisation
TuMV :	turnip mosaic virus	<i>Turnip mosaic virus</i>	En cours d'actualisation
TrV1 :	Trifolium virus 1	<i>Trifolium virus 1</i>	En cours d'actualisation

Liste des abréviations

• Organismes et institutions

3P :	Pôle de protection des Plantes
CBNM :	Conservatoire Bontanique National des Mascarin
CIRAD :	Centre de coopération Internationale en Recherche Agronomique pour le Développement
FAO :	Organisation des Nations Unies pour l'Alimentation et l'Agriculture
GenBank :	Banque de données génomiques des États-Unis d'Amérique
ICTV :	International Committee on Taxonomy of Viruses
NCBI :	National Center of Biotechnology Information
UMR/PVBMT :	Unité Mixte de Recherche/Peuplements Végétaux et Bio-agresseurs en Milieu Tropical

• Autres abréviations

AAP :	Acquisition access period ou période d'accès pour l'acquisition
AE :	Agro-écosystèmes
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 à polarité positive
ADNsb(-) :	Acide désoxyribonucléique simple brin à polarité négative
ADNsb :	Acide désoxyribonucléique simple brin
AfSV :	African streak virus
<i>ApaLI</i> :	Enzyme de restriction produit par la souche bactérienne <i>Acetobacter pasteurianus</i>
ARN :	Acide ribonucléique
ARNi :	Acide ribonucléique interférence
ARNdb :	Acide ribonucléique double brin
ARNsb :	Acide ribonucléique simple brin
BP :	Site de Bassin Plat
bp :	Bases paire ou paire de bases

cccDNA :	Covalently closed circular DNA ou forme super-enroulée de l'ADN viral
CG:	Coastal grassland ou prairie littorale
CH:	Site de Cap la Houssaye
CL:	Site des Colimaçons
CP :	Protéine de capsid
Covid-19 :	Maladie à coronavirus 2019
CRESS-DNA :	Circular Rep-encoding ssDNA
DPI :	Day post inoculation ou jour après inoculation
EcoRI :	Enzyme de restriction produit par la souche bactérienne <i>Escherichia coli</i>
ELISA :	Enzyme Linked Immuno Sorbent Assay
HLV :	Homopolymer length variation ou variation de longueur d'homopolymères
HTS :	High throughput sequencing
IAP :	Inoculation access period ou période d'accès pour l'inoculation
INDEL :	Insertion / délétion
IR :	Intergenic Region ou Région Intergénique
<i>KpnI</i> :	Enzyme de restriction produit par la souche bactérienne <i>Klebsiella pneumoniae</i>
LB :	Milieu de culture Luria-Bertani
LIR :	Long intergenic region
LM:	Site de La Mare
LP:	Site de la Ligne d'Equerre
LP:	Site de Ligne Paradis
MID :	Multiplex identifier
ML :	Maximum-likelihood
MP :	Protéine de mouvement
MV :	Site de l'Espace Naturel Sensible des Hauts de Mont-Vert
<i>NcoI</i> :	Enzyme de restriction produit par la souche bactérienne <i>Gordonia rubripertincta</i>
NGS :	Next generation sequencing
NSP :	Protéine de navette nucléaire
nt :	Nucléotides
ONT :	Oxford Nanopore Technologies

ORF :	Open Reading Frames ou Cadre ouverte de lecture
Ori :	Origine de réplication
OTU :	Operational Taxonomic Unit
PacBio :	Pacific bioscience
PC:	Site de Piton de Caille
PCR :	Réaction de Polymérisation en Chaîne
PE :	Paired-End
PEE :	Plantes Exotiques Envahissantes
PS:	Site de la Pointe au Sel
PTGS :	post transcriptional gene silencing ou suppression génique post-transcriptionnelle
RCA :	Rolling Circle Amplification
RCR :	Rolling Circle Replication ou Réplication en Cercle Roulant
RDR :	Recombination-Dependent Replication ou Réplication dépendante de la Recombinaison
RdRp :	ARN polymérase ARN dépendante
Rep :	Protéine associée à la réplication
RFLP :	Restriction fragment length polymorphisme ou polymorphisme de longueur des fragments de restriction
RT-ADNsb :	Acide désoxyribonucléique simple brin rétro-transcrit
RT-ARN :	Acide ribonucléique rétro-transcrit
Sall :	Enzyme de restriction produit par la souche bactérienne <i>Streptomyces albus</i>
SARS-CoV- :	Coronavirus du syndrome respiratoire aigu sévère
SG:	Subalpine grassland ou prairie altimontaine
SGS :	Second Generation Sequencing
SIR :	Short intergenic region
SJR :	Protéine de capsid de type jelly roll simple
siRNA :	Small interfering Ribonucleic Acid
SOOI :	Sud-ouest de l'océan Indien
ssDNA :	Single-stranded circular DNA
TGS :	Third generation sequencing
TR:	Site de Terre Rouge
VANA :	Virion-associated nucleic acids

VEM : Vector enabled metagenomics

XbaI : Enzyme de restriction produit par la souche bactérienne *Xanthomonas badrii*

Liste des figures

- Introduction

Figure 1 Répartition de la gamme d'hôtes des virus des sept classes de Baltimore (BC). .3	3
Figure 2 Les classes de Baltimore (BC) et les royaumes monophylétiques des virus.4	4
Figure 3 Rangs utilisés dans la taxonomie des virus.5	5
Figure 4 Scénario évolutif de l'origine des phytovirus à ADNsb.7	7
Figure 5 Représentation de la recombinaison et du réassortiment. 11	11
Figure 6 Représentation du paysage de <i>fitness</i> virale par mutations et échange génétiques. 13	13
Figure 7 Représentation schématique d'une cellule végétale décrivant les principaux processus viraux. 16	16
Figure 8 Représentation graphique du continuum de <i>fitness</i> des plantes et des virus. 17	17
Figure 9 Représentation schématique des interactions interspécifiques à différentes échelles et des processus dans lesquels les populations sont intégrées.....24	24
Figure 10 Frise chronologique des principales avancées en écologie des phytovirus.26	26
Figure 11 Méthodes de caractérisation des viromes des communautés végétales.27	27
Figure 12 Schématisation des relations entre le risque de maladie, les activités humaines et la biodiversité.29	29
Figure 13 Représentation des quatre principaux modèles expliquant l'impact des phytovirus suite à l'introduction de plantes hôtes exotiques dans un milieu.....33	33

Figure 14 Arbre phylogénétique non enraciné des protéines associées à la réplication des <i>Cressdnaviricota</i>	34
Figure 15 Organisation schématique du génome des virus à ADNsb des huit types d'architectures génomiques décrites.	35
Figure 16 Structure des particules de géminivirus.	36
Figure 17 Illustration de l'organisation du génome des genres de la famille des <i>Geminiviridae</i>	39
Figure 18 Co-divergence entre les séquences des protéines de capsid des différents genres de <i>Geminiviridae</i> et leurs insectes vecteurs.....	40
Figure 19 Vue latérale d'un adulte et d'une larve de <i>Cicadulina mbila</i>	42
Figure 20 Symptômes de la maladie de la striure du maïs (maize streak disease) causées par le MSV.	44

- **Chapitre 1**

- **Article1**

Figure 1 Schematic representation of the three distinct sequencing methodologies used, the bio-informatics procedure and the method comparison.	56
Figure 2 Maximum likelihood phylogenetic tree of all the sequences of the TrV1-B and TrV1-C strains.....	59
Figure 3 Diagram presenting the positions of the mutations along the genome of the TrV1-B and TrV1-C isolates and assemblies.	60
Figure 4 Coverage plot along the reference genome of TrV1-B and TrV1-C.....	62

- **Chapitre 2**

- Article 2

Figure 1 *Cenchrus purpureus* mild streak virus (CPMSV) and *Cenchrus purpureus* associated virus (CPAV) symptoms.75

Figure 2 Schematic genome organizations of *Cenchrus purpureus* associated virus (CPAV) and *Cenchrus* mild streak virus (CPMSV).....80

Figure 3 Unrooted maximum likelihood tree of (A) replication associated protein of alphasatellites (B) capsid protein and movement protein of mastreviruses.82

Supplementary Figure 1 Annotated nucleotide sequence of *Cenchrus purpureus* associated virus98

- Article 3

Figure 1 Summary of the Poales sampling in Reunion. 110

Figure 2 Phylogenetic tree of all the 94 complete genomes of monocot-infecting mastreviruses from Reunion. 116

Figure 3 Bipartite interaction graph representing the association between Poales species and viral species and strains..... 120

Figure 4 Proportion of realized *Mastrevirus* species richness as a function of plant host richness. 123

Supplementary Figure 1 Principal coordinate analyses (PCoA) of the diversity of Poales sampled. 152

Supplementary Figure 2 (A) Unrooted maximum likelihood tree of various *Mastrevirus* species complete genome, (B) replication associated protein (Rep) capsid protein (CP) amino acid sequences. 153

Supplementary Figure 3 Scatter plots of pairs of variable relative to plant and viruses richness 155

• **Chapitre 3**

- Article 4

Figure 1 Summary of the herbaceous plant species sampling in Reunion..... 170

Figure 2 (A) Pie chart of the richness of viral families and (B) phytoviral families detected in term of OTUs 172

Figure 3 Heatmap of the number of reads of each viral genus detected in each plant sample. 175

Figure 4 (A) Principal component analysis (PCoA) of the diversity of viral families in each sampled pool and colored by more probable clusters. (B-H) Barplots represent for each cluster and count of each viral and plant categories 181

Figure 5 Correlation analysis results, with the fourth corner method, between the plant and virus characteristics. 182

Supplementary Figure 1 Phylogenetic tree of the *Genomoviridae* family 205

Supplementary Figure 2 Representation of the two principal axes of the RLQ analysis of voral family life history traits with (A) plant families and (B) plant sampling site, origin and life cycle..... 206

Liste des tableaux

- **Chapitre 1**

- **Article 1**

Table 1 Long read sequencing statistics.....	58
---	----

- **Chapitre 2**

- **Article 2**

Table 1 Infectivity of CPMSV, CPAV and MSV in single or mixed association after agroinoculation of <i>Cenchrus purpureus</i> and <i>Zea mays</i> , at 35 days post inoculation.	79
---	----

Table 2 Transmission rates of CPMSV, CPAV and MSV by <i>Cicadulina mbila</i> on four host species.....	79
---	----

Supplementary Table 1 List of the <i>Alphasatellitidae</i> species with the accession numbers of the nucleotide and amino acid replication associated protein (Rep) sequences.....	91
---	----

Supplementary Table 2 List of the Mastrevirus species with the accession numbers of the nucleotide complete, amino acid movement protein (MP) and capsid associated protein (CP) sequences.	94
---	----

Supplementary Table 3 Primers used for virus indexing.....	96
---	----

Supplementary Table 4 Long read sequencing statistics.....	97
---	----

- **Article 3**

Table 1 Sampling sites and summary of sampled pools for each campaign.	108
--	-----

Table 2 Summary of taxonomic classification of viral sequences per Poales species for each type of site and site.....	111
--	-----

Table 3 Summary of sampled Hemiptera families in each type of site and site in an altitudinal transect.	118
---	-----

Table 4 Analysis of assortativity of Poales-Mastrevirus interaction matrices.	122
Table 5 Analysis of nestedness and modularity of Poales-Mastrevirus interaction matrices with respectively NODF2 and Newman algorithm.....	122
Supplementary Table 1 Summary of sampled pool in each type site and site for each campaign.	134
Supplementary Table 2 Summary of all the plant species.	142
Supplementary Table 3 List of the no Geminiviridae species detected.....	145
Supplementary Table 4 Summary of the sequenced Mastrevirus species in each plant pool.	146
Supplementary Table 5 List of the Mastrevirus species with the accession numbers of the nucleotide complete, amino acid replication associated protein (Rep) and capsid associated protein (CP) sequences.	150
• Chapitre 3	
- Article 4	
Supplementary Table 1 Sampling sites and summary of sampled pools.	191
Supplementary Table 2 Summary of sampled pool in each type site and site.	192
Supplementary Table 3 HTS sequencing statistics.....	194
Supplementary Table 4 Summary of taxonomic classification and assignation results of viral sequences per plant species for each site.	195
Supplementary Table 5 List of the viral species with the accession numbers of the nucleotide complete sequences, and the outgroups used for the phylogenetic analysis.	201

Introduction

Sommaire de l'introduction

1. Contexte général	1
2. Les virus et les phytovirus	2
2.1. Les virus, des entités ubiquistes.....	2
2.2. Les virus de plantes ou phytovirus.....	5
2.2.1. Les mécanismes moléculaires à l'origine de la diversité.....	8
2.2.1.1. La mutation.....	8
2.2.1.2. La recombinaison.....	9
2.2.1.3. Le réassortiment.....	10
2.2.2. Les forces évolutives modulant la variabilité génétique.....	11
3. Les transmissions et gammes d'hôtes des phytovirus	14
3.1. Les différents modes de transmission des phytovirus.....	14
3.1.1. La transmission verticale.....	14
3.1.2. La transmission horizontale.....	15
3.2. Les interactions plante - virus.....	16
3.2.1. Le continuum symbiotique.....	16
3.2.2. Généralistes et spécialistes : les gammes d'hôtes.....	19
3.2.3. Les co-infections.....	21
3.2.4. La coévolution du triptyque phytovirus-hôte-vecteur.....	22
4. L'écologie des phytovirus	23
4.1. La métagénomique au service de l'écologie virale.....	24
4.2. L'influence de l'environnement biotique et abiotique.....	28
4.2.1. Le rôle des communautés végétales.....	28
4.2.2. L'impact des perturbations anthropiques.....	30
4.2.2.1. Une diversité phytovirale sauvage largement sous-estimée.....	30
4.2.2.2. L'interface agroécologique comme lieu d'émergence.....	30
4.2.3. Les impacts de la dynamique des PEE sur les communautés phytovirales.....	31
5. Les géminivirus	33
5.1. Les particularités morphologiques et génétiques des géminivirus.....	34
5.2. La réplication et la transcription des géminivirus.....	36
5.3. La diversité et la taxonomie.....	37
6. Les mastrévirus	41
6.1. Généralités.....	41
6.2. La diversité des African streak virus.....	42
6.2.1. Les AfSVs du maïs.....	42
6.2.2. Les AfSVs de la canne à sucre.....	44
6.2.3. Les AfSVs des poacées non cultivées.....	44
7. Contexte et objectifs de la thèse	46

Introduction

1. Contexte général

« *There are about 10 million-fold more viruses in the ocean than there are stars in the universe.* » (Curtis A. Suttle, 2013)

"Il y a environ 10 millions de fois plus de virus dans l'océan que d'étoiles dans l'univers."
(Curtis A. Suttle, 2013)

Alors que leur taille microscopique, l'apparente simplicité de leur génome et leur dépendance absolue à leur hôte pourraient suggérer un rôle minime dans le monde vivant, les virus figurent parmi les acteurs majeurs de la biosphère (French and Holmes, 2020). Ils sont présents partout sur Terre, dans tous les types d'écosystèmes à des abondances records (Breitbart *et al.*, 2002 ; Breitbart and Rohwer, 2005). Leur large diversité génétique, biologique et plasticité phénotypique leur confèrent la capacité d'infecter un immense éventail d'organismes appartenant à tous les règnes du vivant (Hull, 2014 ; Choua *et al.*, 2020).

Alors que l'intérêt initial porté aux virus a été principalement motivé par le désir de comprendre et de lutter contre les pathologies qu'ils causaient chez l'être humain, les espèces animales domestiquées ou les cultures, il est aussi ensuite apparu que ces virus avaient largement émergés depuis des hôtes sauvages. Ainsi, l'activité humaine avec le défrichage et la déforestation a favorisé la promiscuité entre le monde domestique et le monde sauvage et exacerbé les sauts d'hôte entre les hôtes sauvages et domestiques. La pandémie de Covid-19 qui a frappé l'humanité entre 2020 et 2022 est un exemple criant (Zhou *et al.*, 2020). Cette maladie humaine d'origine zoonotique est provoquée par un coronavirus (SARS-CoV-2) génétiquement apparenté aux coronavirus de chauve-souris (Zhou *et al.*, 2020). Les données actuelles suggèrent que la chauve-souris est le réservoir sauvage de ce virus et qu'il aurait été transmis à l'humain après une adaptation chez un hôte intermédiaire encore non identifié, probablement un animal sauvage.

Chez les plantes, ce type d'événements de saut d'hôte entre plantes sauvages et plantes cultivées est à l'origine de près de la moitié des émergences d'agents pathogènes des plantes (Anderson *et al.*, 2004). Ainsi, en 2020, l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) estimait qu'environ 38 % (4,7 milliards d'hectares) de la surface émergée était utilisée pour des activités agricoles. Par conséquent, plus d'un tiers des écosystèmes terrestres ont été transformés en agroécosystèmes, conduisant principalement à la réduction de la richesse spécifique (faible nombre d'espèces végétales cultivées), au remplacement des espèces sauvages et à l'introduction d'espèces exotiques (Anderson *et al.*, 2004), voire envahissantes (Tylianakis *et al.*, 2017 ; Roossinck and García-Arenal 2015). De plus, la gestion continue de ces systèmes, l'hétérogénéité spatiale et la fragmentation des habitats (Fargette *et al.*, 2006 ; Garrett *et al.*, 2006), impactent la transmission, les taux de rencontres entre plantes et virus (Parratt *et al.*, 2016) ainsi que le risque de maladie sur les cultures (Bass *et al.*, 2019 ; Stukenbrock and McDonald, 2008 ; Jones *et al.*, 2013 ; Roossinck and García-Arenal, 2015).

Dans ce contexte, il apparaît alors nécessaire d'obtenir une vision holistique de la diversité, de la structure des communautés (*i.e.* assemblage des espèces virales) et des dynamiques écologiques (*i.e.* interactions, évolution et adaptation) des populations virales afin de tenter de comprendre le fonctionnement de ces communautés virales et les facteurs environnementaux (*e.g.* communautés d'hôtes et de vecteurs, perturbations anthropiques) qui les influencent.

2. Les virus et les phytovirus

2.1. Les virus, des entités ubiquistes

Un virus est une entité biologique non cellulaire définie comme "un ensemble d'une ou plusieurs molécules d'acide nucléique, soit de l'ARN soit de l'ADN, le plus souvent enfermées dans une capsidie protéique, et dans certains cas une ou plusieurs enveloppes lipoprotéiques protectrices, capable d'organiser sa propre réplication uniquement à l'intérieur de cellules hôtes appropriées" (Hull, 2014). Les virus sont les entités biologiques les plus répandues et abondantes dans la nature (Wasik and Turner, 2013), avec une estimation atteignant les 10^{31} particules virales sur l'ensemble de la planète (Hendrix *et al.*,

2002 ; Mushegian, 2020). De plus, ils infectent des organismes de toutes les branches de l'arbre du vivant (eucaryotes, bactéries, archées ; Rodriguez-Valera *et al.*, 2009 ; Krupovic *et al.*, 2018 ; Koonin *et al.*, 2021 ; Crespo-Bellido and Duffy, 2023 ; **Figure 1**), faisant d'eux des organismes ubiquistes.

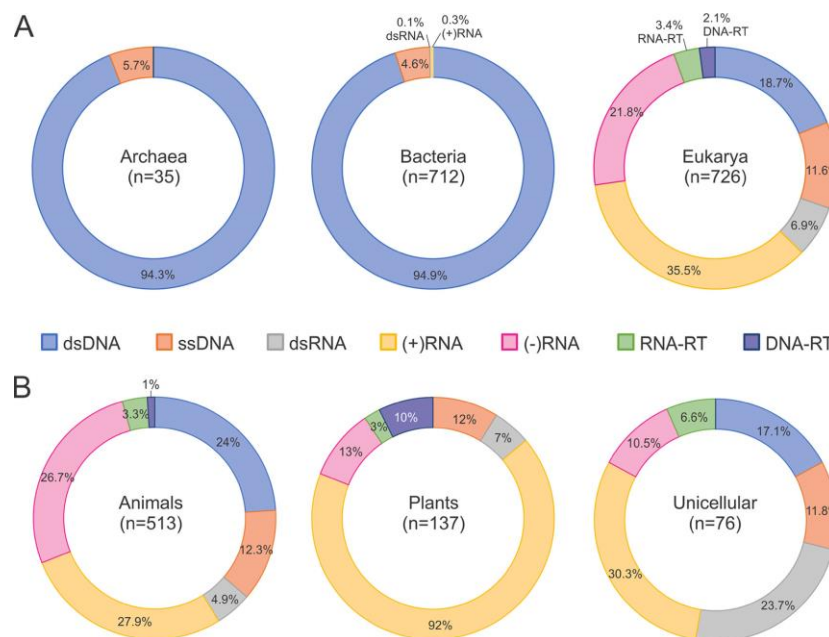


Figure 1 : Répartition de la gamme d'hôtes des virus des sept classes de Baltimore (BC). (A) Répartition des BC dans les trois domaines cellulaires. Le schéma illustre la prédominance des virus à ADNdb chez les bactéries et les archées, ce qui contraste avec la prédominance des virus à ARN chez les eucaryotes. (B) Répartition des BC dans les trois grands groupes eucaryotes. Chaque cercle représente la répartition des genres de virus associés au groupe d'hôtes indiqué. Le nombre de genres de virus (n) est indiqué à l'intérieur de chaque cercle. Les BC sont indiquées par le type d'acide nucléique du virion et sont codées par un code couleur (d'après Koonin *et al.*, 2021).

Dans la seconde moitié du XX^{ème} siècle, Baltimore a classé les virus en six catégories basées sur la nature du génome et les voies d'expression (Baltimore, 1971). Nommées classes de Baltimore (BC), elles couvrent la diversité des génomes viraux et les stratégies d'expression des gènes. Les six classes comprennent les virus à ADN double brin (ADNdb), à ADN simple brin (ADNsb), à ARN double brin (ARNsb), à ARN simple brin de polarité positive (ARNsb(+)), à ARN simple brin de polarité négative (ARNsb(-)) et les virus à ARN rétro-transcrit (RT-ARN) (Baltimore, 1971). Par la suite, une septième classe comprenant les virus à ADN simple brin rétro-transcrit (RT-ADNsb) qui se répliquent à travers un ARN intermédiaire, a été ajoutée (Galibert *et al.*, 1979 ; **Figure 2**). Des caractéristiques virales supplémentaires ont été utilisées pour la classification et la nomenclature des virus, avec la morphologie des virions, la présence ou non d'une enveloppe lipoprotéique, la gamme

d'hôtes, les propriétés sérologiques, la pathogénicité et les similarités de séquence du génome (Simmonds, 2015 ; Lefkowitz *et al.*, 2018). Le Comité International de la Taxonomie des Virus (ICTV) en charge des questions liées à la taxonomie virale proposait jusqu'alors un système de taxonomie hiérarchique à cinq rangs, comprenant l'ordre, la famille (sous-famille pour certains), le genre et l'espèce.

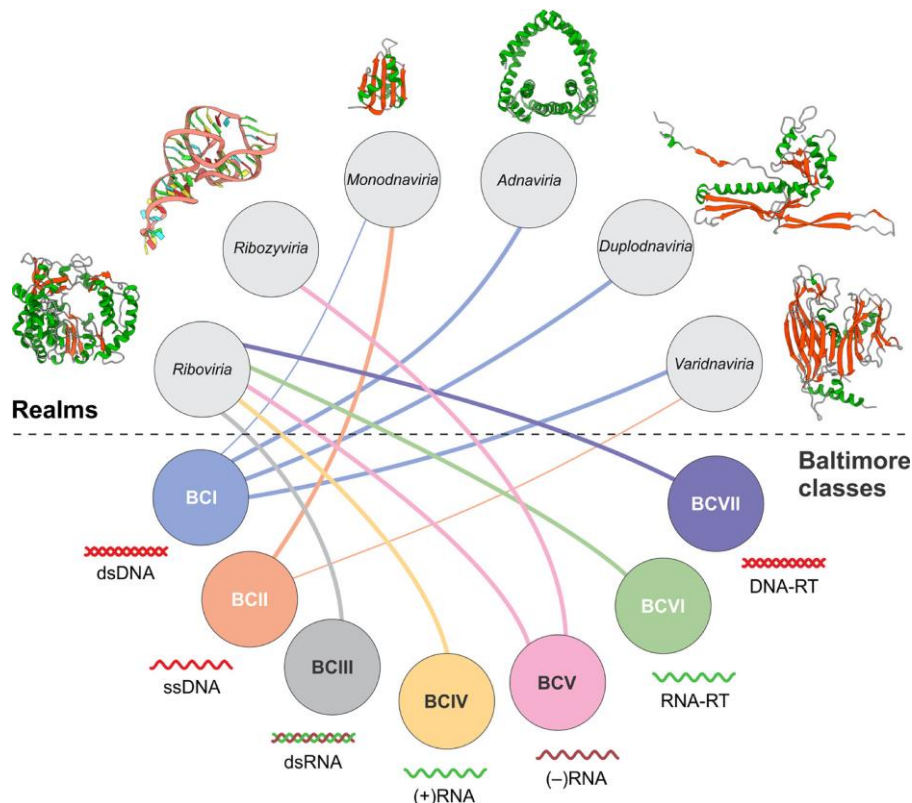


Figure 2 : Les classes de Baltimore (BC) et les royaumes monophylétiques des virus. Les connexions entre les BC et les règnes de virus sont indiquées par des liens colorés. L'épaisseur des liens est fonction de la fréquence des associations. Pour chaque règne viral, une structure protéique caractéristique est présentée (*Riboviria* : ARN polymérase dépendante de l'ARN ; *Ribozviria* : ribozyme ; *Monodnaviria* : endonucléase d'initiation de la réplication en cercle roulant ; *Adnaviria* : protéine de capsid ; *Duplodnaviria* : protéine de capsid ; *Varidnaviria* : protéine de capsid) et est colorée selon leur structure secondaire : les hélices α en vert ; les feuillets β en rouge (d'après Koonin *et al.*, 2021).

Récemment, l'adoption du système taxonomique *Systema naturae* de Linné (Linneaus, 1735) par la communauté des virologistes a été l'opportunité de mettre à jour la taxonomie virale au vue des nouvelles connaissances sur leur origine (Simmonds *et al.*, 2017 ; Gorbalenya *et al.*, 2020 ; Simmonds *et al.*, 2023 ; **Figure 3**) aboutissant à la première mégataxonomie complète des virus (Koonin *et al.*, 2020). Cette systématique a été obtenue à partir d'analyses phylogénétiques de gènes caractéristiques des virus, combinées à celles de réseaux de partage de gènes (Koonin *et al.*, 2020 ; Koonin *et al.*, 2021). Elle démontre au passage que les classes de Baltimore présentent dans la majorité des cas de la

polyphylie (*i.e.* regroupement de taxons non apparentés présentant des ressemblances ; **Figure 2**). En effet, certains règnes viraux monophylétiques présentent des taxons appartenant à différentes classes de Baltimore. Par exemple, des membres des *Ribovira* se retrouve dans cinq des sept classes de Baltimore (Koonin *et al.*, 2021). Un autre changement important approuvé par l'ICTV en 2021 a été l'adoption d'un format binomial uniforme (Genre espèce) pour la désignation des espèces de virus (Zerbini *et al.*, 2022).

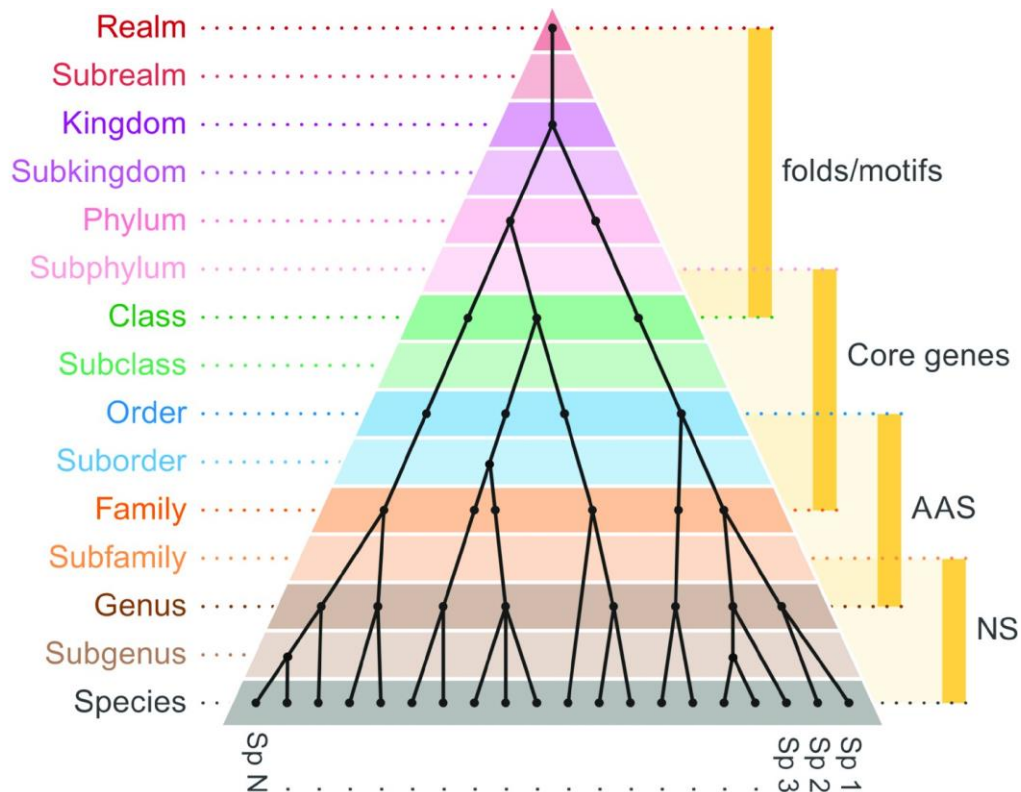


Figure 3 : Rangs utilisés dans la taxonomie des virus. Représentation schématique de la structure taxonomique en 15 rangs utilisée par l'ICTV. Cette structure inclut les méthodologies pouvant être utilisées pour déterminer les relations évolutives entre les virus et faire des assignations à chacun des niveaux. La forme pyramidale indique que le nombre de taxons augmente du rang supérieur (règne) au rang le plus bas (espèce, Sp.). Les noms des 15 rangs sont indiqués à gauche de la pyramide, et les méthodologies à droite (*folds/motifs*, structures/motifs ; *core genes*, gènes nucléaires ; AAS, *amino acid sequence similarity*, similarité des séquences d'acides aminés ; NS, *nucleotide sequence similarity*, similarité des séquences nucléotidiques). La pyramide inclut un exemple hypothétique de la taxonomie d'un règne, indiquant le nombre de taxons à chaque niveau (d'après Simmonds *et al.*, 2023).

2.2. Les virus de plantes ou phytovirus

Les virus capables d'infecter les végétaux sont nommés phytovirus (phyto- du grec φυτόν / *phutón*, « plantes ») et constituent un groupe polyphylétique très divers (Koonin *et al.*, 2015). Le plus ancien témoignage d'une maladie virale des plantes se trouve probablement dans un poème composé par l'impératrice Koken du Japon, il y a plus de 1200 ans, en 752 après J.-C. (Inouye and Osaki, 1980 ; Saunders *et al.*, 2003) :

« Dans ce village,
Il semble que le givre persiste continuellement
Pour la plante que j'ai vue
Dans le champ d'été
La couleur des feuilles jaunissait. »

On a suggéré que ce poème décrivait les symptômes de jaunissement des feuilles chez une eupatoire (*Eupatorium lindleyanum*) causés par un virus de la famille des *Geminiviridae*, l'Eupatorium yellow vein virus. C'est un millénaire plus tard que Charles de L'Écluse a observé des tulipes présentant des motifs panachés, certainement causés par une espèce de *Potyvirus* (*tulip breaking virus*, Lesnaw and Ghabrial, 2000). C'est plus de deux siècles plus tard, à la fin du XIX^{ème}, que la virologie végétale et la virologie en général ont émergé en tant que discipline, avec les premiers travaux sur le tobacco mosaic virus (TMV, famille *Tobamoviridae*) et l'observation que les ultra-filtrats d'extraits de plantes infectées restaient infectieux (Fraenkel-Conrat, 1986 ; Ivanowski, 1892).

Comme pour les virus en général, les phytovirus présentent une large diversité de supports génétiques : ces derniers sont répartis dans six des sept classes de Baltimore (Koonin *et al.*, 2021 ; **Figure 1**) et dans deux règnes viraux : *Riboviria* et *Monodnaviria* (Xavier and Whitfield, 2023). Ainsi, en plus de leurs origines polyphylétiques, le fait que de nombreux phytovirus infectent aussi leurs vecteurs, oblige à ne pas considérer le terme phytovirus comme décrivant un ensemble monolithique (Lefeuvre *et al.*, 2019). Parmi les meilleurs supports à ce constat, reste l'observation de nombreux transferts de gènes au cours de l'évolution entre familles virales (par exemple, les membres des espèces de *Genomoviridae* et *Geminiviridae* ; Shi *et al.*, 2016 ; Zhang *et al.*, 2019 ; Dolja *et al.*, 2020 ; **Figure 4**), parmi lesquels figurent des virus qui infectent exclusivement des hôtes appartenant à différents règnes eucaryotes (Dolja and Koonin, 2011 ; Dolja and Koonin, 2018 ; Krupovic and Koonin, 2017).

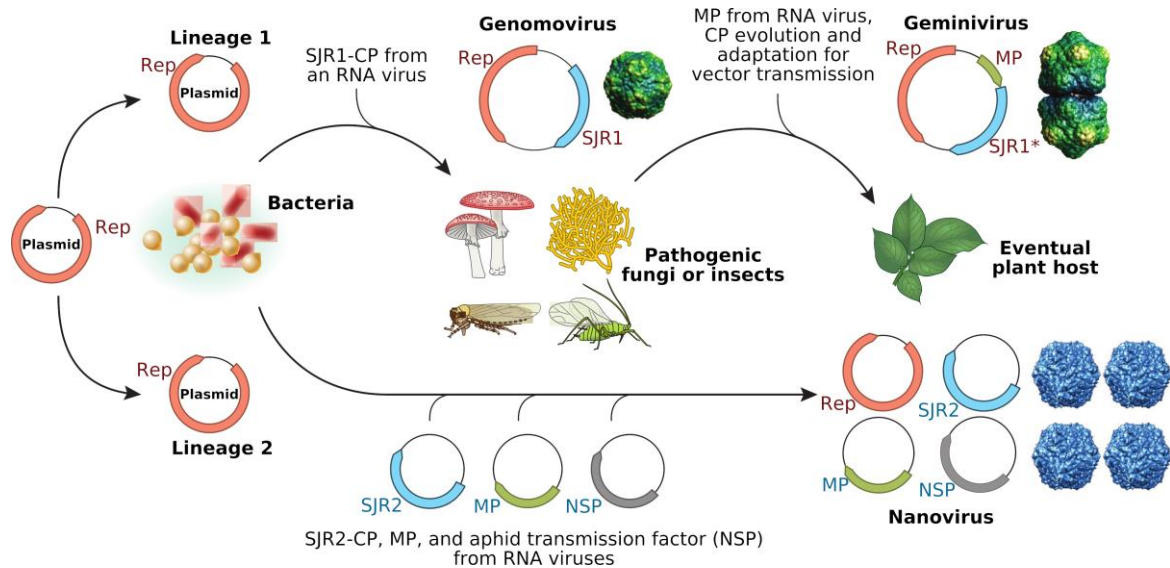


Figure 4 : Scénario évolutif de l'origine des phytovirus à ADNsb. Les géminivirus (en haut) et les nanovirus (en bas) ont évolué à partir de deux lignées de plasmides bactériens apparentés par l'acquisition des gènes codant pour les protéines de capsides (SJR1-CP et SJR2-CP) à partir de virus à ARN lors de deux événements indépendants. L'évolution des géminivirus s'est probablement déroulée à partir d'un ancêtre de type genomovirus infectant des champignons phytopathogènes ou des insectes. Abréviations : CP, protéine de capside ; MP, protéine de mouvement ; NSP, protéine de navette nucléaire ; Rep, protéine d'initiation de la réplication en cercle roulant ; SJR, protéine de capside de type *jelly roll* simple (d'après Dolja *et al.*, 2020).

Malgré une diversité du nombre de gènes, le génome des phytovirus comprend dans la majorité des cas des gènes codant pour des protéines associées à la réplication, des protéines de capside (CP), des protéines de mouvement (MP ; qui, chez les *Riboviria* et *Monodnaviria*, seraient issues de la duplication ou de l'acquisition horizontale de CP ; Butkovic *et al.*, 2023) permettant au virus de se déplacer à l'intérieur de la plante (de cellule à cellule ou à longue distance entre les étages foliaires et les organes), ainsi que des protéines permettant de contourner les systèmes de défense des plantes (Hull, 2014). Leur génome peut être linéaire ou circulaire et composé d'une ou de plusieurs molécules d'acides nucléiques sous forme de composants. Ces virus sont qualifiés de multipartites si chaque composant est encapsidé dans un virion séparé, ou segmentés si les composants sont tous encapsidés ensemble dans un seul virion. La taille des composants du génome des phytovirus peut varier de 1 kb pour les *Nanoviridae* à 20 kb pour les *Closteroviridae* (Hull, 2014 ; Varsani *et al.*, 2018). De plus, des molécules d'ADN satellites ont été décrits en association avec certains phytovirus monopartites. Ces molécules satellites peuvent avoir un effet sur le pouvoir pathogène du virus, en stimulant sa virulence et sa capacité de multiplication et de transmission (Lozano *et al.*, 2016 ; Gnanasekaran and Chakraborty, 2018 ; Varsani *et al.*, 2021 ; Guyot *et al.*, 2022).

2.2.1. Les mécanismes moléculaires à l'origine de la diversité

2.2.1.1. La mutation

Les mutations sont à l'origine des variations qui permettent l'évolution des organismes vivants. Elles se présentent sous trois formes : l'insertion et la délétion, qui correspondent respectivement à l'ajout ou à la perte d'un ou plusieurs nucléotides, et la substitution, qui implique le remplacement d'un nucléotide par un autre. Les mutations peuvent être bénéfiques (avantage sélectif), neutres (sans effet) ou délétères (désavantage sélectif). En raison de la compacité des génomes viraux et de l'importance de l'épistasie (c'est-à-dire l'effet d'une mutation sur plusieurs gènes au sein du même génome ; Holmes, 2003), il est largement considéré que la majorité des mutations sont délétères, voire létales (Domingo-Calap *et al.*, 2009 ; Sanjuán *et al.*, 2004). Cependant, certaines mutations dans le génome viral peuvent être désavantageuses pour l'infection d'un hôte et avantageuses pour un autre hôte (Duffy *et al.*, 2006). Ce phénomène est appelé pléiotropie antagoniste (Whitlock, 1996). Le taux de substitution d'un virus représente le nombre de mutations par position nucléotidique et par unité de temps, et reflète la vitesse à laquelle un virus évolue. Ce taux de substitution est étroitement lié au taux de mutation, qui représente la probabilité qu'une modification de l'information génétique soit transmise à la génération suivante (Sanjuán and Domingo-Calap, 2016). Les taux de mutation élevés des virus semblent leur conférer la capacité de s'adapter rapidement et efficacement aux changements des paysages adaptatifs et aux changements environnementaux (Duffy *et al.*, 2008 ; Crespo-Bellido and Duffy, 2023). Si le taux de mutation varie considérablement parmi les virus et est principalement lié au taux d'erreur de la polymérase, il peut également être influencé par les mécanismes de réplication, la nature, la taille et la structure du génome, l'environnement cellulaire et l'accès à la réparation post-réplivative (Sanjuan *et al.*, 2010 ; Sanjuán and Domingo-Calap, 2016).

Pendant longtemps, les virus à ADN ont été globalement considérés comme des virus à évolution lente en comparaison des virus à ARN. Les taux de mutation rapportés pour ces deux groupes de virus varient de 10^{-4} à 10^{-6} erreurs par nucléotide et par réplication pour les virus à ARN, contre 10^{-6} à 10^{-7} pour les virus à ADNsb et 10^{-7} à 10^{-8} pour les virus à ADNdb (Garcia-Diaz and Bebenek, 2007 ; Jenkins *et al.*, 2002 ; Sanjuán and Domingo-Calap, 2016). Cependant, certaines espèces virales ADNsb ont été décrites comme

évoluant aussi rapidement que les virus à ARN (Duffy *et al.*, 2008). C'est le cas, par exemple, des espèces des familles *Geminiviridae* et *Parvoviridae*, qui présentent des taux de substitution compris entre 1×10^{-4} et 3×10^{-4} substitutions par site et par an (Duffy and Holmes, 2008 ; Harkins *et al.*, 2009). Il est important de préciser que ces taux d'évolutions sont à considérer au regard du pas de temps sur lequel ils ont été mesurés (Aiewsakun and Katzourakis, 2016 ; Simmonds *et al.*, 2019). Les analyses incluant des séquences virales anciennes (telles que celles provenant d'herbiers) ont en effet révélé des taux d'évolution 10 à 20 fois moins importants que ceux estimés uniquement à partir de séquences contemporaines (Smith *et al.*, 2014 ; Rieux *et al.*, 2021). Ces différences d'estimation entre les mesures à court terme et à long terme sont probablement dues à la purge à moyen ou long terme de mutations marginalement délétères ou bénéfiques à court terme (Lythgoe *et al.*, 2017).

2.2.1.2. La recombinaison

La recombinaison est un processus évolutif majeur qui conduit à la formation de molécules nucléotidiques chimères par échange de matériel génétique provenant de génomes parentaux (Pérez-Losada *et al.*, 2015 ; **Figure 5**). On distingue la recombinaison homologue durant laquelle une portion d'un génome est remplacée par une portion homologue d'un autre génome et la recombinaison non-homologue qui implique le réarrangement génomique par duplication, insertion ou délétion (Lai, 1992). Dans le cas des virus de la famille des *Geminiviridae*, des conflits entre les complexes enzymatiques de transcription et de réplication ou des cassures monocaténaïres dans le brin matrice pourraient être impliqués (Owor *et al.*, 2007). La recombinaison peut alors se produire par le mécanisme connu sous le nom de "réplication dépendante de la recombinaison" (RDR), qui est l'un des deux mécanismes de réplication chez les espèces de la famille *Geminiviridae* ou encore le bactériophage T4 (Kreuzer *et al.*, 1995 ; Jeske *et al.*, 2001 ; Roux *et al.*, 2013).

Contrairement à l'évolution liée à l'accumulation progressive de mutations au cours des cycles de réplication, l'échange génétique par recombinaison donne accès à un polymorphisme plus important que par mutation seule (Cramer *et al.*, 1998). Les taux d'échange génétique sont déterminés par différents facteurs tels que les taux de co-infection, le mode de réplication, la fidélité de la polymérase, les mécanismes d'exclusion

de la surinfection et l'encapsidation (Crespo-Bellido and Duffy, 2023). Les données actuelles suggèrent que la recombinaison rapportée au nucléotide est moins fréquente ou équivalente à la fréquence de mutation (Froissart *et al.*, 2005 ; Tromas *et al.*, 2014 ; Crespo-Bellido and Duffy, 2023).

La recombinaison a été largement décrite entre des virus de la même espèce (Zhou *et al.*, 1997 ; Bousalem *et al.*, 2000 ; Kumar *et al.*, 2022 ; Claverie *et al.*, 2023), d'espèces différentes (Monci *et al.*, 2002 ; Chare and Holmes, 2006 ; Vigne *et al.*, 2008), de genres différents (Hernández-Zepeda *et al.*, 2013 ; Moonan *et al.*, 2000) et même de familles différentes (Saunders and Stanley, 1999 ; Miras *et al.*, 2014). Elle peut alors aboutir à de nouvelles souches (comme la souche sévère de l'East African cassava mosaic virus, genre *Begomovirus* ; Zhou *et al.*, 1997), de nouvelles espèces (comme l'African cassava mosaic Burkina Faso virus, genre *Begomovirus* ; Tiendrébéogo *et al.*, 2012), et même de nouveaux genres (comme les becurtovirus issus de la recombinaison entre bégomovirus et curtovirus ; Hernández-Zepeda *et al.*, 2013).

2.2.1.3. Le réassortiment

Les virus segmentés et les virus multipartites ont accès à un mode d'évolution unique appelé réassortiment (également qualifié de "pseudo-recombinaison"). Le réassortiment implique l'échange de composants génomiques entre souches ou espèces, soit par co-encapsidation (pour les virus segmentés), soit par co-transmission (pour les virus multipartites), créant ainsi de nouvelles combinaisons de composants génomiques dérivés d'au moins deux virus parentaux (Varsani *et al.*, 2018 ; **Figure 5**). Le réassortiment des composants génomiques est courant chez divers phytovirus à ARN (Gu *et al.*, 2007 ; Qiu and Moyer, 1999) et à ADN (Stainton *et al.*, 2012 ; Song *et al.*, 2022). Ce processus a été associé à l'acquisition de nouveaux traits biologiques comme la capacité à contourner la résistance (Qiu and Moyer, 1999), l'adaptation à l'hôte (Hou *et al.*, 1998) ou encore associé à une virulence accrue (Gu *et al.*, 2007 ; Pita *et al.*, 2001). Contrairement à la recombinaison et au réassortiment entre virus segmentés, le réassortiment entre virus multipartites peut se produire par mélange des composants lors de la transmission et ne nécessite pas la co-infection de la même cellule des virus parentaux (Varsani *et al.*, 2018 ; Crespo-Bellido and Duffy, 2023 ; Marshall *et al.*, 2013).

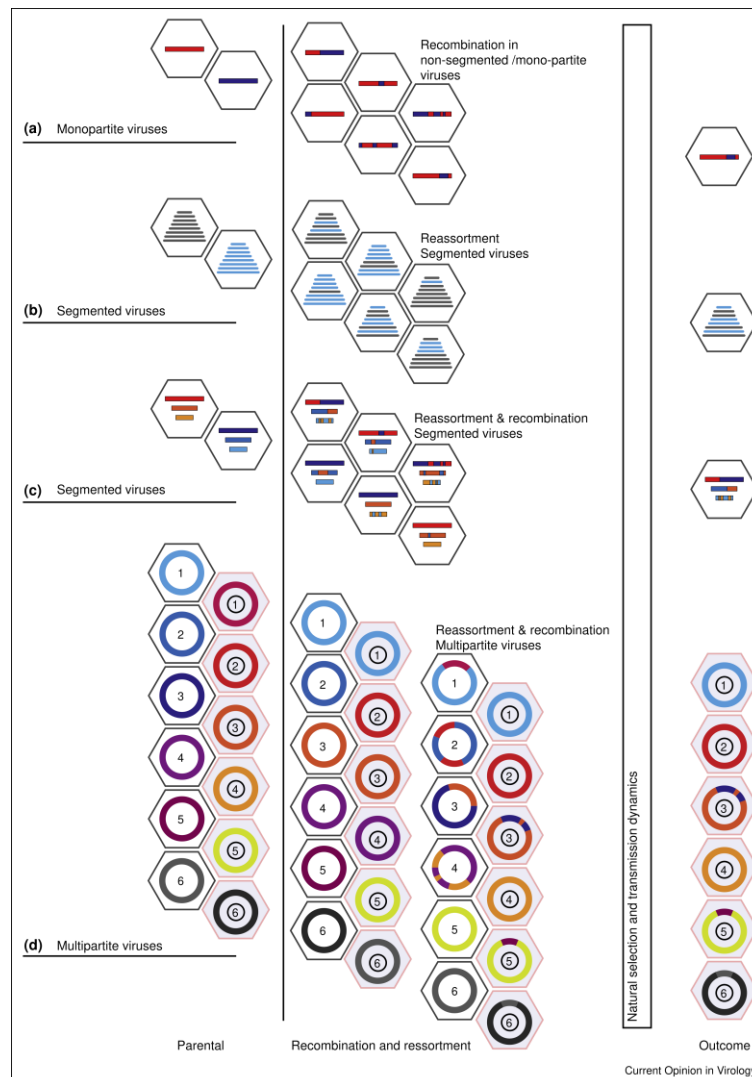


Figure 5 : Représentation (a) de la recombinaison entre virus monopartites, (b) du réassortiment entre virus segmentés, et de la recombinaison et du réassortiment entre (c) virus segmentés et (d) virus multipartites (d'après Varsani *et al.*, 2018).

2.2.2. Les forces évolutives modulant la variabilité génétique

Les variants viraux sont soumis à trois forces évolutives majeures qui déterminent leur fréquence de distribution au sein de la population virale : la dérive génétique, la sélection naturelle et la migration. (**Figure 6**).

La dérive génétique est un processus qui entraîne des fluctuations aléatoires dans les fréquences des variants au sein d'une population d'une génération à l'autre (Charlesworth, 2009). Ce processus évolutif agit de manière indépendante sur tous les variants de la population et ne dépend pas de la *fitness* des individus. L'intensité de la dérive génétique

INTRODUCTION

dépend de la présence de goulots d'étranglement, c'est-à-dire une réduction de la taille de la population. Dans des cas extrêmes, une dérive génétique très intense peut entraîner la perte de certains variants à la génération suivante, voire même l'extinction d'une population (Zwart and Elena, 2015). Les phytovirus, par exemple, subissent la dérive génétique à différentes étapes de l'infection virale, telles que : (i) la colonisation des organes de la plante (French and Stenger, 2003 ; Li and Roossinck, 2004 ; Sacristán *et al.*, 2003), (ii) la transmission par des vecteurs (Ali *et al.*, 2006 ; Betancourt *et al.*, 2008) ou par contact entre des feuilles infectées et non infectées (Sacristán *et al.*, 2011), et (iii) lors des interactions entre virus co-infectant la même plante (Fraile *et al.*, 1997). Les estimations actuelles portant sur les virus transmis par le mode circulant non propagatif (comme les souches Israël et Mild du TYLCV transmis par *Bemisia tabaci* ; Péréfarres *et al.*, 2014) et par le mode non circulant (comme le potyvirus Y (PVY) et le cucumber mosaic virus (CMV) transmis par les pucerons ; Betancourt *et al.*, 2008 ; Moury *et al.*, 2004) suggèrent que seules une à deux particules virales contribuent à la génération suivante. L'étude récente de Ren *et al.* (2023) suggère que l'initiation de la réplication virale du TYLCV dans une cellule de son hôte n'excède pas trois génomes. Ces études mettent en évidence l'importance du goulot d'étranglement et de la dérive génétique à chaque événement de transmission horizontale par les insectes vecteurs et de contamination d'une nouvelle cellule végétale.

Contrairement à la dérive génétique, la sélection naturelle n'agit pas de manière indifférenciée sur tous les variants d'une population, mais est un processus par lequel les fréquences des variants les plus adaptés à un environnement donné vont augmenter dans la population (sélection positive), tandis que celles des variants moins bien adaptés vont diminuer (sélection négative). Toutes les modifications génétiques ne sont pas systématiquement transmises à la génération suivante. Seules celles permettant de maintenir ou d'augmenter la *fitness* (i.e. valeur adaptative ou sélective), c'est-à-dire la capacité de reproduction d'un individu et de contribution à la génération suivante dans un environnement donné, perdurent dans la population virale (Garcia-Arenal *et al.*, 2003). Cette capacité d'adaptation du virus dans un paysage adaptatif peut être représentée par une carte topographique avec des pics et des vallées où les valeurs sélectives associées à des coordonnées qui représentent les génotypes d'individus sont représentées (Crespo-Bellido and Duffy, 2023 ; **Figure 6**). Le concept de paysage adaptatif permet de se faire une

représentation imagée des forces évolutives à la base de l'adaptation. Ainsi, les individus peuvent par mutation atteindre le sommet d'un pic en augmentant leur *fitness* de manière progressive (**Figure 6**). Seulement, dans la théorie des « quasi-espèces » (réplicats apparentés de génomes viraux ; Eigen, 1971) si l'accumulation de mutations dépasse un seuil d'erreurs critiques (ou « seuil de catastrophe génétique ») la diversité génétique excessive peut conduire à l'accumulation de mutations délétères qui compromettent la *fitness* globale de la population virale (Domingo and Holland, 1997 ; Manrubia *et al.*, 2010). En outre, l'échange génétique par recombinaison ou réassortiment offre l'opportunité de réaliser de grands sauts dans les paysages adaptatifs et de passer d'un pic de *fitness* à un autre (Crespo-Bellido and Duffy, 2023 ; **Figure 6**).

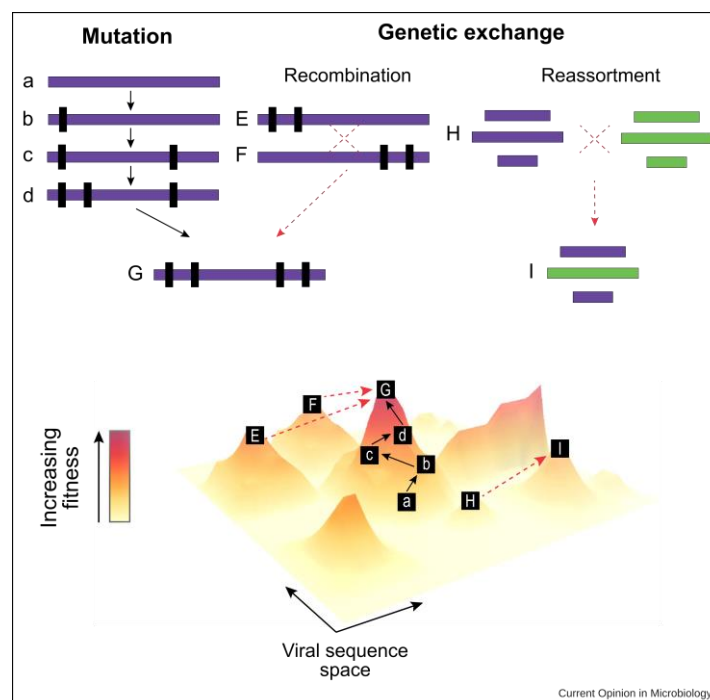


Figure 6 : Exploration du paysage de *fitness* viral par mutation et échange génétique. (En haut) Représentation schématique de génomes viraux (segments violets) évoluant par mutation (en haut à gauche) et par échange génétique (en haut à droite). Les barres noires sur les schémas des génomes représentent des mutations sous forme de substitutions, d'insertions et de délétions. Pour le réassortiment, les segments violets et verts représentent des génomes viraux segmentés distincts. Les lettres représentent différents génotypes qui sont superposés sur le paysage de *fitness* (en bas). L'évolution adaptative par mutation est indiquée par les flèches noires, tandis que les flèches rouges indiquent les grands sauts évolutifs en termes de *fitness* facilités par l'échange génétique. Le paysage de *fitness* est arbitraire et est utilisé uniquement à des fins d'illustration (d'après Crespo-Bellido and Duffy, 2023).

Enfin, la migration est un processus évolutif qui permet le flux de gènes au sein d'une métapopulation. Une métapopulation est un concept écologique développé par Levins en 1969 qui décrit un ensemble de populations d'individus d'une même espèce séparées dans l'espace ou dans le temps, mais interconnectées par des migrations d'individus. La migration virale est le processus par lequel les populations virales se déplacent d'un

écosystème à un autre, permettant ainsi leurs propagations et la proximité avec d'autres populations virales. Cela se solde par de nouvelles interactions avec pouvant aboutir à des phénomènes de recombinaisons ou de réassortiment par exemple.

3. Les transmissions et gammes d'hôtes des phytovirus

Du fait de l'immobilité de leurs hôtes, les phytovirus ont dû développer diverses stratégies au cours de leur évolution pour maximiser leur transmission (Astier *et al.*, 2001).

3.1. Les différents modes de transmission des phytovirus

3.1.1. La transmission verticale

Lorsque les virus sont transmis d'une génération à l'autre par l'intermédiaire du matériel reproducteur végétal, on parle de transmission verticale (**Figure 7**). Cela signifie que le virus a infecté les cellules germinales de son hôte (le pollen ou l'ovule et donc la graine après la fécondation) et pourra se multiplier dans les cellules des individus de la génération suivante (cas des membres des familles *Chrysoviridae*, *Endornavirus*, *Nepovirus*, *Partitiviridae* et des *Totiviridae*). Certains phytovirus peuvent aussi être transmis verticalement par multiplication végétative de leur hôte, infectant les boutures, tubercules ou bulbes (Mink, 1993).

Dans le cas des transmissions verticales, la grande majorité des phytovirus décrits sont persistants (*i.e.* qui subsistent dans leur plante hôte à long terme) et asymptomatiques et leurs effets semblent neutres, voire positifs pour leurs hôtes (Roossinck, 2010). Il est intéressant de noter que les phytovirus à transmission verticale stricte ne présentent d'ailleurs pas de protéine de mouvement leur permettant de se déplacer de cellule en cellule chez la plante (Roossinck, 2010). Pour ces virus, seule une relation mutualiste avec l'hôte est d'ailleurs théorisée (Hamelin *et al.*, 2017). La plupart des virus persistants décrits ont des génomes à ARNdb ne codant que pour une ARN polymérase ARN dépendante (RdRp) et une protéine de capsid, comme les virus de la famille *Partitiviridae* (Roossinck, 2010). À titre d'exception, les virus persistants de la famille *Endornaviridae*, précédemment considérés comme des virus à ARNdb, ont récemment été reclassés en tant que virus à ARNsb sur la base de leur phylogénie (Siddell *et al.*, 2019). Les virus persistants se sont

révélés être abondants chez les plantes, mais ils n'ont été jusqu'à présent que peu étudiés (Roossinck, 2010 ; Roossinck, 2012a ; Roossinck, 2015).

3.1.2. La transmission horizontale

La transmission horizontale est l'ensemble des processus par lesquels les phytovirus passent d'une plante à une autre soit par contact direct, soit en faisant intervenir un élément extérieur à la plante (**Figure 7**). Cet élément peut être abiotique telle que les eaux de ruissellement ou la litière du sol (Li *et al.*, 2016 ; Mehle *et al.*, 2018), ou biotique telles que des organismes vivants telluriques (champignons, protistes et nématodes) ou aériens (insectes piqueurs suceurs tels que les pucerons, aleurodes, cicadelles, cochenilles, thrips ; Jones, 2018). Ces espèces sont qualifiés de vecteurs et capables de transmettre les phytovirus en se dispersant sur des échelles spatiales très variables allant du centimètre à plusieurs kilomètres (Bragard *et al.*, 2013). Plus de 70 % des transmissions horizontales de phytovirus décrites sont réalisées par des insectes de l'ordre Hemiptera alors que seul 10 % sont réalisées par des vecteurs telluriques (Hogenhout *et al.*, 2008 ; Whitfield and Rotenberg, 2015). Cependant, les vecteurs naturels d'une majorité de phytovirus restent inconnus.

Les modes de transmission par des insectes vecteurs sont caractérisés par quatre paramètres : (i) la phase d'acquisition : temps d'alimentation nécessaire à l'insecte pour acquérir le virus ; (ii) la période de latence : période entre l'acquisition du virus et le moment à partir duquel le vecteur est apte à le transmettre ; (iii) la phase d'inoculation : temps d'alimentation nécessaire à la transmission du virus à une nouvelle plante ; et (iv) la période de rétention : temps pendant lequel l'insecte peut transmettre le virus (Gray and Banerjee, 1999 ; Bragard *et al.*, 2013 ; Hull, 2014).

Sur la base de ces paramètres, différents modes de transmission ont été caractérisés pour les insectes piqueurs-suceurs (Bragard *et al.*, 2013). D'une part, la transmission non circulante se caractérise par la localisation des virus au niveau des pièces buccales et plus précisément au niveau des stylets de l'insecte pour les virus non persistant, avec des phases d'acquisition et d'inoculation courtes allant de quelques secondes à quelques minutes (transmission non persistante) ou de quelques heures à quelques jours

(transmission semi-persistante) (Uzest *et al.*, 2007). D'autre part, la transmission circulante persistante concerne les virus ingérés par l'insecte qui ont réussi à traverser le tractus intestinal pour migrer dans l'hémolymphe puis les glandes salivaires. Cette deuxième modalité de transmission peut être propagative ou non propagative, selon que la réplication du virus est respectivement possible ou non au sein de l'organisme du vecteur (Ng and Falk, 2006 ; Bragard *et al.*, 2013).

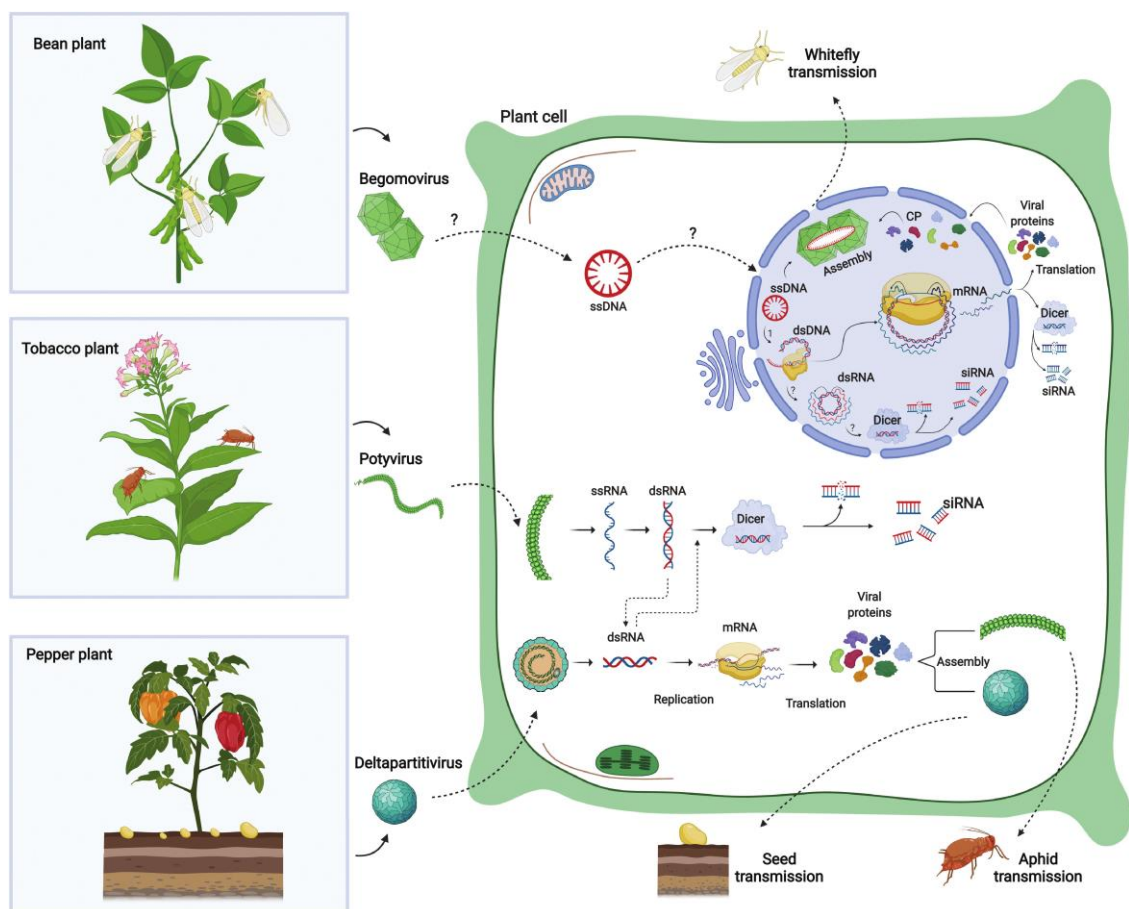


Figure 7 : Représentation schématique d'une cellule végétale, décrivant les principaux processus viraux (la réplication, le *silencing* par ARNi et les transmissions verticales et horizontales) et les compartiments cellulaires où ces processus se déroulent (d'après Moubset *et al.*, 2022).

3.2. Les interactions plante - virus

3.2.1. Le continuum symbiotique

Les interactions symbiotiques réalisées entre les virus et leurs plantes hôtes varient de l'antagonisme (parasitisme et pathogénicité, négatif pour l'hôte) au commensalisme (neutre pour l'hôte) jusqu'au mutualisme (bénéfique pour l'hôte ; Suttle, 2007 ; Rohwer and Thurber, 2009 ; Roossinck, 2011 ; Fraile and García-Arenal, 2016 ; Hamelin *et al.*, 2017 ; **Figure 8**).

Dans le cas des interactions antagonistes entre les phytovirus et leurs hôtes, l'effet négatif de l'infection virale sur la *fitness* de la plante hôte peut être mesurée qualitativement par la pathogénicité, c'est-à-dire la capacité du pathogène à infecter et à causer des maladies sur un hôte spécifique, dans différentes conditions environnementales ; et quantitativement par la virulence, c'est-à-dire le degré de dommages causés à l'hôte (Fraile and García-Arenal, 2010 ; Woolhouse *et al.*, 2002). Au niveau cellulaire, les phytovirus détournent la machinerie métabolique de l'hôte et se multiplie jusqu'à l'altération de la structure et de l'intégrité des membranes intracellulaires et des organites, et dans certains cas extrêmes jusqu'à la destruction de la cellule végétale (Laliberté and Sanfaçon, 2010). Ainsi, les symptômes chlorotiques viraux ont pu être associés à des changements dans le nombre ou la taille des chloroplastes, à des altérations structurales (Li *et al.*, 2016 ; Zhao *et al.*, 2016 ; Bhattacharyya and Chakraborty, 2018), voire à une diminution globale de l'activité des gènes photosynthétiques (Bilgin *et al.*, 2010). A une échelle macroscopique, les dommages causés peuvent se manifester par l'expression de différents symptômes (chlorose, nécrose, mosaïque, nanisme, déformations foliaires) et des pertes de *fitness* de l'hôte (diminution de la survie, de la croissance, de la reproduction, du rendement ; Fraser, 1992).

A l'échelle environnementale, dans les communautés de plantes sauvages, les infections virales sont fréquentes mais souvent asymptomatiques (Prendeville *et al.*, 2012 ; Roossinck, 2012a ; Bernardo *et al.*, 2018 ; Maclot *et al.*, 2023). Cependant, l'absence de symptômes visibles ne signifie pas une absence d'effet négatif sur la *fitness* des plantes (Alexander *et al.*, 2020 ; Malmstrom and Alexander, 2016 ; Prendeville *et al.*, 2012 ; Remold, 2002).

En retour, les plantes ont développé de multiples stratégies de défense pour stopper, limiter ou tolérer l'infection virale : la résistance qui consiste à limiter la multiplication et/ou la propagation du virus (Råberg *et al.*, 2009) ; et la tolérance qui correspond à une réduction des dommages causés par l'infection virale (Fraile and García-Arenal, 2010).

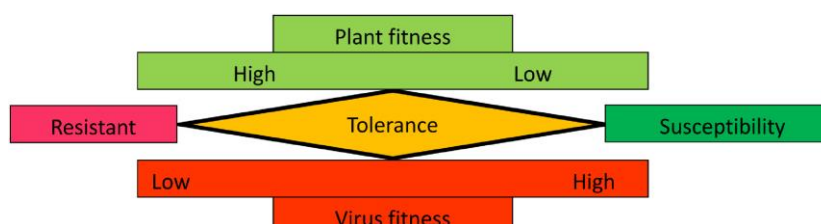


Figure 8 : Représentation graphique du continuum de *fitness* des plantes et des phytovirus selon des interactions résistantes, tolérantes ou sensibles (d'après Paudel *et al.*, 2018)

Les plantes déploient de multiples mécanismes de résistance ou de défense contre les virus assurés par des protéines et des ARN (Carr *et al.*, 2018). L'expression des gènes de défense a généralement des coûts pour la plante, impactant alors sa *fitness* (Heil and Bostock, 2002 ; Tian *et al.*, 2003 ; Yang and Hua, 2004 ; Yi and Richards, 2007 ; Paudel and Sanfaçon, 2018). Parmi ces mécanismes de défense, le mécanisme d'extinction de l'ARN, également appelée *RNA silencing*, est considéré comme la plus importante réponse basale de défense antivirale des plantes (Moon and Park, 2016). C'est un mécanisme de régulation génique, qui repose sur la génération de petits ARN interférents (ARNi) qui entravent le message porté par les ARN messagers en les dégradant ce qui diminue la traduction des protéines virales (de Alba *et al.*, 2013). En guise de réponse, les virus ont développé des mesures contre-défensives sophistiquées avec notamment des supprimeurs du *RNA silencing* pour contourner les défenses de la plante et permettre l'infection (**Figure 7**). En effet, la plupart des virus codent pour un ou plusieurs supprimeurs du *RNA silencing* (Carr *et al.*, 2018). Cette course à l'armement suggère une longue coévolution entre les plantes hôtes et les virus jusqu'à atteindre un équilibre finement réglé entre défense de l'hôte et contre-défense virale, permettant souvent au virus de persister sans causer trop de dommages à son hôte (Csorba *et al.*, 2015).

La tolérance est définie comme une relation de commensalisme dans laquelle les virus s'accumulent dans une certaine mesure sans causer de perte significative de *fitness* à leurs hôtes (Paudel and Sanfaçon, 2018). Cette relation peut être expliquée comme un compromis évolutif acceptable entre les phytovirus et les plantes pour une coexistence à long terme. Bien que la tolérance à l'infection virale soit un phénotype bien connu dans le contexte de l'agriculture, la base génétique et les mécanismes moléculaires qui régissent ces interactions n'ont pas encore été bien caractérisés. Des croisements génétiques et des études de cartographie ont identifié un certain nombre de locus de caractères quantitatifs (QTL) ou de gènes associés à la tolérance (McKenzie *et al.*, 1985 ; Singh *et al.*, 1993 ; Jin *et al.*, 1998 ; Riedel *et al.*, 2011 ; Del Blanco *et al.*, 2014 ; Foresman *et al.*, 2016). Dans tous les cas, la tolérance est un compromis associé à une réduction de l'accumulation des protéines virales. L'hôte peut diriger ses ressources énergétiques vers son développement et sa reproduction alors que les virus peuvent se multiplier et envahir l'intégralité de la plante

(Ghoshal and Sanfaçon, 2015). De plus, chaque partenaire peut bénéficier en retour de certains avantages comme la protection de la plante contre la super-infection par des virus virulents ou l'invasion des tissus méristématiques par le virus permettant la transmission verticale par exemple, on parle alors de mutualisme (Paudel and Sanfaçon, 2018). De même, dans certaines conditions environnementales, un phytovirus pathogène peut devenir bénéfique en cas de stress de l'hôte, améliorant par exemple la tolérance à la sécheresse (*i.e.* turnip mosaic virus (TuMV, famille *Potyviridae*) chez *Arabidopsis thaliana* ; González *et al.*, 2021) ou au gel de la plante hôte (*i.e.* cucumber mosaic virus, (CMV, famille *Bromoviridae*) chez les betteraves ; Xu *et al.*, 2008), ou renforçant la résistance aux maladies bactériennes (*i.e.* zucchini yellow mosaic virus (ZYMV, famille *Bromoviridae*) chez les courges sauvages ; Shapiro *et al.*, 2013).

Cependant, la tolérance et la résistance ne sont pas nécessairement mutuellement exclusives et les mécanismes qui régissent les deux peuvent se chevaucher de manière significative. D'ailleurs, de nombreux gènes de réponse de défense activés lors de mécanismes de résistance sont également induits dans les interactions tolérantes (Bengyella *et al.*, 2015 ; Pagan and Garcia-Arenal, 2018 ; Paudel and Sanfaçon, 2018). Enfin, il est important de noter que la réponse de l'hôte à l'infection virale (traits biologiques, tolérance, etc.) est également influencée par les génotypes du virus et de l'hôte, comme cela a été documenté chez *Arabidopsis* pour le CMV (Hily *et al.*, 2016 ; Montes *et al.*, 2019 ; Pagán *et al.*, 2008) et le TuMV (González *et al.*, 2021).

3.2.2. Généralistes et spécialistes : les gammes d'hôtes

La gamme d'hôtes des phytovirus est très variable, allant par exemple de quelques espèces hôtes pour le barley stripe mosaic virus (BSMV, famille *Virgaviridae*) à plus de 1000 espèces de 92 familles pour le tomato spotted wilt virus (TSWV, famille *Tospoviridae* ; Chatzivassiliou *et al.*, 2001 ; Maris *et al.*, 2003). Dans le spectre de spécificité des interactions plantes virus, deux extrêmes sont décrits avec les virus dits « généralistes » infectant et se répliquant de façon optimale dans un large éventail d'espèces et les virus dits « spécialistes » n'infectant qu'une ou peu d'espèces végétales apparentées. Certains traits liés au génome viral ou à la transmission ont d'ailleurs été associés à ces variations de gammes d'hôtes (Moury *et al.*, 2017). Chaque stratégie présente des avantages et des inconvénients : un virus

généraliste présente une *fitness* plus élevée à l'échelle de l'ensemble des espèces infectées mais en ayant une adaptation plus faible pour chaque espèce (« *Jack of all trades, master of none* » ; Elena *et al.*, 2009) ; tandis qu'un virus spécialiste optimise l'exploitation de ses quelques hôtes au prix de l'adaptation à toutes les espèces disponibles dans le milieu (« *Master of some* » ; Futuyma and Moreno, 1988).

Même si les virus généralistes ont plus d'opportunités de transmission de par leur large gamme d'hôtes (Woolhouse *et al.*, 2001), en raison de l'hypothèse du compromis adaptatif nécessaire pour permettre l'infection de plusieurs espèces hôtes, la théorie suggère que la sélection naturelle favoriserait la spécialisation plutôt que le généralisme (Futuyma and Moreno, 1988). En effet, des expériences menées dans des environnements homogènes en laboratoire ont montré que la *fitness* des généralistes tend à être plus faible que celle des spécialistes (Kawecki, 1994 ; Bedhomme *et al.*, 2015). Une des raisons avancées est la forte compacité des génomes viraux et l'importance des phénomènes d'épistasie qui font que les mutations favorisant la *fitness* d'un virus dépendent de mutations survenant ailleurs sur le génome viral (Ashby *et al.*, 2014). De plus, les mutations adaptatives qui augmentent la *fitness* virale chez un hôte peuvent diminuer la *fitness* chez un autre - c'est la pléiotropie antagoniste (Moury and Simon, 2011 ; Remold, 2012 ; Whitlock, 1996). Cependant, il existe des cas contraires où les mutations adaptatives pour l'infection d'un hôte pourraient également être adaptatives pour d'autres, ainsi certains virus généralistes seraient capables d'avoir une *fitness* égale voire supérieure aux spécialistes et sont décrits comme des généralistes « sans-coûts » (Remold, 2012). La majorité des phytovirus sont actuellement considérés comme généralistes (Power and Flecker, 2003 ; García-Arenal and Fraile, 2013). Les virus généralistes seraient en effet avantagés dans le contexte général des environnements hétérogènes présentant des variations de disponibilité d'hôtes et de vecteurs dans l'espace et le temps (Elena and Lenski, 2003).

Plus de la moitié des virus transmis par vecteurs (58,4%) le sont par une seule espèce (Power and Flecker, 2003). Par conséquent, la distribution virale pourrait dépendre davantage des interactions du virus avec ses vecteurs plutôt qu'avec ses plantes hôtes, car le vecteur limite obligatoirement la gamme d'hôte du virus par ses choix alimentaires (Power, 2008). Ainsi de nombreuses études expérimentales d'inoculation mécanique ou artificielle

des phytovirus, en contournant les contraintes environnementales liées à leurs vecteurs, ont pu montrer des gammes d'hôtes plus larges que celles décrites dans l'environnement (McLeish *et al.*, 2018).

3.2.3. Les co-infections

Au-delà d'être un prérequis à la recombinaison, les co-infections virales ou infections mixtes (*i.e.* infections simultanées du même hôte par plusieurs virus), courantes chez les plantes, peuvent aboutir à la coexistence entre les différents virus (sans effets positifs ou négatifs liés à la présence d'un virus sur le ou les autres) mais peuvent aussi avoir des effets synergiques ou antagonistes (Power, 2008 ; Syller, 2012 ; Mascia and Gallitelli, 2016 ; Alcaide *et al.*, 2020) et influencer la capacité d'évolution des virus et leur adaptation à de nouveaux hôtes ou environnements (McLeish *et al.*, 2018). Les mécanismes aboutissant à l'un ou l'autre type d'interactions n'étant pas encore totalement élucidés, les conséquences biologiques et épidémiologiques de la co-infection restent difficiles à prévoir (Zhang *et al.*, 2000 ; Jeger *et al.*, 2006).

Les interactions synergiques ont un effet facilitateur sur au moins une des espèces virales en co-infection, proches ou non phylogénétiquement, pouvant entraîner l'augmentation de la réplication d'au moins un des virus et l'induction de symptômes plus graves que chaque virus en infection individuelle. L'importance de la synergie virale a notamment été décrite dans l'épidémie de la forme sévère de la mosaïque du manioc en Afrique de l'Est et l'association entre l'African cassava mosaic virus (ACMV, famille *Geminiviridae*) et la souche Ougandaise de l'East African cassava mosaic virus (EACMV-UG ; Harrison *et al.*, 1997). D'autres interactions facilitatrices asymétriques ont été décrites entre différentes souches de la même espèce virale (*i.e.* souches Mild et IL du TYLCV, famille *Geminiviridae* ; Péréfarres *et al.*, 2014) ou différentes espèces virales appartenant à des familles différentes (*i.e.* barley yellow dwarf virus (BYDV; famille *Luteoviridae*) et le wheat dwarf virus (WDV ; famille *Geminiviridae*) ; Armand *et al.*, 2023) aboutissant à une augmentation ou une diminution de certains traits biologiques (taux d'infection, taux de réplication, efficacité de transmission) selon le virus. Le type d'inoculation mixte, simultanée ou séquentielle, semble également influencer l'impact des co-infections sur les paramètres biologiques et écologiques des maladies (Armand *et al.*, 2023).

À l'inverse, en cas d'antagonisme, la co-infection entraîne des effets négatifs sur les traits biologiques des virus aboutissant à une atténuation des symptômes. Deux types d'interactions antagonistes ont été décrites : la protection croisée (aussi appelée « interférence homologue ») se produit lorsqu'une infection est exclue ou réduite en raison d'une infection antérieure par un autre virus proche (dit « protecteur »); et l'exclusion mutuelle, qui est l'exclusion spatiale de virus étroitement apparentés au sein de la plante (aussi appelée « suppression concurrentielle mutuelle » ; Dietrich and Maiss, 2003 ; Syller, 2012).

3.2.4. La coévolution du triptyque phytovirus-hôte-vecteur

Si la diversité des modes de transmission des phytovirus ou l'existence d'interactions mutualiste entre plantes et virus sont les témoins de longue co-évolution, le meilleur exemple de la complexité et de la finesse associées aux interactions phytovirus-hôte-vecteur réside sûrement dans la manipulation par les virus des vecteurs ou de l'hôte afin de maximiser leur transmission. Ainsi, les phytovirus peuvent directement manipuler les vecteurs et modifier leurs comportements alimentaires et leurs performances (e.g. le TYLCV provoque une dégénérescence neuronale de son vecteur *Bemisia tabaci* avec des modifications de son comportement alimentaire en faveur de plantes saines ; Liu *et al.*, 2013 ; Wang *et al.*, 2020) ; ou indirectement en modifiant la morphologie et/ou le métabolisme de la plante (couleur, nutriments, composés volatils, dépôt de callose, obstruction du phloème, etc.) afin d'attirer/dissuader les vecteurs de se nourrir sur la plante en question (Mauck *et al.*, 2012 ; Dáder *et al.*, 2017). Un exemple bien caractérisé de manipulation indirecte est le système d'interaction barley yellow dwarf virus (BYDV, famille *Tombusviridae*) – blé – *Rhopalosiphum padi* (*Aphididae*). Grâce à des changements dans la composition relative des composés organiques volatils émis par la plante, des préférences alimentaires différentes ont été observées. Ainsi, les pucerons virulifères préfèrent se nourrir sur des plants de blé non infectés, tandis que les pucerons non-virulifères ont une préférence pour les plantes infectées par le BYDV (Ingwell *et al.*, 2012). Les phytovirus peuvent également influencer la durée d'alimentation des vecteurs en fonction du mode de transmission (circulant ou non) afin d'optimiser l'acquisition virale. Les virus non circulants auraient tendance à réduire la qualité nutritionnelle des plantes hôtes,

encourageant ainsi les vecteurs à se nourrir pendant de courtes périodes et induisant une dispersion rapide des vecteurs vers d'autres plantes alors que la situation inverse a été observée pour les virus circulants (Mauck *et al.*, 2012). Outre l'influence sur les comportements alimentaires et la dispersion des virus (Mauck *et al.*, 2014), certains phytovirus ont été décrits comme améliorant la croissance, la fécondité, la survie ainsi que la longévité des vecteurs, et donc leur *fitness*, ce qui, *in fine*, favorise la transmission des phytovirus (Casteel *et al.*, 2014 ; Fereres and Moreno, 2009 ; Mauck *et al.*, 2012).

4. L'écologie des phytovirus

L'écologie virale s'intéresse à l'ensemble des acteurs du cycle d'infection virale (virus, vecteur, hôte, et environnement biotique et abiotique) et à leurs interactions (Malmstrom *et al.*, 2011). Plus particulièrement, elle permet d'étudier les patrons de distribution, de la dynamique et des effets des communautés virales dans un écosystème donné et réciproquement des effets de la biodiversité et de l'environnement sur la dynamique et l'évolution des communautés virales (Gibbs, 1983 ; Jones, 2014 ; Lefeuvre *et al.*, 2019 ; **Figure 9**). Ainsi, il faut considérer les phytovirus intégrés dans un système incluant les interactions multi-spécifiques (virus-vecteurs, virus-hôtes et virus-vecteurs-hôtes), les processus évolutifs (la recombinaison, la mutation, le flux génétique, la sélection et la dérive génétique) et écologiques (les goulots d'étranglement et les événements fondateurs) qui modifient les associations virus-hôte (Stinchcombe and Hoekstra, 2008) et à différentes échelles spatio-temporelles (Elena *et al.*, 2014 ; McLeish *et al.*, 2021 ; **Figure 9**). Dans le cas des phytovirus, cette discipline cherche spécifiquement à examiner "(i) les rôles écologiques des virus associés aux plantes et de leurs vecteurs dans les écosystèmes cultivés et non cultivés ; et (ii) l'influence réciproque des propriétés des écosystèmes sur la distribution et l'évolution des virus, de leurs insectes vecteurs et de leurs plantes hôtes" (Malmstrom *et al.*, 2011).

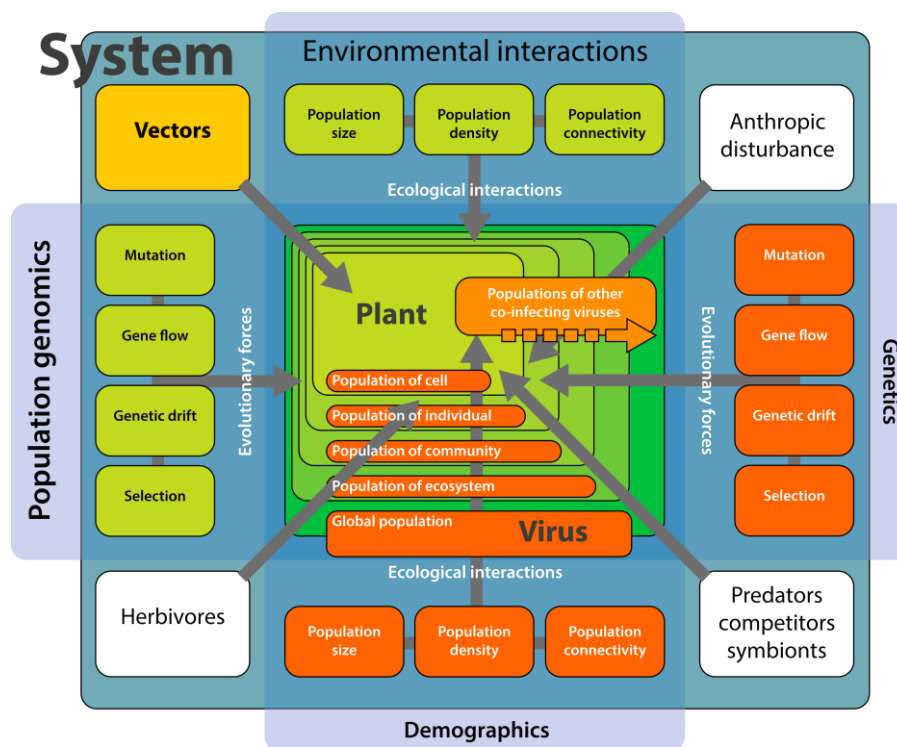


Figure 9 : Représentation schématique des interactions interspécifiques à différentes échelles et des processus dans lesquels les populations sont intégrées. Les compartiments verts représentent les populations végétales et les compartiments oranges représentent les populations virales, définies à différentes échelles, chacune étant représentée de manière hiérarchique, de la cellule (niveau supérieur) au niveau mondial (niveau inférieur). Les flèches grises relient les facteurs internes (plantes et phytovirus) ou externes (vecteurs, herbivores, activités humaines) qui peuvent influencer les populations végétales ou virales à une échelle donnée (d'après McLeish *et al.*, 2021).

4.1. La métagénomique au service de l'écologie virale

Depuis la deuxième moitié du XX^{ème} siècle, des progrès importants ont été réalisés dans la caractérisation biologique et moléculaire des virus grâce au développement d'une gamme d'outils moléculaires tels que l'électrophorèse des protéines et des acides nucléiques, les techniques sérologiques, notamment l'*enzyme-linked immunosorbent assay* (ELISA) (Clark and Adams, 1977), et les méthodes basées sur les acides nucléiques, telles que les techniques d'hybridation moléculaire (Maule *et al.*, 1983). Enfin, la technique de la réaction en chaîne par polymérase (PCR) a été essentielle dans la détection, la caractérisation et l'étude de la diversité des virus (Maclot *et al.*, 2020 ; **Figure 10**). Toutefois, ces dernières années, le domaine de la virologie a subi une révolution majeure suite aux avancées de la métagénomique virale (Koonin and Dolja, 2018), c'est à dire l'étude du contenu génétique d'échantillons complexes. L'importance de cette révolution fait qu'il est aujourd'hui difficile de ne pas associer l'écologie virale et l'évolution de cette discipline avec la métagénomique.

Si pour de nombreux organismes les études basées sur la métagénomique, reposant sur le séquençage de régions universelles du génome (e.g. gènes universels de l'ARN ribosomique 16S et 18S présents respectivement chez les bactéries et les champignons ; Lane *et al.*, 1985 ; Simon *et al.*, 2019), se sont avérées fructueuses (Stackebrandt and Goebel, 1994 ; Seifert, 2009), l'absence de gène canonique chez les virus ont poussés au développement de méthodes spécifiques (Roossinck, 2012b). Leur utilisation s'est rapidement révélée être une alternative convaincante aux tests sérologiques et moléculaires classiques tels que l'ELISA ou la PCR. Le diagnostic métagénomique sans *a priori* a notamment permis de résoudre un certain nombre d'énigmes étiologiques (Adams *et al.*, 2009 ; Al Rwahnih *et al.*, 2009 ; Kreuze *et al.*, 2009 ; Donaire *et al.*, 2009). Toutefois, c'est au niveau conceptuel que la métagénomique a apporté une véritable révolution. Ces méthodologies ont permis un changement de paradigme (Breitbart *et al.*, 2002 ; Wang *et al.*, 2002 ; Roossinck *et al.*, 2010) et à une prise de conscience que les virus faisaient partie intégrante de l'environnement et qu'ils devaient être considérés en tant que tels si l'on souhaitait comprendre les mécanismes structurant leur diversité et distribution.

La métagénomique virale basée sur les nouvelles technologies de séquençage, appelée séquençage de nouvelle génération (*Next Generation Sequencing*, NGS), ou séquençage à haut débit (*high throughput sequencing*, HTS) a alors grandement contribué à une nouvelle compréhension de la diversité virale dans la nature (Maclot *et al.*, 2019 ; **Figure 10 & 11**). L'une des premières études métagénomiques des virus des plantes a porté sur la diversité des virus dans les plantes sauvages natives d'un espace naturel protégé, le *Tallgrass Prairie Preserve* dans le nord-est de l'Oklahoma (Wren *et al.*, 2006 ; Melcher *et al.*, 2008 ; Muthukumar *et al.*, 2009). Ces premiers travaux ont mis à jour une diversité virale largement sous-estimée dans les écosystèmes naturels et ont initié de nombreuses autres études sur la diversité des virus dans les plantes non cultivées (Roossinck *et al.*, 2010, Lebouvier *et al.*, 2011 ; Roossinck *et al.*, 2015 ; Maclot *et al.*, 2021 ; Hasiów-Jaroszewska *et al.*, 2021 ; Maclot *et al.*, 2023). Celles-ci ont notamment permis d'étudier le rôle des virus dans le fonctionnement des écosystèmes et leur distribution au niveau des interfaces agricoles et sauvages (Malmstrom *et al.*, 2011 ; Alexander *et al.*, 2014 ; Bernardo *et al.*, 2018 ; Ma *et al.*, 2020 ; Maachi *et al.*, 2022).

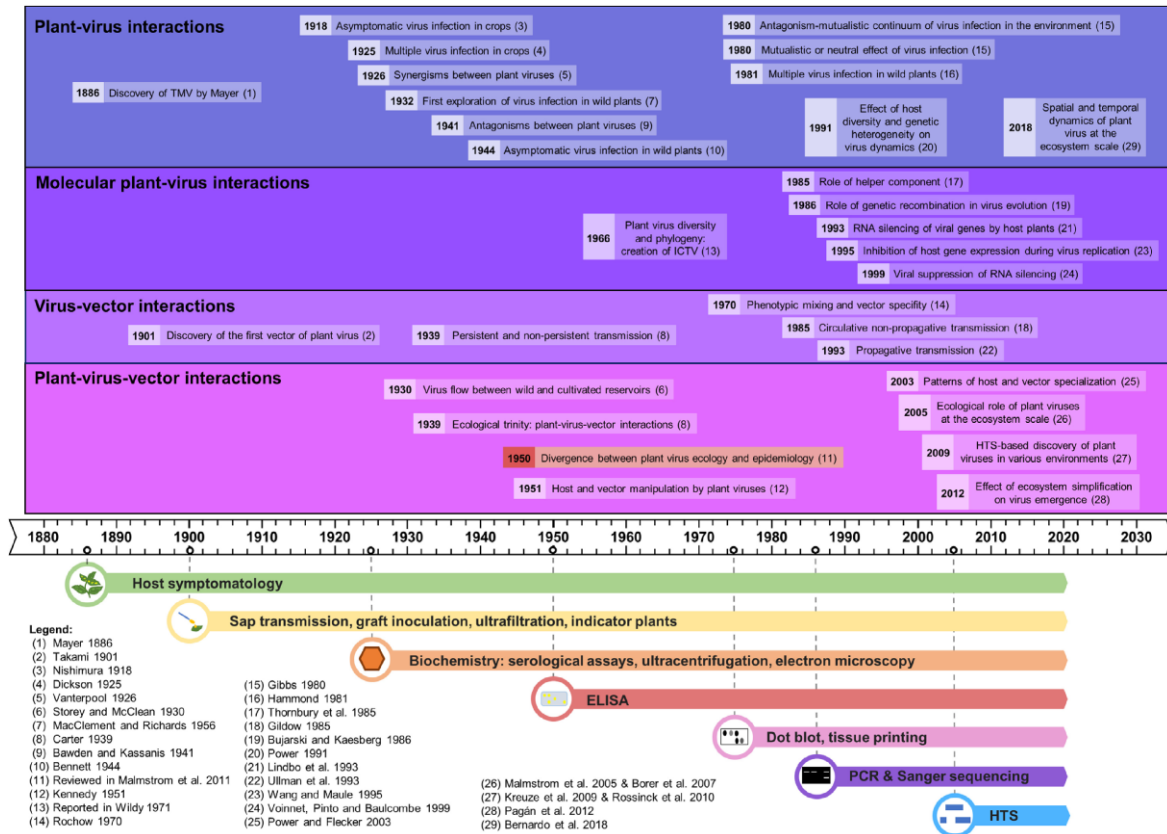


Figure 10 : Frise chronologique des principales avancées en écologie des phytovirus, liée à l'évolution des méthodes de détection (D'après Maclot *et al.*, 2020).

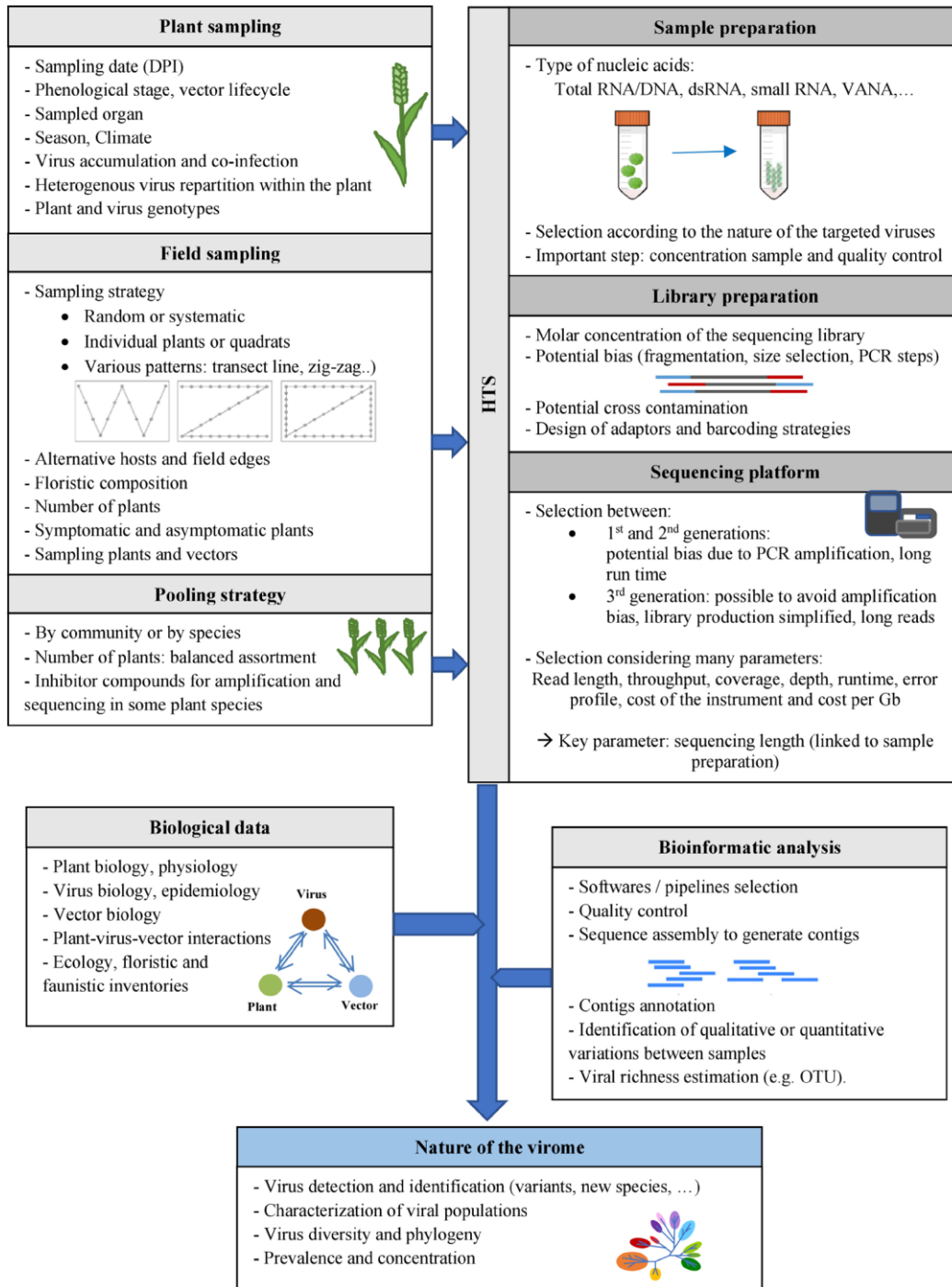


Figure 11 : Méthodes de caractérisation des viromes des communautés végétales. Divers facteurs peuvent potentiellement affecter les études écologiques virales basées sur le séquençage à haut débit (HTS) à différentes étapes : nombre de jours après inoculation (DPI), stratégie d'échantillonnage sur le terrain, stratégies de *pooling*, méthodes d'échantillonnage et de préparation des bibliothèques, sélection des technologies de séquençage et analyses bio-informatiques. Des données biologiques sur les virus, leurs plantes hôtes et leurs vecteurs (si connus) pourraient également être utiles pour les études écologiques virales (d'après Maclot *et al.*, 2020).

4.2. L'influence de l'environnement biotique et abiotique

Un défi majeur pour la compréhension de l'épidémiologie des maladies virales et plus largement l'écologie virale consiste à dépasser la démarche classique focalisée sur les interactions entre un seul hôte et un seul pathogène et à s'intéresser aux interactions entre les hôtes et les pathogènes au sein de communautés complexes et en conditions abiotiques variées (Woolhouse and Gowtage-Sequeria, 2005 ; Johnson *et al.*, 2013 ; McLeish *et al.*, 2022). En effet, à l'échelle des paysages, la composition, la distribution, l'abondance et la dynamique des communautés végétales au cours des saisons peuvent impacter la richesse en phytovirus et les infections réalisées (Alexander, 1998 ; Alexander *et al.*, 2014 ; Borer *et al.*, 2007 ; Dunn *et al.*, 2012 ; Malmstrom *et al.*, 2005 ; Power and Mitchell, 2004).

4.2.1. Le rôle des communautés végétales

La richesse spécifique et la densité végétale du milieu peuvent influencer la prévalence (*i.e.* proportion d'individus dans une population donnée infectés par un virus) et/ou la diversité des agents pathogènes. Deux hypothèses majeures ont été proposées pour expliquer ce constat. D'abord, l'effet d'amplification qui prédit une corrélation positive entre la diversité des espèces de plantes hôtes et le risque d'infection virale et de maladie, due à un plus grand nombre d'hôtes potentiels disponibles pour les virus (Keesing *et al.*, 2006). Cet effet serait lié du côté viral à un trait généraliste d'infection (Pagán *et al.*, 2012) et du côté des communautés végétales à une offre potentiellement plus grande du nombre de niches à exploiter par les phytovirus (Lafferty, 2012 ; Kamiya *et al.*, 2014 ; Johnson *et al.*, 2016). Inversement, l'effet de dilution prédit une diminution du risque d'infection et de maladie avec une augmentation de la biodiversité de plantes (Keesing *et al.*, 2006 ; **Figure 12**). Ce deuxième effet serait lié à un trait spécialiste de l'espèce virale (Pagán *et al.*, 2012) et à une raréfaction de la disponibilité des hôtes potentiels présents dans le milieu (Keesing *et al.*, 2010). Si des éléments de preuves ont été présentés en faveur des deux hypothèses (Haas *et al.*, 2011 ; Lacroix *et al.*, 2014), c'est toutefois l'effet de dilution, avec une corrélation négative entre la biodiversité et le risque de maladie, qui est le plus souvent observé (Pagán *et al.*, 2012 ; Civitello *et al.*, 2015 ; Fraile *et al.*, 2017 ; Liu *et al.*, 2020 ; Susi *et al.*, 2021 ; McLeish *et al.*, 2022).



Figure 12 : Schématisation des relations entre le risque de maladie, les activités humaines et la biodiversité (d'après Roossinck and Garcia-Arenal, 2015).

Par ailleurs, les risques d'infection virale varient largement dans l'espace et le temps (Handel *et al.*, 2014 ; Laine, 2006 ; Hamer *et al.*, 2011 ; Viana *et al.*, 2014 ; Blumenthal *et al.*, 2009 ; McLeish *et al.*, 2017 ; Carlsson-Graner and Thrall, 2015 ; Valverde *et al.*, 2020). D'un point de vue spatial, la distribution des virus est intrinsèquement dépendante de la présence d'hôtes et de vecteurs et est par conséquent, largement dépendante de l'habitat (McLeish *et al.*, 2017 ; Zhao *et al.*, 2018 ; Valverde *et al.*, 2020). D'un point de vue temporel, il existe une abondance de preuves de la saisonnalité de l'infection (Martinez, 2018). Les fluctuations spatiales et temporelles dans la distribution et l'abondance des espèces végétales déterminent la probabilité des rencontres avec les agents pathogènes et façonnent la structure des interactions hôtes-pathogènes (McLeish *et al.*, 2018 ; Ostfeld and Keesing, 2000 ; Gurarie and Seto, 2009 ; Webster *et al.*, 2007). Dans ce contexte, les plantes pérennes semblent jouer un rôle majeur dans le maintien et l'évolution des communautés virales dans les écosystèmes (Christian and Willis, 1993 ; Huggett *et al.*, 1999 ; Pointeau *et al.*, 2014 ; Gilabert *et al.*, 2017). En effet, les espèces végétales pérennes non cultivées peuvent jouer le rôle de réservoir et de source (*i.e. hub*) de phytovirus dans les écosystèmes d'une année à l'autre. Ainsi, contrairement aux espèces annuelles ou cultivées, qui ne sont présentes qu'à court terme, les espèces pérennes permettent de maintenir la présence des phytovirus dans les écosystèmes à long terme.

4.2.2. L'impact des perturbations anthropiques

4.2.2.1. Une diversité phytovirale sauvage largement sous-estimée

Jusqu'à récemment, les études de diversité et de distribution des phytovirus se sont principalement focalisées sur les infections pathogènes et symptomatiques des espèces végétales cultivées et à intérêt économique. Ceci a conduit à sous-estimer la diversité globale des phytovirus et le rôle des réservoirs sauvages dans l'émergence des maladies virales (Cooper and Jones, 2006). Alors qu'on a longtemps supposé qu'une absence de symptômes visuels indiquait une absence d'infection virale chez les plantes (Prendeville *et al.*, 2012), les données récentes semblent indiquer que les infections virales sont très courantes dans la nature et souvent asymptomatiques, impliquant même des infections mixtes (Roossinck, 2010 ; Johnson *et al.*, 2013 ; Tollenaere *et al.*, 2016 ; Susi *et al.*, 2019). En effet, les travaux pionniers de ces dernières années s'accordent sur l'importante richesse virale et les fortes fréquences d'infections dans la sphère naturelle (Roossinck, 2010, Muthukumar *et al.*, 2009 ; Bernardo *et al.*, 2018). Ces études ont aussi montrées que le virome (*i.e.* ensemble des espèces virales présentes dans un environnement donné) des compartiments sauvages et cultivés pouvaient différer significativement (Roossinck *et al.*, 2015 ; Ma *et al.*, 2020 ; Maachi *et al.*, 2022).

Les connaissances actuelles proposent (1) que la coévolution à long terme entre les plantes sauvages et les phytovirus conduirait à un équilibre et à une situation de commensalisme, voire de mutualisme, et (2) que l'agriculture aurait favorisé l'émergence de virus pathogènes depuis les réservoirs sauvages vers les zones cultivées (Anderson *et al.*, 1991 ; Sherwood, 1997 ; Jones, 2009 ; Roossinck and García-Arenal, 2015). Ainsi, dans les zones cultivées, l'abondance d'hôtes et de vecteurs spécialisés et adaptés à des cultures spécifiques permettent aux phytovirus de se propager.

4.2.2.2. L'interface agroécologique comme lieu d'émergence

Les activités humaines telles que les pratiques agricoles ont conduit à la création d'interface agroécologique, zones dans lesquelles les milieux cultivés jouxtent des zones sauvages ou non gérées. Ces frontières représentent une des zones peuplées de plantes cultivées, d'adventices spontanées, de plantes sauvages non cultivées ainsi que d'animaux et de micro-organismes divers (Alexander *et al.*, 2014). Cette interface agroécologique

comportant plante sauvage et culture, offre alors au virus, entités biologiques hautement adaptables et variables, de nombreuses possibilités de saut d'hôte et d'émergence (Elena *et al.*, 2014). L'émergence s'accompagne souvent de changements dans la pathogénèse et d'une gravité accrue des maladies, entraînant des épidémies dévastatrices (Fargette *et al.*, 2006 ; Elena *et al.*, 2014 ; Roossinck and Garcia-Arenal, 2015). Si ce sont généralement les émergences depuis les plantes non cultivées vers les plantes cultivées qui sont considérées, il faut noter que l'inverse est aussi possible et peut fortement impacter les écosystèmes (Maskell *et al.*, 1999 ; Malmstrom *et al.*, 2005 ; Malmstrom and Alexander, 2016). A ce titre, ce sont probablement les plantes introduites par l'homme, et devenue envahissante qui ont l'impact le plus important.

4.2.3. Les impacts de la dynamique des PEE sur les communautés phytovirales

Une plante exotique est définie comme une plante qui n'est pas indigène (ou native) d'une région géographique et qui a été introduite intentionnellement ou accidentellement. D'autre part, une plante invasive (ou envahissante) est une espèce dont la prolifération entraîne des changements significatifs dans la structure ou le fonctionnement des écosystèmes (Richardson *et al.*, 2000). Une plante invasive peut être soit une plante exotique naturalisée, c'est-à-dire qu'elle se reproduit régulièrement dans sa nouvelle région géographique et persiste à long terme, soit une plante indigène (Richardson *et al.*, 2000). Une des principales voies d'introduction de plantes exotiques envahissantes (PEE) est l'introduction via les cultures agricoles en tant que plantes adventices ou plantes qualifiées de « mauvaises herbes » (Wilson *et al.*, 2016 ; Pyšek *et al.*, 2017). Ces espèces exotiques peuvent rapidement devenir envahissantes et avoir des impacts importants sur les écosystèmes locaux (Pyšek *et al.*, 2020). En particulier, les PEE peuvent être une source d'agents pathogènes (Pimentel *et al.*, 2001 ; Marsberg *et al.*, 2017 ; Pyšek *et al.*, 2020) qui peuvent avoir un impact sur les invasions en diminuant ou au contraire amplifiant les dynamiques de leurs hôtes (Faillace *et al.*, 2017 ; Rúa *et al.*, 2011). Quatre principaux modèles ont été proposés par Faillace *et al.* (2017) pour décrire l'impact des phytovirus sur l'introduction d'hôtes exotiques dans un milieu. Dans le cas de phytovirus indigènes, l'installation d'espèces hôtes exotiques susceptibles aux infections virales sera fortement réduite, c'est la résistance biotique (**Figure 13a**). Si les espèces hôtes exotiques sont introduites sans leur cortège d'agents pathogènes et ne sont pas impactées par les pathogènes indigènes,

elles pourront proliférer et se propager sans contraintes, envahissant les niches des populations d'espèces hôtes indigènes continuant d'être impactées par leurs agents pathogènes. C'est l'hypothèse du relâchement de la pression parasitaire ou *enemy release* (Mitchell and Power, 2003 ; **Figure 13b**). Si la transmission d'un phytovirus depuis une espèce hôte primaire (exotique ou indigène) vers une espèce hôte secondaire (indigène ou exotique, respectivement) susceptible entraîne la réduction des populations d'hôtes secondaires au bénéfice des hôtes primaires, on appelle ce processus *pathogen spillover* (Power and Mitchell, 2004 ; Xu *et al.*, 2022 ; **Figure 13c**). Enfin, si les espèces hôtes exotiques introduites sont tolérantes aux infections virales par des phytovirus indigènes, elles peuvent jouer le rôle de réservoir et amplifier l'impact des phytovirus sur les hôtes indigènes, ce qui aura pour conséquence la réduction des populations d'hôtes indigènes au bénéfice des hôtes exotiques, c'est le *pathogen spillback* (*i.e.* amplification des parasites natifs ; **Figure 13d**). Les études menées par Malmstrom *et al.* (2005) sur des *Poaceae* annuelles invasives introduites dans des prairies californiennes ont mis en évidence le phénomène de *pathogen spillback*. Les espèces introduites, tolérantes au BYDV et au CYDV ont agi comme réservoirs de ces pathogènes, entraînant une augmentation globale de leur prévalence et impactant principalement les espèces indigènes. Ces dernières ont alors vu leurs populations se réduire drastiquement au profit des espèces exotiques (Malmstrom *et al.*, 2005 ; Borer *et al.*, 2007). Ces études ont par ailleurs mis en lumière l'influence des plantes invasives sur l'écologie des vecteurs. En effet, les pucerons s'alimentant préférentiellement sur les espèces introduites, l'augmentation de la population des *Poaceae* exotiques a entraîné une augmentation de la fécondité des insectes, favorisant ainsi la transmission du BYDV et du CYDV (Malmstrom *et al.*, 2005). Une fois de plus ce constat souligne l'importance de considérer la dynamique des populations de vecteurs pour comprendre la distribution des phytovirus.

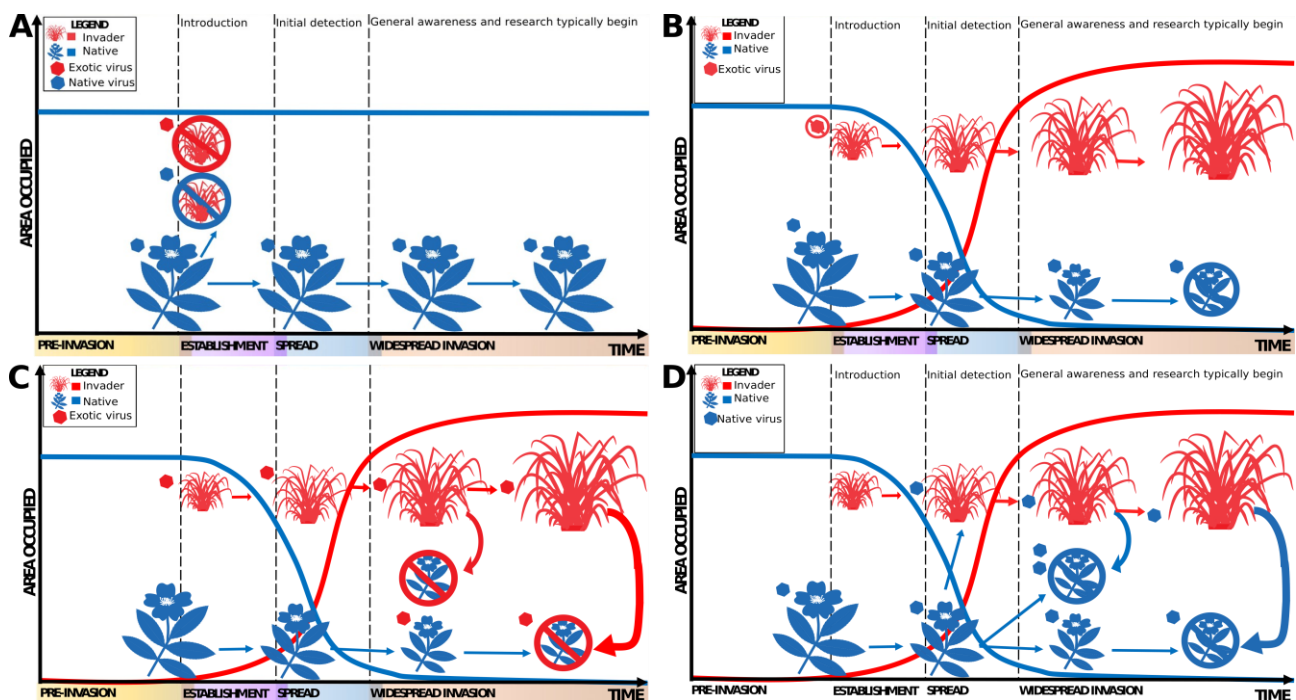


Figure 13 : Représentation des quatre principaux modèles expliquant l'impact des phytovirus suite à l'introduction de plantes hôtes exotiques dans un milieu. (a) la résistance biotique ; (b) l'*enemy release* (relâchement de la pression des pathogènes) ; (c) le *spillover* (débordement) ; (d) le *spillback* (retour de débordement). Les courbes rouges (exotiques) et bleues (indigène) montrent l'aire d'occupation des espèces dans le temps. Les plantes exotiques et le virus associé (hexagone) sont représentés en rouge et l'espèce indigène et le virus en bleu. La taille des illustrations de plantes indique la taille de la population, tandis que les flèches colorées indiquent la transmission des virus. Les populations localement disparues sont indiquées par un cercle barré. La couleur du cercle barré indique la source du pathogène affectant la population disparue. (d'après Faillace *et al.*, 2017).

5. Les géminivirus

Les géminivirus sont des phytovirus appartenant au phylum des *Cressdnaviricota* (*Circular Replication-associated protein Encoding Single Stranded (CRESS) DNA viruses* ; Rosario *et al.*, 2012 ; **Figure 14**). Ces pathogènes infectent à la fois les plantes monocotylédones et dicotylédones et causent des maladies économiquement importantes impactant les plantes cultivées comme la tomate (Moriones and Navas-Castillo, 2000), le maïs (Bosque-Pérez, 2000), le coton (Sattar *et al.*, 2013), le pois chiche (Kanakala and Kuria, 2018) et le manioc (Varma and Malathi, 2003), dans la plupart des régions du monde, avec la plus grande incidence et diversité dans les régions tropicales et subtropicales (Rojas *et al.*, 2018). Les émergences de géminivirus ont souvent été des cas d'écoles avec la démonstration d'effets synergiques entre souches (Zhou *et al.*, 1997), de l'implication d'évènements de recombinaisons génétiques directement liés à l'émergence (Varsani *et al.*, 2009), de la

pullulation d'insectes vecteurs (Varma and Malathi, 2003 ; Colvin *et al.*, 2006 ; Jiu *et al.*, 2007), de la propagation de matériel végétal infecté (Jones, 2020) ou encore de l'expansion de l'agriculture dans de nouvelles zones de culture (Garcia-Arenal and Zerbini, 2019).

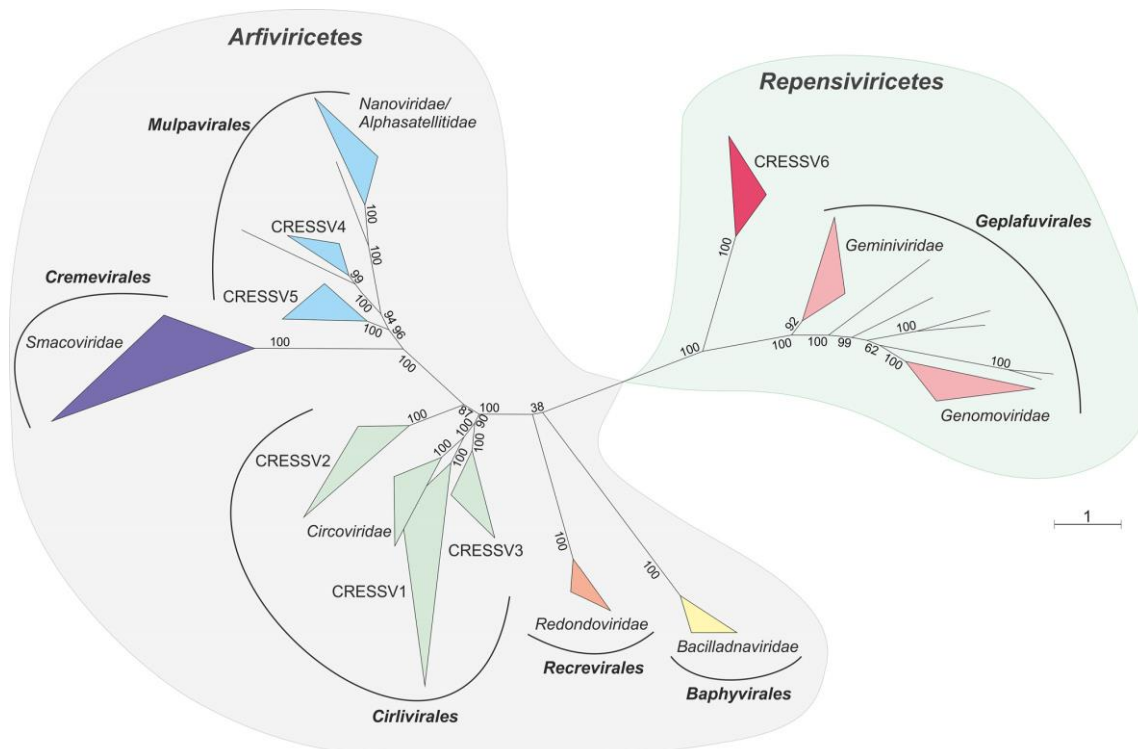


Figure 14 : Arbre phylogénétique non enraciné, basé sur la méthode du maximum de vraisemblance, des protéines associées à la réplication des *Cressdnaviricota* (Krupovic, 2020).

5.1. Les particularités morphologiques et génétiques des géminivirus

Parmi la grande diversité de CRESS-DNA virus décrits par des études métagénomiques (Rosario *et al.*, 2012 ; Dayaram *et al.*, 2016), huit architectures génomiques ont été mises en évidence (**Figure 15**), résultant de divers événements de recombinaison. Les géminivirus appartiennent aux génomes de type II.

Les génomes des géminivirus peuvent être monopartites, avec une seule molécule d'ADN d'environ 2 600 à 3 600 nucléotides (nt), ou bipartites, avec deux molécules d'ADN d'environ 2 600 nt (appelées ADN-A et ADN-B), pour une taille totale du génome d'environ 5 200 nt (Zerbini *et al.*, 2017). Le génome des géminivirus est encapsidé dans des virions quasi-icosaédriques en doublets (jumeaux ou "gémisés", d'où ils tirent leur nom) de 22 × 38 nm, composés de 110 sous-unités de protéines capsidiques organisées en 22 capsomères

pentamériques (Böttcher *et al.*, 2004 ; Hesketh *et al.*, 2018 ; Hipp *et al.*, 2017 ; Zhang *et al.*, 2001 ; **Figure 16**).

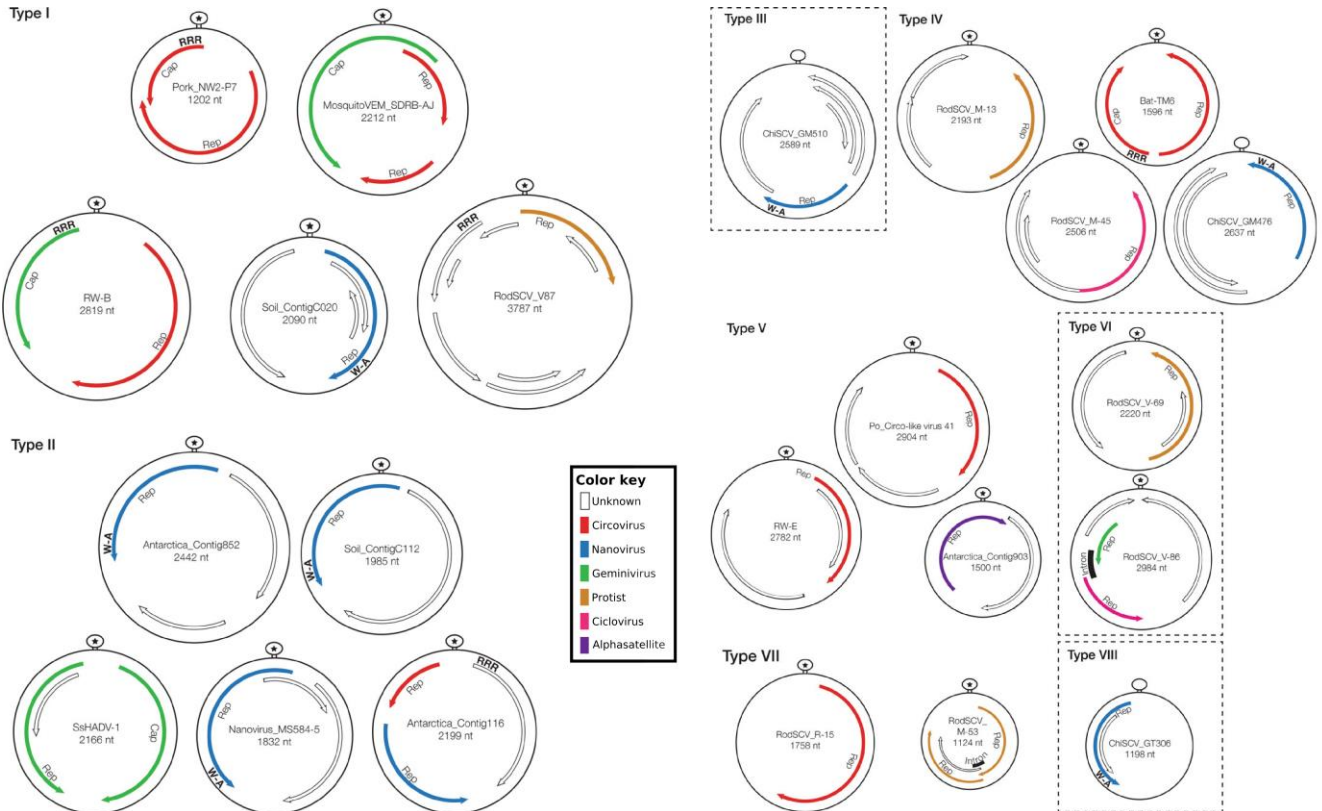


Figure 15 : Organisation schématique du génome des virus à ADNsb des huit types d'architectures génomiques décrites. Les types de génomes ont été classés en fonction des caractéristiques répertoriés. La taille de chaque cercle est proportionnelle à la taille du génome. Les gènes sont colorés suivant les assignations les plus proches. Les fonctions putatives de protéines de réplication (Rep) et des protéine de capsid (CP) sont indiquées. Les génomes contenant un motif de nonanucléotide ('NANTATTAC') en amont d'une structure en tige boucle potentielle sont indiqués par une étoile. 'RRR' indique la présence d'une région N-terminale riche en acides aminés basiques, caractéristique de la CP des circovirus. La présence d'un motif Walker-A unique de l'hélicase SF3 dans les gènes putatifs de la Rep est indiqués par 'W-A' (d'après Rosario *et al.*, 2012).

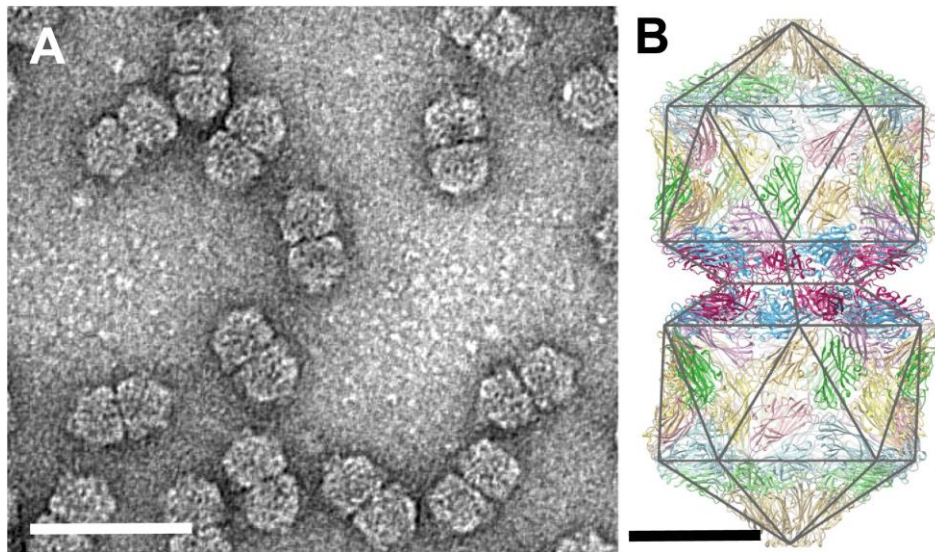


Figure 16 : Structure des particules de géminivirus. (A) Observation au microscope électronique en transmission de particules purifiées du maïze streak virus (MSV) colorées à l'acétate d'uranyle montrant des sous-unités quasi-isométriques jumelées typiques (échelle 50 nm ; ICTV ; Zhang *et al.*, 2001). (B) Modèle atomique complet pour les 110 sous-unités de la capsid du virus de l'Ageratum yellow vein virus (AYVV). L'axe de symétrie à cinq pôles est vertical et passe par le centre de la particule dans cette vue (échelle 10 nm ; ICTV ; Hesketh *et al.*, 2018).

Différents scénarios ont été proposés pour expliquer l'origine des *Cressdnaviricota* (Koonin and Ilyina, 1992 ; Rojas *et al.*, 2005 ; Kazlauskas *et al.*, 2019). Le scénario évolutif le plus récent décrit par Kazlauskas *et al.* (2019) propose que les géminivirus, comme les CRESS-DNA virus en général, sont très vraisemblablement issus de la recombinaison entre des plasmides procaryotes portant des gènes impliqués dans la réplication en cercle roulant des plasmides bactériens et des gènes de capsid de virus à ARN(+) de cellules eucaryotes (**Figure 4** ; Dolja *et al.*, 2020). De plus, Krupovic *et al.* (2009) ont souligné la similarité structurelle marquée de type "*jelly-roll*" (repliement des protéines en huit feuilletts β antiparallèles arrangés en deux feuilletts de quatre brins) entre la protéine de capsid des géminivirus et celle de certains virus à ARNs.

5.2. La réplication et la transcription des géminivirus

Les géminivirus sont dépendants de l'utilisation de l'ADN polymérase cellulaire de leur hôte pour leur réplication. La réplication a lieu selon deux mécanismes (Jeske *et al.*, 2001) connus sous le nom de réplication en cercle roulant (*rolling circle replication*, RCR) et réplication dépendante de la recombinaison (*recombination-dependent replication*, RDR).

La RCR se déroule en quatre étapes (Jeske *et al.*, 2001). D'abord un brin complémentaire est synthétisé permettant la production d'un ADNdb circulaire (Saunders *et al.*, 1991). Puis, la Rep initie la réplication en clivant l'ADNdb viral au niveau de l'origine de réplication, plus précisément la tige-boucle contenant le nonanucléotide hautement conservé (5'-TAATATT[↓]AC-3 ; Hanley-Bowdoin *et al.*, 2000). La Rep se lie ensuite de façon covalente au niveau de l'extrémité 5' de l'ADNdb clivé. L'ADNdb viral, dans un état circulaire ouvert, devient alors une matrice pour la production en continu d'un nouvel ADNsb viral. Une fois la synthèse complète du nouvel ADNsb, celui-ci est alors circularisé avant relargage de la Rep (Jeske *et al.*, 2001). La RDR repose sur l'interaction entre un ADNsb et la forme super-enroulée de l'ADN viral (*covalently closed circular DNA* – cccDNA) à des sites homologues induisant une recombinaison homologue. Par la suite l'élongation de l'ADNsb se produit suivi de l'obtention d'un ADNdb par synthèse complémentaire (Mosig *et al.*, 2001 ; Preiss and Jeske, 2003).

Les génomes de géminivirus sont ensuite transcrits dans le noyau de manière bidirectionnelle, c'est à dire que les ARN messenger (ARNm) résultent à la fois des ORFs du sens viral que du sens complémentaire (Blevins *et al.*, 2006 ; Pooggin, 2016 ; Pooggin, 2018). Toutefois l'étude récente de Lin *et al.* (2023) suggère l'existence de 28 sites de départ de transcription différents chez le tomato yellow leaf curl virus (genre *Begomovirus*), soutenant un mécanisme plus complexe de la transcription que le mécanisme bidirectionnel décrit jusqu'à présent.

5.3. La diversité et la taxonomie

Avec l'essor des nouvelles approches moléculaires (amplification en cercle roulant ; Inoue-Nagata *et al.*, 2004) et de séquençage à haut débit (Chalupowicz *et al.*, 2019 ; Naito *et al.*, 2019 ; Leiva *et al.*, 2020 ; Roumagnac *et al.*, 2022), une légion de géminivirus a été décrite ces deux dernières décennies. L'ensemble de cette diversité a été classé en 14 genres de *Geminiviridae* : *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrilevirus*, *Opunvirus*, *Topilevirus*, *Topocuvirus* et *Turncurtovirus* (Roumagnac *et al.*, 2022 ; **Figure 17**). Le classement taxonomique est fondé sur leurs insectes vecteurs, appartenant à quatre familles d'hémiptères (*Aleyrodidae*, *Aphididae*, *Cicadellidae* et *Membracidae* ; Hogenhout *et al.*,

2008 ; **Figure 18**), leur organisation génomique, leurs gammes d'hôtes et leurs identités de séquences par paires à l'échelle du génome (Roumagnac *et al.*, 2022). Outre pour ces 14 genres, des données sur les vecteurs et la capacité à former des particules virales en doublets (caractéristiques des géminivirus) manquent pour de nombreux isolats très divergents et encore non classés (Varsani *et al.*, 2017). Leur grand nombre et diversité soulignent que la classification de cette famille devrait subir de nombreuses évolutions dans les années à venir.

Parmi les différents genres de *Geminiviridae*, le genre *Begomovirus* dont l'espèce type est le ***Bean golden mosaic virus*** comporte le plus grand nombre d'espèces décrites avec plus de 445 espèces reconnues par l'ICTV (<https://ictv.global/>), selon le seuil de distinction taxonomique de 91 % d'identité nucléotidique (Brown *et al.*, 2015). Il est suivi du genre *Mastrevirus* dont l'espèce type est le ***Maize streak virus*** avec plus de 45 espèces reconnues (Zerbini *et al.*, 2017) sur la base d'un seuil de distinction taxonomique de 78 % d'identité nucléotidique (Muhire *et al.*, 2013). Le genre *Begomovirus* est le seul décrit présentant des espèces dont le génome est soit constitué d'un seul composant circulaire d'ADNsb (monopartite), soit de deux composants circulaires d'ADNsb (bipartite), d'une longueur de 2,6 à 3,2 kb, tandis que tous les autres *Geminiviridae* ont été décrits comme monopartites. Les *Begomovirus* monopartites sont souvent associés à des ADN satellites (alphasatellites, betasatellites et deltasatellites ; Briddon and Stanley, 2006 ; Zhou, 2013 ; Lozano *et al.*, 2016). Certains de ces satellites ont été associés à des gains de virulence (Amin *et al.*, 2011 ; Nawaz-ul-Rehman *et al.*, 2010). Des associations avec des satellites ont été aussi décrites chez des espèces du genre *Mastrevirus* (Kumar *et al.*, 2014 ; Claverie *et al.*, 2020).



Figure 17 : Illustration de l'organisation du génome des genres de la famille *Geminiviridae*. LIR, région intergénique longue ; SIR, région intergénique courte ; CR, région commune ; *cp*, gène codant pour la protéine de capsid ; *mp*, gène codant pour la protéine de mouvement ; *nsp*, gène codant pour la protéine de transport nucléaire ; *reg*, gène régulateur ; *ren*, gène codant pour la protéine activatrice de la réplication ; *rep*, gène codant pour la protéine associée à la réplication ; *sd*, gène codant pour un déterminant de symptôme ; *ss*, gène codant pour un suppresseur de *silencing* ; *trap*, gène codant pour une protéine activatrice de la transcription. Les génomes des membres représentatifs des derniers genres décrits sont indiqués par des astérisques (Roumagnac *et al.*, 2022).

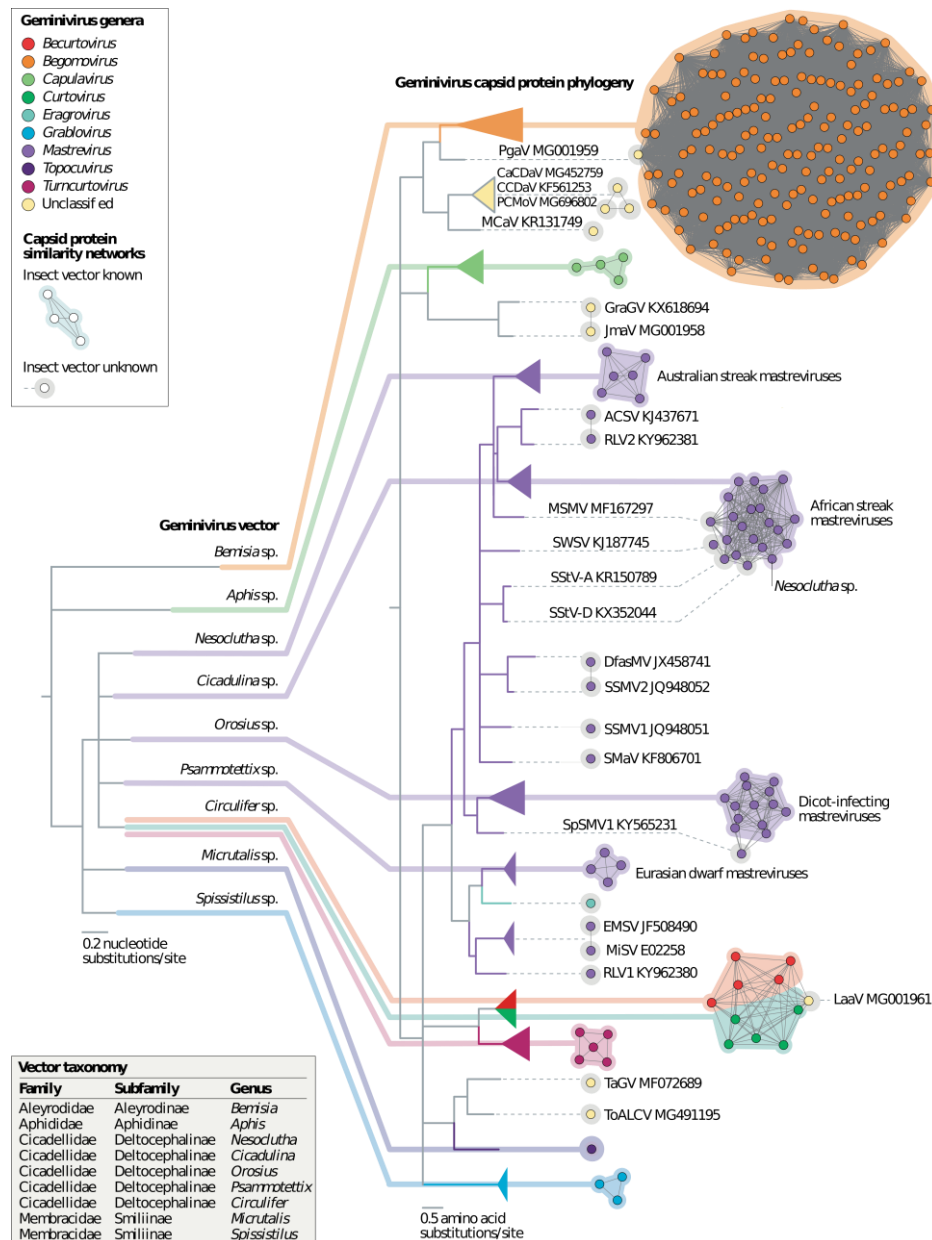


Figure 18 : Arbres phylogénétiques des séquences protéiques de capside des différents genres de *Geminiviridae* et celle de la cytochrome oxydase I de leurs insectes vecteurs. Les branches soutenues avec un ratio de vraisemblance inférieur à 0.8 ont été collapsées. Les réseaux de similarité représentent les différents groupes de séquences protéiques de capside de *Geminiviridae* et sont colorés par genre viral. ACSV, *Axonopus compressus* streak virus ; CaCDaV, *Camellia chlorotic dwarf-associated virus*; CCDaV, *Camellia citrus chlorotic dwarf-associated virus*; DfasMV, dragonfly-associated mastrevirus; EMSV, *Eragrostis minor* streak virus; GraGV, grapevine geminivirus; JmaV, *Juncus maritimus*-associated virus; LaaV, *Limeum africanum*-associated virus; MCaV, mulberry crinkle-associated virus; MiSV, *Miscanthus* streak virus; MSMV, maize streak Reunion virus; PCMoV, passion fruit chlorotic mottle virus; PgaV, *Polygala garcinii*-associated virus; RLV1, rice latent virus 1; RLV2, rice latent virus 2; SMaV, switchgrass mosaic-associated virus; SpSMV1, sweetpotato symptomless mastrevirus 1; SSMV1, *Sporobolus striate* mosaic virus 1; SSMV2, *Sporobolus striate* mosaic virus 2; SStV-A, sugarcane striate virus A; SStV-D, sugarcane striate virus D; SWSV, sugarcane white streak virus; TaGV, tomato-associated geminivirus; ToALCV, tomato apical leaf curl virus. (Lefeuvre et al., 2019).

6. Les mastrévirus

6.1. Généralités

Le génome des mastrévirus est composé d'un seul composant d'ADNsb circulaire de 2,6 à 2,8 kb qui code pour quatre protéines. Les deux protéines codées dans le sens viral sont la protéine de capsid (CP, gène V1) qui encapside l'ADNsb et agit comme une protéine de navette nucléaire (NSP) pour l'ADN viral, et la protéine de mouvement (MP, gène V2), qui permet le déplacement de cellule en cellule. Dans le sens complémentaire sont codées les protéines associées à la réplication (Rep et RepA), exprimées à partir des gènes C1 pour la RepA et l'épissage des deux ARN transcrits des ORF C1 et C2 pour la Rep. La protéine Rep initie la réplication en cercle roulant en introduisant une coupure dans la séquence conservée 5'-TAATA/GTT↓AC-3' dans le sens viral (Dekker *et al.*, 1991 ; Mullineaux *et al.*, 1990 ; Schalk *et al.*, 1989 ; Wright *et al.*, 1997). Ces quatre ORFs sont nécessaires pour l'infection systémique (Liu *et al.*, 1998). Deux régions intergéniques séparent les deux paires d'ORFs, une grande appelée *Large Intergenic Region* (LIR) qui porte l'origine de réplication et le promoteur bidirectionnel et une petite appelée *Small Intergenic Region* (SIR) qui contient les deux signaux de terminaison des transcrits.

Les mastrévirus infectent des plantes monocotylédones (majoritairement) ou dicotylédones de par le monde (Shepherd *et al.*, 2010 ; Kraberger *et al.*, 2013 ; Kraberger *et al.*, 2017 ; Claverie *et al.*, 2019 ; Agindotan *et al.*, 2015 ; Lappe *et al.*, 2022). Parmi les familles de plantes hôtes des mastrévirus, la famille des *Poaceae* (monocotylédones) représente l'une des familles les plus largement étudiées dans le cadre des infections virales en raison de leur importance économique et dans la sécurité alimentaire (O'Mara, 2012 ; Campbell, 2020).

Les espèces du genre *Mastrevirus* sont transmis par des cicadelles de la famille des *Cicadellidae*. Le MSV est transmis par des cicadelles du genre *Cicadulina*, regroupant 22 espèces parmi lesquelles 18 espèces sont présentes en Afrique et 14 en sont endémiques (Webb, 1987). Neuf de ces espèces africaines ont été identifiées comme étant capables de transmettre le MSV (Bosque-Pérez, 2000) mais l'espèce la plus répandue en Afrique, *Cicadulina mbila* (**Figure 19**), semble être un des vecteurs le plus efficace (Storey, 1928 ; Reynaud 1988).



Figure 19 : Vue latérale d'un adulte et d'une larve de *Cicadulina mbila* (Crédit Antoine Franck – Cirad).

Parmi les 45 espèces du genre, les mieux étudiées sont le maïs streak virus (MSV ; *Mastrevirus storeyi*), faisant des ravages sur le maïs en Afrique (Shepherd *et al.*, 2010), et le wheat dwarf virus (WDV ; *Mastrevirus hordei*), dont les membres causent une maladie du blé dans le nord de l'Europe, le Moyen-Orient et l'Asie (Kvarnheden *et al.*, 2002). Ils infectent des céréales tropicales et tempérées ainsi que certaines cultures maraîchères. Les symptômes induits par les mastrevirus infectant les monocotylédones sont majoritairement des chloroses sous forme de striures et/ou de motifs en mosaïque, parfois associées au nanisme des plantes. De nombreux mastrevirus infectant les céréales cultivées, infectent également des espèces de *Poaceae* sauvages, qui sont souvent considérées comme les hôtes naturels de ces virus.

6.2. La diversité des African streak virus

Parmi les 45 espèces de mastrevirus décrites à ce jour, 36 ont été décrites sur poacées (www.ictv/global/taxonomy), dont 19, décrites en Afrique et dans les îles du sud-ouest de l'océan Indien (SOOI) et appelés les African streak viruses (AfSVs), forment un groupe monophylétique. Si ces virus ont été principalement identifiés sur maïs et canne à sucre, il est toutefois important de rappeler d'abord, que ces deux hôtes ne représentent très probablement pas leur hôte de prédilection.

6.2.1. Les AfSVs du maïs

Au sein des AfSVs, le MSV a été le plus largement étudié. Cela s'explique en grande partie par l'impact dévastateur de ce virus sur le maïs, responsable de la maladie de la striure, une des plus importantes viroses du maïs en Afrique subsaharienne (Shepherd *et al.*, 2010). Le maïs a été introduit sur le continent africain au XVI^e siècle où il est devenu une culture de

base. La striure du maïs a été décrite pour la première fois en Afrique du Sud par Claude Fuller en 1901, mais ce n'est qu'en 1925 que l'infection virale et la transmission par des cicadelles ont été démontrées par Storey (pour revue Shepherd *et al.*, 2010). Le MSV infecte non seulement le maïs, mais aussi la canne à sucre, le millet, l'orge, le blé, l'avoine et le seigle ainsi qu'une large gamme d'espèces de *Poaceae* sauvages (Shepherd *et al.*, 2010). A partir du seuil d'identité nucléotidique de 94 % (Muhire *et al.*, 2013), un total de onze souches de MSV a été décrite, du MSV-A au MSV-K, infectant un large éventail d'espèces de *Poaceae* mais seule la souche A est connue pour causer des symptômes sévères sur le maïs (**Figure 20** ; Monjane *et al.*, 2011 ; Oluwafemi *et al.*, 2011 ; Shepherd *et al.*, 2010 ; Varsani *et al.*, 2008). De manière intéressante, le MSV-A adapté au maïs, serait issu de la recombinaison entre des variants ancestraux du MSV-B et MSV-G/-F (Varsani *et al.*, 2008). Une analyse de datation de cet événement indique qu'il s'est probablement produit au milieu du XIXe siècle, environ 20 ans avant les premiers rapports documentés de la maladie. La large gamme de plantes hôtes du recombinant serait associé à sa diffusion rapide à travers le continent africain (Monjane *et al.*, 2011). En plus du MSV, d'autres mastrevirus ont été identifiés sur le maïs, tels que le maize streak Reunion virus (MSRV ; *Mastrevirus bourbonense* ; Pande *et al.*, 2012), le maize streak dwarfing virus (MSDV ; *Mastrevirus fulleri* ; Guadie *et al.*, 2019), le Panicum streak virus (PanSV ; *Mastrevirus panici* ; Varsani *et al.*, 2008) et le North American maize-associated mastrevirus (NAMaMV ; Lappe *et al.*, 2022). Le MSRV a été initialement découvert sur des plants de maïs symptomatiques à La Réunion (Pande *et al.*, 2012) et a depuis été trouvé en Chine et en Éthiopie (Chen *et al.*, 2015 ; Guadie *et al.*, 2019) et sur des poacées sauvages (*Setaria barbata* et *Rottboellia sp.*) au Nigeria (Oluwafemi *et al.*, 2014). Le MSDV a été récemment identifié en Éthiopie sur des plants de maïs atteint de striure et de nanisme (Guadie *et al.*, 2019). Le PanSV, initialement décrit sur des espèces de plantes de la tribu des *Paniceae* venant du continent africain (Varsani *et al.*, 2009) a été décrit récemment sur des plants de maïs en Afrique du Sud (Krabberger *et al.*, 2017). Le NAMaMV est la première espèce de *Mastrevirus* identifiée sur maïs en Amérique du Nord (Mexique ; Lappe *et al.*, 2022). Le PanSV, également isolé sur l'orge et des *Poaceae* sauvages (Krabberger *et al.*, 2017), est la deuxième espèce la mieux caractérisée après le MSV avec dix souches identifiées (PanSV-A à PanSV-J). Cette espèce présente un niveau de diversité similaire au MSV (Varsani *et al.*, 2008 ; Varsani *et al.*, 2009) et est également transmis par la cicadelle *Cicadulina mbila* (Briddon *et al.*, 1992). Le PanSV

a été identifié à ce jour sur environ 17 espèces de poacées et a été décrit au sein de sept pays du continent africain ainsi qu'à Mayotte et à La Réunion.



Figure 20 : Symptômes de la maladie de la striure du maïs (*maize streak disease*) causées par le MSV (Shepherd *et al.*, 2010)

6.2.2. Les AfSVs de la canne à sucre

Six espèces de mastrevirus ont été décrites sur la canne à sucre : Saccharum streak virus (SacSV ; *Mastrevirus sachari* ; Lawry *et al.*, 2009), sugarcane chlorotic streak virus (SCSV ; *Mastrevirus saccharumpallidi* ; Yahaya *et al.*, 2017), sugarcane streak Egypt virus (SSEV ; *Mastrevirus saccharumegyptense* ; Bigarré *et al.*, 1999), sugarcane streak Reunion virus (SSRV ; *Mastrevirus saccharumbourbonense* ; Shepherd *et al.*, 2008), sugarcane streak virus (SSV ; *Mastrevirus saccharofficinari* ; Hughes *et al.*, 1992) et sugarcane white streak virus (SWSV ; *Mastrevirus saccharumalbusvenae* ; Candresse *et al.*, 2014). Ces espèces ont été isolées en Afrique et dans les îles de l'océan Indien et sont principalement associées à la canne à sucre. Certaines de ces espèces ont également été décrites sur des poacées sauvages comme le SSRV à La Réunion, au Zimbabwe et au Nigeria sur trois espèces sauvages, *Eleusine coracana*, *Paspalum conjugatum* et *Setaria barbata* ou le SSV sur *Cenchrus echinatus* à La Réunion (Oluwafemi *et al.*, 2014 ; Shepherd *et al.*, 2008).

6.2.3. Les AfSVs des poacées non cultivées

Huit autres espèces d'AfSVs ont été identifiées sur des plantes exclusivement sauvages (Kraberger *et al.*, 2017 ; Martin *et al.*, 2011 ; Oluwafemi *et al.*, 2014 ; Shepherd *et al.*, 2008), notamment l'*Axonopus compressus* streak virus (ACSV ; *Mastrevirus axonopi* ; Oluwafemi *et al.*, 2014), l'*Eleusine indica* associated virus (EIAV ; *Mastrevirus eleusinae* ; Claverie *et*

INTRODUCTION

al., 2019), l'Eragrostis minor streak virus (EMSV ; *Mastrevirus eragrostisminoris* ; Martin *et al.*, 2011), l'Eragrostis streak virus (ESV ; *Mastrevirus eragrostis* ; Shepherd *et al.*, 2008), le Melinis repens associated virus (MeRAV ; *Mastrevirus melinis* ; Claverie *et al.*, 2019), le Sorghum arundinaceum associated virus (SAAV ; *Mastrevirus sorghi* ; Claverie *et al.*, 2019) et l'Urochloa streak virus (USV ; *Mastrevirus urochloae* ; Oluwafemi *et al.*, 2008).

7. Contexte et objectifs de la thèse

Au regard des connaissances actuelles, il apparaît que les interactions réalisées par les phytovirus avec leurs hôtes, les autres phytovirus, et éventuellement leurs vecteurs sont complexes et subtils. L'élargissement du champ d'investigation de la virologie à l'échelle d'écosystèmes entiers apporte un degré supplémentaire de complexité. Ainsi, la connaissance de l'ensemble des interactions hôte-virus réalisées dans les différents écosystèmes en constante évolution semble nécessaire pour comprendre les patrons de diversité viraux observés et prédire leurs évolutions face aux changements. Cependant, de nombreuses zones d'ombre persistent sur les spécificités de l'écologie virale.

L'île volcanique de La Réunion, située dans l'archipel des Mascareignes fait partie du point chaud de biodiversité du Sud-Ouest de l'Océan Indien (SOOI) (Myers *et al.*, 2000 ; Mittermeier *et al.*, 2011) où 30 % de la flore (indigène et endémique) est en voie de disparition. La principale cause de cette perte en diversité est l'installation et l'extension de parcelles agricoles qui couvrent la majeure partie des surfaces anthropisées et se retrouvent en contact direct avec les limites des espaces naturels (Fenouillas *et al.*, 2021). Récemment, Claverie *et al.* (2023) ont réalisé un inventaire de la richesse et une analyse de la structure des communautés des espèces du genre *Mastrevirus* (famille des *Geminiviridae*), à l'échelle d'un de ces agro-écosystèmes réunionnais. Cette étude a démontré l'existence d'une diversité élevée de mastrevirus à l'échelle d'un site unique. Elle a aussi mis en lumière le caractère généraliste de la plupart des espèces de *Mastrevirus*, infectant majoritairement des plantes non cultivées, annuelles et indigènes de La Réunion (Claverie *et al.*, 2023). Toutefois, ces résultats interrogent sur la diversité potentielle de ces virus au-delà des agro-écosystèmes, sur l'impact des perturbations du milieu sur la structure du réseau d'interactions réalisées entre virus et hôtes et sur la contribution des invasions biologiques à la diffusion des virus et inversement.

Pour répondre à ces interrogations, nous nous sommes proposés d'étudier et de comprendre l'écologie des phytovirus à travers le prisme des hôtes et de différents écosystèmes de La Réunion.

INTRODUCTION

L'objectif principal de cette thèse a été d'étendre le travail précurseur de Claverie *et al.* (2023) et d'étudier dans un premier temps (i) la richesse qui compose les communautés de mastrévirus infectant les monocotylédones en explorant les gammes d'hôtes potentielles à l'échelle de l'ordre des Poales (comprenant les *Poaceae*, les *Cyperaceae* et les *Juncaceae*), dans trois types d'écosystèmes réunionnais : des agro-écosystèmes, des prairies littorales et des prairies altimontaines. Puis, (ii) en analysant la structure des réseaux d'interactions obtenus, nous nous sommes intéressés aux facteurs associés aux hôtes et aux écosystèmes qui influencent la richesse et la structure de ces communautés virales révélées. Enfin, (iii) nous avons exploré la diversité de l'ensemble des phytovirus présents le long d'un gradient altitudinal de La Réunion au regard du statut indigène ou exotique envahissant de leurs hôtes.

Pour se faire, le premier chapitre de la thèse a été consacré au développement et à la mise au point de nouvelles méthodologies de séquençage haut débit de troisième génération permettant de détecter les virus à ADN circulaires, dont les mastrévirus font partie. Les résultats obtenus ont été comparés à ceux obtenus par l'application d'un protocole classique de clonage - séquençage Sanger sur deux variants d'une même espèce de *Capulavirus* (famille des *Geminiviridae*) présents dans un échantillon de luzerne (*Medicago arborea*) asymptomatique.

Sur la base de la validation de ces protocoles, les travaux présentés dans le deuxième chapitre de la thèse ont visé à déterminer la richesse et la répartition des communautés de *Mastrevirus*, puis à analyser les facteurs (hôtes / écosystèmes) ayant pu façonner la structure des espèces de ce genre viral circulant dans les différents écosystèmes de La Réunion. Ces travaux ont permis d'approfondir les connaissances sur les relations entre la richesse des espèces du phylum *Cressdnaviricota* et les gammes d'hôtes. Le lien étroit mis en lumière entre mastrévirus et plantes cultivées et adventices des cultures dans les agro-écosystèmes soulève la question des voies d'introduction privilégiées de ces virus dans un environnement insulaire océanique.

Enfin, dans le troisième chapitre de la thèse, nous avons choisi d'étendre l'étude précédente à l'ensemble des phytovirus décrits sur les espèces herbacées indigènes et exotiques les

INTRODUCTION

plus répandues et appartenant aux familles les plus communes le long d'un transect altitudinal qui présentait un gradient d'invasion et de perturbation des écosystèmes. Ainsi, grâce à des approches métagénomiques holistiques, nous avons cherché à compléter les connaissances de la diversité virale circulant à La Réunion et à étudier les facteurs influençant leurs diversités et leurs distributions dans le contexte des invasions végétales. Alors que certaines espèces virales semblent être associées aux espèces végétales indigènes, d'autres, connues comme des pathogènes associés aux cultures, ont été détectées sur des espèces non cultivées en milieu sauvage, ce qui soulève la question d'un impact potentiel de ces virus sur le succès des invasions végétales.

Chapitre 1

Chapitre 1

Mise au point et optimisation d'un protocole de séquençage haut débit pour la détection des virus à ADN simple brin circulaire

En 2004, la publication d'un protocole d'amplification et de clonage du génome de bégomovirus reposant sur l'utilisation de la polymérase *phi29* par Inoue-Nagata *et al.* (2004) a été à l'origine d'une première révolution dans l'étude des virus à ADN circulaire de petite taille. Présentant le grand avantage de ne pas nécessiter de connaissances préalables des séquences virales à amplifier, le protocole RCA-RFLP basé sur l'amplification en cercle roulant (*rolling circle amplification* ou RCA), la restriction enzymatique (*restriction fragment length polymorphism* ou RFLP), le clonage et le séquençage Sanger est rapidement devenu le protocole standard pour l'étude du génome des virus à ADN circulaire.

Alors que l'application de ce protocole a largement contribué à l'établissement des *Cressdnaviricota* comme un groupe viral d'intérêt majeur (Rosario *et al.*, 2012), l'utilisation de méthodes de séquençage à haut débit dites de seconde génération (*second generation sequencing* ou SGS) a été plus récemment à l'origine d'un bouleversement total de notre connaissance de la diversité des virus en général. En effet, les découvertes associées à l'utilisation d'approches de séquençage massif ont mené à un changement radical de la perception par la communauté scientifique de l'importance des virus dans le vivant.

Afin de combiner les avantages de ces deux approches, des protocoles associant RCA et SGS ont rapidement émergé (Rosario *et al.*, 2015 ; Jeske, 2018). Si de nombreuses découvertes ont suivies (Delwart and Li, 2012 ; Tisza *et al.*, 2020 ; Crespo-Bellido and Duffy, 2023), des limitations de natures très différentes restent encore associées aux protocoles utilisés. Tout d'abord, une première limitation réside dans la complexité du multiplexage d'échantillons (*i.e.* mélange de plusieurs échantillons) dans une même bibliothèque de séquençage. En effet, la profondeur de séquençage des plateformes actuelles est adaptée au séquençage d'organisme dont le génome est de plusieurs ordre

de grandeur plus important que ceux des CRESS DNA virus. Ensuite, la faible longueur des séquences brutes obtenues entraîne à la fois des difficultés d'assemblage des génomes complets et d'assignation taxonomique. Enfin, la complexité de la mise en œuvre du séquençage et le coût élevé des appareillages requis rendent la méthode peu accessible en routine pour la majorité des laboratoires.

Les méthodes de séquençage *long reads* de troisième génération (*third generation sequencing* ou TGS) avec en particulier le dispositif portable MinION (Oxford Nanopore Technologies ; ONT) sont apparues comme une alternative à fort potentiel. Ainsi, cette méthode a été appliquée avec succès pour détecter et séquencer les génomes quasi-complets de plusieurs espèces de phytovirus à ARN dans le cadre de l'étude du virome de plantes (Chalupowicz *et al.*, 2019) ou d'énigmes étiologiques (Adams *et al.*, 2017 ; Badial *et al.*, 2018 ; Filloux *et al.*, 2018).

Dans le but d'étendre cette méthodologie aux virus à ADN circulaire, nous avons évalué la performance d'approches TGS en comparaison d'un protocole plus classique de RCA-RFLP et de séquençage Sanger. A partir d'un même échantillon de luzerne (*Medicago arborea*) asymptomatique mais infecté par deux souches de *Trifolium virus 1* (TrV1-B et TrV1-C ; genre *Capulavirus* ; Ma *et al.*, 2021), le premier chapitre de cette thèse a consisté à la mise au point et à l'évaluation de deux protocoles de séquençage *long reads* sur MinION adaptés aux virus CRESS DNA.

Le premier protocole, nommé « RCA-MinION » repose sur une approche sans *a priori* (absence de connaissances préalables du génome viral) initiée par une étape RCA. Le second protocole, nommé « PCR-MinION », repose sur une étape de PCR avec l'utilisation d'amorces spécifiques pour cibler les matrices virales recherchées. Les matrices amplifiées (*i.e.* produits RCA et PCR) ont ensuite été utilisées pour la construction de bibliothèques avant séquençage sur MinION. Des *pipelines* bioinformatiques (*i.e.* chaîne de traitement informatisée) spécifiques ont été mis au point afin de filtrer et assembler les séquences obtenues. Les génomes assemblés par les deux méthodologies ont ensuite été comparés avec ceux obtenus en utilisant la méthode classique de RCA-RFLP et de séquençage Sanger. Cette comparaison a permis de montrer que les

approches de séquençage sur MinION permettaient l'assemblage de séquences nucléotidiques très proches du consensus des séquences Sanger. Seules des erreurs sur les régions homopolymères (*i.e.* répétition de nucléotides identiques) ont été constatées, confirmant un manque de précision spécifique pour ces zones avec les approches MinION (McNaughton *et al.*, 2019 ; Viehweger *et al.*, 2019). Globalement, la qualité des génomes assemblés, la facilité de multiplexage de plusieurs échantillons lors du séquençage et la rapidité de mise en œuvre de ces approches de séquençage sur MinION, font de celles-ci des approches convaincantes pour la caractérisation des génomes de virus à ADN circulaire de petites tailles. Si le protocole devra être adapté aux évolutions à venir des TGS, son utilisation apportera une plus grande flexibilité pour les études de métagénomique des virus de l'ordre des *Cressdnaviricota*.

Article 1

**Nanopore sequencing is a credible alternative to recover
complete genomes of geminiviruses**



Article

Nanopore Sequencing Is a Credible Alternative to Recover Complete Genomes of Geminiviruses

Selim Ben Chehida ¹, Denis Filloux ^{2,3}, Emmanuel Fernandez ^{2,3}, Oumaima Moubset ^{2,3}, Murielle Hoareau ¹, Charlotte Julian ^{2,3}, Laurence Blondin ^{2,3}, Jean-Michel Lett ¹, Philippe Roumagnac ^{2,3} and Pierre Lefeuvre ^{1,*}

¹ CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France; selim.ben_chehida@cirad.fr (S.B.C.); murielle.hoareau@cirad.fr (M.H.); jean-michel.lett@cirad.fr (J.-M.L.)

² CIRAD, PHIM, F-34398 Montpellier, France; Denis.Filloux@Cirad.Fr (D.F.); emmanuel.fernandez@cirad.fr (E.F.); oumaima.moubset@cirad.fr (O.M.); charlotte.julian@cirad.fr (C.J.); laurence.blondin@cirad.fr (L.B.); philippe.roumagnac@cirad.fr (P.R.)

³ PHIM Plant Health Institute, University Montpellier, CIRAD, INRAE, Institut Agro, IRD, F-34398 Montpellier, France

* Correspondence: pierre.lefeuvre@cirad.fr

Abstract: Next-generation sequencing (NGS), through the implementation of metagenomic protocols, has led to the discovery of thousands of new viruses in the last decade. Nevertheless, these protocols are still laborious and costly to implement, and the technique has not yet become routine for everyday virus characterization. Within the context of CRESS DNA virus studies, we implemented two alternative long-read NGS protocols, one that is agnostic to the sequence (without a priori knowledge of the viral genome) and the other that use specific primers to target a virus (with a priori). Agnostic and specific long read NGS-based assembled genomes of two capulavirus strains were compared to those obtained using the gold standard technique of Sanger sequencing. Both protocols allowed the detection and accurate full genome characterization of both strains. Globally, the assembled genomes were very similar (99.5–99.7% identity) to the Sanger sequences consensus, but differences in the homopolymeric tracks of these sequences indicated a specific lack of accuracy of the long reads NGS approach that has yet to be improved. Nevertheless, the use of the bench-top sequencer has proven to be a credible alternative in the context of CRESS DNA virus study and could offer a new range of applications not previously accessible.

Keywords: MinION; nanopore sequencing; rolling circle amplification; viral metagenomics; CRESS DNA; capulavirus; homopolymer



Citation: Ben Chehida, S.; Filloux, D.; Fernandez, E.; Moubset, O.; Hoareau, M.; Julian, C.; Blondin, L.; Lett, J.-M.; Roumagnac, P.; Lefeuvre, P. Nanopore Sequencing Is a Credible Alternative to Recover Complete Genomes of Geminiviruses. *Microorganisms* **2021**, *9*, 903. <https://doi.org/10.3390/microorganisms9050903>

Academic Editor: Fred Asiegbu

Received: 24 March 2021

Accepted: 21 April 2021

Published: 23 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recent advances in metagenomics applied to viruses has fostered a greater inventory of the viral diversity [1–4]. Hence, the large scale sampling of oceanic water [5,6], plants, animals, and humans [7–10], extreme environments [11], or the mining of genomic and transcriptomic data [12–15] have completely shifted our understanding of viral diversity and the function of viruses in host populations or even at the global ecosystem scale. However, these inventories remain largely incomplete, and the current knowledge of the virus diversity probably only represents the contour of the extant diversity [4]. The collection of hundreds of new genome sequences with, sometimes only remote resemblance to known viruses led to the acceptance of genome from metagenomic studies as genuine and legitimate taxonomic material for the description of new viruses [16–21], even without knowledge of the phenotype or the host associated to these viruses [17,22].

Yet, as access to metagenomics is usually costly and requires sophisticated technical expertise in data management and analysis [23,24]; for day-to-day analysis, classical Sanger sequencing remains more common [25]. Also, despite the potential of third generation sequencing technique to provide, in real time, hundreds of sequences with read length

of more than 15 kb [26,27], the resulting assembled genomes remains tainted with doubt. Indeed, assembly of the low base quality reads is required to obtain more accurate contigs but may result in chimeric genomes, or miss less frequent variants [17,28,29]. It would thus be beneficial to evaluate the use of third generation sequencing as an alternative to the gold standard of cloning and Sanger sequencing for everyday virus characterization. Indeed, several studies have successfully employed third-generation sequencing for virus detection or full genome recovery. It has been largely used to detect and sequence full genomes of a range of animal and human viruses, including influenza [30], Ebola [31], Dengue [32], Zika [33], or SRAS-Cov-2 [34]. Recently the method was also successfully applied to plant viruses for some yam infecting viruses [29], maize yellow mosaic virus [35] and plum pox virus [36].

Here, within the context of the study of circular replication-associated protein encoding single stranded (CRESS) DNA viruses (from the *Cressdnaviricota* phylum), we harness the power of third-generation sequencing for routine full genome assembly of viruses. A breakthrough in CRESS DNA virus studies was associated with the development of protocols using isothermal rolling circle amplification (RCA) in frequent association with enzymatic restriction, cloning and Sanger sequencing [37]. Metagenomic protocols using RCA were soon developed and greatly improved our knowledge of the CRESS DNA virus diversity [38,39]. Yet, the use of Illumina based protocols remains laborious and expensive for day-to-day viral genome characterization and sequences analysis usually requires the use of complex bioinformatics tools. The field would thus benefit from the development of a more convenient protocol on the MinION platform from the Oxford Nanopore Technologies (ONT). Indeed, recent studies have paved the way towards using nanopore sequencing, either from direct sequencing of total DNA extracts or after the application of RCA, to achieve full genome sequencing of CRESS-DNA viruses in general [40–42] or viruses from the *Geminiviridae* family in particular [43–45].

Two alternative protocols for use on the MinION bench-top sequencing device were designed in this study and the assemblies of the nanopore-sequenced reads of multiple strains of capulaviruses infecting a *Medicago arborea* plant [46] were compared to the Sanger genome sequences. *Capulavirus* is one of the nine genera of the *Geminiviridae* family [47]. This family is composed of plant-infecting viruses with genomes comprising one or two circular ssDNA of 2.5–5.2 kb encapsidated in twinned icosahedral (geminata) particles. They are transmitted by a high range of hemipterans (whiteflies, leafhoppers, aphids, and treehoppers) [48,49]. Whereas members of the *Geminivirus* family were first described in 1993 (ICTV), the standardization of RCA based protocols led to an explosion of its known diversity and the family regularly counts new genus-level lineage addition [47,50–52]. Following the description of the *Euphorbia caput-medusae* latent virus (EcmLV), the genus capulavirus has been proposed [53]. Their genome length ranged between 2.7 and 2.8 kb with two intergenic regions. The replication-associated protein (Rep) is expressed from a spliced complementary-strand transcript. A unique feature of capulavirus genomes is a complex arrangement of possible MP-encoding ORFs located in the 5' direction from the coat protein gene (*cp*) [47]. It is known to be transmitted by only two species of aphids: *Aphis Craccivora* [49] and *Dysaphis plantaginea* [54].

Our analysis revealed that MinION sequencing followed by read assembly results in genome sequences mostly similar to the consensus of the virus population that was obtained after Sanger sequencing of multiple isolates. The two alternate protocols used, one that does not require knowledge of the viruses present within the sample (without a priori) and the other designed to specifically amplify a given virus (with a priori), were successful for full genome assembly of the two capulavirus strains. MinION assembled consensus sequences present with more differences in the homopolymeric tracks but remain closely related to any sequence of the sample than any Sanger sequence to one another. Overall, both nanopore-based protocols are adapted to the genome size of the CRESS DNA viruses and the cost of the method is on par with the Sanger sequencing approach.

2. Material and Methods

2.1. Sampling and DNA Extraction

Leaf samples of an apparently asymptomatic *Medicago arborea* were collected in November 2019 at Montferrier-sur-Lez (France). The sample was stored at $-20\text{ }^{\circ}\text{C}$ before use. Total DNA was extracted using the DNeasy Plant DNA extraction kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. DNA extract was stored at $-20\text{ }^{\circ}\text{C}$ before use. From a previous analysis [46], using a PCR amplification and Sanger sequencing, two strains of the capulavirus Trifolium virus 1 (TrV1-B and TrV1-C) were identified into the sample.

2.2. Full Genome Cloning and Sanger Sequencing

Pairs of abutting primers were designed to recover the full-length genome of TrV1-B and TrV1-C isolates. A two-step amplification was achieved, including a first amplification step using either the primer pair 3580F-CAPULUZARB-1F: 5'-ACT TGC CTG TCG CTC TAT CTT CTC CCT TGG AGA TGT AAT CTG CCA CGT CAG-3', and PR2-CAPULUZARB-2R: 5'-TTT CTG TTG GTG CTG ATA TTG CGG AGT TTT TGA GGA ACG AGG AAT ACT TAG AGC TTC A-3' for amplifying TrV1-B genomes or the primer pair 3580F-CAPUCORO-1F: 5'-ACT TGC CTG TCG CTC TAT CTT CAA CTG TCC TCC CTT TGC AAT GTA GTC AGC C-3' and PR2-CAPUCORO-2R: 5'-TTT CTG TTG GTG CTG ATA TTG CCG AGG AGC GAG GAC TTC TTA AGG CAA GT-3' for amplifying TrV1-C genomes. Amplification was carried out using the GoTaq[®] Master Mix Kit (Promega Corporation, Madison, WI, USA) and the following conditions: an initial denaturation at $95\text{ }^{\circ}\text{C}$ for 5 min, 8 cycles at $94\text{ }^{\circ}\text{C}$ for 30 s, $60\text{ }^{\circ}\text{C}$ for 30 s, $72\text{ }^{\circ}\text{C}$ for 3 min, and a final extension step at $72\text{ }^{\circ}\text{C}$ for 10 min. A common second amplification step was then performed using the primer pair (3580F: 5'-ACT TGC CTG TCG CTC TAT CTT C-3' and PR2: 5'-TTT CTG TTG GTG CTG ATA TTG C-3') and the GoTaq[®] Master Mix Kit (Promega Corporation, Madison, WI, USA). The amplification conditions were as followed: an initial denaturation at $95\text{ }^{\circ}\text{C}$ for 5 min, 25 cycles at $94\text{ }^{\circ}\text{C}$ for 30 s, $60\text{ }^{\circ}\text{C}$ for 30 s, $72\text{ }^{\circ}\text{C}$ for 3 min, and a final extension step at $72\text{ }^{\circ}\text{C}$ for 10 min. Amplification products of approximately 2.7–2.8 kb were gel purified, ligated to pGEM-T (Promega, Madison, WI, USA) and sequenced by standard Sanger sequencing using a primer walking approach.

2.3. Minion Sequencing

Two alternative protocols for MinION sequencing were used. In the first protocol, called hereafter the RCA-MinION protocol, a RCA amplification was performed using *Phi29* DNA polymerase (Illustra TempliPhi Amplification Kit, GE Healthcare, Chicago, IL, USA) by mixing 2 μL of total plant DNA extract with 5 μL of Sample Buffer before incubation at $95\text{ }^{\circ}\text{C}$ during 3 min. After cooling at room temperature, 0.2 μL of enzyme mix and 5 μL of Reaction Buffer were added before incubation at $30\text{ }^{\circ}\text{C}$ for 6 h followed by 20 min of polymerase deactivation at $65\text{ }^{\circ}\text{C}$. RCA products were cleaned-up using 2 \times of Sera-Mag Select Size Selection beads (GE Healthcare) and the 10 μL eluate were digested with 1 μL of T7 Endonuclease I (NEB), 2 μL of 5 \times buffer in a 10 μL reaction volume at $37\text{ }^{\circ}\text{C}$ during 1 h. The fragments were purified with a 1 \times Sera-Mag Select Size Selection beads and eluate with 10 μL of purified water. Library construction for MinION sequencing was performed using the PCR Barcoding Kit (SQK-PBK004), following the manufacturer's instructions but using SeraMag Select Size Selection beads (GE Healthcare) for DNA purification. Sequencing was performed as described below for the PCR-MinION procedure. Two Flongles (flow cell dongle) FLO-FLG001 were used for sequencing. Whereas in the first Flongle, a single RCA amplicon was sequenced, in the second, three distinct RCA amplicons were multiplexed.

In the second protocol, called hereafter the PCR-MinION protocol, a two-step amplification was carried out. In the first PCR, both sets of abutting primers (3580F-CAPULUZARB-1F/PR2-CAPULUZARB-2R and 3580F-CAPUCORO-1F/PR2-CAPUCORO-2R) described above for respectively amplifying the genomes of strains TrV1-B and TrV1-C and the GoTaq[®] Master Mix Kit (Promega Corporation, Madison WI, USA) were employed. The amplification conditions were as follows: an initial denaturation at 95 °C for 5 min, 15 cycles at 94 °C for 30 s, 60 °C for 30 s, 72 °C for 3 min, and a final extension step at 72 °C for 10 min. The second amplification step was performed using the cDNA Primer (cPRM) supplied in the PCR-cDNA Sequencing Kit (SQK-PCS109) and the LongAmp Taq 2X Master Mix Kit (New England Biolabs, Evry, France). The amplification conditions were as followed: an initial denaturation at 95 °C for 30 s, 20 cycles at 95 °C for 15 s, 62 °C for 15 s, 65 °C for 3 min, and a final extension step at 65 °C for 6 min. The amplicons were purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA) and MinION sequencing library was constructed using the PCR-cDNA Sequencing Kit (SQK-PCS109), following manufacturer's instructions. Sequencing was then performed on Flongle (FLO-FLG001) using MinKNOW software 19.06.8. Three Flongles were used, two for TrV1-B (Flongle 13 and Flongle 14) and another one for TrV1-C (Flongle 15).

2.4. MinION Sequence Assembly

For the reads obtained through the two MinION protocols, accurate basecalling was performed using Guppy (v4.09 or 4.015; [55]). Demultiplexing and adapter removal was then performed using Porechop v0.2.4 [56]. Quality of the reads was investigated using NanoPlot v1.33.0 [57].

The demultiplexed reads obtained from RCA-MinION procedure were assembled with FLYE 2.6 [58], using the "meta" and "plasmid" parameters and when possible circularized as monomers using a homemade script. Contigs were then subjected to a BLASTn search against a CRESS DNA reference sequence database obtained from GenBank. CRESS DNA contigs were then polished using Medaka v1.2.2 [59]. PCR-MinION sequences higher than 1500 nt for one run (TrV1-C) and 2000 nt for the two other run (TrV1-B), were assembled using Canu v1.8 [60]. CRESS DNA contigs were filtered using BLAST as described above. Contigs coverage was estimated after mapping the raw reads back to the assembled CRESS DNA sequences using Minimap2 v2.17 [61] (Figure 1).

2.5. Sequence Comparison and Phylogenetic Analysis

All the available capulavirus full genome sequences were downloaded from GenBank on 4 March 2021 and aligned with the sequences of strains TrV1-B and TrV1-C obtained after Sanger sequencing using MAFFT v7.475 [62]. A maximum likelihood (ML) phylogenetic tree was inferred using FastTree2 [63] using the "gtr" and "gamma" parameters. Branch supports were tested using SH-like local supports. Tree edition was performed using the ape R package [64]. In order to properly classify the sequences obtained, an analysis that include a subset of representative capulavirus from GenBank and the Sanger sequence obtained in this study was performed using SDT1.2 [65].

Sequences obtained using the three distinct protocols (i.e., Sanger, PCR-MinION, and RCA-MinION procedures) were aligned together with MAFFT v7.475 before manual edit of the alignment. A home-made R script was used for sequence comparison and mutation count. Mutations were classified in three categories: substitution, insertion/deletion (INDEL) and homopolymer length variation (HLV) (Figure 1). ML trees were inferred from these alignments using FastTree2 as described above.

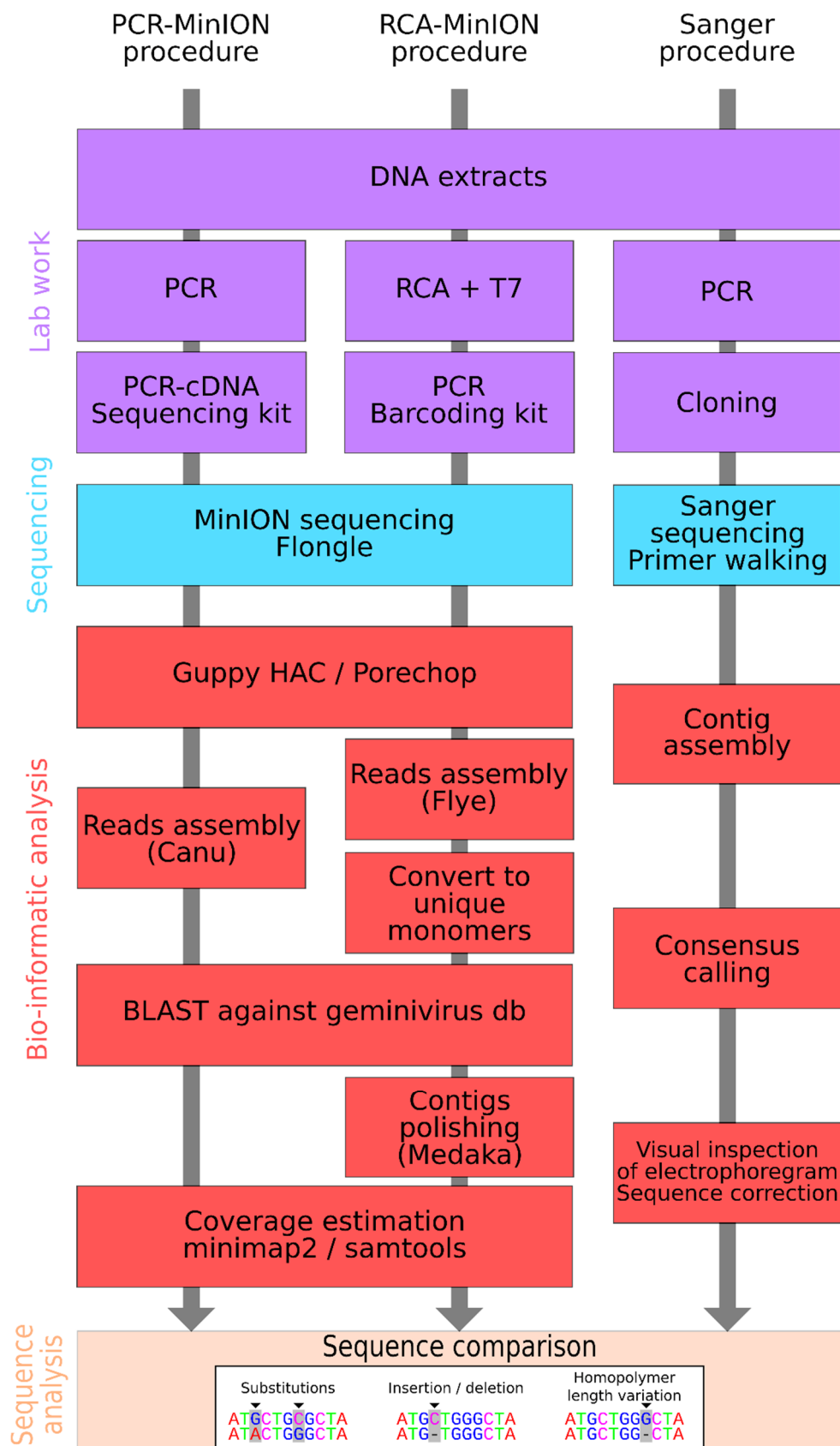


Figure 1. Schematic representation of the three distinct sequencing methodologies used including the molecular biology procedure (purple boxes) the sequencing procedure (light blue boxes), the bio-informatics procedure (red boxes) and the method comparison (orange boxes).

3. Results and Discussion

3.1. Sanger Sequences

A total of 46 sequences were obtained after cloning and Sanger sequencing. These sequence groups in two distinct clades (Figure S1) that share a mean identity of 90.3%. Within the first clade ($n = 19$), sequences present with identity ranging from 99.5 to 100% with each other and were most similar to the isolate BG2_capuz_47 of the B strain of TrV1 (GenBank accession number MW698819) with a minimum identity of 99.7%. Within the second clade ($n = 27$), sequences present identity ranging from 99.0 to 100% with each other and were most closely related to the isolate BG2_coro_02-2 of the C strain of TrV1 (GenBank accession number MW698820) with a minimum identity of 99.1%. Therefore, our two groups of sequences belong to two distinct strains (TrV1-B and TrV1-C) of the capulavirus *Trifolium virus 1* species (Figure S1). It is important to notice here that among the 46 isolates, five of the TrV1-B isolates and seven of the TrV1-C isolates present with defective genomes. Two of the TrV1-B defective isolates have deletions that encompass a fraction of the V3 gene, two other isolates have deletions within the gene of the replication-associated protein and one last has a deletion that encompasses the *rep* gene. Four sequences of the TrV1-C isolates have deletions that encompass a fraction of the CP gene. Among these, the deletions also span a fraction of the V3 gene or the *rep* gene, depending on the isolate. One isolate has a deletion that encompasses a fraction of all the gene encoded in the complement strand. These twelve defective sequences were excluded from downstream analysis. The genomic organization of the remaining sequences confirms the presence of a short intergenic region (SIR) and a long intergenic region (LIR), a characteristic inverted repeat potentially capable of forming a stem-loop structure that included a conserved nonanucleotide sequence TAATATTAC present at almost all geminivirus virion-strand replication origins, the *cp*, a spliced complementary-strand intron-containing transcript which expresses replication-associated protein gene (*rep*), a large complementary-sense ORF (C3) that is completely embedded within the *rep* gene, and a complex arrangement of possible MP-encoding ORFs located in the 5' direction from the *cp* gene, which is a unique feature of *Capulavirus* genomes [47]. Four and three sequences of the TrV1-B and -C strains, present truncated ORFs. The full genome sequences of the isolate without ORFs truncation are available on GenBank under the accession numbers MW698819–MW698821 and MW768713–MW768736.

3.2. Long Read Sequencing and Assembly

The RCA-MinION generated reads that confirmed the presence of both TrV1-B and TrV1-C strains in the *Medicago arborea* sample (Figure 2). The raw sequencing statistics are available in Table 1. From Flongle 1 and Flongle 2, 188,123 and 273,088 raw reads were obtained from which 110,830 (59%) and 152,076 (56%) barcoded reads passed the quality control (Figure S2), respectively. The median read length was 1154 bp with a longest read of 10,491 bp. From Flongle 1, only 27 reads (0.02%) mapped with the capulavirus references. From Flongle 2, barcodes were retrieved from 65,413 reads, 40,242 reads and 46,421 reads for each of the three barcodes, respectively. From 6 to 8.5% of those reads mapped with the capulavirus sequences. Although it was performed on the same DNA extract, RCA amplification yielded more than two order of magnitude less viral sequence for the Flongle 1 amplification than those performed for Flongle 2. It highlights known bias of the RCA amplification [66,67] that were already evidenced in the context of CRESS virus amplification [68]. All the four distinct sequence sets (one barcode sample from Flongle 1 and three for Flongle 2) were then submitted to the assembly and circularization pipeline. For each of the four barcodes, unique contigs corresponding to the full genome sequences of the two strains were obtained. The four TrV1-B sequence lengths ranged from 2745 to 2769 bp and were at least 99.5% similar to any Sanger sequence. One of the TrV1-B sequences (RCA-Minion_10_2, Flongle 2 barcode 10, Table 1) present with a region that seems to be mis-assembled (100 nt in length, see grey tracks in Figure 3A). The four TrV1-C sequences length ranged from 2763 to 2771 bp and were at least 99.1% similar to any

Sanger sequence. Again, a probable mis-assembly (68 nt in length, grey tracks in Figure 3B) was present within one sequence (RCA-MinION_01_1). As none of the raw minion reads supported the presence of this putative recombinant region, it has not been taken into consideration for further analyses.

Table 1. Long read sequencing statistics.

Flongle Id	Raw Reads	Passed Reads	Barcode ID	Trimmed Reads	Capulavirus Reads	TrV1-B Reads	TrV1-C Reads	Length TrV1-B Assembly	Length TrV1-C Assembly
1	188,123	162,263	1	110,830	27	14	13	2745	2771
			10	65,413	4809	3230	1579	2769	2764
2	273,088	215,143	11	40,242	3423	2244	1179	2748	2763
			12	46,421	2803	1585	1218	2746	2765
13	492,922	386,099	-	385,029	380,933	380,933	-	2731	-
14	414,665	371,700	-	370,755	367,381	367,381	-	2739	-
15	768,144	714,730	-	602,579	337,338	-	337,338	-	2754

The PCR-MinION generated reads that also confirmed the presence of TrV1-B and TrV1-C strains after RCA-MinION and Sanger sequencing. From Flongle 13, Flongle 14, and Flongle 15, 492,922, 414,665, and 768,144 raw reads were obtained from which 386,099 (78.3%), 371,700 (89.6%), and 714,730 (93.0%) reads passed the quality control (Table 1, Figure S2). The median length of the passed reads was 643 bp with a longest read of 9816 bp. From Flongle 13, 14 and 15, 380,933 reads (98.7%), 367,381 reads (98.8%) and 337,338 reads (47.2%) mapped with the TrV1-B and -C reference sequences. All the three read sets were then submitted to the assembly. For every Flongle, contigs corresponding to the full genome sequence of the TrV1-B and TrV1-C Sanger references were obtained. The two TrV1-B sequences length ranged from 2731 to 2739 bp and were at least 99.7% similar to any Sanger sequence. The TrV1-C sequence is 2754 bp length and were at least 99.3% similar to any Sanger sequence (Figure 2).

3.3. Sequences Comparison

In order to more precisely determine the nature of the differences between the sequences obtained through the different procedures, per capulavirus strains, all the full genome sequences obtained were compared to the consensus of the Sanger sequences. With more than 99.1% nucleotide identity for both the RCA-MinION and PCR-MinION sequences, the two methods demonstrate their ability to recover full genomes that are accurately assigned to strains TrV1-B and -C and whose sequence are representative of the viral population they originate. However, none of the assemblies obtained from MinION sequences was 100% identical to the Sanger sequence.

Beside mis-assemblies (grey tracks on Figure 3), the differences between sequences were classified in three categories with substitution (green ticks on Figure 3), INDEL (blue ticks) and HLV (red ticks). It must be noticed here that HLVs are a category of INDELs but were count separately as HLVs are recognized as the main source of errors in nanopore sequencing [69–72].

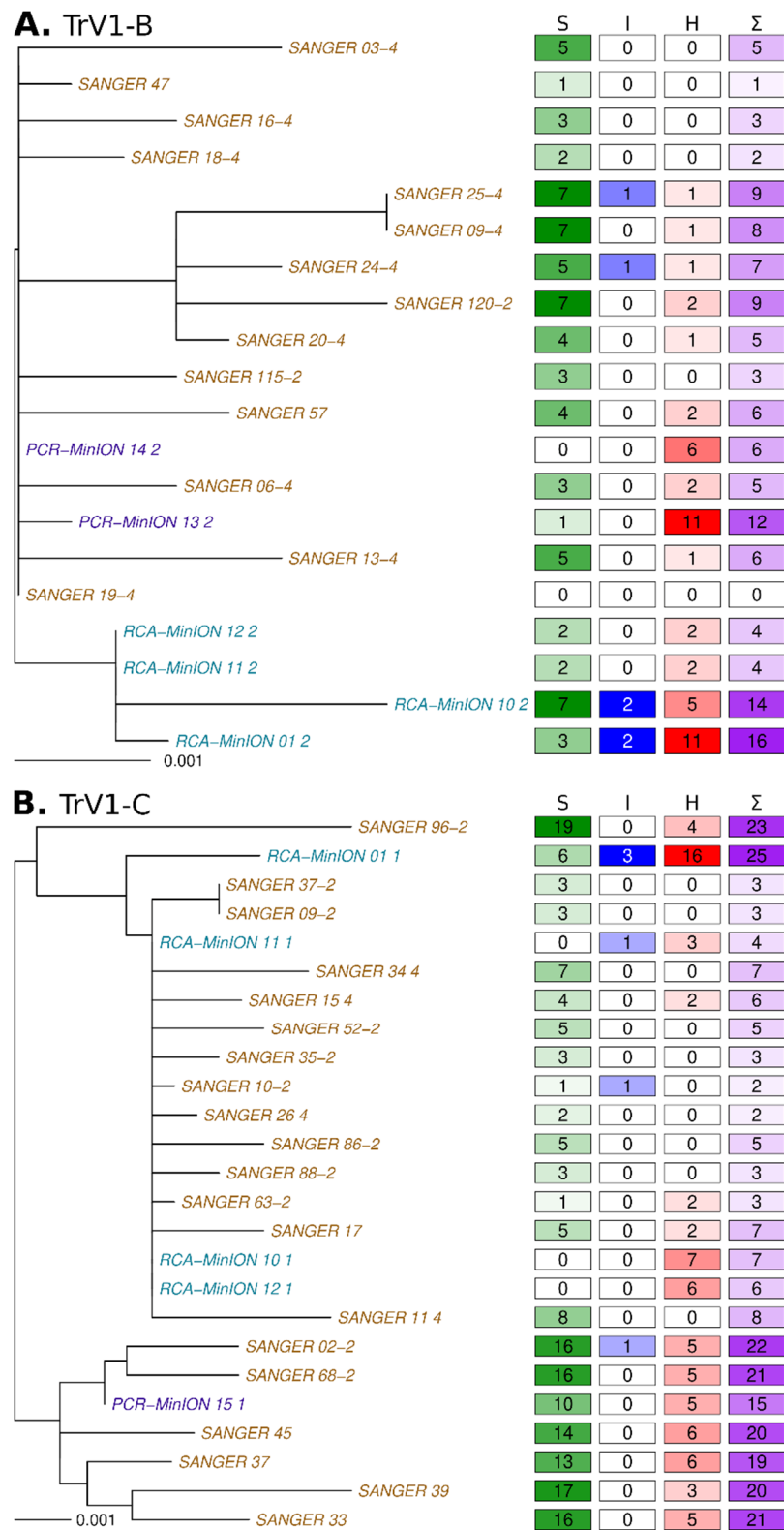


Figure 2. Maximum likelihood phylogenetic tree of all the sequences of the TrV1-B (A) and TrV1-C (B) strains obtained after Sanger sequencing (brown tips) or MinION sequencing followed by read assemblies (blue and purple tips for the RCA-MinION and PCR-MinION procedures respectively) on the left along with a matrix presenting the number of mutation relative to the Sanger consensus on the right. The numbers in the four columns present the substitutions (S), INDELs (I), HLVs (H) and the sum (Σ) of all these variations.

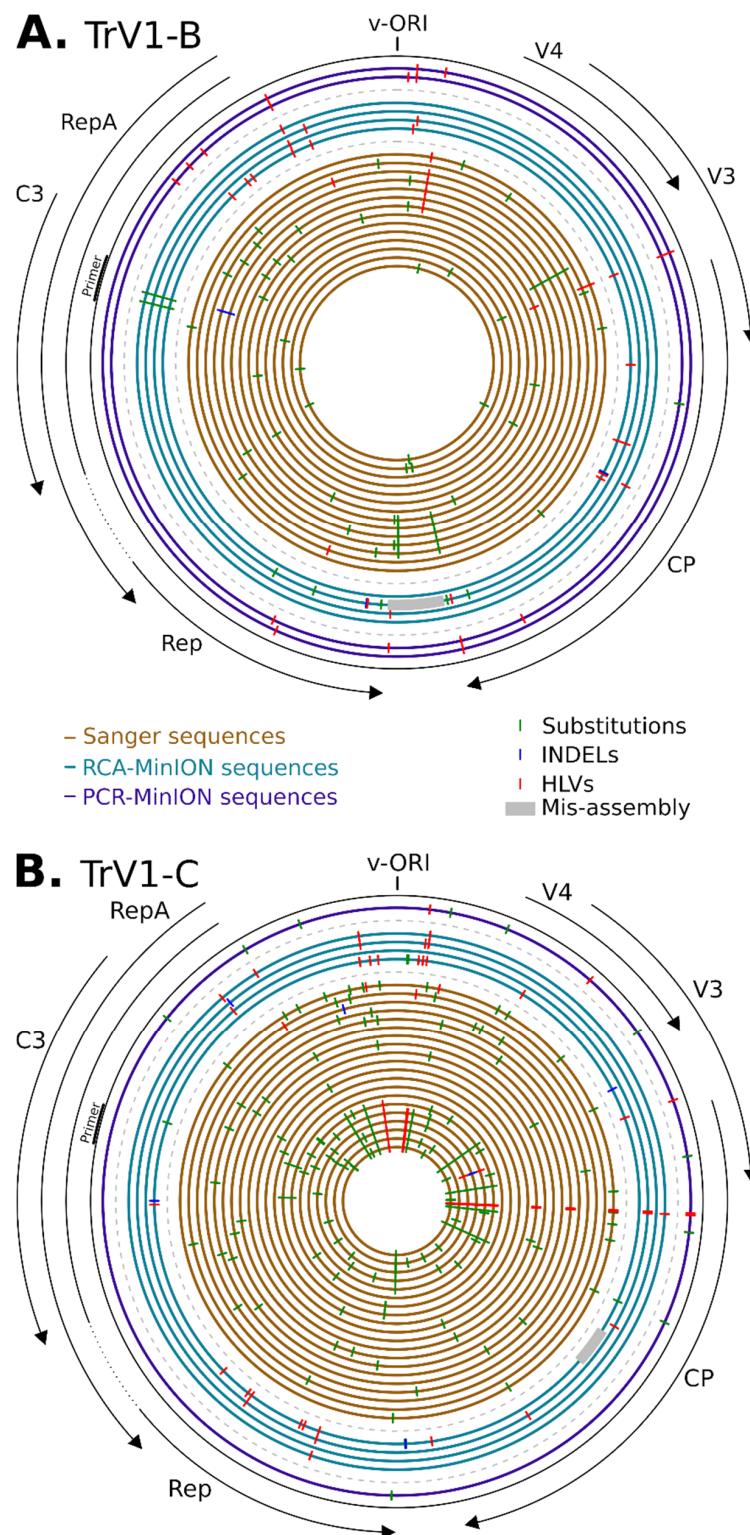


Figure 3. Diagram presenting the positions of the mutations along the genome of the TrV1-B (A) and TrV1-C (B) isolates and assemblies. Concentric circles represent each complete genome sequence obtained, with from the center to the outer, Sanger sequences in brown, RCA-MinION in blue and PCR-MinION in purple. Substitutions, INDELs and homopolymer length variations are represented with green, blue and red ticks respectively. Grey areas represent large deletions. Positions are relative to reference sequences MW698819 for TrV1-B and MW698820 for TrV1-C. The origin of replication (v-ORI) is indicated on top and the ORFs are represented on the outside of the figure.

First, Sanger sequences differ from each other (and to the consensus) mostly with substitutions (accounting for 79.5% of the variations) and more marginally with HLVs (19.4% of the variations) and INDELs (1.5% of the variations) (Figure 2). Except for two of the TrV1-C sequences, all the other Sanger sequences were unique and the number of variations between all these sequences was up to 12 and 27 for TrV1-B and TrV1-C, respectively. A total of 46 and 113 polymorphic sites were present in the Sanger sequences for TrV1-B and TrV1-C, respectively (Figure 3). Whereas these polymorphic sites tend to be more frequently present within non-coding regions (binomial test p -value of 0.090 and 0.016 for TrV1-B and TrV1-C, respectively), these sites had globally more mutations among all the Sanger sequences (binomial test p -value of 3.7×10^{-3} and 7.7×10^{-6} for TrV1-B and TrV1-C, respectively). Whereas sequencing errors can explain a fraction of the mutation, with the high accuracy of Sanger sequencing in mind (per base Phred quality score of 50 [73]), one can expect that most of the variations uncovered during the analysis are genuine and represent the biological variation associated with the diversity of the viral population infecting the plant [74–76].

The analysis of the RCA-MinION and PCR-MinION sequences revealed a different pattern of polymorphism. Two sequences, one of TrV1-B and one of TrV1-C, both assembled from a restricted number of reads (Flongle 1, RCA-MinION_01-1 and RCA-MinION_01-2) were, as expected, less accurate. Coverage of the assembly, (e.g., representing the mean number of times every position of an assembly was read) is therefore a good indicator of the reliability of the resulting assembly (Figure 4). Among the nine other MinION sequences assembled, seven present with no substitution to the consensus and one with ten substitutions (Figure 2). Through the TrV1-B sequences there were few common mutations: only a single HLV was common between Sanger and MinION assemblies (Figure 3A). PCR-MinION and RCA-MinION sequences presented four common HLVs (Figure 3A). For TrV1-C, Sanger sequences presented one substitution, no INDEL and three HLVs in common with MinION sequences (Figure 3B). PCR-MinION and RCA-MinION sequences presented one common substitution, three common HLVs and no common INDELs (Figure 3B). As these sequences were obtained after the assembly of reads, we were unable to catch the diversity of the distinct variants forming the viral population but rather to obtain a sequence very similar to the consensus of that population. Conversely to the reduced number of substitutions in comparison to the consensus, the assemblies present with larger numbers of INDELs (from 1 to 3) and HLVs (2 to 16). Some of these variations were also found in the Sanger sequences and it is probable that the assemblies actually represented some of the variability within the population. Nevertheless, for six INDELs and 38 HLVs, no corresponding mutations were found, and most would induce frameshift or protein truncation (Figure 3). Despite the use of a dedicated sequence correction program, multiple HLVs remained in the assembled sequences.

3.4. Defective Genome and Sequence Coverage

Defective genomes are frequently detected within geminivirus populations [77]. For instance, twelve of the Sanger sequences displayed large deletions in comparison to full reference genomes. As MinION raw sequences are obtained after direct reads of either PCR or RCA amplicons, they should capture the diversity of defective genomes of the viral populations. Indeed, to obtain full genome assemblies using the PCR-MinION procedure, a selection of the reads approaching full genome size was required. The analysis of coverage of the reads (Figure 4) confirms the pervasive nature of defective genomes for both the TrV1-B and -C strains. The highest coverage was obtained for regions encompassing the stem-loop and reads were frequently missing most of the coding regions (see the blue lines for RCA-MinION). Importantly, it must be noticed here that the coverage inferred using the PCR-MinION procedure are not representative of the global population but rather represent the subset of virus that contains the priming site of the abutting primers used in the PCR (indicated with the verticals red dotted lines in Figure 4). The results indicate that every position of the TrV1-B strains present with a high coverage, most of the amplicons of

TrV1-C were defective for a region encompassing the whole CP gene (from position 147 to 1559).

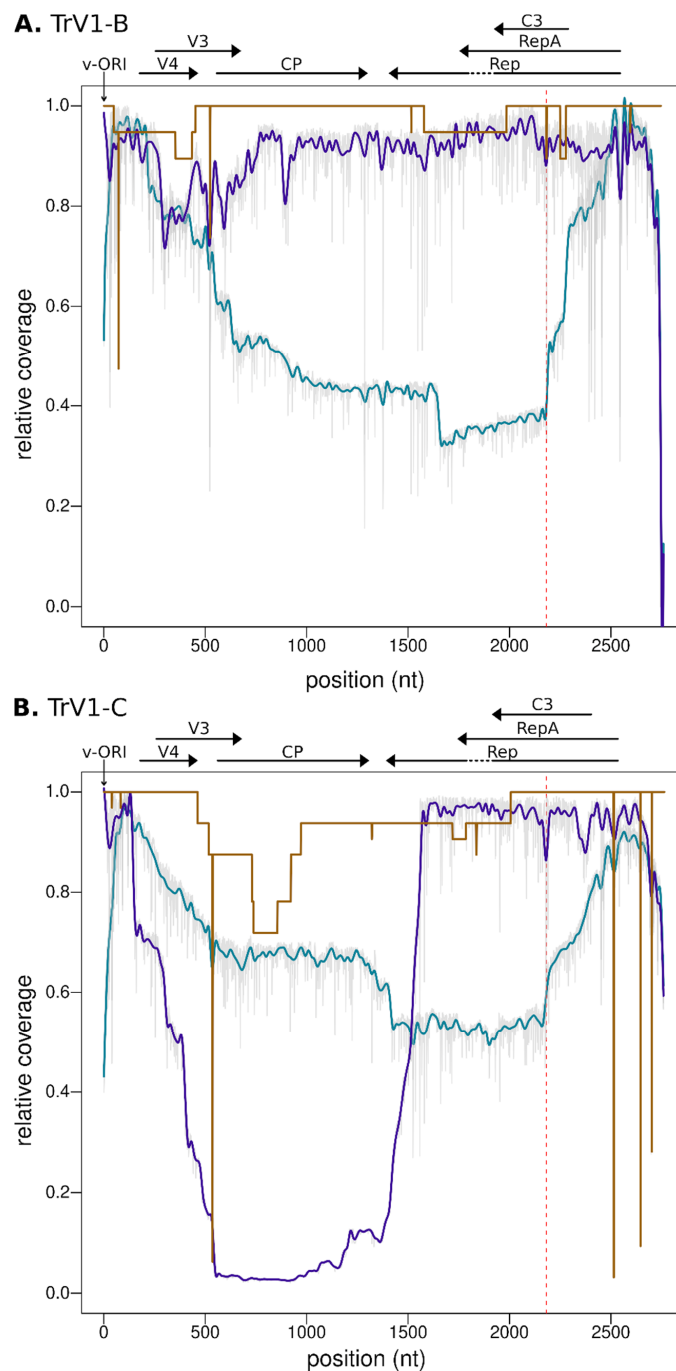


Figure 4. Coverage plot along the reference genome of TrV1-B (A) and TrV1-C (B). The coverage, interpreted here as the fold coverage of a given position divided by the maximum coverage of the whole genome, is represented with a brown line for Sanger sequences, a blue line for RCA-MinION sequences and a purple line for PCR-MinION sequences. MinION coverage are interpolations of the raw coverages presented here in grey on the background. The position of the abutting primers used in the PCR-MinION procedure is indicated by the vertical red dotted line. Positions are relative to the reference genome MW698819 for TrV1-B and MW698820 for TrV1-C. ORFs of these reference genomes are symbolized with horizontal arrows on top of the figure. The Origin of replication is indicated with an arrow. The position 1 of each sequence correspond to the nick-site within the stem loop structure.

4. Conclusions

From an asymptomatic sample of *Medicago arborea*, two distinct strains of the capulavirus TrV1 have been cloned and sequenced by the Sanger methodology. Using both the *a priori* and the agnostic nanopore-based procedures, both TrV1-B and TrV1-C strains were detected, and full genome sequences were assembled. Despite being very similar to the consensus of Sanger sequences, mutations specific to the MinION assemblies were detected mostly within homopolymeric regions of the genomes, which is in agreement with other studies that have also pinpointed higher number of errors associated with homopolymer lengths [78,79]. Whereas it could be argued that the HLV errors would be reduced with the increasing accuracy of sequencing, the development of a new base caller technologies and correction algorithm [80], current MinION sequences assemblies may be avoided for some specific applications where the exact nucleotide sequence is required. Otherwise, for other applications, such as virus discovery, virus classification, or recombination analysis, MinION assemblies represent a competitive alternative to Sanger sequencing. Although, similarly to other NGS protocols, MinION based studies require the use of sophisticated bioinformatics tools for data management and analysis and despite the specific drawbacks on sequence quality, bench top sequencer such as the MinION would probably become routinely used in the laboratory. It allows a real-time detection and diagnostic of multiple viral strains or virus species in a single run. For niche applications, such as the exploration of geminivirus genomes, which do not exceed 10 kb, nanopore sequencing is poised to push the cost and performance limits of sequencing technologies. The high reactivity offered by the platform would make it more and more democratized as a mobile real-time plant disease diagnostic tool.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/microorganisms9050903/s1>. Figure S1: Maximum likelihood phylogenetic tree of all the *Capulavirus* genus full genome sequences obtained from NCBI along with the two TrV1-B and TrV1-C genotypes obtained in this study; Supplementary Figure S2: Reads length (*x*-axis) against quality score (*y*-axis) for all passed reads (A and C) and capulavirus only passed reads (B and D) obtained using the RCA-MinION and PCR-MinION procedures.

Author Contributions: D.F., P.L. and P.R. conceived and designed the experiments. S.B.C., L.B., E.F., M.H., C.J. and O.M. performed the experiments. S.B.C., D.F. and P.L. analyzed and interpreted the data. S.B.C. and P.L. wrote original draft. D.F., P.L., J.-M.L. and P.R. revised and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the European Union (ERDF, contract GURDT I2016-1731-0006632), the *Conseil Régional de la Réunion* and CIRAD. SBC is a recipient of a PhD fellowship from CIRAD and ANR (Phytovirus project number: ANR-19-CE35-0008-02).

Data Availability Statement: Publicly available datasets were analyzed in this study. Sanger sequences accession numbers [MW698819–MW698821, MW768713–MW768736] are available on NCBI Genbank. MinION data are available at the NCBI Short read archive under the BioProject [PRJNA715304].

Acknowledgments: This work was conducted on the Plant Protection Platform (3P, IBISA).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dolja, V.V.; Koonin, E.V. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* **2018**, *244*, 36–52. [[CrossRef](#)] [[PubMed](#)]
2. Greninger, A.L. A decade of RNA virus metagenomics is (not) enough. *Virus Res.* **2018**, *244*, 218–229. [[CrossRef](#)]
3. Koonin, E.V.; Dolja, V.V. Metaviromics: A tectonic shift in understanding virus evolution. *Virus Res.* **2018**, *246*, A1–A3. [[CrossRef](#)] [[PubMed](#)]
4. Roux, S.; Matthijnsens, J.; Dutilh, B.E. Metagenomics in Virology. *Encycl. Virol.* **2021**, *1*, 133–140. [[CrossRef](#)]
5. Roux, S.; Brum, J.R.; Dutilh, B.E.; Sunagawa, S.; Duhaime, M.B.; Loy, A.; Poulos, B.T.; Solonenko, N.; Lara, E.; Poulain, J.; et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nat. Cell Biol.* **2016**, *537*, 689–693. [[CrossRef](#)]

6. McMullen, A.; Martinez-Hernandez, F.; Martinez-Garcia, M. Absolute quantification of infecting viral particles by chip-based digital polymerase chain reaction. *Environ. Microbiol. Rep.* **2019**, *11*, 855–860. [[CrossRef](#)]
7. Breitbart, M.; Hewson, I.; Felts, B.; Mahaffy, J.M.; Nulton, J.; Salamon, P.; Rohwer, F. Metagenomic Analyses of an Uncultured Viral Community from Human Feces. *J. Bacteriol.* **2003**, *185*, 6220–6223. [[CrossRef](#)]
8. Wylie, K.M.; Weinstock, G.M.; Storch, G.A. Emerging view of the human virome. *Transl. Res.* **2012**, *160*, 283–290. [[CrossRef](#)] [[PubMed](#)]
9. Lecuit, M.; Eloit, M. The diagnosis of infectious diseases by whole genome next generation sequencing: A new era is opening. *Front. Cell. Infect. Microbiol.* **2014**, *4*, 25. [[CrossRef](#)] [[PubMed](#)]
10. Maclot, F.; Candresse, T.; Filloux, D.; Malmstrom, C.M.; Roumagnac, P.; Van Der Vlugt, R.; Massart, S. Illuminating an Ecological Blackbox: Using High Throughput Sequencing to Characterize the Plant Virome Across Scales. *Front. Microbiol.* **2020**, *11*, 578064. [[CrossRef](#)]
11. Dávila-Ramos, S.; Castelán-Sánchez, H.G.; Martínez-Ávila, L.; Sánchez-Carbente, M.D.R.; Peralta, R.; Hernández-Mendoza, A.; Dobson, A.D.W.; Gonzalez, R.A.; Pastor, N.; Batista-García, R.A. A Review on Viral Metagenomics in Extreme Environments. *Front. Microbiol.* **2019**, *10*, 2403. [[CrossRef](#)] [[PubMed](#)]
12. Goodwin, S.; McPherson, J.D.; McCombie, W.R. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **2016**, *17*, 333–351. [[CrossRef](#)]
13. Shi, M.; Lin, X.-D.; Tian, J.-H.; Chen, L.-J.; Chen, X.; Li, C.-X.; Qin, X.-C.; Li, J.; Cao, J.-P.; Eden, J.-S.; et al. Redefining the invertebrate RNA virosphere. *Nat. Cell Biol.* **2016**, *540*, 539–543. [[CrossRef](#)]
14. Shi, M.; Zhang, Y.-Z.; Holmes, E.C. Meta-transcriptomics and the evolutionary biology of RNA viruses. *Virus Res.* **2018**, *243*, 83–90. [[CrossRef](#)]
15. Zhang, Y.-Z.; Shi, M.; Holmes, E.C. Using Metagenomics to Characterize an Expanding Virosphere. *Cell* **2018**, *172*, 1168–1172. [[CrossRef](#)] [[PubMed](#)]
16. Rosario, K.; Duffy, S.; Breitbart, M. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Arch. Virol.* **2012**, *157*, 1851–1871. [[CrossRef](#)] [[PubMed](#)]
17. Simmonds, P.; Adams, M.J.; Benkő, M.; Breitbart, M.; Brister, J.R.; Carstens, E.B.; Davison, A.J.; Delwart, E.; Gorbalenya, A.E.; Harrach, M.B.B.; et al. Virus taxonomy in the age of metagenomics. *Nat. Rev. Genet.* **2017**, *15*, 161–168. [[CrossRef](#)]
18. Batovska, J.; Lynch, S.E.; Rodoni, B.C.; Sawbridge, T.I.; Cogan, N.O. Metagenomic arbovirus detection using MinION nanopore sequencing. *J. Virol. Methods* **2017**, *249*, 79–84. [[CrossRef](#)] [[PubMed](#)]
19. Roux, S.; Adriaenssens, E.M.; Dutilh, B.E.; Koonin, E.V.; Kropinski, A.M.; Krupovic, M.; Kuhn, J.H.; Lavigne, R.; Brister, J.R.; Varsani, A.; et al. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat. Biotechnol.* **2019**, *37*, 29–37. [[CrossRef](#)] [[PubMed](#)]
20. Claverie, S.; Ouattara, A.; Hoareau, M.; Filloux, D.; Varsani, A.; Roumagnac, P.; Martin, D.P.; Lett, J.-M.; Lefeuvre, P. Exploring the diversity of Poaceae-infecting mastreviruses on Reunion Island using a viral metagenomics-based approach. *Sci. Rep.* **2019**, *9*, 1–11. [[CrossRef](#)]
21. International Committee on Taxonomy of Viruses Executive Committee. The new scope of virus taxonomy: Partitioning the virosphere into 15 hierarchical ranks. *Nat. Microbiol.* **2020**, *5*, 668–674. [[CrossRef](#)]
22. Roossinck, M.J.; Martin, D.P.; Roumagnac, P. Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology* **2015**, *105*, 716–727. [[CrossRef](#)]
23. Quiñones-Mateu, M.E.; Avila, S.; Reyes-Teran, G.; Martinez, M.A. Deep sequencing: Becoming a critical tool in clinical virology. *J. Clin. Virol.* **2014**, *61*, 9–19. [[CrossRef](#)] [[PubMed](#)]
24. Mikheyev, A.S.; Tin, M.M.Y. A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* **2014**, *14*, 1097–1102. [[CrossRef](#)]
25. Sanger, F.; Coulson, A. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **1975**, *94*, 441–448. [[CrossRef](#)]
26. Jain, M.; Olsen, H.E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biol.* **2016**, *17*, 1–11. [[CrossRef](#)]
27. Pomerantz, A.; Peñafiel, N.; Arteaga, A.; Bustamante, L.; Pichardo, F.; Coloma, L.A.; Barrio-Amorós, C.L.; Salazar-Valenzuela, D.; Prost, S. Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. *GigaScience* **2018**, *7*, 1–14. [[CrossRef](#)] [[PubMed](#)]
28. Xu, Y.; Lewandowski, K.; Lumley, S.; Pullan, S.; Vipond, R.; Carroll, M.; Foster, D.; Matthews, P.C.; Peto, T.; Crook, D. Detection of Viral Pathogens with Multiplex Nanopore MinION Sequencing: Be Careful With Cross-Talk. *Front. Microbiol.* **2018**, *9*, 2225. [[CrossRef](#)]
29. Filloux, D.; Fernandez, E.; Loire, E.; Claude, L.; Galzi, S.; Candresse, T.; Winter, S.; Jeeva, M.L.; Makesh Kumar, T.; Martin, D.P.; et al. Nanopore-based detection and characterization of yam viruses. *Sci. Rep.* **2018**, *8*, 1–11. [[CrossRef](#)] [[PubMed](#)]
30. Ewang, J.; Moore, N.E.; Edeng, Y.-M.; Eccles, D.A.; Hall, R.J. MinION nanopore sequencing of an influenza genome. *Front. Microbiol.* **2015**, *6*, 766. [[CrossRef](#)]
31. Quick, J.; Loman, N.J.; Duraffour, S.; Simpson, J.T.; Severi, E.; Cowley, L.; Bore, J.A.; Koundouno, R.; Dudas, G.; Mikhail, A.; et al. Real-time, portable genome sequencing for Ebola surveillance. *Nat. Cell Biol.* **2016**, *530*, 228–232. [[CrossRef](#)]

32. Yamagishi, J.; Runtuwene, L.R.; Hayashida, K.; Mongan, A.E.; Thi, L.A.N.; Thuy, L.N.; Nhat, C.N.; Limkittikul, K.; Sirivichayakul, C.; Sathirapongsasuti, N.; et al. Serotyping dengue virus with isothermal amplification and a portable sequencer. *Sci. Rep.* **2017**, *7*, 1–10. [[CrossRef](#)]
33. Quick, J.; Grubaugh, N.D.; Pullan, S.T.; Claro, I.M.; Smith, A.D.; Gangavarapu, K.; Oliveira, G.; Robles-Sikisaka, R.; Rogers, T.F.; Beutler, N.A.; et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **2017**, *12*, 1261–1276. [[CrossRef](#)] [[PubMed](#)]
34. Kim, D.; Lee, J.-Y.; Yang, J.-S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914–921.e10. [[CrossRef](#)]
35. Adams, I.P.; Braidwood, L.A.; Stomeo, F.; Phiri, N.; Uwumukiza, B.; Feyissa, B.; Mahuku, G.; Wangi, A.; Smith, J.; Mumford, R.; et al. Characterising Maize Viruses Associated with Maize Lethal Necrosis Symptoms in Sub Saharan Africa. *bioRxiv* **2017**. preprint. [[CrossRef](#)]
36. Badial, A.B.; Sherman, D.; Stone, A.; Gopakumar, A.; Wilson, V.; Schneider, W.; King, J. Nanopore Sequencing as a Surveillance Tool for Plant Pathogens in Plant and Insect Tissues. *Plant Dis.* **2018**, *102*, 1648–1652. [[CrossRef](#)] [[PubMed](#)]
37. Inoue-Nagata, A.K.; Albuquerque, L.C.; Rocha, W.B.; Nagata, T. A simple method for cloning the complete begomovirus genome using the bacteriophage ϕ 29 DNA polymerase. *J. Virol. Methods* **2004**, *116*, 209–211. [[CrossRef](#)] [[PubMed](#)]
38. Rosario, K.; Dayaram, A.; Marinov, M.; Ware, J.; Kraberg, S.; Stainton, D.; Breitbart, M.; Varsani, A. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J. Gen. Virol.* **2012**, *93*, 2668–2681. [[CrossRef](#)]
39. Rosario, K.; Schenck, R.O.; Harbeitner, R.C.; Lawler, S.N.; Breitbart, M. Novel circular single-stranded DNA viruses identified in marine invertebrates reveal high sequence diversity and consistent predicted intrinsic disorder patterns within putative structural proteins. *Front. Microbiol.* **2015**, *6*, 696. [[CrossRef](#)]
40. Theuns, S.; Vanmechelen, B.; Bernaert, Q.; Deboutte, W.; Vandenhole, M.; Beller, L.; Matthijnsens, J.; Maes, P.; Nauwynck, H.J. Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobovirus as an important enteric virus. *Sci. Rep.* **2018**, *8*, 1–13. [[CrossRef](#)]
41. Boykin, L.M.; Sseruwagi, P.; Alicai, T.; Ateka, E.; Mohammed, I.U.; Stanton, J.-A.L.; Kayuki, C.; Mark, D.; Fute, T.; Erasto, J.; et al. Tree Lab: Portable genomics for Early Detection of Plant Viruses and Pests in Sub-Saharan Africa. *Genes* **2019**, *10*, 632. [[CrossRef](#)] [[PubMed](#)]
42. Cao, J.; Zhang, Y.; Dai, M.; Xu, J.; Chen, L.; Zhang, F.; Zhao, N.; Wang, J. Profiling of Human Gut Virome with Oxford Nanopore Technology. *Med. Microecol.* **2020**, *4*, 100012. [[CrossRef](#)]
43. Chalupowicz, L.; Dombrowsky, A.; Gaba, V.; Luria, N.; Reuven, M.; Beerman, A.; Lachman, O.; Dror, O.; Nissan, G.; Manulis-Sasson, S. Diagnosis of plant diseases using the Nanopore sequencing platform. *Plant Pathol.* **2018**, *68*, 229–238. [[CrossRef](#)]
44. Naito, F.Y.B.; Melo, F.L.; Fonseca, M.E.N.; Santos, C.A.F.; Chanes, C.R.; Ribeiro, B.M.; Gilbertson, R.L.; Boiteux, L.S.; Pereira-Carvalho, R.D.C. Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. *Arch. Virol.* **2019**, *164*, 1907–1910. [[CrossRef](#)]
45. Leiva, A.M.; Siriwan, W.; Alvarez, D.L.; Barrantes, I.; Hemniam, N.; Saokham, K.; Cuellar, W.J. Nanopore-Based Complete Genome Sequence of a Sri Lankan Cassava Mosaic Virus (Geminivirus) Strain from Thailand. *Microbiol. Resour. Announc.* **2020**, *9*, 9. [[CrossRef](#)]
46. Ma, Y.; Svanella-Dumas, L.; Julian, C.; Galzi, S.; Fernandez, E.; Yvon, M.; Pirolles, E.; Lefebvre, M.; Filloux, D.; Roumagnac, P.; et al. Genome characterization and diversity of trifolium virus 1: Identification of a novel legume-infective capulavirus. *Arch. Virol.* **2021**, in press.
47. Varsani, A.; Roumagnac, P.; Fuchs, M.; Navas-Castillo, J.; Moriones, E.; Idris, A.; Briddon, R.W.; Rivera-Bustamante, R.; Zerbini, F.M.; Martin, D.P. Capulavirus and Grabovirus: Two new genera in the family Geminiviridae. *Arch. Virol.* **2017**, *162*, 1819–1831. [[CrossRef](#)]
48. Fauquet, C.M.; Briddon, R.W.; Brown, J.K.; Moriones, E.; Stanley, J.; Zerbini, M.; Zhou, X. Geminivirus strain demarcation and nomenclature. *Arch. Virol.* **2008**, *153*, 783–821. [[CrossRef](#)]
49. Roumagnac, P.; Granier, M.; Bernardo, P.; Deshoux, M.; Ferdinand, R.; Galzi, S.; Fernandez, E.; Julian, C.; Abt, I.; Filloux, D.; et al. Alfalfa Leaf Curl Virus: An Aphid-Transmitted Geminivirus. *J. Virol.* **2015**, *89*, 9683–9688. [[CrossRef](#)]
50. Varsani, A.; Navas-Castillo, J.; Moriones, E.; Hernández-Zepeda, C.; Idris, A.; Brown, J.K.; Zerbini, F.M.; Martin, D.P. Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. *Arch. Virol.* **2014**, *159*, 2193–2203. [[CrossRef](#)]
51. Zerbini, F.M.; Briddon, R.W.; Idris, A.; Martin, D.P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Roumagnac, P.; Varsani, A. ICTV Report Consortium ICTV Virus Taxonomy Profile: Geminiviridae. *J. Gen. Virol.* **2017**, *98*, 131–133. [[CrossRef](#)] [[PubMed](#)]
52. Claverie, S.; Bernardo, P.; Kraberg, S.; Hartnady, P.; Lefebvre, P.; Lett, J.-M.; Galzi, S.; Filloux, D.; Harkins, G.W.; Varsani, A.; et al. From Spatial Metagenomics to Molecular Characterization of Plant Viruses: A Geminivirus Case Study. In *Advances in Clinical Chemistry*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 101, pp. 55–83.
53. Bernardo, P.; Golden, M.; Akram, M.; Naimuddin; Nadarajan, N.; Fernandez, E.; Granier, M.; Rebelo, A.G.; Peterschmitt, M.; Martin, D.P.; et al. Identification and characterisation of a highly divergent geminivirus: Evolutionary and taxonomic implications. *Virus Res.* **2013**, *177*, 35–45. [[CrossRef](#)]

54. Susi, H.; Filloux, D.; Frilander, M.J.; Roumagnac, P.; Laine, A.-L. Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ* **2019**, *7*, e6140. [[CrossRef](#)] [[PubMed](#)]
55. Oxford Nanopore Technologies. Available online: <https://nanoporetech.com/> (accessed on 18 March 2021).
56. Porechop. Available online: <https://github.com/rrwick/Porechop> (accessed on 18 March 2021).
57. De Coster, W.; D’Hert, S.; Schultz, D.T.; Cruets, M.; Van Broeckhoven, C. NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics* **2018**, *34*, 2666–2669. [[CrossRef](#)] [[PubMed](#)]
58. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [[CrossRef](#)]
59. Medaka. Available online: <https://github.com/nanoporetech/medaka> (accessed on 18 March 2021).
60. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptivek-mer weighting and repeat separation. *Genome Res.* **2017**, *27*, 722–736. [[CrossRef](#)] [[PubMed](#)]
61. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)] [[PubMed](#)]
62. Rozewicki, J.; Li, S.; Amada, K.M.; Standley, D.M.; Katoh, K. MAFFT-DASH: Integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5–W10. [[CrossRef](#)]
63. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, *5*, e9490. [[CrossRef](#)] [[PubMed](#)]
64. Paradis, E.; Schliep, K. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **2018**, *35*, 526–528. [[CrossRef](#)]
65. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* **2014**, *9*, e108277. [[CrossRef](#)]
66. Yilmaz, S.; Allgaier, M.; Hugenholtz, P. Multiple displacement amplification compromises quantitative analysis of metagenomes. *Nat. Methods* **2010**, *7*, 943–944. [[CrossRef](#)]
67. Kim, K.-H.; Bae, J.-W. Amplification Methods Bias Metagenomic Libraries of Uncultured Single-Stranded and Double-Stranded DNA Viruses. *Appl. Environ. Microbiol.* **2011**, *77*, 7663–7668. [[CrossRef](#)]
68. Gallet, R.; Fabre, F.; Michalakakis, Y.; Blanc, S. The Number of Target Molecules of the Amplification Step Limits Accuracy and Sensitivity in Ultradeep-Sequencing Viral Population Studies. *J. Virol.* **2017**, *91*, 91. [[CrossRef](#)]
69. Laehnemann, D.; Borkhardt, A.; McHardy, A.C. Denoising DNA deep sequencing data—high-throughput sequencing errors and their correction. *Brief. Bioinform.* **2016**, *17*, 154–179. [[CrossRef](#)] [[PubMed](#)]
70. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **2019**, *20*, 1–10. [[CrossRef](#)] [[PubMed](#)]
71. Gargis, A.S.; Cherney, B.; Conley, A.B.; McLaughlin, H.P.; Sue, D. Rapid Detection of Genetic Engineering, Structural Variation, and Antimicrobial Resistance Markers in Bacterial Biothreat Pathogens by Nanopore Sequencing. *Sci. Rep.* **2019**, *9*, 13501–13514. [[CrossRef](#)]
72. Seah, A.; Lim, M.C.; McAloose, D.; Prost, S.; Seimon, T.A. MinION-Based DNA Barcoding of Preserved and Non-Invasively Collected Wildlife Samples. *Genes* **2020**, *11*, 445. [[CrossRef](#)] [[PubMed](#)]
73. Shendure, J.; Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **2008**, *26*, 1135–1145. [[CrossRef](#)]
74. Ge, L.; Zhang, J.; Zhou, X.; Li, H. Genetic Structure and Population Variability of Tomato Yellow Leaf Curl China Virus. *J. Virol.* **2007**, *81*, 5902–5907. [[CrossRef](#)]
75. Harkins, G.W.; Delpont, W.; Duffy, S.; Wood, N.; Monjane, A.L.; Owor, B.E.; Donaldson, L.; Saumtally, S.; Triton, G.; Briddon, R.W.; et al. Experimental evidence indicating that mastreviruses probably did not co-diverge with their hosts. *Virol. J.* **2009**, *6*, 104. [[CrossRef](#)]
76. Domingo, E.; Sheldon, J.; Perales, C. Viral Quasispecies Evolution. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 159–216. [[CrossRef](#)]
77. Jeske, H. Barcoding of Plant Viruses with Circular Single-Stranded DNA Based on Rolling Circle Amplification. *Viruses* **2018**, *10*, 469. [[CrossRef](#)] [[PubMed](#)]
78. McNaughton, A.L.; Roberts, H.E.; Bonsall, D.; De Cesare, M.; Mokaya, J.; Lumley, S.F.; Golubchik, T.; Piazza, P.; Martin, J.B.; De Lara, C.; et al. Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). *Sci. Rep.* **2019**, *9*, 1–14. [[CrossRef](#)] [[PubMed](#)]
79. Viehweger, A.; Krautwurst, S.; Lamkiewicz, K.; Madhugiri, R.; Ziebuhr, J.; Hölzer, M.; Marz, M. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res.* **2019**, *29*, 1545–1554. [[CrossRef](#)] [[PubMed](#)]
80. Chang, J.J.M.; Ip, Y.C.A.; Ng, C.S.L.; Huang, D. Takeaways from Mobile DNA Barcoding with BentoLab and MinION. *Genes* **2020**, *11*, 1121. [[CrossRef](#)] [[PubMed](#)]

Chapitre 2

Chapitre 2

Exploration de la diversité génomique et de la structure des communautés phytovirales à ADN circulaire à La Réunion : Association étroite des mastrévirus avec les milieux agricoles

Le développement des approches métagénomiques a récemment permis d'apporter des éléments de réponses aux questions fondamentales ayant trait à la variation des patrons de diversité virale face aux changements globaux, tels que ceux associés aux activités anthropiques, à la perte de biodiversité ou aux invasions biologiques (Faust *et al.*, 2017). En particulier, au regard de ces changements et en s'appuyant sur les cadres théoriques en écologie, plusieurs études ont été réalisées afin de déterminer si la richesse virale (*i.e.* nombre d'entités du même niveau taxonomique) augmente (hypothèse d'amplification) ou diminue (hypothèse de dilution) avec l'augmentation de la diversité des plantes hôtes (Keesing *et al.*, 2006 ; Faust *et al.*, 2017). Si, dans certains contextes, une diversité d'hôtes plus élevée dans les zones naturelles que dans les zones cultivées va de paire avec une plus grande richesse en espèces virales dans les zones naturelles (Bernardo *et al.*, 2018), dans d'autres, l'inverse est observé avec les zones cultivées pauvres en espèces végétales présentant une diversité virale plus élevée que les zones naturelles riches en espèces végétales (Halliday *et al.*, 2017 ; Susi *et al.*, 2021 ; McLeish *et al.*, 2022). La reconnaissance que ces interactions entre diversité d'hôtes et de virus ne pouvaient être triviales mais étaient aussi dépendantes de facteurs variant dans l'espace et le temps, permettait de réconcilier ces observations contradictoires au premier abord. Ce constat de complexité souligne aussi la nécessité d'obtenir une compréhension plus approfondie de la diversité et de la distribution des virus à l'échelle d'écosystèmes entiers afin de comprendre comment leurs diversités respectives peuvent influencer l'une sur l'autre.

Néanmoins, réaliser une telle description détaillée reste souvent illusoire même avec le recours aux avancées méthodologiques de la métagénomique virale. Afin de contourner cette difficulté, les travaux de Claverie *et al.* (2023) s'étaient focalisés sur l'étude spécifique de virus du genre *Mastrevirus* à l'échelle d'un unique agro-écosystème de La Réunion.

Outre la description d'espèces déjà identifiées mais aussi d'espèces précédemment inconnues de mastrévirus, ce travail a permis de révéler la structure d'association hôte-virus réalisée en conditions naturelles, montrant la cohabitation de virus spécialistes et généralistes ainsi que de nombreux chevauchement de gammes d'hôtes.

Dans ce travail, nous avons étendu l'étude de Claverie *et al.* (2023) à neuf autres sites présentant une écologie et un historique d'utilisation des terres diversifiés. Ce faisant, nous avons entrepris de déterminer la diversité et la distribution des espèces de *Mastrevirus* détectées dans ces sites, puis d'analyser les facteurs écologiques influençant leur richesse et leur répartition géographique.

Pour se faire, nous avons réalisé l'échantillonnage de la majorité des espèces de Poales présentes sur dix sites (quatre agro-écosystèmes, trois prairies littorales et trois prairies altimontaines) avec au total, le prélèvement d'échantillons appartenant à 61 espèces de plantes différentes. Après application du protocole RCA-MinION mis au point précédemment (**Chapitre 1 – Article 1**) sur les différents pools de plantes, 92 séquences à ADNsb circulaire ont été identifiées (**Chapitre 2 – Article 3**), parmi lesquelles se distinguaient de nombreuses espèces du genre *Mastrevirus*, ici nouvellement décrites ou non, mais aussi un virus apparenté au phylum *Cressdnaviricota* présentant une organisation génomique originale.

Ce virus, provisoirement nommé *Cenchrus purpureus associated virus* (CPAV) et isolé depuis des échantillons de *Cenchrus purpureus*, présente une architecture génomique inhabituelle vraisemblablement issue de la recombinaison inter-familiale entre un alphasatellite (*Alphasatellitidae*) et un mastrévirus (*Geminiviridae*) (**Chapitre 2 – Article 2**) : le segment du gène de la protéine associé à la réplication apparenté aux alphasatellites est positionné en amont et dans la même orientation que le segment des gènes de la protéine de mouvement et de la capsid apparentés aux mastrévirus. Le CPAV a été détecté systématiquement sur le terrain en association avec une autre espèce de mastrévirus précédemment non décrite dans la littérature, *Cenchrus purpureus mild streak virus* (CPMSV).

Dans le cadre d'expérimentations en conditions contrôlées, nous avons premièrement évalué le pouvoir pathogène du CPAV, en infection simple ou mixte avec le CPMSV, par l'utilisation de clones agroinfectieux. Il apparaît que le CPAV est capable d'initier l'infection et de se propager dans la plante de manière systémique mais que l'assistance du CPMSV semble nécessaire à la persistance de l'infection. Deuxièmement, nous avons démontré la capacité de transmission du CPAV et du CPMSV par la cicadelle *Cicadulina mbila* (Hemiptera : Cicadellidae), principal vecteur connu des mastrévirus et présent à la Réunion. La capacité du CPAV et du CPMSV à réaliser leur cycle infectieux, à persister, à contourner les mécanismes de résistance de *C. purpureus* et à être transmis par *C. mbila* de manière autonome est discutée.

Globalement, au-delà de confirmer la présence de la grande majorité des mastrévirus précédemment décrits à la Réunion et en plus du CPMSV, nous avons caractérisé deux nouvelles espèces de mastrévirus nommées *Cenchrus echinatus associated virus* et *Urochloa decumbens associated virus*. Les douze espèces de *Mastrevirus* ont été identifiées sur 22 espèces de plantes. Parmi toutes les espèces de plantes de l'ordre des Poales évaluées, seules des espèces appartenant à cinq tribus de la famille des *Poaceae* ont été détectées comme positives aux mastrévirus. La majorité d'entre elles sont des espèces annuelles exotiques sauvages ou cultivées échantillonnées dans les agro-écosystèmes (**Chapitre 2 – Article 3**).

L'analyse du réseau d'interactions plantes-virus confirme une imbrication (*nestedness*) significative des gammes d'hôtes qui a notamment été documentée dans le cas de communautés végétales riches et perturbées (Bastolla *et al.*, 2009). Par ailleurs, l'absence de modularité significative du réseau d'interactions semble corroborer l'instabilité de l'écosystème étudié. La présence majoritaire d'espèces cultivées et adventices annuelles et leurs disparitions saisonnières de l'agro-écosystème pourraient ainsi représenter un facteur clef dans l'absence d'interactions modulaires significatives détectées au sein du réseau d'interactions plantes-virus. A ce titre, l'importance des poacées pérennes dans le maintien des communautés virales et des populations d'insectes vecteurs pendant les intersaisons est également soulignée par nos analyses.

Capitalisant sur la connaissance du réseaux global d'interactions observé, l'analyse des réseaux d'interactions réalisés dans chacun des différents sites de collecte a mis en lumière une association positive entre richesse d'hôtes et réalisation de la richesse virale (**Chapitre 2 – Article 3**). Autrement dit, plus nous trouvons d'espèces d'hôtes de mastrévirus sur un site, plus la fraction de virus potentiellement présente sachant la diversité d'hôte est importante. Ce résultat du type « les riches seront plus riches » (« *rich get richer* »), largement compatible avec l'effet d'amplification (Hechinger and Lafferty, 2005 ; Lafferty, 2012 ; Kamiya *et al.*, 2014 ; Halliday *et al.*, 2020) indique la possibilité d'une synergie interspécifique chez les mastrévirus. Toutefois, l'absence de corrélation claire entre le nombre d'interactions réalisées (*i.e.* le nombre d'infections d'hôtes différents par une même espèce virale) et la richesse virale ou la richesse d'hôte, pourrait à l'inverse invalider cette hypothèse de synergie. Les mastrévirus étant transmis horizontalement, il est aussi largement probable que ces résultats soient déterminés par l'impact du comportement d'alimentation des populations de vecteurs dans les sites (Keesing *et al.*, 2006) et des confirmations ultérieures seront nécessaires pour valider nos premières observations.

Enfin, le fait que les mastrévirus à La Réunion soient inféodés au milieu agricole semble suggérer que leur présence sur cette île volcanique soit principalement issue de l'introduction de matériel végétal infecté, telle que des plantes fourragères ou des adventices des cultures depuis le continent Africain. Ces résultats soulignent aussi l'absence d'une communauté résidente de mastrévirus ayant co-évolué avec les Poales indigènes de La Réunion. La généralisation de ce type d'analyse à un plus grand nombre de genres ou familles virales permettra de nous renseigner sur les relations entre les traits d'histoires de vie viraux et les facteurs environnementaux influant sur les patrons de diversité de ces virus.

Article 2

A novel chimeric circular Rep-encoding single stranded DNA virus with interfamilial gene exchange illustrates the considerable evolutionary capacity of ssDNA viruses

Article 2

A novel chimeric circular Rep-encoding single stranded DNA virus with interfamilial gene exchange illustrates the considerable evolutionary capacity of ssDNA viruses

Sélim Ben Chehida¹, Murielle Hoareau¹, Babbitha Fénélon¹, Martial Grondin¹, Michel Peterschmitt², Arvind Varsani^{3,4}, Darren P. Martin⁵, Pierre-Yves Teycheney¹, Pierre Lefeuvre¹, Jean-Michel Lett¹

¹CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

²CIRAD, UMR PHIM, F-34090 Montpellier, France

³The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, USA

⁴Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Rondebosch, Cape Town, South Africa

⁵Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory, South Africa

Keywords: CRESS DNA viruses, recombination, *Mastrevirus*, *Alphasatellitidae*, helper virus.

Abstract

Viruses in the family *Geminiviridae* cause significant economic losses in many crops worldwide. Geminiviruses are often associated with satellite DNA molecules such as alphasatellites (family *Alphasatellidae*) that require the assistance of a helper virus for their transmission. In this study, we report the discovery of a chimeric virus, tentatively named *Cenchrus purpureus* associated virus (CPAV), in *Cenchrus purpureus* plants in Reunion Island. The genome of CPAV is monopartite and displays unique features. Its *rep* gene is phylogenetically related to alphasatellites and positioned in the viral genome upstream of and in the same orientation as the movement and capsid protein genes, which are both phylogenetically related to mastreviruses (family *Geminiviridae*). We found that CPAV is associated with a new mastrevirus, tentatively named *Cenchrus purpureus* mild streak virus (CPMSV). Using agroinfection and insect transmission assays, we showed that CPAV is able to initiate infection in *C. purpureus* but requires the assistance of CPMSV to establish long-term infection. The chimeric nature of CPAV illustrates how interfamily gene exchange between circular ssDNA viruses and satellites shapes the evolution of ssDNA viruses and has the potential of promoting the emergence of new viruses with modified pathogenicity and host range.

Introduction

A large number of novel circular replication single stranded DNA (CRESS) viruses have been discovered and characterized using viral metagenomic approaches (Rosario *et al.*, 2012; Dayaram *et al.*, 2016). These viruses are called cressdnaviruses and are classified in a phylum called *Cressdnaviricota* (Krupovic *et al.*, 2020). Their genomes encode a conserved replication initiator protein (Rep) involved in rolling circle replication (RCR; Fauquet *et al.*, 2005). The most recently proposed evolutionary scenario for the origin of cressdnaviruses hypothesizes that they originated from recombination events between prokaryotic plasmids carrying *rep* genes and capsid genes from eukaryotic (+)RNA viruses (Dolja *et al.*, 2020).

Although cressdnaviruses have now been identified from all kind of environments and hosts (Rosario *et al.*, 2012; Dayaram *et al.*, 2016), they were first described in plants and in particular within the *Geminiviridae* family (Rybicki, 1994), which currently comprises 14

genera (*Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrilevirus*, *Opunvirus*, *Topilevirus*, *Topocuvirus*, and *Turncurtovirus*; Roumagnac *et al.*, 2022). Viruses in this family are prevalent in many tropical and subtropical crops and wild plants (Rojas *et al.*, 2018). Their genomes are either monopartite and ~2,600-3,600 nucleotides (nt) in size, or bipartite, with two DNA molecules of ~2,600 nt each, termed DNA-A and DNA-B (Zerbini *et al.*, 2017). All genomic DNAs share a virion-strand origin of replication nonanucleotide motif 'TAATATT[↓]AC', regardless of the monopartite or bipartite organization (Roumagnac *et al.*, 2022). They are transmitted by insect vectors, such as leafhoppers (mastreviruses, becurtoviruses, turncurtoviruses), aphids (capulaviruses), whiteflies (begomoviruses) or treehoppers (topocuviruses and grabloviruses).

Some viruses in families *Geminiviridae*, *Nanoviridae* and *Metaxyviridae*, which all have ssDNA genomes, are associated with additional circular single stranded DNA molecules called satellite DNAs (Zerbini *et al.*, 2017; Briddon *et al.*, 2018; Claverie *et al.*, 2020; Thomas *et al.*, 2021; Gronenborn *et al.*, 2020; Walker *et al.*, 2021; Varsani *et al.*, 2021). These satellite DNAs are classified in two families, *Alphasatellitidae* and *Tolecusatellitida*. Satellite DNAs can modulate the pathogenicity of helper viruses, by enhancing (betasatellites) and/or reducing (alphasatellites and deltasatellites) both the accumulation of the helper virus and symptoms (Gnanasekaran and Chakraborty, 2018). Alphasatellites encode a Rep and have a predicted stem-loop structure with, in most cases, the nonanucleotide sequence TAGTATT[↓]AC forming part of the loop (Briddon *et al.*, 2018). Alphasatellites rely on their helper virus for encapsidation, movement and vector transmission, but they replicate autonomously. Family *Alphasatellitidae* is subdivided in three subfamilies, *Geminialphasatellitinae*, *Nanoalphasatellitinae* and *Petromalphasatellitinae* to accommodate the geminivirus-, nanovirus- and metaxy-associated alphasatellites, respectively. Satellite DNAs associated with mastreviruses were first described ten years ago (Rosario *et al.*, 2013 ; Kumar *et al.*, 2014). More recently, Claverie *et al.* (2020) described the first natural association between a geminialphasatellite, Sorghum mastrevirus-associated alphasatellite (SMasA), and a mastrevirus, African streak virus (AfSV), in Reunion Island. African streak viruses are widespread in sub-Saharan Africa and surrounding islands (Comoros, Madagascar, Mayotte, Mauritius, Reunion) and they are transmitted by maize leafhopper (*Cicadulina mbila*; Kraberger *et al.*, 2017).

Here, we report a novel cressdnavirus, tentatively named *Cenchrus purpureus* associated virus (CPAV), naturally associated with a mastrevirus (*Cenchrus purpureus* mild streak virus, CPMSV) in elephant grass (*Cenchrus purpureus*, family *Poaceae*) in Reunion Island. We show that CPAV has a chimeric genome related to alphasatellites and mastreviruses. Using agro-inoculation of infectious clones and vector transmission, we also show that CPAV relies on a helper virus or another DNA component to fulfill its replication and transmission cycles, making it potentially a more complex type of satellite than those previously described.

Materials and Methods

Sampling and DNA extraction

Three shoots were collected in March 2021 at Cirad's Colimaçons experimental outstation (Latitude -21.1300; Longitude 55.3050) in Saint-Leu (Reunion Island) from three *Cenchrus purpureus* (elephant grass) plants. They were labelled 21_REU_E0816_3, 21_REU_E0816_4 and 21_REU_E0816_5 and displayed discontinuous parallel chlorotic streak symptoms (**Figure 1**). All three collected shoots were potted in potting soil and maintained in an insect-proof climatic chamber at 25°C, under 80% relative humidity and a 12-h photoperiod. Symptomatic leaves were collected from each potted plant and grouped as one sample per plant, resulting in a total of three samples. Samples were dehydrated at 50°C for 48 hours and stored at room temperature until further use. Dehydrated leaf samples were homogenized using a TissueLyser II (Qiagen, Courtaboeuf, France) and total DNA was extracted from 20mg of leaf material using the DNeasy Plant Pro DNA extraction kit (Qiagen, Courtaboeuf, France), following the manufacturer's instructions. Purified DNA was stored at -20°C until further use.

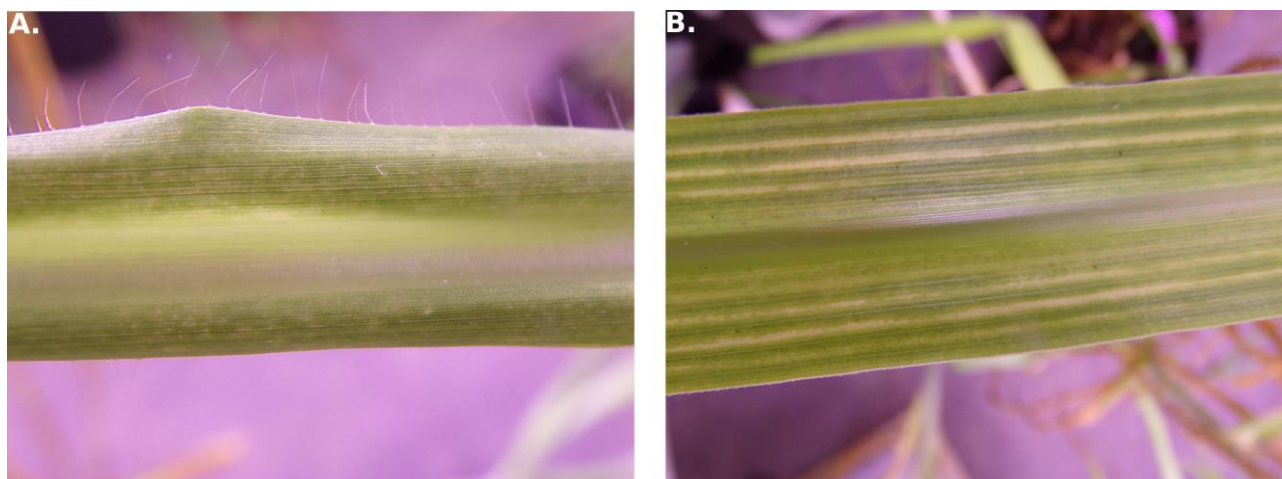


Figure 1. Disease symptoms on *Cenchrus purpureus* leaves collected in March 2021 at Cirad's Colimaçons experimental outstation of (A) *Cenchrus purpureus* mild streak virus (CPMSV) with discontinuous and fine chlorotic streak symptoms and of (B) CPMSV and *Cenchrus purpureus* associated virus (CPAV) co-infection with continuous and large chlorotic streak symptoms.

Virus indexings

Pairs of primers were designed to recover a genomic segment of CPMSV, CPAV and MSV isolates. PCR amplification was achieved using either the primer pair NV3_F185: 5'-TGGAAAGTGGTAATTCGCCC-3' and NV3_R1061: 5'-CGACTCCACTCTGAACTTCC-3' for amplifying a 876 nt-long segment of CPMSV, the primer pair NV4_F241: 5'-AGCTTTGTCCCTCTACTGGT-3' and NV4_R690: 5'-AGGTTGTCTTACCCTCTGCT-3' for amplifying a 449 nt-long segment of CPMSV or the primer pair MSV_F1875: 5'-GGGACTGACCTGGAAGATGT-3' and MSV_R2625: 5'-CTATAAAACAAGGAACGGCGG-3' for amplifying a 750 nt-long segment of MSV. Amplification was carried out using the GoTaq® Master Mix Kit (Promega Corporation, Madison Wisconsin- USA) and the following conditions: an initial denaturation at 95°C for 3 min, 30 cycles at 95°C for 30 sec, 60°C for 30 sec, 72°C for 3 min, and a final extension step at 72°C for 5 min.

MinION sequencing and assembly

Sequence data was produced from the three total DNA samples using the RCA-MinION protocol described by Ben Chehida *et al.* (2021), with the following modification. *Phi29* DNA polymerase was replaced by *EquiPhi29* DNA polymerase (ThermoFisher Scientific, Les Ullis, France) for the rolling circle amplification (RCA) step. The three samples were multiplexed along with a known positive control (DNA extract from a tomato yellow leaf curl virus infected tomato plant) and a known negative control (DNA extracted from a healthy tomato plant) on a single Flongle (FLO-FLG001).

Full cressdnaviruses and alphasatellites genomic components were assembled as described in Ben Chehida *et al.* (2021), following high accuracy basecalling using Guppy v4.2.2 (<https://nanoporetech.com/>) and demultiplexing. Sequence coverage was estimated after raw reads were mapped back to the assembled sequences using minimap2 v2.17 (Li, 2018). All contigs were subjected to a BLAST search against the NCBI nucleotide database for preliminary species assignment.

Full genome cloning and Sanger sequencing

Total genomic DNA extractions were subjected to RCA-RFLP for full genome cloning and Sanger sequencing as described in Inoue-Nagata *et al.* (2004). *Xba*I restriction enzyme was used to linearize monomers of viral genomes.

Phylogenetic analysis

Firstly, an initial BLASTn and BLASTx search (NCBI) were performed for each sequence. Then a research of putative ORFs was done using DNAMAN v.10.0 (Wang, 2016) and results were confirmed by BLASTp. After genome structural annotation, done on UGENE v.43.0 (Okonechnikov *et al.*, 2012), trees were constructed for each ORFs product against reference sequences from the most similar genus (**Supplementary Table 1&2**). All sequence sets were aligned using MAFFT v7.475 (Kato and Standley, 2013). After selecting the best fitting amino acid substitution model using ModelFinder (Kalyaanamoorthy *et al.*, 2017), maximum likelihood (ML) phylogenetic trees were inferred using FastTree2 (Price *et al.*, 2010) with a SH-like approximate likelihood ratio test (Guindon *et al.*, 2010) to test branch supports. Trees were edited using Figtree 1.4.4. (<http://tree.bio.ed.ac.uk/software/figtree/>) and the “ape” R package (Paradis and Schliep, 2019).

Construction of viral agroinfectious clones

CPAV and CPMSV agroinfectious clones were constructed as described by Martin *et al.* (2001). Briefly, head-to-tail partial tandem copies (1.2 mer) of CPAV and CPMSV genomes (3,361 and 3,392 nt, respectively) were synthesized using Epoch Life Science facilities, linearized using EcoRI and cloned into the EcoRI site of the binary vector pCambia 2300 (Hajdukiewicz *et al.*, 1994). Recombinant plasmids were transformed into *Escherichia coli* strain DH5 α by electroporation and then transferred into *Agrobacterium tumefaciens* strain C58 by triparental mating using *E. coli* strain HMB101 harbouring the helper plasmid

pRK2013 (Ditta *et al.*, 1980). The nucleotide sequence of each construct was confirmed by Sanger sequencing performed by Macrogen Europe (Amsterdam, the Netherlands) using primer walking.

Agroinoculation and evaluation of infectivity

Agro inoculations were performed on three-day-old maize seedlings raised from seeds (*Zea mays*, cv. Golden Bantam) and on *C. purpureus* seedlings raised from cuttings (indexed as healthy prior to handling by PCR diagnostic) as described in Martin *et al.* (2001). *A. tumefaciens* containing an agro-infectious clone of MSV-A [R2] (Peterschmitt *et al.*, 1996) was used as a positive control of agroinoculation, whereas LB culture medium and *A. tumefaciens* devoid of recombinant binary vector were used as negative controls of agroinoculation.

Finally, ten individuals of each plant species were used as non-inoculated negative controls (**Table 2**). All plants were maintained in an insect-proof growth climatic chamber as described above. Symptoms were monitored visually at 35 days post inoculation (dpi) and virus indexations were performed by PCR on systemic leaves using specific primer pairs (**Supplementary Table 3**), as described above. Plants indexed positive were indexed again at 70 dpi.

Insect transmission assays

Transmission of CPAV and CPMSV by leafhopper species *Cicadulina mbila* Naudé (Hemiptera: Cicadellidae), which is one of the main vectors of African streak viruses in Africa (Kvarnheden *et al.*, 2016), was evaluated. A non-viruliferous colony of *C. mbila* was reared on pearl millet plants (*Pennisetum glaucum*) in an insect-proof growth climatic chamber at 25°C, under 80% relative humidity and a 12-h photoperiod. Viruliferous insects were obtained after a 72-h acquisition access period (AAP) performed either on field-collected *C. purpureus* plants co-infected by CPAV and CPMSV or on maize plants agroinfected by the MSV-A [R2] infectious clone used as a positive control of agroinoculation (**Table 2**). Following the AAP, insects were transferred on *C. purpureus*, maize, sugarcane (*Sacharum officinarum*, cv. R570) and *Cenchrus echinatus* seedlings (indexed as healthy prior to handling by PCR diagnostic) for a 72-h inoculation access period (IAP), then insects were manually removed and seedlings were sprayed with a systemic insecticide (Confidor®, Bayer). Mock-inoculated plants, using non-viruliferous leafhoppers, were used as negative controls. Plants were maintained in the same growth

conditions as those described above. After 35 days, symptoms of infection were recorded and the presence of viral DNA was assessed in systemic leaves using PCR and specific primer pairs (**Supplementary Table 3**). Plants indexed positive were indexed again at 70 dpi. The experiment was conducted twice (CPMSV+CPAV_A and CPMSV+CPAV_B).

Results

Detection of a novel virus infecting C. purpureus

A total of 176,330 raw reads were obtained from the three analysed *C. purpureus* samples using the RCA-MinION approach (**Supplementary Table 4**). Two complete viral genomes were assembled, that of an undescribed mastrevirus tentatively named *Cenchrus purpureus* mild streak virus (CPMSV) after the symptoms of discontinuous parallel chlorotic streaks observed on the leaves of infected *C. purpureus* plants, and that of a novel cressdnavirus tentatively named *Cenchrus purpureus* associated virus (CPAV). Both viruses were present in samples 21_REU_E0816_3 and 21_REU_E0816_4 whereas only CPMSV was present in sample 21_REU_E0816_5 (**Supplementary Table 4**).

The two CPAV genome sequences assembled from samples 21_REU_E0816_3 and 21_REU_E0816_4 were 2,741 and 2,725 nt long, respectively, whereas the three CPMSV genome sequences assembled from samples 21_REU_E0816_3, 21_REU_E0816_4 and 21_REU_E0816_5 were 2,811, 2,804 and 2,797 nt long, respectively (**Supplementary Table 4**). A functional CPAV genome sequence of 2,734 nt was amplified, cloned into binary vector pCambia 2300 and proved infectious upon agroinoculation of *C. purpureus* (**Table 1**). This 2,734 nt sequence shared 99.4% and 99.6% nucleotide identity with the sequences assembled from samples 21_REU_E0816_3 and 21_REU_E0816_4, respectively. This sequence was considered the sequence of the exemplar isolate of CPAV and used in further molecular characterization of the CPAV genome. Similarly, a CPMSV genome sequence of 2,808 nt was amplified, cloned and sequenced. This sequence shared 99.1%, 99.5% and 99.2% with the sequences assembled from samples 21_REU_E0816_3, 21_REU_E0816_4 and 21_REU_E0816_5, respectively.

Table 1. Infectivity of CPMSV, CPAV and MSV in single or mixed association after agroinoculation of *Cenchrus purpureus* and *Zea mays*, at 35 days post inoculation. Percentage of survival plants infectivity was evaluated based on viral DNA detection by PCR. Proportions are indicated between brackets.

Host species / inoculum	Mock	<i>A. tumefaciens</i>	MSV-A [R2]	CPMSV	CPAV	CPMSV+CPAV
<i>Cenchrus purpureus</i>	0% (0/4)	0% (0/9)	0% (0/9)	0% (0/13)	11% (2/18)	0% (0/15)
<i>Zea mays</i>	0% (0/2)	0% (0/5)	45% (10/22)	0% (0/32)	0% (0/26)	0% (0/23)

Table 2. Transmission rates of CPMSV, CPAV and MSV by *Cicadulina mbila* on four host species. Experiments were conducted with *C. mbila* after viral acquisition on co-infected *Cenchrus purpureus* plants from the field. Percentage of transmission on survival plants was evaluated based on viral DNA detection by PCR. Proportions are indicated between brackets.

Host species	Mock	MSV-A [R2]	CPMSV+CPAV_A			CPMSV+CPAV_B		
			CPMSV	CPAV	CPMSV+CPAV	CPMSV	CPAV	CPMSV+CPAV
<i>Cenchrus purpureus</i>	0% (0/5)	0% (0/13)	6% (1/17)	12% (2/17)	6% (1/17)	0% (0/15)	0% (0/15)	0% (0/15)
<i>Zea mays</i>	0% (0/5)	100% (5/5)	0% (0/8)	0% (0/8)	0% (0/8)	0% (0/6)	0% (0/6)	0% (0/6)
<i>Cenchrus echinatus</i>	0% (0/5)	NA	0% (0/8)	0% (0/5)	0% (0/5)	0% (0/5)	0% (0/5)	0% (0/5)
<i>Saccharum officinarum</i>	0% (0/5)	0% (0/13)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)

CPAV is a novel mastrevirus with a chimeric genome structure

The 2,734 nt genome sequence of CPAV exemplar isolate has three non-overlapping open reading frames (ORFs) oriented in the same orientation (**Figure 2A**), a unique feature among members of the family *Geminiviridae*. ORF1 encodes a putative replication-associated protein (Rep, 309 aa) most closely related to the Rep of melon chlorotic mosaic alphasatellite (YP_003828902), with 64% aa similarity. CPAV putative Rep shares the conserved rolling circle replication (RCR) motifs (motif I : CFTLFE; motif II : KHLQG; motif III : YCTK) and superfamily 3 helicase (SF3) motifs (Walker-A : GSVGAEGKT; Walker-B : CVFDL; Motif C : VFAN) of alphasatellite replication-associated proteins (Rosario *et al.*, 2012; **Supplementary Figure 1**). ORF2 encodes a putative movement protein (MP, 113 aa) most closely related to the MP of maize striate mosaic virus (YP_009551895), a mastrevirus, with 49% aa similarity. ORF3 encodes a putative coat protein (CP, 244 aa) most closely related to the CP of sugarcane white streak virus (NP_620491), a mastrevirus, with 35% aa similarity. ORFs 1 and 3 are separated by a short intergenic region (SIR, 144 nt) and a long intergenic region (LIR, 589 nt) for which we failed to detect any nucleotide sequence similarities with BLASTn (**Figure 3**). However, CPAV LIR includes a stem-loop structure in the putative origin of replication, with the canonical nonanucleotide motif (5'-TAGTATT[↓]AC-3') of geminialphasatellites (Bridson *et al.*, 2018). CPAV LIR also displays the distinctive adenine rich region (A-rich region) of geminialphasatellites (Bridson *et al.*, 2018) at its 5' end (**Supplementary Figure 1**). Lastly,

TATA and GC boxes of the presumed Rep and MP-CP promoters were identified upstream of the ORFs 1 and 2, respectively.

Phylogenetic analysis using Rep amino acid sequences placed CPAV in the geminialphasatellites but did not allow its assignment to a particular genus (**Figure 3A**). However, a similar analysis performed on either CPAV MP or CP amino acid sequences placed CPAV in genus *Mastrevirus* (**Figure 3B**), suggesting that the genome of CPAV has a chimeric structure associating ORF1 from a geminialphasatellite and ORFs 2 and 3 from a mastrevirus.

The complete genome of CPMSV (**Figure 2B**) shared the highest nucleotidic identity (57.9%) with that of *Sorghum arundinaceum* associated virus (MK546381), a mastrevirus. Its organization was typical of mastreviruses with two ORFs on the virion-sense strand coding for a movement protein (55% similar to that of *Sorghum arundinaceum* associated virus YP_010087741) and a capsid protein (52% similar to that of sugarcane striate virus YP_009389275), respectively, and two overlapping ORFs on the complementary-sense strand coding for a replication-associated protein and a replication protein A (50% similar to that of *Sporobolus striate* mosaic virus 2 YP_006666531). However, the genome of CPMSV contains an unfamiliar nonanucleotide sequence TAACATT[↓]GC in the stem-loop located in the large intergenic region and supposed to correspond to the origin of replication.

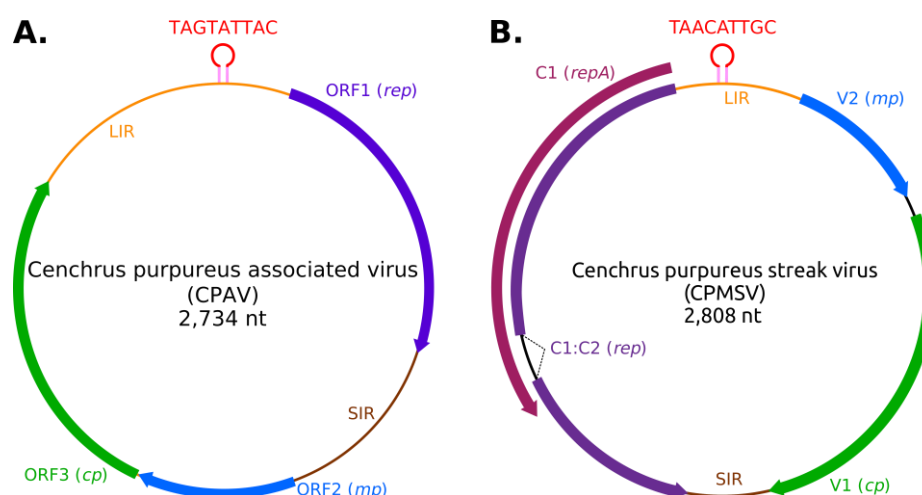


Figure 2. (A) Schematic genome organizations of *Cenchrus purpureus* associated virus (CPAV) and (B) *Cenchrus mild streak virus* (CPMSV). Open reading frames are shown with colored arrows and intergenic regions with thin colored lines. The stem-loop and nonanucleotide motif sequence are also shown.

Infectivity of CPAV infectious clone in the virus natural host

At 35 days post inoculation (dpi), 45% (10/22) of the survival maize plants agroinoculated with the MSV-A [R2] infectious clone were tested positive for MSV-A and presented leaf streak symptoms (**Table 1**). For the *C. purpureus* survival plants agroinoculated with CPAV, 11% (2/18) were tested positive by PCR at 35 dpi and 70 dpi although they remained asymptomatic throughout the experiment. The sequencing of the amplicons confirmed the presence of the viruses and the specificity of the PCR-based detection test. None of the *C. purpureus* survival plants agroinoculated with CPMSV or CPMSV+CPAV were tested positive by PCR.

Leafhopper transmission of CPAV and CPMSV

At 35 dpi, all five *Z. mays* plants used for transmission of MSV-A [R2] by *C. mbila* were indexed positive by PCR and showed severe leaf streak symptoms. On the opposite, none of the *C. echinatus* and sugarcane plants used in a similar insect transmission assay was infected by MSV-A [R2] and/or symptomatic. At 35 dpi, 12.5% (4/32; repetitions A and B) of the *C. purpureus* survival plants used for assessing the transmission of CPMSV+CPAV by *C. mbila* were indexed positive by PCR either for CPMSV (1 plant), CPAV (2 plants) or jointly CPAV and CPMSV (1 plant) (**Table 2**). All four infected plants were asymptomatic. At 70 dpi, the two plants previously indexed positive for CPAV became negative while the plant co-infected by CPAV and CPMSV remained positive for the same viruses. However, the plant found infected by CPMSV alone at 35 dpi was indexed positive by PCR for both CPAV and CPMSV at 70 dpi. All positive PCR results were confirmed by sequencing of amplicons.

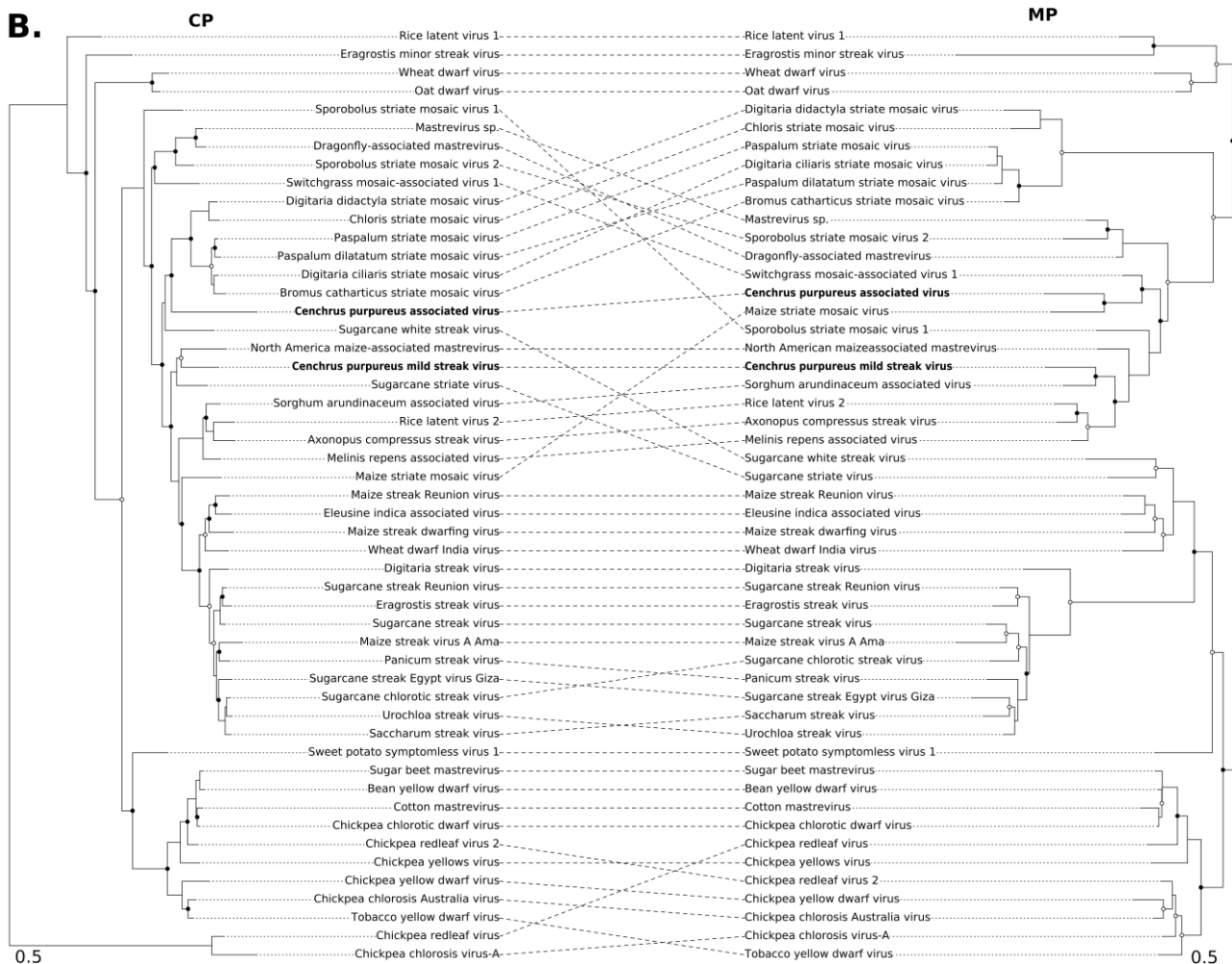
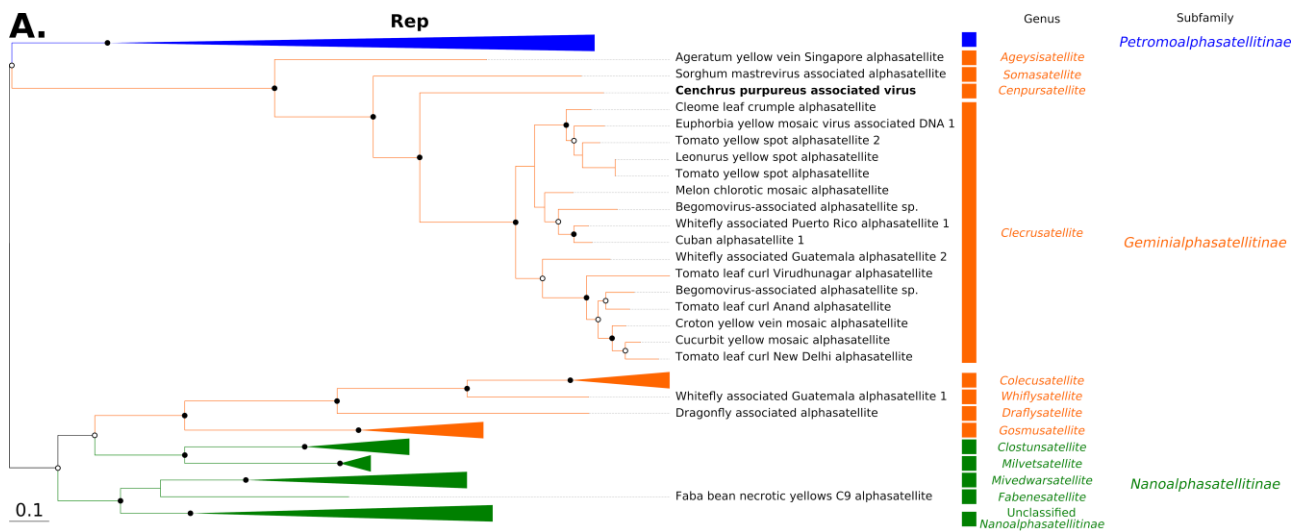


Figure 3. (A) Unrooted maximum likelihood tree inferred from aligned replication associated protein (Rep) sequences of viruses in family *Alphasatellitidae*. **(B)** On the left, unrooted maximum likelihood tree inferred from aligned capsid protein (CP) and on the right, unrooted neighbor-joining tree inferred from aligned movement protein (MP) amino acid sequences of reference sequences from various mastreviruses. CPAV is shown in bold font. Open and closed circles on nodes indicate bootstrap support for the branches to their left of 70-89% and $\geq 90\%$ respectively.

Discussion

Discovery of an original chimeric ssDNA viral genome

Here, we describe a novel virus, *Cenchrus purpureus* associated virus (CPAV), whose genome displays a unique architecture. Its *rep* gene and LIR are most closely related to those of alphasatellites (family *Alphasatellitidae*) although its movement and capsid protein genes are most closely related to those of mastreviruses (family *Geminiviridae*). Interestingly, CPAV *rep* gene and LIR together represent 1,516 nt, well within the size range of alphasatellite genomes, suggesting that this peculiar viral genome could be the progeny of an alphasatellite that acquired and domesticated the *mp* and *cp* genes from a currently undescribed mastrevirus. Inter-family recombination between alphasatellites and begomoviruses has been reported previously (Saunders and Stanley, 1999; Shuja *et al.*, 2017). These findings suggest that both ssDNA viruses and satellites, despite their evolutionary divergence, co-infect the same host cells, have the ability to exchange their genetic information, and that the resulting recombinant architecture carries a functional replication protein, enabling these recombinant viruses to multiply within their hosts.

To accommodate for the diversity of ssDNA viruses, a classification into eight types was proposed by Rosario *et al.* (2012). Geminiviruses belong to type II, based on the ambisense orientation of their major ORFs, whereas alphasatellites are part of the type VII, based on their genome carrying a single ORF encoding a Rep protein. According to this classification, CPAV should be classified in group V, whose members have genomes with all ORFs in the same orientation.

Evolution of circular ssDNA viruses

CPAV adds to a growing list of chimeric cressdnviruses (Roux *et al.*, 2013), providing further evidence that horizontal transfer (HT) of genetic material is paramount in the making of the *Cressdnviricota* phylum. HT is known to play a key role in the evolution of viruses through homologous recombination (HR), another driving force of viral evolution (Dolja and Koonin, 2018 ; Martin *et al.*, 2011). It has been hypothesized that HT drives the adaptation of viruses to new vectors (Rybicki, 1994) and hosts (Lefeuvre and Moriones, 2015), and helps maximizing fitness and transmission rates. Recombination provides the opportunity for larger jumps in the adaptive landscapes than what would be possible by mutation alone (Crameri *et al.*, 1998; Crespo-Bellido and Duffy, 2023). Although the propensity of CRESS DNA viruses to recombine through recombination-dependent

replication (RDR; Preiss and Jeske, 2003 ; Erdmann *et al.*, 2010) is most probably paramount for the emergence of viruses with mixed origin, most of these horizontal exchange would certainly result in detrimental and negatively selected combinations later purged from viral populations (Martin *et al.*, 2005 ; Martin *et al.*, 2011). In the case of CPAV, it is likely that the relative conservation of an alphasatellite backbone would facilitate the integration of other ORFs into a functional ensemble.

Is CPAV a genuine autonomous virus?

The ca. 2.8kb genome size and the presence of a CP similar in sequence and size to those of other mastreviruses would suggest that CPAV forms twinned quasi-icosahedral particles (Frischmuth *et al.*, 2001) and follows a transmission process similar to those of mastreviruses. Indeed, transmission assays demonstrated that CPAV was transmitted by the leafhopper *C. mbila* and was able to initiate infection. Our experimental data also suggest that CPAV is not able to sustain single infections in the long term and may be dependent on CPMSV to complete infection cycles and efficient transmission. However, the absence of detection of CPAV at 70 dpi could result from a decreased viral load over time, below the PCR detection threshold, a situation that was encountered for other geminiviruses (Pooggin, 2018) and attributed to the triggering of post transcriptional gene silencing (PTGS)-based defense mechanisms (Chellappan *et al.*, 2004). Additional analyses, such as investigating the presence of viral short interfering RNAs (Kalantidis *et al.*, 2008; Chiwood and Timmermans, 2010), are needed to determine whether CPAV infection triggers PTGS in *C. purpureus*. Noticeably, CPMSV infection could not be experimentally established in *C. purpureus* following agro inoculation of a CPMSV infectious clone but proved successful following insect transmission from a plant co infected by CPMSV and CPAV. This result suggests that CPAV may also act as helper virus in vector transmission of CPMSV. Mutualistic interactions, which are known to promote viral synergism and host resistance bypassing (Fondong *et al.*, 2000 ; Tollenaere *et al.*, 2016 ; Fontenele *et al.*, 2021 ; Armand *et al.*, 2023), were previously reported for members of family *Geminiviridae* in different situations including coinfections by a luteovirus (barley yellow dwarf virus) and a mastrevirus (wheat dwarf virus) in wheat (Armand *et al.*, 2023), coinfections by two viral components of a begomovirus such as pepper yellow vein Mali virus DNA-A and DNA-B in tomato (Ouattara *et al.*, 2022), or coinfections of a begomovirus (Euphorbia yellow mosaic virus) and an alphasatellite

(Euphorbia yellow mosaic alphasatellite) in *Euphorbia heterophylla* plants (Mar *et al.*, 2017).

Importantly, no part of the CPAV genome displayed a close similarity to any known virus present on GenBank database, including ssDNA viruses with circular genomes that were identified recently in Reunion through circomics approaches (Claverie *et al.*, 2019, Claverie *et al.*, 2023), suggesting that the parental viruses of CPAV are not highly prevalent in Reunion or may even be extinct. The detection of CPAV from multiple plant samples provides a strong indication that CPAV is not a transient viral construct but rather represents an actual circulating virus. Further studies and surveys of *C. purpureus* in Reunion and elsewhere are now needed to investigate the prevalence of this novel virus and its possible invasion routes.

Authors information

Affiliations

CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 chemin de l'Irat, 97410 Saint Pierre, La Réunion, France

Sélim Ben Chehida, Murielle Hoarau, Babbitha Fénélon, Martial Grondin, Jean-Michel Lett & Pierre Lefeuvre

CIRAD, UMR BGPI, TA A-54/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

Michel Peterschmitt

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

Arvind Varsani

Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory 7925, South Africa

Darren P. Martin

Author contributions

S.B.C., P.L. and J.M.L. conceived and designed the experiments. S.B.C., M.H., B.F., M.G., P.L. and J.M.L. performed the experiments. S.B.C., P.L. and J.M.L. analysed the data. S.B.C, M.P., A.V., D.P.M., P.L. and J.M.L wrote the paper. J.M.L. and P.L. secured funding for the project's execution.

Corresponding author

Correspondence to Jean-Michel Lett

Funding

This study was funded by the European Union (ERDF, contract GURDT I2016-1731-0006632), the *Conseil Régional de la Réunion* and CIRAD. S.B.C. is a recipient of a PhD fellowship from CIRAD and ANR (Phytovirus project number: ANR-19-CE35-0008-02).

Data Availability Statement

Publicly available datasets were analyzed in this study.

Acknowledgments

This work was conducted on the Plant Protection Platform (3P, IBISA).

Conflicts of Interest

The authors declare no conflict of interest.

References

- Armand, T., Souquet, M., Korn, L., Gauthier, K., & Jacquot, E. (2023). Asymmetric interactions between barley yellow dwarf virus-PAV and wheat dwarf virus in wheat. *Frontiers in Plant Science*, *14*, 1194622.
- Ben Chehida, S., Filloux, D., Fernandez, E., Moubset, O., Hoareau, M., Julian, C., Blondin, L., Lett, J. M., Roumagnac, P., & Lefeuvre, P. (2021). Nanopore sequencing is a credible alternative to recover complete genomes of Geminiviruses. *Microorganisms*, *9*, 903.
- Briddon, R. W., Martin, D. P., Roumagnac, P., Navas-Castillo, J., Fiallo-Olivé, E., Moriones, E., Lett, J.-M., Zerbini, F. M. & Varsani, A. (2018). *Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus-and nanovirus-associated alphasatellites. *Archives of virology*, *163*, 2587-2600.
- Chellappan, P., Vanitharani, R., & Fauquet, C. M. (2004). Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. *Journal of Virology*, *78*(14), 7465-7477.

- Chitwood, D. H., & Timmermans, M. C. (2010). Small RNAs are on the move. *Nature*, 467(7314), 415-419.
- Claverie, S., Ouattara, A., Hoareau, M., Filloux, D., Varsani, A., Roumagnac, P., Lett, J. M. & Lefeuvre, P. (2019). Exploring the diversity of Poaceae-infecting mastreviruses on Reunion Island using a viral metagenomics-based approach. *Scientific reports*, 9(1), 1-11.
- Claverie, S., Varsani, A., Hoareau, M., Filloux, D., Roumagnac, P., Martin, D. P., Lefeuvre, P. & Lett, J. M. (2020). Sorghum mastrevirus-associated alphasatellites: new geminalphasatellites associated with an African streak mastrevirus infecting wild Poaceae plants on Reunion Island. *Archives of Virology*, 165, 1925-1928.
- Claverie, S., Hoareau, M., Ben Chehida, S., Filloux, D., Varsani, A., Roumagnac, P., Martin, D. P., Lett, J. M. & Lefeuvre, P. (2023). Metagenomics reveals the structure of Mastrevirus-host interaction network within an agro-ecosystem. *Virus Evolution*, vead043.
- Cramer, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391: 288-291.
- Crespo-Bellido, A., & Duffy, S. (2023). The how of counter-defense: viral evolution to combat host immunity. *Current Opinion in Microbiology*, 74, 102320.
- 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.
- Ditta, G., Stanfield, S., Corbin, D., & Helinski, D. R. (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(12), 7347-7351.
- Dolja, V. V., & Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus research*, 244, 36-52.
- Dolja, V. V., Krupovic, M., & Koonin, E. V. (2020). Deep roots and splendid boughs of the global plant virome. *Annual Review of Phytopathology*, 58, 23-53.
- Erdmann, J. B., Shepherd, D. N., Martin, D. P., Varsani, A., Rybicki, E. P., & Jeske, H. (2010). Replicative intermediates of maize streak virus found during leaf development. *Journal of General Virology*, 91(4), 1077-1081.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (Eds.). (2005). *Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses*. Academic Press.
- Figtree. Available online: <http://tree.bio.ed.ac.uk/software/figtree/> (accessed on 31/08/2023).
- Fondong, V. N., Thresh, J. M., & Fauquet, C. (2000). Field experiments in Cameroon on cassava mosaic virus disease and the reversion phenomenon in susceptible and resistant cassava cultivars. *International Journal of Pest Management*, 46(3), 211-217.
- Fontenele, R. S., Bhaskara, A., Cobb, I. N., Majure, L. C., Salywon, A. M., Avalos-Calleros, J. A., Argüello-Astorga G. R., Schmidlin, K., Roumagnac, P., Ribeiro S. G., Kraberger, S., Martin, D. P., Lefeuvre, P. & Varsani, A. (2021). Identification of the begomoviruses squash leaf curl virus and watermelon chlorotic stunt virus in various plant samples in North America. *Viruses*, 13(5), 810.
- Frischmuth, T., Ringel, M., & Kocher, C. (2001). The size of encapsidated single-stranded DNA determines the multiplicity of African cassava mosaic virus particles. *Journal of General Virology*, 82(3), 673-676.
- Gnanasekaran, P., & Chakraborty, S. (2018). Biology of viral satellites and their role in pathogenesis. *Current opinion in virology*, 33, 96-105.

- Gronenborn, B., Randles, J. W., Vetten, H. J., & Thomas, J. E. (2020). Create one new family (Metaxyviridae) with one new genus (Cofodevirus) and one species (Coconut foliar decay virus) moved from the family Nanoviridae (Mulpavirales). *Int. Comm. Taxon. Viruses Proposition*, 2021, 022P.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321.
- Hajdukiewicz, P., Svab, Z., & Maliga, P. (1994). The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Molecular Biology*, 25, 989-994.
- 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 ϕ 29 DNA polymerase. *Journal of Virological Methods*, 116(2), 209-211.
- Kalantidis, K., Schumacher, H. T., Alexiadis, T., & Helm, J. M. (2008). RNA silencing movement in plants. *Biology of the Cell*, 100(1), 13-26.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jeremiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587-589.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.
- Kraberger, S., Saumtally, S., Pande, D., Khoodoo, M. H., Dhayan, S., Dookun-Saumtally, A., Shepherd, D. N., Hartnady, P., Atkinson, R., Lakay, F. M., Hanson, B., Redhi, D., Monjane, A. L., Windram, O. P., Walters, M., Oluwafemi, S., Lett, J. M., 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.
- Krupovic, M., Varsani, A., Kazlauskas, D., Breitbart, M., Delwart, E., Rosario, K., Yutin, N., Wolf, Y. I., Harrach, B., Zerbini, M. F., Dolja, V. V., Kuhn, J. H. & Koonin, E. V. (2020). Cressdnaviricota: a virus phylum unifying seven families of rep-encoding viruses with single-stranded, circular DNA genomes. *Journal of Virology*, 94(12), 10-1128.
- 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(12):7093–7104.
- Kvarnheden, A., Lett, J. M., & Peterschmitt, M. (Eds.) (2016). CHAPTER 16: Mastreviruses: Tropical and Temperate Leafhopper-Borne Geminiviruses. In *Vector-mediated transmission of plant pathogens* (pp. 231-241). The American Phytopathological Society.
- 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.
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094–3100.
- Mar, T. B., Mendes, I. R., Lau, D., Fiallo-Olivé, E., Navas-Castillo, J., Alves, M. S., & Murilo Zerbini, F. (2017). Interaction between the New World begomovirus Euphorbia yellow mosaic virus and its associated alphasatellite: effects on infection and transmission by the whitefly Bemisia tabaci. *Journal of General Virology*, 98(6), 1552-1562.
- Martin, D. P., Willment, J. A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., James, D. & Rybicki, E. P. (2001). Sequence diversity and virulence in *Zea mays* of Maize streak virus isolates. *Virology*, 288(2), 247-255.
- Martin, D. P., Van der Walt, E., Posada, D., & Rybicki, E. P. (2005). The evolutionary value of recombination is constrained by genome modularity. *PLoS Genetics*, 1(4), e51.

- Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., & Varsani, A. (2011). 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. (2011). Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. *PLoS Pathogens*, 7(9), e1002203.
- Okonechnikov, K., Golosova, O., Fursov, M., & Ugene Team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8), 1166-1167.
- Ouattara, A., Tiendrébéogo, F., Becker, N., Urbino, C., Thébaud, G., Hoareau, M., Allibert, A., Chiroleu, F., Vernerey, M.-S., Traoré, E. V., Barro, N., Traoré, O., Lefeuvre, P. & Lett, J. M. (2022). Synergy between an emerging monopartite begomovirus and a DNA-B component. *Scientific Reports*, 12(1), 695.
- Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.
- Peterschmitt, M., Granier, M., Frutos, R., & Reynaud, B. (1996). Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Reunion Island. *Archives of Virology*, 141, 1637-1650.
- Pooggin, M. M. (2018). Small RNA-omics for plant virus identification, virome reconstruction, and antiviral defense characterization. *Frontiers in Microbiology*, 9, 2779.
- Preiss, W., & Jeske, H. (2003). Multitasking in replication is common among geminiviruses. *Journal of Virology*, 77(5), 2972-2980.
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS one*, 5(3), e9490.
- Rojas, M. R., Macedo, M. A., Maliano, M. R., Soto-Aguilar, M., Souza, J. O., Briddon, R. W., Kenyon, L., Rivera Bustamante, R. F., Zerbini, F. M., Adkins S., Legg, J., P., Kvarnheden, A., Wintermantel, W. M., Sudarshana, M. R., Peterschmitt, M., Lapidot, M., Martin, D. P., Moriones, E., Inoue-Nagata, A. K. & Gilbertson, R. L. (2018). World management of geminiviruses. *Annual Review of Phytopathology*, 56, 637-677.
- Rosario, K., Duffy, S., & Breitbart, M. (2012). A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Archives of Virology*, 157, 1851-1871.
- Rosario, K., Padilla-Rodriguez, M., Kraberger, S., Stainton, D., Martin, D. P., Breitbart, M. & Varsani, A. (2013). Discovery of a novel mastrevirus and alphasatellite-like circular DNA in dragonflies (*Epirocta*) from Puerto Rico. *Virus Research* 171(1):231–237.
- Roumagnac, P., Lett, J. M., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F. M., Martin, D. P., & Varsani, A. (2022). Establishment of five new genera in the family Geminiviridae: Citlodavirus, Maldovirus, Mulcrilevirus, Opunvirus, and Topilevirus. *Archives of Virology*, 1-16.
- 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(1), 2700.
- Rybicki, E. P. (1994). A phylogenetic and evolutionary justification for three genera of Geminiviridae. *Archives of Virology*, 139, 49-77.
- 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(1), 142-152.

- Shuja, M. N., Tahir, M., & Briddon, R. W. (2017). Occurrence of a recombinant molecule carrying sequences derived from an alphasatellite and the helper virus in cotton affected with cotton leaf curl disease. *Tropical Plant Pathology*, 42, 397-402.
- Thomas, J. E., Gronenborn, B., Harding, R. M., Mandal, B., Grigoras, I., Randles, J. W., Sano, Y., Timchenko, T., Vetten, H. J., Yeh, H.-H., Ziebell, H. & ICTV Report Consortium. (2021). ICTV virus taxonomy profile: Nanoviridae. *Journal of General Virology*, 102(3), 001544.
- Tollenaere, C., Susi, H., & Laine, A. L. (2016). Evolutionary and epidemiological implications of multiple infection in plants. *Trends in Plant Science*, 21(1), 80-90.
- Varsani, A., Martin, D. P., Randles, J. W., Vetten, H. J., Thomas, J. E., Fiallo-Olivé, E., Navas-Castillo, J., Lett, J.-M., Zerbini, F. M., Roumagnac, P. & Gronenborn, B. (2021). Taxonomy update for the family Alphasatellitidae: new subfamily, genera, and species. *Archives of Virology*, 166, 3503-3511.
- Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., Davidson, A. J., Dempsey, D. M., Dutilh, B. E., Garcia, M. L., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knwoles, N. J., Krupovic, M., Kuhn, J. H., Lambert, A., J., Lobočka M., Nibert, M. L., Oksanen, H. M., Orton, R. J., Robertson, D. L., Rubino, L., Sabanadzovic, S., Simmonds, P., Smith, D. B., Suzuki, N., Van Doorslaer, K., Vandamme, A.-M., Varsani, A. & Zerbini, F. M. (2021). Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). *Archives of Virology*, 166(9), 2633-2648.
- Wang, W. (2016). The molecular detection of *Corynespora Cassiicola* on cucumber by PCR assay using DNAMAN software and NCBI. In *Computer and Computing Technologies in Agriculture IX: 9th IFIP WG 5.14 International Conference, CCTA 2015, Beijing, China, September 27-30, 2015, Revised Selected Papers, Part II 9* (pp. 248-258). Springer International Publishing.
- Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A. & ICTV Report Consortium. (2017). ICTV virus taxonomy profile: Geminiviridae. *Journal of General Virology*, 98(2), 131-133.

Supplementary Table 1. List of the *Alphasatellitidae* species with the accession numbers of the nucleotide complete and amino acid replication associated protein (Rep) sequences.

<i>Alphasatellitidae</i> species	Complete	Rep
Tomato leaf curl Pakistan alphasatellite	NC_012789	YP_002941920
Parthenium leaf curl alphasatellite	NC_030226	YP_009254744
Begomovirus-associated alphasatellite sp. TW	KU923760	AOY34818
Leonurus yellow spot alphasatellite	KX348231	AQT01637
Tomato leaf curl New Delhi alphasatellite	MH550542	QAB08755
Begomovirus-associated alphasatellite sp. UHAsV-1.CD-W	MN203219	QHB15188
Sorghum mastrevirus associated alphasatellite	MN901968	QKK13640
Tomato yellow spot alphasatellite 2	NC_055597	YP_010087853
Parsley severe stunt alphasatellite 4	NC_055566	YP_010087622
Parsley severe stunt alphasatellite 3	NC_055565	YP_010087621
Tomato leaf curl Anand alphasatellite	NC_055563	YP_010087615
Faba bean necrotic yellows virus associated alphasatellite 1	NC_055480	YP_010086795
Tomato leaf curl Virudhunagar alphasatellite	NC_055478	YP_010086787
Eclipta yellow vein alphasatellite	NC_055475	YP_010086775
Coconut foliar decay alphasatellite 5	NC_055553	YP_010087372
Coconut foliar decay alphasatellite 4	NC_055552	YP_010087370
Coconut foliar decay alphasatellite 2	NC_055551	YP_010087364
Coconut foliar decay alphasatellite 7	NC_040812	YP_009553579
Coconut foliar decay alphasatellite 6	NC_040697	YP_009553054
Coconut foliar decay alphasatellite 3	NC_040624	YP_009552699
Tomato yellow leaf curl China alphasatellite	NC_043449	YP_009666132
Banana bunchy top alphasatellite 2	NC_038953	YP_009508281
Whitefly associated Guatemala alphasatellite 1	NC_038974	YP_009508355
Banana bunchy top alphasatellite 3	NC_038955	YP_009508283
Banana bunchy top alphasatellite 1	NC_038892	YP_009508165
Sida leaf curl alphasatellite	NC_039240	YP_009513247
Malvastrum yellow mosaic alphasatellite	NC_039239	YP_009513246
Chilli leaf curl alphasatellite	NC_039234	YP_009513231
Vernonia yellow vein Fujian alphasatellite	NC_039233	YP_009513230
Gossypium darwinii symptomless alphasatellite	NC_039232	YP_009513229
Tomato yellow leaf curl Yunnan alphasatellite	NC_039229	YP_009513233
Hollyhock yellow vein virus associated symptomless alphasatellite	NC_039084	YP_009508886
Bendhi Yellow Vein Mosaic/Mesta Yellow Vein Mosaic alphasatellite	NC_039083	YP_009508885
Nanovirus-like particle Egypt	NC_039081	YP_009508879

CHAPITRE 2

Nanovirus-like particle Singapore	NC_039080	YP_009508878
Nanovirus-like particle Pakistan	NC_039079	YP_009508877
Sophora alopecuroides yellow stunt alphasatellite 9	NC_039000	YP_009508448
Sophora alopecuroides yellow stunt alphasatellite 4	NC_038999	YP_009508447
Sophora alopecuroides yellow stunt alphasatellite 7a	NC_038998	YP_009508446
Sophora alopecuroides yellow stunt alphasatellite 1a	NC_038997	YP_009508445
Sophora alopecuroides yellow stunt alphasatellite 8	NC_038996	YP_009508444
Tomato yellow spot alphasatellite	NC_038993	YP_009508434
Whitefly associated Puerto Rico alphasatellite 1	NC_038975	YP_009508356
Whitefly associated Guatemala alphasatellite 2	NC_038973	YP_009508354
Pea necrotic yellow dwarf alphasatellite 3	NC_038959	YP_009508292
Pea necrotic yellow dwarf alphasatellite 1	NC_038958	YP_009508291
Gossypium mustelinum symptomless alphasatellite	NC_038954	YP_009508282
Tomato leaf curl Cameroon alphasatellite CM:TOS2D1:To:09	NC_038903	YP_009508213
Tomato leaf curl Cameroon alphasatellite CM:OMHD3:Ok:09	NC_038902	YP_009508212
Okra yellow crinkle Cameroon alphasatellite	NC_014906	YP_004123952
Ageratum yellow vein alphasatellite	NC_033555	YP_009344825
Pedilanthus leaf curl alphasatellite	NC_033335	YP_009338004
Tomato leaf curl alphasatellite	NC_033274	YP_009337831
Hibiscus leaf curl alphasatellite	NC_033273	YP_009337830
Ageratum yellow vein Pakistan alphasatellite	NC_033271	YP_009337826
Cucurbit yellow mosaic alphasatellite	NC_029907	YP_009246456
Bhindi yellow vein alphasatellite	NC_029313	YP_009230209
Ageratum conyzoides symptomless alphasatellite	NC_028116	YP_009174978
Sida yellow vein mosaic alphasatellite	NC_028115	YP_009174977
Chilli leaf curl India alphasatellite	NC_027431	YP_009154718
Synedrella leaf curl alphasatellite	NC_025244	YP_009073582
Faba bean necrotic yellows C1 alphasatellite	NC_024886	YP_009058890
Ageratum conyzoides associated symptomless alphasatellite	NC_024710	YP_009052483
Sida yellow vein alphasatellite	NC_024372	YP_009044083
Faba bean necrotic stunt alphasatellite 1	NC_023881	YP_009021881
Faba bean necrotic stunt alphasatellite 2	NC_023880	YP_009021880
Black medic leafroll alphasatellite 1	NC_023875	YP_009021872
Sida yellow vein China alphasatellite	NC_023484	YP_009002582
Ageratum yellow vein China alphasatellite	NC_023443	YP_009001905
Cardamom bushy dwarf alphasatellite	NC_022919	YP_008854133
Cuban alphasatellite 1	NC_021708	YP_008169853
Guar leaf curl alphasatellite	NC_020502	YP_007518505

CHAPITRE 2

Ageratum enation alphasatellite	NC_019547	YP_007011043
Ageratum yellow vein India alphasatellite	NC_019546	YP_007011042
Sunflower leaf curl Karnataka alphasatellite	NC_019499	YP_007004041
Dragonfly associated alphasatellite	NC_019498	YP_007004040
Cassava mosaic Madagascar alphasatellite	NC_018628	YP_006742179
Bhendi yellow vein mosaic virus-associated alphasatellite	NC_018574	YP_006666513
Mesta yellow vein mosaic alphasatellite	NC_018573	YP_006666512
Cotton leaf curl Multan alphasatellite	NC_018082	YP_006488615
Cotton leaf curl Lucknow alphasatellite	NC_015327	YP_004347413
Malvastrum yellow mosaic Cameroon alphasatellite	NC_014907	YP_004123953
Tomato leaf curl Cameroon alphasatellite	NC_014747	YP_004046695
Cleome leaf crumple alphasatellite	NC_014646	YP_003987456
Euphorbia yellow mosaic virus associated DNA 1	NC_014630	YP_003966132
Tobacco leaf curl PUSA alphasatellite	NC_014597	YP_003934916
Melon chlorotic mosaic alphasatellite	NC_014379	YP_003828902
Croton yellow vein mosaic alphasatellite	NC_013801	YP_003433564
Cotton leaf curl Burewala alphasatellite	NC_013803	YP_003433566
Cotton leaf curl Gezira alphasatellite	NC_013593	YP_003334471
Chilli leaf curl Multan alphasatellite	NC_013103	YP_003104752
Gossypium davidsonii symptomless alphasatellite	NC_013011	YP_003082245
Okra leaf curl alphasatellite Mali	NC_010620	YP_001856218
Mimosa yellow leaf curl virus-associated DNA 1	NC_009564	YP_001285943
Sida yellow vein Vietnam alphasatellite	NC_009563	YP_001285945
Okra leaf curl alphasatellite Pakistan	NC_005954	YP_031730
Tobacco curly shoot alphasatellite	NC_005057	NP_878243
Subterranean clover stunt C6 alphasatellite	NC_003818	NP_620700
Subterranean clover stunt C2 alphasatellite	NC_003814	NP_620696
Milk vetch dwarf C10 alphasatellite	NC_003647	NP_619768
Milk vetch dwarf C3 alphasatellite	NC_003640	NP_619761
Milk vetch dwarf C2 alphasatellite	NC_003639	NP_619760
Milk vetch dwarf C1 alphasatellite	NC_003638	NP_619759
Faba bean necrotic yellows C9 alphasatellite	NC_003567	NP_619574
Faba bean necrotic yellows C7 alphasatellite	NC_003565	NP_619572
Faba bean necrotic yellows C11 alphasatellite	NC_003558	NP_619565
Ageratum yellow vein Singapore alphasatellite	NC_003414	NP_579867
Coconut foliar decay alphasatellite	NC_001465	NP_040944

Supplementary Table 2. List of the *Mastrevirus* species with the accession numbers of the nucleotide complete, amino acid movement protein (MP) and capsid associated protein (CP) sequences.

<i>Mastrevirus</i> Species	Complete	CP	MP
Axonopus compressus streak virus	NC_023864	YP_009021762	YP_009021761
Bean yellow dwarf virus	NC_003493	NP_612220	NP_612219
Bromus catharticus striate mosaic virus	NC_014822	YP_004089626	YP_004089625
Chickpea chlorosis Australia virus	NC_022131	YP_008472703	YP_008472702
Chickpea chlorosis virus	NC_014740	YP_004046666	YP_004046665
Chickpea chlorotic dwarf virus	NC_011058	YP_002014711	YP_002014710
Chickpea redleaf virus	NC_014739	YP_004046662	YP_004046661
Chickpea redleaf virus 2	NC_055579	YP_010087771	YP_010087770
Chickpea yellow dwarf virus	NC_025475	YP_009104362	YP_009104361
Chickpea yellows virus	NC_038478	YP_009506576	YP_009506575
Chloris striate mosaic virus	NC_001466	NP_040952	NP_040951
Cotton mastrevirus	HE956706	CCI88379	CCI88378
Digitaria ciliaris striate mosaic virus	NC_018579	YP_006666534	YP_006666533
Digitaria didactyla striate mosaic virus	NC_014547	YP_003915158	YP_003915157
Digitaria streak virus	NC_001478	NP_040963	NP_040962
Dragonfly-associated mastrevirus	NC_019497	YP_007004037	YP_007004036
Eleusine indica associated virus	NC_055572	YP_010087739	YP_010087738
Eragrostis minor streak virus	NC_015553	YP_004465363	YP_004465362
Eragrostis streak virus	NC_010352	YP_001686793	YP_001686792
Maize streak dwarfing virus	NC_055519	YP_010087227	YP_010087226
Maize streak Reunion virus	NC_017917	YP_006331072	YP_006331071
Maize streak virus	NC_001346	YP_009154762	YP_009154761
Maize striate mosaic virus	NC_040541	YP_009551896	YP_009551895
Mastrevirus sp.	MN203180	QHB15156	QHB15157
Melinis repens associated virus	MK546380	QDO73337	QDO73336
North American maize-associated mastrevirus	MZ852895	ULE36139	ULE36138
Oat dwarf virus	NC_010799	YP_001941161	YP_001941160
Panicum streak virus	NC_001647	NP_042589	NP_042588
Paspalum dilatatum striate mosaic virus	NC_018576	YP_006666522	YP_006666521
Paspalum striate mosaic virus	NC_018530	YP_006659973	YP_006659972
Rice latent virus 1	NC_040792	YP_009553474	YP_009553473
Rice latent virus 2	NC_040537	YP_009551799	YP_009551798
Saccharum streak virus	NC_013464	YP_003288767	YP_003288766
Sorghum arundinaceum associated virus	NC_055573	YP_010087742	YP_010087741

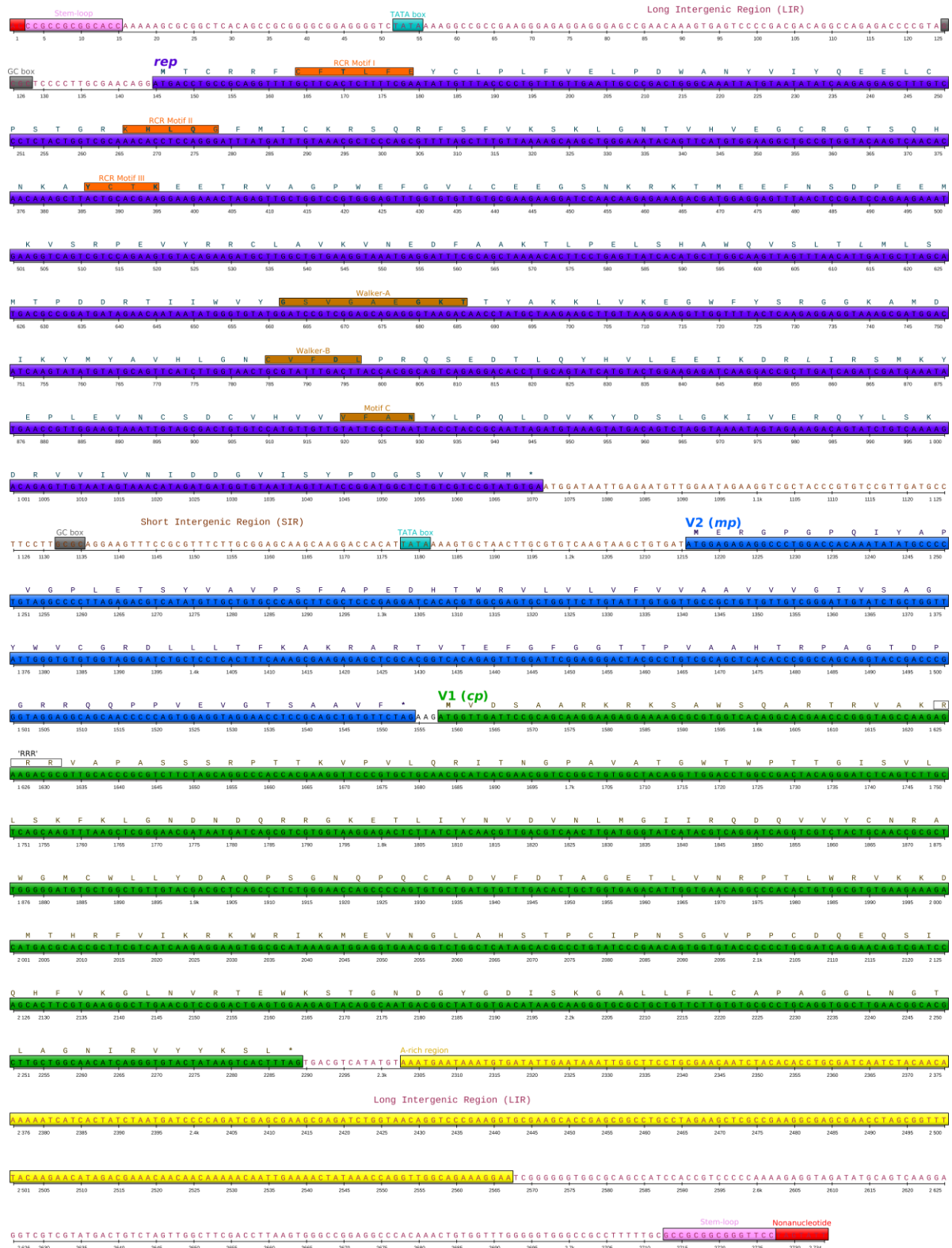
Sporobolus striate mosaic virus 1	NC_018577	YP_006666526	YP_006666525
Sporobolus striate mosaic virus 2	NC_018578	YP_006666530	YP_006666529
Sugar beet mastrevirus	EU034169	ABV82711	ABV82710
Sugarcane chlorotic streak virus	NC_032004	YP_009325924	YP_009325923
Sugarcane streak Egypt virus	NC_001868	NP_045943	NP_045942
Sugarcane streak Reunion virus	NC_004755	NP_840051	NP_840050
Sugarcane streak virus	NC_003744	NP_620491	NP_620490
Sugarcane striate virus	NC_035179	YP_009389275	YP_009389274
Sugarcane white streak virus	NC_023989	YP_009026387	YP_009026386
Sweet potato symptomless virus 1	NC_034630	YP_009362981	YP_009362980
Switchgrass mosaic-associated virus 1	NC_025834	YP_009111308	YP_009111307
Tobacco yellow dwarf virus	NC_003822	NP_620725	NP_620724
Urochloa streak virus	NC_010797	YP_001941154	YP_001941153
Wheat dwarf India virus	NC_017828	YP_006273069	YP_006273068
Wheat dwarf virus	NC_003326	NP_542348	NP_542347

Supplementary Table 3. Primers used for virus indexing

Primer name	Primer sequence	Amplicon size	Hybridization temperature (°C)	Targeted virus
NV3_F185	5'-TGGAAAGTGGTAATTCGCCC-3'	876	60	CPMSV
NV3_R1061	5'-CGACTCCACTCTGAACTTCC-3'			
NV4_F241	5'-AGCTTTGTCCCTCTACTGGT-3'	449	60	CPAV
NV4_R690	5'AGGTTGTCTTACCCTCTGCT-3'			
MSV_F1875	5'-GGGACTGACCTGGAAGATGT-3'	750	60	MSV
MSV_R2625	5'-CTATAAAACAAGGAACGGCGG-3'			

Supplementary Table 4. Long read sequencing statistics.

Raw reads	Passed reads	Sample Id	Trimmed reads	CPMSV reads	CPAV reads	Length CPMSV assembly	Length CPAV assembly
		21_REU_E0816_3	10,616	325	1,543	2,811	2,741
176,330	125,778	21_REU_E0816_4	19,498	455	677	2,804	2,725
		21_REU_E0816_5	10,352	90	-	2,797	-



Supplementary Figure 1. Nucleotide sequence of Cenchrus purpureus associated virus linearised at the nick site (\downarrow) of conserved nonnucleotide sequence TAGTATT \downarrow AC (highlighted in red) located at the stem loop (highlighted in pink). Replication associated protein gene (*rep*), movement protein gene (*V2, mp*) and capsid protein gene (*V1, cp*) are respectively highlighted in purple, blue and green. The amino acid sequence of the three ORFs are shown and the start codon is in bold. Conservative RCR and SF3 helicase amino acid motives are respectively highlighted in orange and brown and the conserved amino acid residues are in bold. 'RRR' indicates the presence of an N-terminal region rich in basic amino acids. TATA and GC boxes are respectively highlighted in light blue and gray. Long intergenic region (LIR) and short intergenic region (SIR) sequences are respectively in dark red and dark brown font. The A-rich region of the LIR is highlighted in yellow.

Article 3

Reconstructing the Mastrevirus communities structure in

Reunion: The tale of agricultural associated pathogens

Article 3

Reconstructing the *Mastrevirus* communities structure in Reunion: The tale of agricultural associated pathogens

Sélim Ben Chehida¹, Heemee Devi Bunwaree^{1,2}, Murielle Hoareau¹, Oumaima Moubset^{3,4}, Charlotte Julian^{3,4}, Laurence Blondin^{3,4}, Denis Filloux^{3,4}, Christophe Lavergne⁵, Philippe Roumagnac^{3,4}, Arvind Varsani^{6,7}, Darren P. Martin⁸, Jean-Michel Lett¹, Pierre Lefeuvre¹

¹CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

²Institut de biologie moléculaire des plantes, CNRS UPR2357, Université de Strasbourg, 67084 Strasbourg, France

³CIRAD, UMR PHIM, F-34090 Montpellier, France

⁴PHIM Plant Health Institute, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

⁵Conservatoire Botanique National de Mascarin, F-97436, St Leu, La Réunion, France

⁶The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, USA

⁷Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Rondebosch, Cape Town, South Africa

⁸Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory, South Africa

Keywords: ecology, viral metagenomics, *Mastrevirus*, host-pathogen interaction network, structure of community, Poales, agroecosystem.

Abstract

After the observations that viruses diversity and distribution was greatly different between densely cultivated areas and their natural borders, questions have emerged regarding the interplay between plant diversity and virus richness and prevalence. As both the amplification and the dilution of viral richness due to increasing host diversity have been theorized and observed, a more in-depth understanding of how plant-virus interact in natural environment was required in order to understand how host availability conditioned viral diversity and distributions. In this study, we explore interactions of viruses from the insect transmitted *Mastrevirus* genus with their natural hosts from the *Poaceae* family collected in ten sites from three contrasted ecosystems from Reunion Island. From a total of 273 plant pools, collectively representing 61 Poales species, 15 *Mastrevirus* species were characterized from 22 hosts. The strong association of mastreviruses with host collected in agro-ecosystems and the absence of mastreviruses in subalpine areas, where most of the native plants were collected, indicates that while there was probably no native population of mastreviruses that co-diverged with the native plants of Reunion, mastreviruses were also most probably introduced through agriculture activities. The analyses of the structure of plant/virus interaction network also revealed a pattern of increasing viral diversity with increasing host availability. Going one step further, the analyses of the realized viral richness in a site, taking into account the diversity of host in that site, indicated an increase of niche occupancies for host rich sites. Put it another way, this results indicates that viruses diversity was conditioned with the global quality of a site to host diverse mastreviruses. Whether this tendency is driven by synergy between viruses or indirectly by interplays between insect vector populations and plant richness remain to be established.

Introduction

Historically, the comprehension that crop associated viruses emerged from weed plants has prompted for a more thorough characterization of viruses associated with non-model and weed plants. Pioneering studies performed either in the natural compartment or at the so-called agro-ecological borders between natural and cultivated areas, have revealed the

extensive diversity of viruses circulating in natural areas (Malmstrom and Alexander, 2016; Bernardo *et al.*, 2018; Susi *et al.*, 2019; Valverde *et al.*, 2020; McLeish *et al.*, 2022; Holmes, 2022). Such ecosystem-scale studies of virus ecology have now firmly demonstrated that viruses are pervasive, and perhaps essential, components of global ecosystems (Sullivan *et al.*, 2017; O'Malley, 2016).

Owing for the stark contrast of disease expression from more dense cultivated areas and diversified natural environments bordering these areas (Shates *et al.*, 2019; García-Arenal and Zerbini, 2019), questions have been raised about the patterns that exist between plant diversity, virus diversity, virus prevalence and disease. Building on the ecological frameworks established for various non-pathogenic organisms, a flurry of studies has been performed to determine whether viral richness (*i.e.* count of entities from the same taxonomic level) increases (the amplification hypothesis) or decreases (the dilution hypothesis) with increasing host diversity (Keesing *et al.*, 2006; Faust *et al.*, 2017). Whereas in some situations higher host diversity in natural areas than in cultivated areas is matched with there being higher viral species richness in the natural areas (Bernardo *et al.*, 2018) in other situations the opposite is the case with species-poor cultivated areas having higher viral diversity than species rich natural areas (Halliday *et al.*, 2017; Susi *et al.*, 2021; McLeish *et al.*, 2022). These divergent observations in the context of the amplification and dilution hypotheses can be potentially reconciled by acknowledging that viral species richness and community species (*i.e.* a group of species occupying the same geographical area at the same time) composition are likely influenced by multiple ecological factors that vary substantially from area to area and from season to season (Martinez, 2018; Valverde *et al.*, 2020). Indeed, considering the complexity of the interactions occurring in host-parasite communities, the interplay between parasite species richness and host species richness might be non-trivial with accurate information of the viral community structure (*i.e.* remarkable interaction patterns between the entities of the community) and functioning being required to properly interpret observed trends (Keesing *et al.*, 2006). Despite the importance of precisely charting global plant-virus interaction networks to decipher such complex relationships, actual data on plant-virus associations from the field remain scarce (McLeish *et al.*, 2019).

Recently Claverie *et al.* (2023) focused on a single phytovirus genus (the *Mastrevirus* in the family, *Geminiviridae*) to achieve such a detailed picture of plant virus interactions. Mastreviruses present with a double icosahedral particle containing a single-stranded circular DNA of ~2.7 kb. These viruses are transmitted by leafhoppers of the genus *Cicadulina* (Muhire *et al.*, 2013). Currently 36 *Mastrevirus* species have been described, among which a group of 17, called the African Streak Viruses (AfSV) have been described almost exclusively from *Poaceae* species growing on the African continent and the adjacent islands of the southwest Indian ocean (AfSV; Shepherd *et al.*, 2008; Kraberger *et al.*, 2017).

The first genomic sequence-confirmed report of *Mastrevirus* species on the island of Reunion (French overseas island in the Indian Ocean) was in 1986 with the description of maize streak virus – the type species of the genus *Mastrevirus* (Peterschmitt *et al.*, 1996) – on symptomatic maize plants. Subsequent studies have increased our knowledge of mastreviruses circulating on Reunion (Claverie *et al.*, 2023), revealing host-pathogen interactions between eight virus species and 19 Poales species, with viruses displaying both diverse degrees of apparent host-specialization, and frequent host-range overlaps. This study also emphasized the importance of perennials in fostering the stability of virus species communities. Yet, as the study was conducted at a single site on the island, it remains undetermined how the specific characteristics of different sites might impact the host-range dynamics and species compositions of mastrevirus populations.

Here, in order to gain a better understanding of the interplay between the diversity of plant species and virus species, we extended the survey carried out by Claverie *et al.*, (2023) to include nine additional sites on Reunion island presenting with diverse ecologies and land-use histories. Doing so, we aimed to expand our knowledge of *Mastrevirus* diversity on Reunion, provide insights into viral community structure and investigate the factors driving the richness and spatial and host-range distributions of viral species. Finally, we explored how the host-ranges distribution(s) of species in this one viral genus in different landscapes helps us comprehend the life-history traits of these viruses.

Materials and Methods

Poales sampling

Poales (the plant order that includes the *Poaceae* family) samples were collected between November 2020 and April 2022, from ten different sites across Reunion (**Figure 1, Table 1**): four agro-ecosystem (AE) sites (Bassin Plat (BP), Ligne Paradis (LP), Colimaçons (CL) and La Mare (LM)), three coastal grassland (CG) sites (Terre Rouge (TR), Pointe au Sel (PS) and Cap la Houssaye (CH)), and three subalpine grassland (SG) sites (Ligne d'Equerre (LE), Espace Naturel Sensible des Hauts de Mont-Vert (MV) and Piton de Caille (PC)). For each site, the most prevalent Poales species were sampled with leaf samples from 50 individual plants of a given species being pooled. Samples were collected regardless of potential infection symptoms (**Supplementary Table 1**). Pools were then dried overnight at 55°C and stored at room temperature before use. The CBNM botanical database (<https://mascarine.cbnm.org/>) was used to classify the collected plant species according to (1) crop status (cultivated or weed); (2) lifespan (annual, perennial or annual/perennial); and (3) origin (native or alien to Reunion).

Insect sampling

To obtain an overview of the diversity of hemiptera insects, the order that includes mastrevirus vectors, five bulk collections of insects were carried out using a vacuum collection device: one collection at each of the TR, BP, LE, MV and PC sites. For each bulk collection, insects from the Hemiptera order were sorted by family and counted. The number of *Cicadulina mbila* specimens was also recorded.

Total DNA extraction, MinION sequencing and read assemblies

After sample homogenization using the TissueLyser II (Qiagen), total DNA was extracted from 20mg of plant material using the DNeasy Plant Pro DNA extraction kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. DNA extracts were stored at -20°C before use. The RCA-MinION protocol was then employed as described in Ben Chehida *et al.* (2021) except that the *EquiPhi29* DNA polymerase (Stepanauskas *et al.*, 2017) kit (ThermoFisher) was used for rolling circle amplification (RCA). For each sequencing run, ten samples were multiplexed along with a known positive control (DNA

extract from a *Tomato yellow leaf curl virus* infected tomato plant) and a known negative control (DNA extracted from a healthy tomato plant) on a single Flongle (FLO-FLG001). High accuracy basecalling and demultiplexing were performed using Guppy v4.2.2 (<https://nanoporetech.com/>). Read mapping statistics on a Poales chloroplast database were first obtained using minimap2 v2.17 (Li, 2018) and Samtools v1.11 (Danecek *et al.*, 2021) to both confirm the ascribed plant species of each sample, and ensure the absence of sample mix-ups during sample preparation and sequencing procedures. Full *Cressdnaviricota* (phylum of the *Mastrevirus* genus), *Alphasatellitidae* and *Tolecusatellitidae* (two families of satellite molecules associated with *Cressdnaviricota* viruses) genomic components were assembled using two complementary pipelines. The first assembly pipeline is described in Ben Chehida *et al.* (2021). In the second assembly pipeline, high quality reads were search against a *Cressdnaviricota*, and satellite molecule reference sequence database (obtained from Genbank on January 2023) using BLASTn. Filtered reads were used for assembly using Flye 2.6 (Kolmogorov *et al.*, 2019), Wtdbg 2 2.5 (Ruan and Li, 2020) and Miniasm v0.3 (<https://github.com/lh3/miniasm>). After first “polishing” using Flye, contigs were circularized using custom scripts when appropriate. Sequences obtained from the three assembly programs were clustered, merged into consensus sequences before a second step of polishing using Flye. A final polishing of the contigs was performed using Medaka v1.2.2 (<https://github.com/nanoporetech/medaka>). All the contigs obtained from both approaches were subjected to a BLASTn search against the NCBI nt database for preliminary species assignment. Coverage statistics were obtained using minimap2 v2.17 (Li, 2018) by mapping the raw reads back to the assembled contigs that presented with similarities to *Cressdnaviricota* and satellite molecule sequences. In each sample, only contigs representing a minimum of 2% of the raw reads were considered for further analysis.

Full Genome Cloning and Sanger Sequencing

Samples presenting with sequences from putative new viral species (following the ICTV guidelines, Roumagnac *et al.*, 2022 ; Briddon *et al.*, 2018) or from viral species newly identified on Reunion were submitted to RCA-RFLP for full genome cloning and Sanger sequencing as described in Inoue-Nagata *et al.* (2004). *Apa*LI, *Kpn*II, *Nco*I, *Sa*II or *Xba*I were used to linearize monomers of mastreviruses genomes.

Phylogenetic and recombination analyses

A phylogenetic tree that includes all the mastrevirus reference sequences (as obtained from GenBank in January 2023) along with the sequences identified in this study was constructed for classification confirmation. To that end, all sequences were aligned using MAFFT v7.475 (Kato and Standley, 2013) and a maximum likelihood (ML) phylogenetic tree was inferred using FastTree2 (Price *et al.*, 2010) with the GTR+G+I substitution model selected as best fitting by ModelFinder (Kalyaanamoorthy *et al.*, 2017) and with 1000 bootstrap replicates for SH-like approximate likelihood ratio branch support test (Guindon *et al.*, 2010). Additional trees for encoded CP and Rep amino acid sequences were similarly obtained with the LG+G+I substitution model. Trees were edited using the “ape” R package (Paradis and Schliep, 2019). Along with the mastreviruses refseq dataset, all *Mastrevirus* sequences obtained from samples collected on Reunion were downloaded from GenBank in January 2023. Notably, this included the 49 sequences described in Claverie *et al.* (2023) and obtained from 2014 to 2017 from the BP sampling site. A phylogenetic tree was constructed from the nucleotide alignment as described above. Recombination events were then searched within the alignment using the RDP (Martin and Rybicki, 2000), GENECONV (Padidam *et al.*, 1999), BOOTSCAN (Salminen *et al.*, 1995), MAXCHI (Smith, 1992), CHIMERA (Posada and Crandall, 2001), SISCAN (Gibbs *et al.*, 2000) and 3SEQ (Boni *et al.*, 2007) methods implemented in the RDP5 program (Martin *et al.*, 2021). As our main interest was to uncover recombination events that have potentially occurred on Reunion island, a “group” analysis was used in RDP5 to focus on recombination events involving sequences closely related to those that have previously been found on Reunion. In other words, only triplets of sequences in which the recombinant sequence was isolated in Reunion, or sequences closely related to either of the two parental sequences of the recombinant were isolated in Reunion, were tested for recombination. In addition to the grouping scheme, default RDP5 settings were used, and recombination events were considered significant when detected by at least four different recombination detection methods.

Viral community diversity and structure analyses

Between site differences in plant sample species richness were calculated with the Jaccard distance on binary data using the *vegdist* function of the “vegan” R package

(Oksanen *et al.*, 2019). The ordination of the sites by host species was projected in two dimensions with principal coordinates analysis (PCoA) using the *cmdscale* function of the “vegan” R package, the optimal number of clusters was calculated with the *fviz_nbclust* function from the “factoextra” R package (Kassambara and Mundt, 2020) and clustering at a 95% confidence level was performed with the *kmeans* function from the “stats” R package.

Plant characteristics either for all the plants at a site or for individual host plants across all sites, were compared between sites using exact Fisher tests. Pairwise comparisons for proportions (“stats” R package) based on the Benjamini and Benjamini & Hochberg Hochberg methods (1995) were calculated when global significant differences were found. Differences in viral richness between both the different sites and the different plant host species, respectively according to the site-type and the characteristics of the plant hosts, were assessed using Kruskal-Wallis rank sum tests (Hollander and Wolfe, 1973) implemented in the “stats” R package. The three tests were performed with a 95% confidence level.

From the binary (presence/absence) plant-virus interactions matrix, a bipartite network was generated using the R “bipartite” package (Dormann *et al.*, 2008). Three measures were used to further characterize the Poales/*Mastrevirus* community. (1) The assortativity, a measure of the tendency for vertices in a network to be connected to other vertices that shares similar characteristics (here, the plant tribes, crop status and plant lifespan ; Newman, 2003), was obtained using the *assortativity_nominal* function of the “igraph” R package (Csardi and Nepusz, 2006). (2) The nestedness, a measure of the tendency of interactions to be hierarchically organized for both viruses and hosts, where the set of hosts of a given virus is a subset of the viruses of broader host ranges, was computed using the “NODF2” method (Almeida-Neto *et al.*, 2008) of the *nested* function implemented in the “bipartite” R package (Dormann *et al.*, 2008). (3) The modularity, a measure of the strength of the division of a network into groups of viruses and hosts that preferentially interact together (Olesen *et al.*, 2007 ; Weitz *et al.*, 2013 ; Dormann and Strauss, 2014 ; Beckett, 2016 ; McLeish *et al.*, 2017 ; Valverde *et al.*, 2020 ; Moury *et al.*, 2021), was

calculated using *computeModules* function with the "Newman" method (Newman, 2004 ; Beckett, 2016), also implemented in the "bipartite" R package (Dormann *et al.*, 2008).

The statistical significance of assortativity, nestedness and modularity were estimated after comparison with 1000 random permutations of the binary contingency matrix using "quasiswap" null models ("none", "rows", "columns", "both") from the "vegan" R package (Oksanen *et al.*, 2019). The "none" model preserves the total number of interactions, the "rows" model preserves the sum of interactions per plant species, the "columns" model preserves the sum of interactions per virus species and the "both" model preserves the sum of interactions per plant species and virus species.

Results and discussion

Poales diversity among different site-types in Reunion

A total of 273 pools of individual plant species (collectively representing more than 13,000 individual plants) were collected from ten sites on Reunion island (Figure 1A) between 2020 and 2022 (**Table 1 ; Supplementary Table 1**). Globally, these pools represented 50 *Poaceae* (nine tribes), ten *Cyperaceae* (four tribes) and one *Juncaceae* species (**Supplementary Table 2 ; Figure 1B**). Whereas a single species, *Cynodon dactylon* (Figure 1C), was collected in all three of the different environment types, 13 (21%), five (8%) and one (2%) of the species were common to the AE and CG, the AE and SG and the CG and SG environments respectively. The other species were found in only one of the environment types with 19 (31%), nine (15%) and 13 species specific to the AE, CG and SG environments, respectively. Confidence ellipses of the PCoA (**Supplementary Figure 1**) confirmed that sites with the most similar plant sample compositions were of the same environmental type.

Table 1. Sampling sites and summary of sampled pools for each campaign. The number in brackets refers to the number of pools detected for which complete *Mastrevirus* genomes were obtained.

Type of Site	Site	Lat	Lon	Alt (m)	Nov 2020	Jan 2021	Apr 2021	Nov 2021	Apr 2022	Total
CG	TR	-21.3498	55.4975	31	0	0	4 (1)	6	7	17 (1)
CG	PS	-21.2024	55.2814	8	0	0	9	6 (1)	11 (2)	26 (3)
CG	CH	-21.0187	55.2352	15	0	0	13 (1)	5	17 (2)	35 (3)
AE	LM	-20.9032	55.5311	70	0	0	0	0	18 (8)	18 (8)
AE	BP	-21.3216	55.4892	166	4	13 (1)	15 (9)	16 (2)	22 (10)	70 (22)
AE	LP	-21.3166	55.4873	180	0	0	0	10 (2)	17 (4)	27 (6)
AE	CL	-21.1304	55.3054	819	0	0	13 (1)	9 (3)	15 (1)	37 (5)
SG	LE	-21.2916	55.5733	1062	5	0	0	8	0	13
SG	MV	-21.2859	55.6030	1586	7	0	0	6	0	13
SG	PC	-21.1902	55.6380	2130	9	0	0	8	0	17
Total					25	13 (1)	54 (12)	74 (8)	107 (27)	273 (48)

In all, 14 (23%) species were weed and native to Reunion (four in AE, five in CG and nine in SG), 41 (67%) were weed alien species (28 in AE, 19 in CG and ten in SG) and six (10%) were cultivated alien species (six in AE and one in SG) (**Figure 1B** ; **Supplementary Table 1 & 2**). In accordance with both degrees of human usage at the distinct sites, and expected decreases in numbers of alien species with increasing altitude (Fenouillas *et al.*, 2021), there was a significant difference in the proportions of native species between the different site-types (exact Fisher test p-value = 1.0×10^{-2}), and the pairwise proportions test was significant for native species being more prevalent at the SG than at the AE sites (p-value = 7.8×10^{-3}).

Regarding plant lifespans, 22 annuals (36%; 16 in AE, twelve in CG and two in SG), 27 perennials (44%; 18 in AE, eight in CG and twelve in SG) and three annual/perennial species (5%; six in AE and one in SG) were found (**Figure 1B** ; **Supplementary Table 1 & 2**). There was only a marginally significant difference were between sites with respect to the proportions of annual, perennial and annual/perennial plant species observed at the sites (exact Fisher test p-value = 5.6×10^{-2}). It must be noted, however, that our life-cycle classification is based on data collected for the plants when growing in their native environment and that alien species may present with different life-cycles than those found in their native environments when they are introduced into foreign environments.

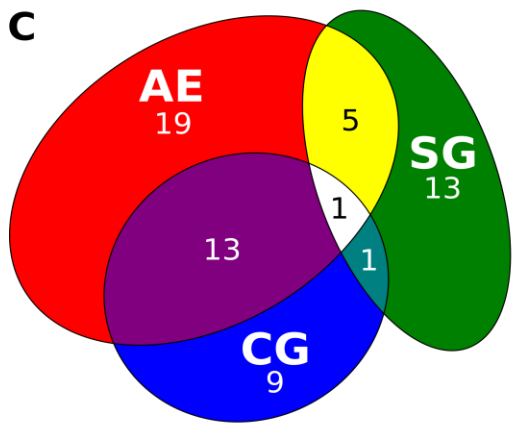
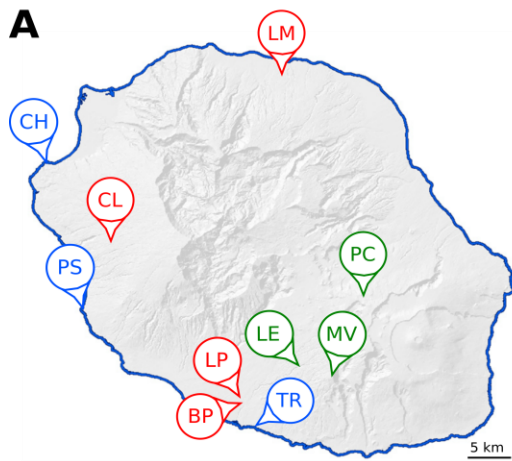
Multiple cressdnaviricota are associated with Poaceae

Of all the 273 pooled plant, cressdnaviruses and/or known satellites of cressdnaviruses were found in 49. These positive samples represented 22 of the 61 collected plant species. It is important to note first that none of the ten *Cyperaceae* and the single *Juncaceae* species (together representing 24 plant sample pools) contained detectable evidence of CRESS DNA viruses.

All the CRESS DNA positive plants belonged to the *Poaceae* family. Whereas most of the genomic sequences obtained were from the *Mastrevirus* genus (65 out of 92), contigs assigned to other members of the *Cressdnaviricota* phylum were detected: *Genomoviridae* (22 contigs), Sorghum mastrevirus associated alphasatellite (three contigs) and a putative new *Cressdnaviricota* member (two contigs) (**Supplementary Table 3**). *Genomoviridae* sequences were isolated from 14 distinct pools (*Cenchrus purpureus*, *Chloris barbata*, *Cynodon dactylon*, *Dactyloctenium sp.*, *Dichanthium aristatum*, *Digitaria ciliaris*, *Eragrostis sp.*, *Stenotaphrum dimidiatum* and *Zea mays*). Genome sequences of the Sorghum mastrevirus associated alphasatellite, a satellite belonging to the *Alphasatellitidae* family, were obtained from three pools of *Sorghum arundinaceum*. Finally, the putative new *Cressdnaviricota* genus member (Ben Chehida *et al.*, in prep) was found in two *C. purpureus* pools.

Extending our knowledge of mastreviruses diversity and host ranges

A total of 65 complete genome assemblies from twelve different *Mastrevirus* species were obtained (**Figure 2 ; Table 2**). *Mastrevirus* genome sequences were recovered from 22 out of the 50 *Poaceae* species collected (representing 48 of the 217 *Poaceae* pools, **Table 1 & 2**). These species were from five distinct tribes: *Andropogoneae*, *Paniceae*, *Eragrostideae*, *Cynodonteae* and *Zoysieae* (**Table 2 ; Supplementary Table 2**). In total, twelve *mastrevirus* species were identified, with eight having been previously characterized, three being putative new *Mastrevirus* species and one presenting with a defective genome of a probable new species.



Site-types

CG: Coastal grasslands

AE: Agroecosystems

SG: Subalpine grasslands

Sites

TR: Terre Rouge
PS: Pointe au Sel
CH: Cap la Houssaye

LM: La Mare
BP: Bassin Plat
LP: Ligne Paradis
CL: Colimaçons

LE: Ligne d'Equerre
MV: Espace Naturel Sensibles
 des Hauts de Mont-Vert
PC: Piton de Caille

Family

Tribes

- Cyperaceae**
 - Abildgaardieae
 - Cariceae
 - Cypereae
 - Schoeneae
- Juncaceae**
 - Juncus
 - Andropogoneae
 - Aristideae
 - Bromeae
 - Cynodontae
- Poaceae**
 - Eragrostideae
 - Panicaceae
 - Paspaleae
 - Poeae
 - Zoysieae

Crop status

- N:** Native
- W:** Alien Weed
- C:** Alien Cultivated

Lifespans

- A:** Annual
- P:** Perennial
- A/P:** Both
- NA:** Unknown

B

	Crop status	Lifespans	CG			AE			SG						
			TR	PS	CH	LM	BP	LP	CL	LE	MV	PC			
<i>Fimbristylis cymosa</i>	N	P													
<i>Carex boryana</i>	N	P													
<i>Carex leporina</i>	W	P													
<i>Carex pilulifera</i>	W	P													
<i>Cyperus aromaticus</i>	N	P													
<i>Cyperus polystachyos</i>	N	P													
<i>Cyperus rotundus</i>	W	P													
<i>Astrochaete nitens</i>	N	P													
<i>Costularia cadetii</i>	N	P													
<i>Costularia mellicoides</i>	N	P													
<i>Juncus effusus</i>	N	P													
<i>Bothriochloa</i> sp.	N	A													
<i>Chrysopogon zizanioides</i>	C	P													
<i>Dichanthium aristatum</i>	W	P													
<i>Heteropogon contortus</i>	N	P													
<i>Pennisetum glaucum</i>	C	A													
<i>Rottboellia cochinchinensis</i>	W	A													
<i>Saccharum</i> spp.	C	P													
<i>Sorghum arundinaceum</i>	W	A/P													
<i>Sorghum bicolor</i>	C	A													
<i>Themeda quadrivalvis</i>	W	A													
<i>Zea mays</i>	C	A													
<i>Aristida adscensionis</i>	W	A													
<i>Bromus catharticus</i>	W	A													
<i>Chloris barbata</i>	W	P													
<i>Chloris gayana</i>	W	P													
<i>Chloris pycnothrix</i>	W	A													
<i>Chloris</i> sp.	W	NA													
<i>Cynodon dactylon</i>	N	P													
<i>Cynodon</i> sp.	W	P													
<i>Dactyloctenium</i> sp.	N	A													
<i>Eleusine indica</i>	W	A													
<i>Tragus mongolorum</i>	W	A													
<i>Enneapogon canchroides</i>	W	A													
<i>Eragrostis</i> sp.	W	A/P													
<i>Cenchrus clandestinus</i>	W	P													
<i>Cenchrus echinatus</i>	W	A													
<i>Cenchrus purpureus</i>	C	P													
<i>Cenchrus</i> sp.	W	P													
<i>Digitaria ciliaris</i>	W	A													
<i>Digitaria radicata</i>	W	A													
<i>Echinochloa colona</i>	W	A													
<i>Megathyrsus maximus</i>	W	P													
<i>Melinis repens</i>	W	A/P													
<i>Panicum juniperinum</i>	N	P													
<i>Setaria pumila</i>	W	A													
<i>Setaria</i> sp.	W	A													
<i>Stenotaphrum dimidiatum</i>	W	P													
<i>Urochloa deflexa</i>	W	A													
<i>Urochloa panicoides</i>	W	A													
<i>Urochloa ramosa</i>	W	A													
<i>Urochloa decumbens</i>	W	A													
<i>Paspalum dilatatum</i>	W	P													
<i>Paspalum notatum</i>	W	P													
<i>Paspalum paniculatum</i>	W	P													
<i>Paspalum virgatum</i>	W	P													
<i>Anthoxanthum odoratum</i>	W	P													
<i>Dactylis glomerata</i>	W	P													
<i>Festuca borbonica</i>	N	P													
<i>Holcus lanatus</i>	W	P													
<i>Sporobolus africanus</i>	W	P													

Figure 1. Summary of the Poales sampling in Reunion. **(A)** Map of Reunion indicating the ten sampling sites. In blue, Coastal grassland (CG) with Terre Rouge (TR), Pointe au Sel (PS) and Cap la Houssaye (CH); in red agroecosystems (AE) with La Mare (LM), Bassin Plat (BP), Ligne Paradis (LP) and Colimaçons (CL); in green Subalpine grassland (SG) with Ligne d'Equerre (LE), Espace Naturel Sensible des Hauts de Mont-Vert (MV) and Piton de Caille (PC). **(B)** Presence/absence matrix of the Poales species (rows) in the different sites (columns). The Poales species are organized by families (*Cyperaceae*, *Juncaceae*, *Poaceae*) and colored and organized by tribes. Crop status (N: native ; W: alien weed ; C: alien cultivated) and lifespan (A: annual ; P: perennial ; A/P: annual or perennial ; NA: unknown) are indicated in the first and second column, respectively. Light colored squares indicate the presence of the Poales species. Dark colored squares indicate the detection of mastreviruses in the corresponding species. **(C)** Euler diagram of the number of Poales species exclusively or commonly sampled in the three site-types.

Table 2. Summary of taxonomic classification of mastreviruses sequences per Poales species for each type of site and site.

Survey	Type of Site	Site	Host Species	Virus Species
Jan 2021	AE	BP	<i>Cenchrus echinatus</i>	MSV-F
Jan 2021	AE	BP	<i>Cenchrus echinatus</i>	USV
Apr 2021	CG	TR	<i>Heteropogon contortus</i>	USV
Apr 2021	CG	CH	<i>Urochloa panicoides</i>	MSV-B
Apr 2021	CG	CH	<i>Urochloa panicoides</i>	USV
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	CEAV
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	MSRV
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	MSV-B
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	SSRV
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	USV
Apr 2021	AE	BP	<i>Chrysopogon zizanioides</i>	MSRV
Apr 2021	AE	BP	<i>Cynodon dactylon</i>	PanSV-J
Apr 2021	AE	BP	<i>Eleusine indica</i>	EIAV
Apr 2021	AE	BP	<i>Megathyrsus maximus</i>	PanSV-J
Apr 2021	AE	BP	<i>Saccharum spp.</i>	SWSV
Apr 2021	AE	BP	<i>Sorghum arundinaceum</i>	SAAV-B
Apr 2021	AE	BP	<i>Stenotaphrum dimidiatum</i>	USV
Apr 2021	AE	CL	<i>Urochloa decumbens</i>	UDAV
Nov 2021	CG	PS	<i>Megathyrsus maximus</i>	PanSV-C
Nov 2021	AE	BP	<i>Digitaria ciliaris</i>	MSV-B
Nov 2021	AE	BP	<i>Saccharum spp.</i>	DefMS
Nov 2021	AE	BP	<i>Saccharum spp.</i>	SWSV
Nov 2021	AE	LP	<i>Saccharum spp.</i>	SWSV
Nov 2021	AE	LP	<i>Sorghum arundinaceum</i>	SAAV-A

Nov 2021	AE	CL	<i>Cenchrus purpureus</i>	CPMSV
Nov 2021	AE	CL	<i>Sporobolus africanus</i>	MSV-B
Nov 2021	AE	CL	<i>Urochloa decumbens</i>	UDAV
Apr 2022	CG	PS	<i>Megathyrsus maximus</i>	PanSV-C
Apr 2022	CG	PS	<i>Urochloa deflexa</i>	PanSV-C
Apr 2022	CG	PS	<i>Urochloa deflexa</i>	USV
Apr 2022	CG	CH	<i>Digitaria ciliaris</i>	MSV-B
Apr 2022	CG	CH	<i>Urochloa ramosa</i>	USV
Apr 2022	AE	LM	<i>Cenchrus echinatus</i>	MSRV
Apr 2022	AE	LM	<i>Cenchrus echinatus</i>	MSV-B
Apr 2022	AE	LM	<i>Cenchrus echinatus</i>	PanSV-J
Apr 2022	AE	LM	<i>Chloris barbata</i>	EIAV
Apr 2022	AE	LM	<i>Chloris barbata</i>	MSRV
Apr 2022	AE	LM	<i>Eleusine indica</i>	EIAV
Apr 2022	AE	LM	<i>Eleusine indica</i>	MSV-B
Apr 2022	AE	LM	<i>Megathyrsus maximus</i>	PanSV-J
Apr 2022	AE	LM	<i>Megathyrsus maximus</i>	SWSV
Apr 2022	AE	LM	<i>Melinis repens</i>	PanSV-J
Apr 2022	AE	LM	<i>Rottboellia cochinchinensis</i>	MSRV
Apr 2022	AE	LM	<i>Rottboellia cochinchinensis</i>	PanSV-J
Apr 2022	AE	LM	<i>Saccharum spp.</i>	DefMS
Apr 2022	AE	LM	<i>Sorghum arundinaceum</i>	SAAV-A
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	CEAV
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	MSRV
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	MSV-B
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	SSRV
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	USV
Apr 2022	AE	BP	<i>Chloris gayana</i>	MSRV
Apr 2022	AE	BP	<i>Cynodon dactylon</i>	MSRV
Apr 2022	AE	BP	<i>Eleusine indica</i>	EIAV
Apr 2022	AE	BP	<i>Megathyrsus maximus</i>	PanSV-J
Apr 2022	AE	BP	<i>Saccharum spp.</i>	SWSV
Apr 2022	AE	BP	<i>Sorghum arundinaceum</i>	SAAV-A
Apr 2022	AE	BP	<i>Sorghum bicolor</i>	SAAV-B
Apr 2022	AE	BP	<i>Urochloa panicoides</i>	MSV-B
Apr 2022	AE	BP	<i>Urochloa panicoides</i>	USV

Apr 2022	AE	BP	<i>Zea mays</i>	MSRV
Apr 2022	AE	LP	<i>Eleusine indica</i>	EIAV
Apr 2022	AE	LP	<i>Sorghum arundinaceum</i>	SAAV-A
Apr 2022	AE	LP	<i>Urochloa panicoides</i>	MSV-B
Apr 2022	AE	LP	<i>Zea mays</i>	MSRV
Apr 2022	AE	CL	<i>Urochloa decumbens</i>	UDAV

The three putative new *Mastrevirus* species were tentatively named Cenchrus echinatus associated virus (CEAV) which was found associated with *Cenchrus echinatus*, Cenchrus purpureus mild streak virus (CPMSV) which was found associated with *C. purpureus* and Urochloa decumbens associated virus (UDAV) which was found associated with *Urochloa decumbens* (**Figure 2 & 3 ; Table 2**). CPMSV was named for the symptoms of discontinuous parallel chlorotic streaks observed on the leaves *C. purpureus* individuals observed within the pooled plant leaves from this species.

Whereas the RCA-RFLP procedure was unsuccessful for CEAV, we obtained one and four full genome sequences of CPMSV and UDAV respectively (**Supplementary Table 4**). The three viruses presented with typical *Mastrevirus* genome organizations with ORFs likely encoding movement protein and capsid protein in the virion sense, and ORFs encoding the replication associated protein and the replication associated protein A in the complementary sense. The genome sizes were 2,808 nt for CPMSV 2,809 to 2,811 nt for CEAV and 2,837 nt for UDAV.

A TAATATTAC nonanucleotide characteristic of geminivirus viron strand origins of replication (V-ori) was identified in CPMSV and CEAV and a less commonly found TAATATTGC nonnucleotide is the likely V-ori of UDAV.

Among the known mastreviruses CEAV is most similar to wheat dwarf India virus (MN273439) with which it shares 58.6% nucleotide sequence identity. UDAV is most similar to Axonopus compressus streak virus (KJ437671) with which it shares 72.3% nucleotide sequence identity; and CPMSV is most similar to Sorghum arundinaceum associated virus (MK546381) with which it shares 57.9% nucleotide sequence identity (**Supplementary Figure 2A ; Supplementary Table 5**). A putative defective *Mastrevirus* genome (DefMS) was obtained from a *Saccharum spp.* pool and four sequences DefMS sequences were determined using RCA-RFLP and Sanger sequencing. The size of the DefMS genomes ranged from 1,505 to 1,508 nt and possessed a mastrevirus-like V-ori

nonanucleotide (TAATATTAC) and two ORFs likely encoding a movement protein and capsid protein. Among the known mastreviruses It was most similar to the North American maize-associated mastrevirus (MZ852895) with which it shared 57.2% nucleotide sequence identity over their homologous genome regions. The absence in DefMS of an ORF encoding a replication-associated protein and the presence of DefMS in the same pool as SWSV suggests that SWSV may act as a helper for DefMS replication (Simon *et al.*, 2004).

Among the eight previously characterized AfSV that we identified in the plant pools , only Urochloa streak virus (USV) has never previously been reported on Reunion. It was found infecting *C. echinatus*, *Heteropogon contortus*, *U. deflexa*, *U. panicoides*, *U. ramosa* and *S. dimidiatum* (**Figure 2 ; Table 2**). Three full genome sequences of USV were cloned and sequenced using the Sanger procedure from pools of *C. echinatus* (21_REU_E0111 ; **Supplementary Table 4**).

Of all the AfSV species previously identified on Reunion, all but the *Melinis repens* associated virus and Sugarcane streak virus were identified here. All seven of the AfSV species identified here were found at least once in BP, the same sampling site studied in Claverie *et al.* (2023). Of the eleven maize streak virus (MSV) sequences obtained from eleven different pools, ten were belonged to the MSV-B strain and one to the MSV-F strain. MSV-B sequences were obtained from *C. echinatus*, *Digitaria ciliaris*, *Eleusine indica*, *Sporobolus africanus* and *U. panicoides* (**Table 2**). The single MSV-F sequence was obtained from *C. echinatus*. Maize streak Reunion virus isolates (MSRV, n=10) were obtained from ten pools representing *C. echinatus*, *C. barbata*, *Chloris gayana*, *Chrysopogon zizanioides*, *C. dactylon*, *Rottboellia cochinchinensis* and *Z. mays*. Two of the nine known Panicum streak virus strains (PanSV), were identified PanSV-C (n = 3) found in two *Megathyrsus maximus* pools and one *U. deflexa* pool – the first detection of this strain on Reunion – and PanSV-J (n = 7) found in one *C. echinatus* pool, one *C. dactylon* pool, three *M. maximus* pools, one *M. repens* pool and one *R. cochinchinensis* pool. SAAV (n = 6) was detected in five *S. arundinaceum* pools and one *Sorghum bicolor* pools. Two SAAV strains could be distinguished if one applies the informal *Mastrevirus* strain demarcation criterion of 94% identity (Muhire *et al.*, 2013): SAAV-A that includes previously described sequences (Claverie *et al.*, 2019) and four sequences described here from *S. arundinaceum*, and SAAV-B obtained from *S. arundinaceum* and *S. bicolor*.

Genomic sequences of pairs of viruses belonging to the different strains shared ~91.9% nucleotide sequence identity. Sorghum mastrevirus associated alphasatellite was also detected in one pool together with SAAV-A sequences and in two pools together with SAAV-B sequences (**Supplementary Table 3**). Eleusine indica associated virus was detected in a *E. indica* pool and in a *C. barbata* pool. Complete sugarcane white streak genomes were obtained from a *Saccharum spp.* pool and a *M. maximus* pool, the latter representing the first report of this *Mastrevirus* species being associated with a host other than *Saccharum spp.* Sugarcane streak Reunion virus was detected in two *C. echinatus* pools.

Globally, of the 20 plant species collected both here and in Claverie *et al.* (2023), eleven were detected as hosts of mastreviruses in both studies. Whereas four species identified as mastrevirus hosts by Claverie *et al.* (2023; *Cyperus rotundus*, *Dactyloctenium sp.*, *Eragrostis sp.* and *Setaria pumila*) were not identified as mastrevirus hosts here, one species was identified here as a mastrevirus host that was not identified as one by Claverie *et al.* (2023; *C. zizanioides*). Among the Poales species that were collected in this study but not in Claverie *et al.* (2023) (N=36), eleven were found that likely host mastreviruses ; extending the known natural Reunion host-ranges of viruses in this genus to include *C. purpureus*, *C. barbata*, *C. zizanioides*, *H. contortus*, *S. bicolor*, *S. dimidiatum*, *U. decumbens*, *U. panicoides*, *U. ramosa*, *U. deflexa*, and *R. cochinchinensis*). In total, 2/14 (14%) native weed species, 15/41 (37%) alien weed species and 5/6 (83%) alien cultivated species were hosts of mastreviruses (**Figure 1**). Whereas most of the hosts were alien weed species, there was no significant differences between the viral species richness obtained from the cultivated and weed species (Kruskal-Wallis test p-value = 1.9×10^{-1}) or from the native and alien species (Kruskal-Wallis test p-value = 1). When comparing annual vs perennial *Poaceae* species, 10/22 (45%) of the annual, 10/33 (30%) of the perennial and 2/3 (67%) of the annual/perennial species that were collected likely host mastrevirus, but no significant difference was observed between the groups either in the proportion of pooled samples (exact Fisher test p-value = 2.6×10^{-1}) harbouring mastreviruses nor in the degree of viral species richness that they harboured (Kruskal-Wallis test 9.1×10^{-1}). A previous comparison of infection frequencies between different categories of plant species at one of the sites studied here (site BP; Claverie *et al.*, 2023) revealed that mastrevirus infections at that site were significantly more frequent in weed

plants than cultivated plants, more frequent in annuals than perennials and more frequent in native weeds than alien weeds. Whereas we found no strong evidence of such trends, it should be emphasized that, unlike in the Claverie *et al.* (2023) study, our analyses were performed on pools of plants and, as such, on binary virus presence vs absence counts, rather than on infection frequencies.



Figure 2. Phylogenetic tree of all the 94 complete genomes of monocot-infecting mastreviruses from Reunion. Sequences determined in this study are indicated in colored bold font with sequences from agroecosystem (AE) samples in red and from coastal grassland (CG) samples in blue. Sequences obtained using the RCA-MinION procedure are underlined. Grey rectangles refer to the mastreviruses species and strains. Open and closed circles on nodes indicate bootstrap support for the branches to their left of 70-89% and $\geq 90\%$ respectively.

Environment type impacts proportions of mastrevirus-infected Poales species

Whereas 34/56 (61%) of the Poales species at AE sites and 15/24 (63%) at CG sites were known mastrevirus hosts, 32/34 (94%) of the pools of known hosts were detectably infected with mastreviruses at the AE sites and only 6/15 (40%) at the CG sites. Regardless of the presence of mastreviruses, considering all 35 of the mastrevirus host species known to occur in Reunion, 17/35 (49%) were exclusively collected at AE sites and 1/35 (3%) at CG sites, 13/35 (37%) were collected at both the AE and CG sites, 3/35 (9%) were collected at both the AE and SG sites and 1/35 (3%; *Cynodon dactylon*) were collected at all three site-types.

It is noteworthy that no mastreviruses were found in any plant pools collected from the SG sites. As these sites are at high elevation (ranging from 1,062 to 2,130 m), and it is expected that the lower temperatures at these elevations would reduce potential insect vectors population sizes. MSV (Nielson, 1968), PanSV (Briddon *et al.*, 1992), SSRV (Bigarré *et al.*, 1999) and SSV (Briddon *et al.*, 1996) are all transmitted by leafhopper species in the genus *Cicadulina* (Shepherd *et al.*, 2010), with *C. mbila* (Bonfils *et al.*, 1994) being one of the most important vector species (Webb, 1987). The vectors of the other mastrevirus species identified here remain unknown. Insect captures performed during the study demonstrated the presence of large numbers of hemiptera species belonging to at least eight families: *Cicadellidae*, *Delphacidae*, *Membracidae*, *Aphidae*, *Miridae*, *Psyllidae*, *Tropiduchidae* and *Aleyrodidae* (**Table 3**). Insects from four, five and six families of hemiptera were sampled at the LE, MV and PC sites respectively. Whereas *Cicadellidae* were present at all the SG site-types the only SG site where *C. mbila* was recovered was the LE site (where it was abundantly present). The absence of sampling of *C. mbila* at the TR site (a CG site-type) despite the presence of plants infected with mastreviruses known to be transmitted by this species may indicate either that more thorough sampling would be required to recover *C. mbila* at this site, or that other *Cicadellidae* are effective vectors of AfSVs in Reunion. The presence of abundant *C. mbila* at the LE site and the presence of *Cicadallidae* would suggest that AfSV transmission by these insects is possible at this

site. Indeed, MSV transmission has been proven at altitudes up to 1,300m in Reunion (Reynaud, 1988). On reunion Maize plots are scarce at altitudes as high as this and the maximum reach of MSV transmission is thus difficult to assess. Although seasonal temperature variations might result in pauses in transmission cycles at the high altitude SG site-types, the presence of several perennials grasses at the LE, MV and PC sites (18 out of the 20 Poales collected at these SG site-types) implies the plausibility of that inter-seasonal persistence of *C. mbila* transmitted AfSVs within these environments. Combining this observation with the higher proportion of native Poales at the SG site-types (eight out of 20), it would indicate that there is no resident AfSV diversity in Reunion, *i.e.* AfSV adapted and co-evolved with the native (frequently endemic) Poales sampled at high altitude sites. Conversely, AfSV species found in Reunion were largely associated with exotic plants. The observation that AfSV species present in both Reunion and on mainland Africa is largely consistent with the hypothesis that mastrevirus species on Reunion that have a preference for infecting alien hosts were introduced to the island together with their alien hosts by humans during the colonial and post-colonial eras (Kopp, 1930 ; Monjane *et al.*, 2011).

Table 3. Summary of sampled Hemiptera families in each type of site and site in an altitudinal transect.

Type of Site	Site	Aleyrodidae	Aphididae	<i>Cicadulina mbila</i>	Other Cicadellidae	Delphacidae	Membracidae	Miridae	Psyllidae	Tropiduchidae
CG	TR	5	6	0	121	269	6	27	4	1
AE	BP	8	45	101	4	111	11	72	40	11
SG	LE	0	39	>1000	177	52	1	221	7	0
SG	MV	0	0	0	4	3	0	35	2	0
SG	PC	0	3	0	112	9	0	18	1	0

The minimum host-virus interaction network

Building on previous data obtained from Reunion (all from AE site-types) and that obtained here from a wider-variety of site-types we attempted to construct a more complete network of mastrevirus-host species interactions (**Figure 3**). We understood that that: (1) this interaction network would represent the minimal set of actual mastrevirus-host species interactions in that it is very likely that our sampling has been too sparse to detect all these interactions ; (2) it is very unlikely that even with exhaustive sampling and perfect detection

methods examples of all possible interactions would likely not be present at a given time at a given site; and (3) the actual structure of the network might not be static, but instead change over time as new viral and host lineages emerge, spread and intermingle across the sampling sites (Valverde *et al.*, 2020). Indeed as the network we constructed was assembled from data accumulated by multiple studies, neither the timings of collections nor the precise sampling sites were taken into account. We would therefore consider this network to represent the overall minimal host ranges of the mastrevirus species found on Reunion.

The virus-host species interaction network represents 74 interactions between 35 plants species belonging to seven tribes and 14 viral species (or 19 viral entities when considering strains). Among all the *Mastrevirus* species and strains, five have only ever been sampled on Reunion from a single host species (DefMS, UDAV, CPMSV, SSV, MSV-E and CEAV): we refer to these here as specialists. Thirteen 13 other virus entities all found infecting two or more different host species (MeRAV, SAAV-A, SAAV-B, SSRV, USV, PanSV-C, PansV-J, MSV-B, MSV-A, MSV-F, SWSV, EIAV, MSRV) with MSV-B infecting the highest number of different host species: 23 of the 35 sampled host species. For classification purposes we refer to the virus species found infecting two or more host species as generalists. Among these 13 virus species, one could consider as generalists the virus species that were found infecting two or more hosts belonging to two or more different plant tribes (or genera) - some of the plant species are really closely related (*i.e.* the plant species are not all equally different from one another). In that way EIAV, SAAV-B, PanSV-C and MSV-F could be considered as specialists.

Of the 14 host species sampled both here and in Claverie *et al.* (2023), five were found to be infected with MSV-B here, and 13 were found to be infected by MSV-B in the Claverie *et al.* (2023) study. Individual sampled plants species were found hosting between one (for 14 plant species) and eight virus species (for *Cenchrus echinatus*) with a mean number of two virus species per host (**Table 2**).

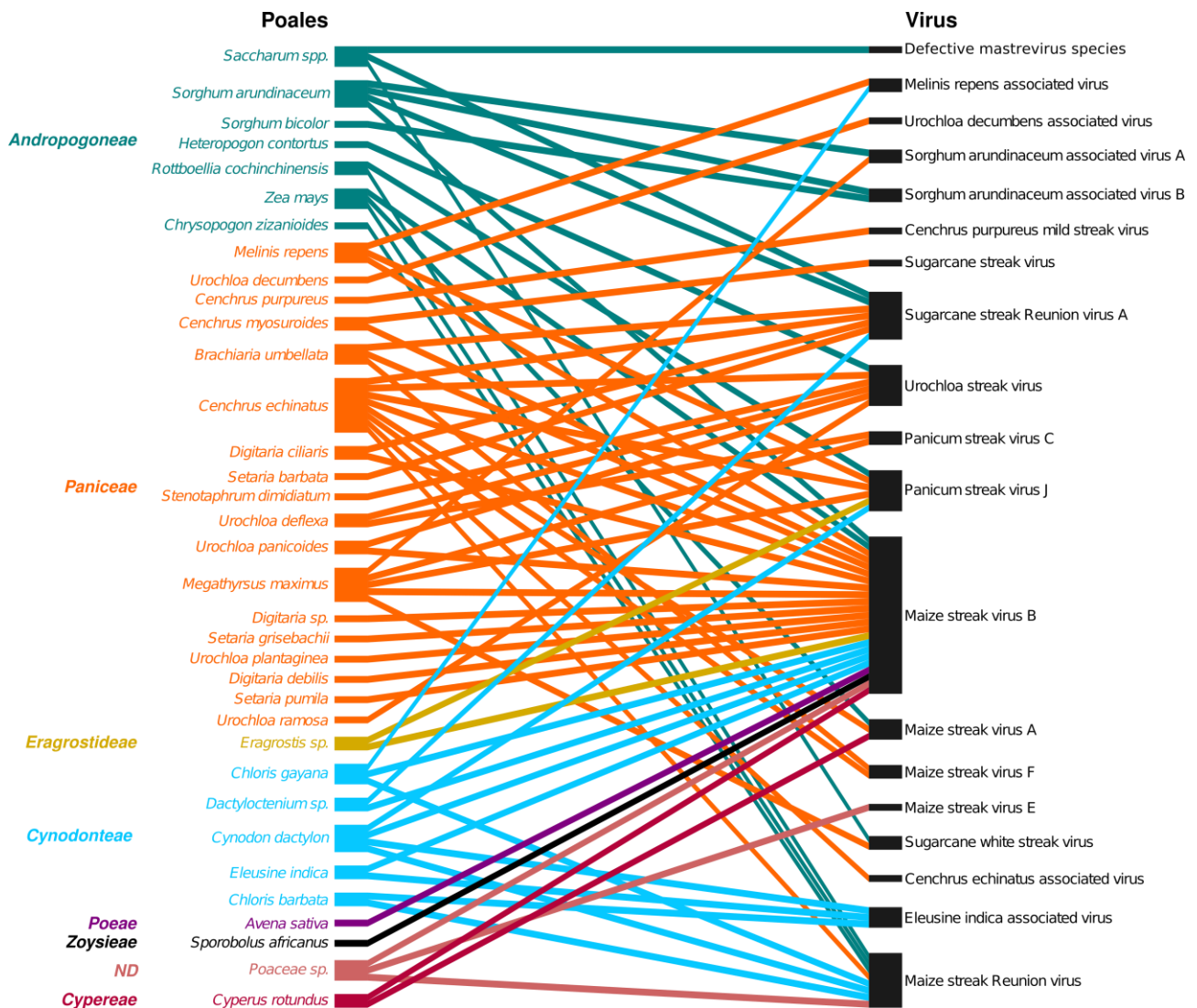


Figure 3. Bipartite interaction graph representing the association between Poales species (left side of the diagram) and viral species and strains (right side of the diagram). Poales species are colored and organized by tribe. ND: no determined.

Three different structural features of the virus-host species interaction network – assortativity, nestedness and modularity – were assessed to determine whether the interactions represented by the network deviated from expectations under a model of random virus-host species interactions. Assortativity analysis revealed no association between the network structure and the classification of hosts according to whether they were native or alien species (p-values of [0.314 ; 0.511]) or whether they were annuals, perennials or annual/perennials (p-values of [0.189 ; 0.471]). Conversely, the tribes to which host species belonged contributed substantially to defining the structure of the network with higher assortativity (meaning a higher tendency for plants from the same tribe to host similar complements of virus entities) obtained in comparison to the majority of null

models employed for permutation (p -value ≥ 0.990 ; **Table 4**). Whereas this result would go the way of modularity in the network (*i.e.* plants from similar tribe tend to harbor similar viruses), no significant modularity was observed (**Table 5**). Conversely, significant degrees of nestedness were detected within the network (p -value ≥ 0.981 ; **Table 5**) implying that the host-ranges of the viruses tended to be embedded in one another. Nestedness is frequently a feature in host-parasites communities (Flores *et al.*, 2011 ; Moury *et al.*, 2021). It has been suggested that nestedness would make for robust (Bascompte *et al.*, 2006 ; Memmott *et al.*, 2004) and species rich networks (Bastolla *et al.*, 2009). Conversely, significant degrees of modularity within the network would have suggested a pattern of virus-host species interactions that has arisen through parasites having become specialized in some hosts at the expense of losing their ability to infect other hosts: *i.e.* a host-shifting pattern of evolution rather than a host-range expansion (Moury *et al.*, 2017). It is a signal of strong host-parasite associations that would lead to less stable species-poor networks (Grilli *et al.*, 2016 ; Dormann *et al.*, 2017 ; Valverde *et al.*, 2020). It has been hypothesized that community stability is a prerequisite for the emergence of modularity (Valverde *et al.*, 2020 ; Moury *et al.*, 2021). Here, the plant community is composed of 57% (20/35) of cultivated or weed annual species (6/35 cultivated and 14/35 weed annual species ; **Figure 3** ; **Table 2** ; **Supplementary Table 2**). These cultivated and weed annual species account for 62% (46/74) of the interactions within the network. Whereas the cultivated annual species are periodically harvested and removed from the ecosystem, so too are the annual weed species by natural seasonal changes: and it is plausible that this factor may prevent the emergence of modular interactions within the virus-host species interaction network. The 20 cultivated and weed annual plant species considered here include the unique hosts of four of the “specialist” mastrevirus species (CEAV, CPMSV, UDAV and DefMS): all of which have been described here for the first time. CPMSV and DefMS are associated with vegetatively propagated plants which may preclude the requirement for an insect vector and/or alternative hosts for these species. However, the hosts of CEAV and UDAV are not vegetatively propagated which suggests that, in the absence of seed transmission, these viruses might require either alternative currently unidentified host species or a long-lived vector species within which they can overwinter. The ten other *Mastrevirus* species that are represented within the network were each found infecting two or more different host species among which at least one was an weed

perennial host that could, conceivably, facilitate the persistence of the viruses across the seasons.

Table 4. Analysis of assortativity of *Poales-Mastrevirus* interaction matrices. The probability value (p-value) indicates the frequency of the actual matrix showing a strictly higher assortativity score than that of null-model matrices. P-values > 0.95 (significant assortativity) are in white on black cells and p-values ≤ 0.05 (significant disassortativity) are in bold on grey cells.

Plant characteristic	Assortativity score	Null model			
		none	rows	columns	both
Tribe	-1.4x10 ⁻³	0.864	0.806	0.990	0.992
Invasion status	-4.6x10 ⁻²	0.511	0.473	0.332	0.314
Life cycle	-5.1x10 ⁻²	0.471	0.423	0.202	0.189

Table 5. Analysis of nestedness and modularity of *Poales-Mastrevirus* interaction matrices with respectively *NODF2* and *Newman* algorithm. The probability value (p-value) indicates the frequency of the actual matrix showing a strictly higher nestedness or modularity score than that of null-model matrices. P-values > 0.95 (significant nestedness or modularity) are in white on black cells and p-values ≤ 0.05 (significant anti-nestedness or anti-modularity) are in bold on grey cells.

Matrix	Nestedness score	<i>NODF2</i>				Modularity score	<i>Newman algorithm</i>			
		Null model					Null model			
		none	rows	columns	both		none	rows	columns	both
CG	38.89	0.644	0.644	0.44	0.502	0.44	0.930	0.752	0.830	0.287
AE	34.58	1	1	0.981	0.460	0.45	0.005	0	0.283	0.011
Reunion	30.52	1	1	0.991	0.430	0.48	0.061	0	0.737	0.256

Not all interactions are achieved in the realized network

Globally, 24, 15 and four hosts were found at the AE, CG and SG site-types respectively. Apart from the MV and PC sites (both SG site types), where no known mastrevirus hosts were sampled, the other eight sampling sites had between three to 18 known mastrevirus hosts present (**Figure 4**). The analysis of viruses richness among sites revealed that there was a positive correlation (Pearson's $R^2 = 0.81$; p-value=4.2x10⁻⁴) between the numbers of mastrevirus host species present and the numbers of mastrevirus species detected at the sampling sites (**Supplementary Figure 3**). Going one step further and using information on the known natural host-ranges of the mastreviruses detected, we calculated the maximum numbers of viral species that could have been detected at each sampling site – which are contingent on the actual host-species that were actually present at each

site – and calculated the proportions of these viruses that were actually observed at each of the sites. There was a significant correlation (Pearson's $R^2 = 0.77$; $p\text{-value}=4.0\times 10^{-3}$) between the realized mastrevirus species richness and the richness of known mastrevirus host species found at those sites (**Figure 4**). Beyond indicating an association between host species diversity and virus species diversity, this significant correlation indicates a tendency for sites at which mastrevirus hosts are most abundant to present higher niche fillings than sites where hosts are less abundant (**Supplementary Figure 3**). Indeed, whereas at the BP sampling site, all ten of the virus species that might have been expected given the plant hosts that were present, were detected, at the LE sampling site where four known mastrevirus hosts were sampled (*C. purpureus*, *C. dactylon*, *S. africanus* and *C. rotundus*) and abundant leaf-hoppers were found (including *C. mbila*) no *Mastrevirus* species were detected. However the fact that *C. purpureus*, *C. dactylon*, *S. africanus* and *C. rotundus* were all potential mastrevirus hosts but remained negative for the presence of mastreviruses across multiple sampling sites where they were found (including some of the AE sites) may indicate that mastreviruses in Reunion are not particularly well adapted to infecting these particular plant species. Increasing host diversity and potentially “optimal” host count would in turn increase potential maintenance of higher viral richness.

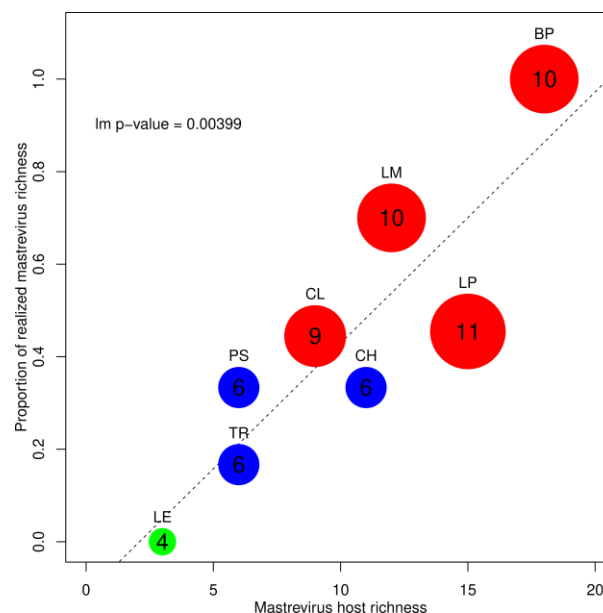


Figure 4. Proportion of realized *Mastrevirus* richness as a function of plant host richness. Sites are presented as colored dots (red for AE sites, blue for CG sites and green for SG sites) whose size represent the maximum number of *Mastrevirus* species that could have been achieved knowing the plant virus diversity of the site. The maximum number of species is also indicated on each dot. For each site, the proportion is obtained by dividing the realized number of *Mastrevirus* species with the maximum number of *Mastrevirus*

species achievable. Linear regression is indicated with the dotted line. The p-value for the linear regression slope to be different from zero is indicated on the top left.

Going one step further, we analyzed the numbers of mastrevirus-host species interactions that were observed at each of the sampling sites as a function of the fraction of the maximum number of interactions that could have been observed at the sites (**Supplementary Figure 3**). Doing so, we accounted for the fact that some of the mastreviruses can infect multiple host species. These “realized interaction proportions” varied from 0% for the LE sampling site to 100% for the CL sampling site. At the BP sampling site (the one at which the highest number of mastrevirus species was observed), 67% of all the mastrevirus-host species interactions that could have been observed at the site, were actually observed. While the numbers of interaction were significantly correlated with mastrevirus species and known mastrevirus host species counts (*i.e.* virus and host richness) there was no detectable correlation between the proportions of the potential virus-plant host interactions that were actually observed and either counts of observed mastrevirus species or counts of known mastrevirus host species. This implies that the number of host species that a given virus was observed to infect at a site was not obviously influenced by either the numbers of known mastrevirus host species found at the site or the numbers of different mastrevirus species found at that site.

Bearing in mind that the virus species accounted for in our network display distinct (although in some cases overlapping) host ranges, it was anticipated that more host species at a sampling site would be conducive to the discovery of more mastrevirus species at the site. This relationship would be consistent with the amplification hypothesis (Rottstock *et al.*, 2014; Kamiya *et al.*, 2014; Halliday *et al.*, 2020) which posits that disease risks and/or pathogen frequencies increase with increasing host diversity. Consistent with this hypothesis and similar to the observation that we have made here, it has been observed that, irrespective of the particular viral genus being considered, numbers of detectable *Poaceae* infecting virus species increase with increasing numbers sampled *Poaceae* species (Maclot *et al.*, 2023), tending towards that regardless of the viral genus or family studied, the tendency may be consistently observed.

In this light the correlation between the proportion of the possible virus-plant host interactions that were actually observed and counts of observed mastrevirus species and known host species was somewhat unexpected. Indeed there seems to be a “Rich get richer” relationship, wherein more host species or virus species would yield an increase in the proportion of the potentially observable virus-host species interactions that will actually be observed. While this could indicate that synergism between virus species might occur, the absence of an obvious correlation between the observed numbers of virus-plant host species interactions and the observed counts of plant host species or virus species would conversely be indicative of an absence of synergism.

It is entirely possible that increases in observed virus species counts are directly and uniquely associated with environmental conditions that are more favorable for virus transmission (*i.e.* conditions that foster the proliferation and survival of vector species). The fact that no mastrevirus species were detected at the LE sampling site despite the occurrence at this site of abundant *C. mbila* would be counter-intuitive. Given the high altitude of the LE sampling site, large variations of temperature throughout the year could greatly limit local viruses transmissions. Both dilution and amplification effects have been theorized, and observed, for vector-borne pathogens (Rohr *et al.*, 2020) and the relationship between host and parasite species richness is non trivial and dependent of vector feeding behavior and the specifics of parasite transmission (Keesing *et al.*, 2006). For mastreviruses, the absence of definitive information of the vectors in general, and in the context of Reunion in particular, makes it difficult to interpret the observed pattern. The positive correlation obtained between observed host species counts and observed mastrevirus species counts at sampling sites on Reunion may imply the absence of substantial variations between the observed mastrevirus species with respect to vector specificity and/or complex variations in the feeding behaviors of the various leafhopper species that presumably transmit these viruses on the island.

Here we find that the distribution of *Poaceae*-infecting mastreviruses found in different ecological contexts on the island of Reunion is consistent with the viruses having originated from multiple, likely human-mediated, introductions to the island of African *Poaceae* species and their associated complements of viral species. The observed

patterns of mastrevirus and host species diversity and the of host-pathogen interactions that are observed to occur on the island, imply that the mastreviruses, their plant host species and their insect vector species form a community that is likely most suitable for the survival and propagation of viruses within agro-ecological contexts – *i.e.* low altitude environments with large numbers of alien *Poaceae* species that are capable of hosting mastreviruses that grow in close proximity to native *Poaceae* species that are also capable of hosting mastreviruses. Our analyses of mastrevirus-host species interaction patterns and their associations with mastrevirus species richness suggest that it may be necessary to obtain a better understanding of the impacts of environmental factors on viral transmission dynamics to more fully explain why mastrevirus-host species interactions that are known to be possible are frequently undetectable at field sites that are otherwise apparently conducive to the occurrence of these interactions.

Authors information

Affiliations

CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 chemin de l'Irat, 97410 Saint Pierre, La Réunion, France

Sélim Ben Chehida, Murielle Hoarau, Jean-Michel Lett & Pierre Lefeuvre

CIRAD, UMR BGPI, TA A-54/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

Oumaima Moubset, Charlotte Julian, Laurence Blondin, Denis Filloux & Philippe Roumagnac

Conservatoire Botanique National de Mascarin, F-97436, St Leu, La Réunion, France
Christophe Lavergne

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

Arvind Varsani

Computational Biology Division, Departement of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory 7925, South Africa

Darren P. Martin

Author contributions

S.B.C., J.M.L and P.L. conceived and designed the experiments. S.B.C., H.D.B., M.H., O.M., C.J., L.B., C.L., J.M.L and P.L. performed the experiments. S.B.C., D.P.M., J.M.L and P.L. analysed the data. S.B.C, P.R., A.V., D.P.M., J.M.L and P.L. wrote the paper. J.M.L. and P.L. secured funding for the project's execution.

Corresponding author

Correspondence to Pierre Lefeuvre

Funding

This study was funded by the European Union (ERDF, contract GURDT I2016-1731-0006632), the Conseil Régional de la Réunion and CIRAD. S.B.C. is a recipient of a PhD fellowship from CIRAD and ANR (Phytovirus project number: ANR-19-CE35-0008-02).

Data Availability Statement

Publicly available datasets were analyzed in this study.

Acknowledgments

This work was conducted on the Plant Protection Platform (3P, IBISA). A. Ouattara and D. Otron helped with collecting the samples.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Almeida-Neto, M., Guimarães, P., Guimarães, P.R., Loyola, R.D. & Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement. *Oikos*, 117 (8), pp.1227-1239
- Bascompte, J., Jordano, P., & Olesen, J. M. (2006). Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, 312(5772), 431-433.
- Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458(7241), 1018-1020.
- Beckett, S.J. (2016). Improved community detection in weighted bipartite networks. *Royal Society open sciences* 3: 140536.
- Ben Chehida, S., Filloux, D., Fernandez, E., Moubset, O., Hoareau, M., Julian, C., Blondin, L., Lett, J. M., Roumagnac, P. & Lefeuvre, P. (2021). Nanopore Sequencing Is a Credible Alternative to Recover Complete Genomes of Geminiviruses. *Microorganisms*, 9, 903.
- Ben Chehida, S., Hoareau, M., Fénélon, B., Grondin, M., Peterschmitt, M., Varsani, A., Martin, D. P., Teycheney, P. Y., Lefeuvre, P. & Lett, J.-M. (in preparation). A novel chimeric circular Rep-encoding single stranded DNA virus with interfamilial gene exchange illustrates the considerable evolutionary capacity of ssDNA viruses.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: series B (Methodological)*, 57(1), 289-300.
- 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. (2018). Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME journal*, 12(1), 173-184.
- Bigarré, L., Salah, M., Granier, M., Frutos, R., Thouvenel, J. C., & Peterschmitt, M. (1999). Nucleotide sequence evidence for three distinct sugarcane streak mastreviruses. *Archives of Virology*, 144(12), 2331-2344.
- Bonfils, J., Quilici, S., & Reynaud, B. (1994). Les Hémiptères Auchénorhynques de l'île de la Réunion. *Bulletin de la Société entomologique de France*, 99 (3), 1994 : 227-240.
- Boni, M. F., Posada, D., & Feldman, M. W. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*, 176(2), 1035-1047.
- Briddon, R. W., Lunness, P., Chamberlin, L. C. L., Pinner, M. S., Brundish, H., & Markham, P. G. (1992). The nucleotide sequence of an infectious insect-transmissible clone of the geminivirus Panicum streak virus. *Journal of General Virology*, 73(5), 1041-1047.
- Briddon, R. W., Lunness, P., Bedford, I. D., Chamberlin, L. C., Mesfin, T., & Markham, P. G. (1996). A streak disease of pearl millet caused by a leafhopper-transmitted geminivirus. *European Journal of Plant Pathology*, 102, 397-400.
- Briddon, R. W., Martin, D. P., Roumagnac, P., Navas-Castillo, J., Fiallo-Olivé, E., Moriones, E., Lett, J.-M., Zerbini, F. M. & Varsani, A. (2018). *Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus-and nanovirus-associated alphasatellites. *Archives of virology*, 163, 2587-2600.
- Claverie, S., Ouattara, A., Hoareau, M., Filloux, D., Varsani, A., Roumagnac, P., Lett, J. M. & Lefeuvre, P. (2019). Exploring the diversity of Poaceae-infecting mastreviruses on Reunion Island using a viral metagenomics-based approach. *Scientific reports*, 9(1), 1-11.

- Claverie, S., Hoareau, M., Ben Chehida, S., Filloux, D., Varsani, A., Roumagnac, P., Martin, D. P., Lett, J. M. & Lefeuvre, P. (2023). Metagenomics reveals the structure of Mastrevirus-host interaction network within an agro-ecosystem. *Virus Evolution*, vead043.
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *International Journal of Complex Systems*, 1695(5), 1-9.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2), giab008.
- Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite package: analysing ecological networks. *interaction*, 1(0.2413793), 8-11.
- Dormann, C. F., & Strauss, R. (2014). A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, 5(1), 90-98.
- Dormann, C. F., Fründ, J., & Schaefer, H. M. (2017). Identifying causes of patterns in ecological networks: opportunities and limitations. *Annual Review of Ecology, Evolution, and Systematics*, 48, 559-584.
- Faillace, C. A., Lorusso, N. S., & Duffy, S. (2017). Overlooking the smallest matter: viruses impact biological invasions. *Ecology Letters*, 20(4), 524-538.
- Faust, C. L., Dobson, A. P., Gottdenker, N., Bloomfield, L. S., McCallum, H. I., Gillespie, T. R., Diuk-Wasser, M. & Plowright, R. K. (2017). Null expectations for disease dynamics in shrinking habitat: dilution or amplification? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1722), 20160173.
- Fenouillas, P., Ah-Peng, C., Amy, E., Bracco, I., Dafreville, S., Gosset, M., Ingrassa, F., Lavergne, C., Lequette, B., Notter, J.-C., Pausé, J.-M., Payet, G., Payet, N., Picot, F., Pougavanon, N., Strasberg, D., Thomas, H., Triolo, J., Turquet, V. & Rouget, M. (2021). Quantifying invasion degree by alien plants species in Reunion Island. *Austral Ecology*, 46(7), 1025-1037.
- Flores, C. O., Meyer, J. R., Valverde, S., Farr, L., & Weitz, J. S. (2011). Statistical structure of host-phage interactions. *Proceedings of the National Academy of Sciences*, 108(28), E288-E297.
- Fraile, A., McLeish, M. J., Pagán, I., González-Jara, P., Piñero, D., & García-Arenal, F. (2017). Environmental heterogeneity and the evolution of plant-virus interactions: viruses in wild pepper populations. *Virus Research*, 241, 68-76.
- García-Arenal, F., & Zerbini, F. M. (2019). Life on the edge: geminiviruses at the interface between crops and wild plant hosts. *Annual Review of Virology*, 6, 411-433.
- Gibbs, M. J., Armstrong, J. S., & Gibbs, A. J. (2000). Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*, 16(7), 573-582.
- Grilli, J., Rogers, T., & Allesina, S. (2016). Modularity and stability in ecological communities. *Nature Communications*, 7(1), 12031.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321.
- Halliday, F. W., Heckman, R. W., Wilfahrt, P. A., & Mitchell, C. E. (2017). A multivariate test of disease risk reveals conditions leading to disease amplification. *Proceedings of the Royal Society B: Biological Sciences*, 284(1865), 20171340.
- Hollander, M., & Wolfe, D. A. (1973). Kruskal-Wallis: a distribution-free test. *Nonparametric Statistical Methods*, 114-137.

- Holmes, E. C. (2022). The ecology of viral emergence. *Annual Review of Virology*, 9, 173-192.
- 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 ϕ 29 DNA polymerase. *Journal of Virological Methods*, 116(2), 209-211.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587-589.
- Kamiya, T., O'Dwyer, K., Nakagawa, S., & Poulin, R. (2014). Host diversity drives parasite diversity: meta-analytical insights into patterns and causal mechanisms. *Ecography*, 37(7), 689-697.
- Kassambara, A., & Mundt, F. (2020). factoextra: Extract and visualize the results of multivariate data analyses, R package v. 1.0. 7. *factoextra: Extract and visualize the results of multivariate data analyses, R package v. 1.0. 7.*
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.
- Keesing, F., Holt, R. D., & Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters*, 9(4), 485-498.
- Kolmogorov, M.; Yuan, J.; Lin, Y. & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Natural Biotechnology* 37,540–546.
- Kopp, A. (1930). Les maladies des plantes a La Réunion. *Rev. Bot. Appl. Agric. Trop.* 105:1–7. (In French.)
- Kraberger, S., Saumtally, S., Pande, D., Khoodoo, M. H., Dhayan, S., Dookun-Saumtally, A., Shepherd, D. N., Hartnady, P., Atkinson, R., Lakay, F. M., Hanson, B., Redhi, D., Monjane, A. L., Windram, O. P., Walters, M., Oluwafemi, S., Lett, J. M., 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.
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094–3100.
- Maclot, F., Debue, V., Malmstrom, C. M., Filloux, D., Roumagnac, P., Eck, M., Tamisier, L., Blouin, A. G., Candresse, T. & Massart, S. (2023). Long-Term Anthropogenic Management and Associated Loss of Plant Diversity Deeply Impact Virome Richness and Composition of Poaceae Communities. *Microbiology Spectrum*, 11(2), e04850-22.
- Malmstrom, C. M., & Alexander, H. M. (2016). Effects of crop viruses on wild plants. *Current Opinion in Virology*, 19, 30-36.
- Martin, D., & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics*, 16(6), 562-563.
- Martin, D. P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T., Kelz, Z., Kumar, V. & Murrell, B. (2021). RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution*, 7(1), veaa087.
- Martinez, M. E. (2018). The calendar of epidemics: Seasonal cycles of infectious diseases. *PLoS pathogens*, 14(11), e1007327.
- Medaka. Available online: <https://github.com/nanoporetech/medaka> (accessed on 09 February 2023)

- Memmott, Jane, Nickolas M. Waser, and Mary V. Price. "Tolerance of pollination networks to species extinctions." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271.1557 (2004): 2605-2611.
- McLeish, M., Sacristán, S., Fraile, A., & García-Arenal, F. (2017). Scale dependencies and generalism in host use shape virus prevalence. *Proceedings of the Royal Society B: Biological Sciences*, 284(1869), 20172066.
- McLeish, M. J., Fraile, A., & García-Arenal, F. (2019). Evolution of plant–virus interactions: host range and virus emergence. *Current Opinion in Virology*, 34, 50-55.
- McLeish, M. J., Zamfir, A. D., Babalola, B. M., Peláez, A., Fraile, A., & García-Arenal, F. (2022). Metagenomics show high spatiotemporal virus diversity and ecological compartmentalisation: Virus infections of melon, Cucumis melo, crops, and adjacent wild communities. *Virus Evolution*, 8(2), veac095.
- Miniasm. Available online: <https://github.com/lh3/miniasm> (accessed on 09 February 2023)
- Monjane, A. L., Harkins, G. W., Martin, D. P., Lemey, P., Lefeuvre, P., Shepherd, D. N., Oluwafemi, S., Simuyandi, M., Zinga, I., Komba, E. K., Lakoutene, D. P., Mandakombo, N., Mboukoulida, J., Semballa, S., Tagne, A., Tiendrébéogo, F., Erdmann, J. B., van Antwerp, T., Owaor, B. E., Flett, B., Ramusi, M., Windram, O. P., Syed, R., Lett, J.-M., Briddon, R. W., Markham, P. G., Rybicki, E. P. & Varsani, A. (2011). Reconstructing the history of maize streak virus strain a dispersal to reveal diversification hot spots and its origin in southern Africa. *Journal of Virology*, 85(18), 9623-9636.
- Moury, B., Fabre, F., Hébrard, E., & Froissart, R. (2017). Determinants of host species range in plant viruses. *Journal of General Virology*, 98(4), 862-873.
- Moury, B., Audergon, J. M., Baudracco-Arnas, S., Ben Krima, S., Bertrand, F., Boissot, N., Buisson, M., Caffier, V., Cantet, M., Chanéac, S., Constant, C., Delmotte, F., Dogimont, C., Doumayrou, J., Fabre, F., Fournet, S., Grimault, V., Jaunet, T., Justafré, I., Lefebvre, V., Losdat, D., Marcel, T., Montarry, J., Morris, C. E., Omrani, M., Paineau, M., Perrot, S., Pilet-Nayel, M. L. & Ruellan, Y. (2021). The quasi-universality of nestedness in the structure of quantitative plant-parasite interactions. *Peer Community Journal*, 1.
- Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., Rivera-Bustamante, R., Malathi, V. G., Briddon, R. B. & 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.
- Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, 67(2), 026126.
- Newman, M. E. (2004). Detecting community structure in networks. *The European Physical Journal B*, 38, 321-330.
- Nielson, M. W. (1968). *The leafhopper vectors of phytopathogenic viruses (Homoptera, Cicadellidae): taxonomy, biology, and virus transmission* (No. 1382). US Agricultural Research Service.
- O'Malley, M. A. (2016). The ecological virus. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 59, 71-79.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P; R., O'hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. & Maintainer, H. W. (2019). Vegan: community ecology package. *Ordination methods, diversity analysis and other functions for community and vegetation ecologists*, 05–26. Version 2.5-1.
- Olesen, J. M., Bascompte, J., Dupont, Y. L., & Jordano, P. (2007). The modularity of pollination networks. *Proceedings of the National Academy of Sciences*, 104(50), 19891-19896.

- Oxford Nanopore Technologies. Available online: <https://nanoporetech.com/> (accessed on 09 February 2023).
- Padidam, M., Sawyer, S., & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology*, 265(2), 218-225.
- Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.
- Peterschmitt, M., Granier, M., Frutos, R., & Reynaud, B. (1996). Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Reunion Island. *Archives of Virology*, 141, 1637-1650.
- Posada, D., & Crandall, K. A. (2001). Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences*, 98(24), 13757-13762.
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS one*, 5(3), e9490.
- 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): Approches biologique, génétique et épidémiologique de la relation vecteur-virus-plante (Doctoral dissertation).
- Rohr, J. R., Civitello, D. J., Halliday, F. W., Hudson, P. J., Lafferty, K. D., Wood, C. L., & Mordecai, E. A. (2020). Towards common ground in the biodiversity–disease debate. *Nature Ecology & Evolution*, 4(1), 24-33.
- Rottstock, T., Joshi, J., Kummer, V., & Fischer, M. (2014). Higher plant diversity promotes higher diversity of fungal pathogens, while it decreases pathogen infection per plant. *Ecology*, 95(7), 1907-1917.
- Roumagnac, P., Lett, J. M., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F. M., Martin, D. P., & Varsani, A. (2022). Establishment of five new genera in the family Geminiviridae: Citlodavirus, Maldovirus, Mulcrilevirus, Opunvirus, and Topilevirus. *Archives of Virology*, 1-16.
- Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, 17(2), 155-158.
- Salminen, M. O., Carr, J. K., Burke, D. S., & McCUTCHAN, F. E. (1995). Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS research and human retroviruses*, 11(11), 1423-1425.
- Shates, T. M., Sun, P., Malmstrom, C. M., Dominguez, C., & Mauck, K. E. (2019). Addressing research needs in the field of plant virus ecology by defining knowledge gaps and developing wild dicot study systems. *Frontiers in Microbiology*, 9, 3305.
- Shepherd, D. N., Varsani, A., Windram, O. P., Lefeuvre, P., Monjane, A. L., Owor, B. E., & Martin, D. P. (2008). Novel sugarcane streak and sugarcane streak Reunion mastreviruses from southern Africa and La Reunion. *Archives of Virology*, 153, 605-609.
- Shepherd, D. N., Martin, D. P., Van der Walt, E., Dent, K., Varsani, A., & Rybicki, E. P. (2010). Maize streak virus: an old and complex 'emerging' pathogen. *Molecular Plant Pathology*, 11(1), 1-12.
- Simon, A. E., Roossinck, M. J., & Havelda, Z. (2004). Plant virus satellite and defective interfering RNAs: new paradigms for a new century. *Annual Review Phytopathol.*, 42, 415-437.
- Smith, J. M. (1992). Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, 34, 126-129.

- Stepanauskas, R., Fergusson, E. A., Brown, J., Poulton, N. J., Tupper, B., Labonté, J. M., Becraft, E. D., Brown J. M., Pachiadaki, M. G., Povilaitis, T., Thompson, B. P., Mascena, C. J., Bellows, W. K. & Lubys, A. (2017). Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. *Nature Communications*, 8(1), 84.
- Sullivan, M. B., Weitz, J. S., & Wilhelm, S. (2017). Viral ecology comes of age. *Environmental Microbiology Reports*, 9(1), 33-35.
- Susi, H., Filloux, D., Frilander, M. J., Roumagnac, P., & Laine, A. L. (2019). Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ*, 7, e6140.
- Susi, H., & Laine, A. L. (2021). Agricultural land use disrupts biodiversity mediation of virus infections in wild plant populations. *New Phytologist*, 230(6), 2447-2458.
- Valverde, S., Vidiella, B., Montañez, R., Fraile, A., Sacristán, S., & García-Arenal, F. (2020). Coexistence of nestedness and modularity in host–pathogen infection networks. *Nature Ecology & Evolution*, 4(4), 568-577.
- Wasik, B. R., & Turner, P. E. (2013). On the biological success of viruses. *Annual Review of Microbiology*, 67, 519-541.
- Webb, M. D. (1987). Species recognition in Cicadulina leafhoppers (Hemiptera: Cicadellidae), vectors of pathogens of Gramineae. *Bulletin of Entomological Research*, 77(4), 683-712.
- Weitz, J. S., Poisot, T., Meyer, J. R., Flores, C. O., Valverde, S., Sullivan, M. B., & Hochberg, M. E. (2013). Phage–bacteria infection networks. *Trends in Microbiology*, 21(2), 82-91.

Supplementary Table 1. Summary of sampled pool in each type site and site for each campaign.

Date Sampling	Type of Site	Site	Poales Species	Pool Id	Number of Samples
Nov 2020	AE	BP	<i>Chloris gayana</i>	20_REU_E0102	50
Nov 2020	AE	BP	<i>Cyperus rotundus</i>	20_REU_E0103	50
Nov 2020	AE	BP	<i>Cynodon dactylon</i>	20_REU_E0109	50
Nov 2020	AE	BP	<i>Bromus catharticus</i>	20_REU_E0111	50
Nov 2020	SG	LE	<i>Bromus catharticus</i>	20_REU_E0303	50
Nov 2020	SG	LE	<i>Cyperus rotundus</i>	20_REU_E0304	50
Nov 2020	SG	LE	<i>Holcus lanatus</i>	20_REU_E0306	50
Nov 2020	SG	LE	<i>Cynodon dactylon</i>	20_REU_E0311	50
Nov 2020	SG	LE	<i>Cenchrus clandestinus</i>	20_REU_E0312	50
Nov 2020	SG	MV	<i>Dactylis glomerata</i>	20_REU_E0401	50
Nov 2020	SG	MV	<i>Holcus lanatus</i>	20_REU_E0406	50
Nov 2020	SG	MV	<i>Carex leporina</i>	20_REU_E0408	50
Nov 2020	SG	MV	<i>Juncus effusus</i>	20_REU_E0410	50
Nov 2020	SG	MV	<i>Anthoxanthum odoratum</i>	20_REU_E0412	50
Nov 2020	SG	MV	<i>Cenchrus clandestinus</i>	20_REU_E0415	50
Nov 2020	SG	MV	<i>Carex boryana</i>	20_REU_E0418	50
Nov 2020	SG	PC	<i>Juncus effusus</i>	20_REU_E0502	50
Nov 2020	SG	PC	<i>Holcus lanatus</i>	20_REU_E0505	50
Nov 2020	SG	PC	<i>Anthoxanthum odoratum</i>	20_REU_E0509	50
Nov 2020	SG	PC	<i>Carex leporina</i>	20_REU_E0510	50
Nov 2020	SG	PC	<i>Panicum juniperinum</i>	20_REU_E0513	50
Nov 2020	SG	PC	<i>Festuca borbonica</i>	20_REU_E0515	50
Nov 2020	SG	PC	<i>Costularia melicoides</i>	20_REU_E0516	50
Nov 2020	SG	PC	<i>Costularia cadetii</i>	20_REU_E0517	50
Nov 2020	SG	PC	<i>Carex boryana</i>	20_REU_E0518	50
Jan 2021	AE	BP	<i>Eragrostis sp.</i>	21_REU_E0101	50
Jan 2021	AE	BP	<i>Digitaria ciliaris</i>	21_REU_E0102	50
Jan 2021	AE	BP	<i>Cyperus rotundus</i>	21_REU_E0103	50
Jan 2021	AE	BP	<i>Paspalum dilatatum</i>	21_REU_E0104	50
Jan 2021	AE	BP	<i>Dactyloctenium sp.</i>	21_REU_E0105	50
Jan 2021	AE	BP	<i>Chloris gayana</i>	21_REU_E0106	50
Jan 2021	AE	BP	<i>Megathyrsus maximus</i>	21_REU_E0107	50
Jan 2021	AE	BP	<i>Melinis repens</i>	21_REU_E0108	50
Jan 2021	AE	BP	<i>Cynodon dactylon</i>	21_REU_E0109	50
Jan 2021	AE	BP	<i>Sorghum arundinaceum</i>	21_REU_E0110	50

CHAPITRE 2

Jan 2021	AE	BP	<i>Cenchrus echinatus</i>	21_REU_E0111	50
Jan 2021	AE	BP	<i>Eleusine indica</i>	21_REU_E0112	50
Jan 2021	AE	BP	<i>Urochloa deflexa</i>	21_REU_E0113	50
Apr 2021	CG	TR	<i>Cynodon dactylon</i>	21_REU_E0201	50
Apr 2021	CG	TR	<i>Dactyloctenium sp.</i>	21_REU_E0202	50
Apr 2021	CG	TR	<i>Heteropogon contortus</i>	21_REU_E0203	50
Apr 2021	CG	TR	<i>Megathyrsus maximus</i>	21_REU_E0204	50
Apr 2021	CG	PS	<i>Chloris barbata</i>	21_REU_E0701	50
Apr 2021	CG	PS	<i>Cynodon dactylon</i>	21_REU_E0702	50
Apr 2021	CG	PS	<i>Dactyloctenium sp.</i>	21_REU_E0703	50
Apr 2021	CG	PS	<i>Heteropogon contortus</i>	21_REU_E0704	50
Apr 2021	CG	PS	<i>Megathyrsus maximus</i>	21_REU_E0705	50
Apr 2021	CG	PS	<i>Setaria pumila</i>	21_REU_E0706	50
Apr 2021	CG	PS	<i>Urochloa panicoides</i>	21_REU_E0707	50
Apr 2021	CG	PS	<i>Bothriochloa sp.</i>	21_REU_E0708	50
Apr 2021	CG	PS	<i>Tragus mongolorum</i>	21_REU_E0709	50
Apr 2021	CG	CH	<i>Aristida adscensionis</i>	21_REU_E0601	50
Apr 2021	CG	CH	<i>Chloris barbata</i>	21_REU_E0602	50
Apr 2021	CG	CH	<i>Cynodon dactylon</i>	21_REU_E0603	50
Apr 2021	CG	CH	<i>Dactyloctenium sp.</i>	21_REU_E0604	50
Apr 2021	CG	CH	<i>Digitaria ciliaris</i>	21_REU_E0605	50
Apr 2021	CG	CH	<i>Eleusine indica</i>	21_REU_E0606	50
Apr 2021	CG	CH	<i>Enneapogon cenchroides</i>	21_REU_E0607	50
Apr 2021	CG	CH	<i>Heteropogon contortus</i>	21_REU_E0608	50
Apr 2021	CG	CH	<i>Megathyrsus maximus</i>	21_REU_E0609	50
Apr 2021	CG	CH	<i>Melinis repens</i>	21_REU_E0610	50
Apr 2021	CG	CH	<i>Themeda quadrivalvis</i>	21_REU_E0611	50
Apr 2021	CG	CH	<i>Urochloa panicoides</i>	21_REU_E0612	50
Apr 2021	CG	CH	<i>Urochloa ramosa</i>	21_REU_E0613	50
Apr 2021	AE	BP	<i>Stenotaphrum dimidiatum</i>	21_REU_E0116	50
Apr 2021	AE	BP	<i>Cenchrus echinatus</i>	21_REU_E0117	50
Apr 2021	AE	BP	<i>Chloris gayana</i>	21_REU_E0118	50
Apr 2021	AE	BP	<i>Chrysopogon zizanioides</i>	21_REU_E0119	50
Apr 2021	AE	BP	<i>Cynodon dactylon</i>	21_REU_E0120	50
Apr 2021	AE	BP	<i>Dactyloctenium sp.</i>	21_REU_E0121	50
Apr 2021	AE	BP	<i>Eleusine indica</i>	21_REU_E0122	50
Apr 2021	AE	BP	<i>Eragrostis sp.</i>	21_REU_E0123	50
Apr 2021	AE	BP	<i>Megathyrsus maximus</i>	21_REU_E0124	50
Apr 2021	AE	BP	<i>Melinis repens</i>	21_REU_E0125	50

CHAPITRE 2

Apr 2021	AE	BP	<i>Paspalum dilatatum</i>	21_REU_E0126	50
Apr 2021	AE	BP	<i>Saccharum officinarum</i>	21_REU_E0127	50
Apr 2021	AE	BP	<i>Sorghum arundinaceum</i>	21_REU_E0128	50
Apr 2021	AE	BP	<i>Zea mays</i>	21_REU_E0129	50
Apr 2021	AE	BP	<i>Urochloa panicoides</i>	21_REU_E0130	50
Apr 2021	AE	CL	<i>Bromus catharticus</i>	21_REU_E0801	50
Apr 2021	AE	CL	<i>Cenchrus purpureus</i>	21_REU_E0802	50
Apr 2021	AE	CL	<i>Chloris gayana</i>	21_REU_E0803	50
Apr 2021	AE	CL	<i>Digitaria radicata</i>	21_REU_E0804	50
Apr 2021	AE	CL	<i>Eleusine indica</i>	21_REU_E0805	50
Apr 2021	AE	CL	<i>Megathyrsus maximus</i>	21_REU_E0806	50
Apr 2021	AE	CL	<i>Urochloa decumbens</i>	21_REU_E0807	50
Apr 2021	AE	CL	<i>Paspalum paniculatum</i>	21_REU_E0808	50
Apr 2021	AE	CL	<i>Cenchrus purpureus</i>	21_REU_E0809	50
Apr 2021	AE	CL	<i>Sporobolus africanus</i>	21_REU_E0810	50
Apr 2021	AE	CL	<i>Sorghum arundinaceum</i>	21_REU_E0811	50
Apr 2021	AE	CL	<i>Stenotaphrum dimidiatum</i>	21_REU_E0812	50
Apr 2021	AE	CL	<i>Zea mays</i>	21_REU_E0813	50
Nov 2021	CG	TR	<i>Bothriochloa sp.</i>	21_REU_E0205	50
Nov 2021	CG	TR	<i>Chloris barbata</i>	21_REU_E0206	50
Nov 2021	CG	TR	<i>Cynodon dactylon</i>	21_REU_E0207	50
Nov 2021	CG	TR	<i>Dactyloctenium sp.</i>	21_REU_E0208	15
Nov 2021	CG	TR	<i>Dichanthium aristatum</i>	21_REU_E0209	50
Nov 2021	CG	TR	<i>Megathyrsus maximus</i>	21_REU_E0210	50
Nov 2021	CG	PS	<i>Chloris barbata</i>	21_REU_E0710	50
Nov 2021	CG	PS	<i>Cynodon dactylon</i>	21_REU_E0711	50
Nov 2021	CG	PS	<i>Dichanthium aristatum</i>	21_REU_E0712	50
Nov 2021	CG	PS	<i>Heteropogon contortus</i>	21_REU_E0713	50
Nov 2021	CG	PS	<i>Fimbristylis cymosa</i>	21_REU_E0714	50
Nov 2021	CG	PS	<i>Megathyrsus maximus</i>	21_REU_E0715	50
Nov 2021	CG	CH	<i>Cynodon dactylon</i>	21_REU_E0614	50
Nov 2021	CG	CH	<i>Dichanthium aristatum</i>	21_REU_E0615	50
Nov 2021	CG	CH	<i>Enneapogon cenchroides</i>	21_REU_E0616	50
Nov 2021	CG	CH	<i>Heteropogon contortus</i>	21_REU_E0617	50
Nov 2021	CG	CH	<i>Megathyrsus maximus</i>	21_REU_E0618	50
Nov 2021	AE	BP	<i>Bromus catharticus</i>	21_REU_E0132	50
Nov 2021	AE	BP	<i>Cenchrus echinatus</i>	21_REU_E0133	5
Nov 2021	AE	BP	<i>Chloris gayana</i>	21_REU_E0134	50
Nov 2021	AE	BP	<i>Chrysopogon zizanioides</i>	21_REU_E0135	50

CHAPITRE 2

Nov 2021	AE	BP	<i>Cynodon dactylon</i>	21_REU_E0136	50
Nov 2021	AE	BP	<i>Cyperus rotundus</i>	21_REU_E0137	50
Nov 2021	AE	BP	<i>Digitaria ciliaris</i>	21_REU_E0138	50
Nov 2021	AE	BP	<i>Eleusine indica</i>	21_REU_E0139	15
Nov 2021	AE	BP	<i>Eragrostis sp.</i>	21_REU_E0140	50
Nov 2021	AE	BP	<i>Megathyrsus maximus</i>	21_REU_E0141	50
Nov 2021	AE	BP	<i>Melinis repens</i>	21_REU_E0142	50
Nov 2021	AE	BP	<i>Saccharum officinarum</i>	21_REU_E0143	50
Nov 2021	AE	BP	<i>Sorghum arundinaceum</i>	21_REU_E0144	50
Nov 2021	AE	BP	<i>Stenotaphrum dimidiatum</i>	21_REU_E0145	50
Nov 2021	AE	BP	<i>Urochloa ramosa</i>	21_REU_E0146	50
Nov 2021	AE	BP	<i>Zea mays</i>	21_REU_E0147	50
Nov 2021	AE	LP	<i>Bromus catharticus</i>	21_REU_E1001	50
Nov 2021	AE	LP	<i>Cynodon dactylon</i>	21_REU_E1002	50
Nov 2021	AE	LP	<i>Eleusine indica</i>	21_REU_E1003	30
Nov 2021	AE	LP	<i>Eragrostis sp.</i>	21_REU_E1004	45
Nov 2021	AE	LP	<i>Megathyrsus maximus</i>	21_REU_E1005	22
Nov 2021	AE	LP	<i>Melinis repens</i>	21_REU_E1006	32
Nov 2021	AE	LP	<i>Saccharum officinarum</i>	21_REU_E1007	50
Nov 2021	AE	LP	<i>Sorghum arundinaceum</i>	21_REU_E1008	27
Nov 2021	AE	LP	<i>Chloris sp.</i>	21_REU_E1009	50
Nov 2021	AE	LP	<i>Zea mays</i>	21_REU_E1010	50
Nov 2021	AE	CL	<i>Bromus catharticus</i>	21_REU_E0815	50
Nov 2021	AE	CL	<i>Cenchrus purpureus</i>	21_REU_E0816	50
Nov 2021	AE	CL	<i>Chloris gayana</i>	21_REU_E0817	50
Nov 2021	AE	CL	<i>Eragrostis sp.</i>	21_REU_E0818	50
Nov 2021	AE	CL	<i>Urochloa decumbens</i>	21_REU_E0819	50
Nov 2021	AE	CL	<i>Paspalum paniculatum</i>	21_REU_E0820	50
Nov 2021	AE	CL	<i>Sporobolus africanus</i>	21_REU_E0821	50
Nov 2021	AE	CL	<i>Stenotaphrum dimidiatum</i>	21_REU_E0822	50
Nov 2021	AE	CL	<i>Cynodon sp.</i>	21_REU_E0823	50
Nov 2021	SG	LE	<i>Bromus catharticus</i>	21_REU_E0301	50
Nov 2021	SG	LE	<i>Cenchrus clandestinus</i>	21_REU_E0302	50
Nov 2021	SG	LE	<i>Cenchrus purpureus</i>	21_REU_E0303	50
Nov 2021	SG	LE	<i>Cynodon dactylon</i>	21_REU_E0304	50
Nov 2021	SG	LE	<i>Cyperus aromaticus</i>	21_REU_E0305	50
Nov 2021	SG	LE	<i>Holcus lanatus</i>	21_REU_E0306	50
Nov 2021	SG	LE	<i>Setaria sp.</i>	21_REU_E0307	50
Nov 2021	SG	LE	<i>Sporobolus africanus</i>	21_REU_E0308	50

CHAPITRE 2

Nov 2021	SG	MV	<i>Anthoxanthum odoratum</i>	21_REU_E0401	50
Nov 2021	SG	MV	<i>Carex leporina</i>	21_REU_E0402	50
Nov 2021	SG	MV	<i>Cenchrus clandestinus</i>	21_REU_E0403	50
Nov 2021	SG	MV	<i>Dactylis glomerata</i>	21_REU_E0404	50
Nov 2021	SG	MV	<i>Holcus lanatus</i>	21_REU_E0405	50
Nov 2021	SG	MV	<i>Juncus effusus</i>	21_REU_E0406	50
Nov 2021	SG	PC	<i>Anthoxanthum odoratum</i>	21_REU_E0501	50
Nov 2021	SG	PC	<i>Asterochaete nitens</i>	21_REU_E0502	50
Nov 2021	SG	PC	<i>Carex boryana</i>	21_REU_E0503	50
Nov 2021	SG	PC	<i>Carex pilulifera</i>	21_REU_E0504	50
Nov 2021	SG	PC	<i>Festuca borbonica</i>	21_REU_E0505	50
Nov 2021	SG	PC	<i>Holcus lanatus</i>	21_REU_E0506	50
Nov 2021	SG	PC	<i>Juncus effusus</i>	21_REU_E0507	50
Nov 2021	SG	PC	<i>Panicum juniperinum</i>	21_REU_E0508	50
Apr 2022	CG	TR	<i>Chloris barbata</i>	22_REU_E0201	50
Apr 2022	CG	TR	<i>Cynodon dactylon</i>	22_REU_E0202	50
Apr 2022	CG	TR	<i>Dactyloctenium sp.</i>	22_REU_E0203	50
Apr 2022	CG	TR	<i>Heteropogon contortus</i>	22_REU_E0204	50
Apr 2022	CG	TR	<i>Megathyrsus maximus</i>	22_REU_E0205	50
Apr 2022	CG	TR	<i>Stenotaphrum dimidiatum</i>	22_REU_E0206	50
Apr 2022	CG	TR	<i>Urochloa panicoides</i>	22_REU_E0207	50
Apr 2022	CG	PS	<i>Cenchrus clandestinus</i>	22_REU_E0701	50
Apr 2022	CG	PS	<i>Chloris barbata</i>	22_REU_E0702	50
Apr 2022	CG	PS	<i>Cynodon dactylon</i>	22_REU_E0703	50
Apr 2022	CG	PS	<i>Dactyloctenium sp.</i>	22_REU_E0704	50
Apr 2022	CG	PS	<i>Dichanthium aristatum</i>	22_REU_E0705	50
Apr 2022	CG	PS	<i>Heteropogon contortus</i>	22_REU_E0706	50
Apr 2022	CG	PS	<i>Fimbristylis cymosa</i>	22_REU_E0707	50
Apr 2022	CG	PS	<i>Megathyrsus maximus</i>	22_REU_E0708	50
Apr 2022	CG	PS	<i>Setaria pumila</i>	22_REU_E0709	50
Apr 2022	CG	PS	<i>Tragus mongolorum</i>	22_REU_E0710	50
Apr 2022	CG	PS	<i>Urochloa deflexa</i>	22_REU_E0711	50
Apr 2022	CG	CH	<i>Aristida adscensionis</i>	22_REU_E0601	50
Apr 2022	CG	CH	<i>Chloris barbata</i>	22_REU_E0602	50
Apr 2022	CG	CH	<i>Chloris gayana</i>	22_REU_E0603	50
Apr 2022	CG	CH	<i>Cynodon dactylon</i>	22_REU_E0604	50
Apr 2022	CG	CH	<i>Dactyloctenium sp.</i>	22_REU_E0605	50
Apr 2022	CG	CH	<i>Dichanthium aristatum</i>	22_REU_E0606	50
Apr 2022	CG	CH	<i>Digitaria ciliaris</i>	22_REU_E0607	50

CHAPITRE 2

Apr 2022	CG	CH	<i>Echinochloa colona</i>	22_REU_E0608	50
Apr 2022	CG	CH	<i>Eleusine indica</i>	22_REU_E0609	50
Apr 2022	CG	CH	<i>Enneapogon cenchroides</i>	22_REU_E0610	50
Apr 2022	CG	CH	<i>Eragrostis sp.</i>	22_REU_E0611	50
Apr 2022	CG	CH	<i>Heteropogon contortus</i>	22_REU_E0612	50
Apr 2022	CG	CH	<i>Megathyrsus maximus</i>	22_REU_E0613	50
Apr 2022	CG	CH	<i>Melinis repens</i>	22_REU_E0614	50
Apr 2022	CG	CH	<i>Themeda quadrivalvis</i>	22_REU_E0615	50
Apr 2022	CG	CH	<i>Urochloa deflexa</i>	22_REU_E0616	50
Apr 2022	CG	CH	<i>Urochloa ramosa</i>	22_REU_E0617	50
Apr 2022	AE	LM	<i>Cenchrus echinatus</i>	22_REU_E1201	50
Apr 2022	AE	LM	<i>Chloris barbata</i>	22_REU_E1202	50
Apr 2022	AE	LM	<i>Chloris pycnothrix</i>	22_REU_E1203	50
Apr 2022	AE	LM	<i>Cynodon dactylon</i>	22_REU_E1204	50
Apr 2022	AE	LM	<i>Dactyloctenium sp.</i>	22_REU_E1205	50
Apr 2022	AE	LM	<i>Echinochloa colona</i>	22_REU_E1206	50
Apr 2022	AE	LM	<i>Eleusine indica</i>	22_REU_E1207	23
Apr 2022	AE	LM	<i>Eragrostis sp.</i>	22_REU_E1208	50
Apr 2022	AE	LM	<i>Megathyrsus maximus</i>	22_REU_E1209	50
Apr 2022	AE	LM	<i>Melinis repens</i>	22_REU_E1210	50
Apr 2022	AE	LM	<i>Paspalum dilatatum</i>	22_REU_E1211	50
Apr 2022	AE	LM	<i>Paspalum paniculatum</i>	22_REU_E1212	50
Apr 2022	AE	LM	<i>Rottboellia cochinchinensis</i>	22_REU_E1213	50
Apr 2022	AE	LM	<i>Saccharum officinarum</i>	22_REU_E1214	50
Apr 2022	AE	LM	<i>Sorghum arundinaceum</i>	22_REU_E1215	50
Apr 2022	AE	LM	<i>Sporobolus africanus</i>	22_REU_E1216	50
Apr 2022	AE	LM	<i>Stenotaphrum dimidiatum</i>	22_REU_E1217	50
Apr 2022	AE	LM	<i>Urochloa panicoides</i>	22_REU_E1218	6
Apr 2022	AE	BP	<i>Cenchrus echinatus</i>	22_REU_E0101	50
Apr 2022	AE	BP	<i>Chloris barbata</i>	22_REU_E0102	50
Apr 2022	AE	BP	<i>Chloris gayana</i>	22_REU_E0103	50
Apr 2022	AE	BP	<i>Chrysopogon zizanioides</i>	22_REU_E0104	50
Apr 2022	AE	BP	<i>Cynodon dactylon</i>	22_REU_E0105	50
Apr 2022	AE	BP	<i>Cyperus aromaticus</i>	22_REU_E0106	50
Apr 2022	AE	BP	<i>Dactyloctenium sp.</i>	22_REU_E0107	50
Apr 2022	AE	BP	<i>Digitaria ciliaris</i>	22_REU_E0108	50
Apr 2022	AE	BP	<i>Eleusine indica</i>	22_REU_E0109	50
Apr 2022	AE	BP	<i>Eragrostis sp.</i>	22_REU_E0110	50
Apr 2022	AE	BP	<i>Megathyrsus maximus</i>	22_REU_E0111	50

CHAPITRE 2

Apr 2022	AE	BP	<i>Melinis repens</i>	22_REU_E0112	50
Apr 2022	AE	BP	<i>Paspalum dilatatum</i>	22_REU_E0113	50
Apr 2022	AE	BP	<i>Paspalum paniculatum</i>	22_REU_E0114	50
Apr 2022	AE	BP	<i>Pennisetum glaucum</i>	22_REU_E0115	50
Apr 2022	AE	BP	<i>Saccharum officinarum</i>	22_REU_E0116	50
Apr 2022	AE	BP	<i>Sorghum arundinaceum</i>	22_REU_E0117	50
Apr 2022	AE	BP	<i>Sorghum bicolor</i>	22_REU_E0118	50
Apr 2022	AE	BP	<i>Sporobolus africanus</i>	22_REU_E0119	50
Apr 2022	AE	BP	<i>Stenotaphrum dimidiatum</i>	22_REU_E0120	50
Apr 2022	AE	BP	<i>Urochloa panicoides</i>	22_REU_E0121	50
Apr 2022	AE	BP	<i>Zea mays</i>	22_REU_E0122	50
Apr 2022	AE	LP	<i>Cenchrus echinatus</i>	22_REU_E1001	21
Apr 2022	AE	LP	<i>Chloris barbata</i>	22_REU_E1002	50
Apr 2022	AE	LP	<i>Chloris gayana</i>	22_REU_E1003	50
Apr 2022	AE	LP	<i>Chloris pycnothrix</i>	22_REU_E1004	50
Apr 2022	AE	LP	<i>Cynodon dactylon</i>	22_REU_E1005	50
Apr 2022	AE	LP	<i>Digitaria ciliaris</i>	22_REU_E1006	50
Apr 2022	AE	LP	<i>Eleusine indica</i>	22_REU_E1007	50
Apr 2022	AE	LP	<i>Eragrostis sp.</i>	22_REU_E1008	50
Apr 2022	AE	LP	<i>Megathyrsus maximus</i>	22_REU_E1009	50
Apr 2022	AE	LP	<i>Melinis repens</i>	22_REU_E1010	50
Apr 2022	AE	LP	<i>Urochloa decumbens</i>	22_REU_E1011	50
Apr 2022	AE	LP	<i>Saccharum officinarum</i>	22_REU_E1012	50
Apr 2022	AE	LP	<i>Sporobolus africanus</i>	22_REU_E1013	50
Apr 2022	AE	LP	<i>Sorghum arundinaceum</i>	22_REU_E1014	50
Apr 2022	AE	LP	<i>Urochloa deflexa</i>	22_REU_E1015	50
Apr 2022	AE	LP	<i>Urochloa panicoides</i>	22_REU_E1016	50
Apr 2022	AE	LP	<i>Zea mays</i>	22_REU_E1017	50
Apr 2022	AE	CL	<i>Bromus catharticus</i>	22_REU_E0801	50
Apr 2022	AE	CL	<i>Cenchrus purpureus</i>	22_REU_E0802	50
Apr 2022	AE	CL	<i>Chloris gayana</i>	22_REU_E0803	50
Apr 2022	AE	CL	<i>Cyperus aromaticus</i>	22_REU_E0804	50
Apr 2022	AE	CL	<i>Cyperus polystachyos</i>	22_REU_E0805	50
Apr 2022	AE	CL	<i>Eleusine indica</i>	22_REU_E0806	50
Apr 2022	AE	CL	<i>Eragrostis sp.</i>	22_REU_E0807	50
Apr 2022	AE	CL	<i>Megathyrsus maximus</i>	22_REU_E0808	50
Apr 2022	AE	CL	<i>Urochloa decumbens</i>	22_REU_E0809	50
Apr 2022	AE	CL	<i>Paspalum notatum</i>	22_REU_E0810	50
Apr 2022	AE	CL	<i>Cenchrus sp.</i>	22_REU_E0811	50

CHAPITRE 2

Apr 2022	AE	CL	<i>Paspalum virgatum</i>	22_REU_E0812	50
Apr 2022	AE	CL	<i>Sporobolus africanus</i>	22_REU_E0813	50
Apr 2022	AE	CL	<i>Sorghum arundinaceum</i>	22_REU_E0814	50
Apr 2022	AE	CL	<i>Stenotaphrum dimidiatum</i>	22_REU_E0815	50
Total			61 species	273 pools	13341

Supplementary Table 2. Summary of all the plant species.

Host Family	Host Tribe	Host Genus	Host Species	Invasion Status	Crop Status	Life Cycle
<i>Cyperaceae</i>	<i>Abildgaardieae</i>	<i>Fimbristylis</i>	<i>Fimbristylis cymosa</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Cariceae</i>	<i>Carex</i>	<i>Carex boryana</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Cariceae</i>	<i>Carex</i>	<i>Carex leporina</i>	Alien	Weed	Perennial
<i>Cyperaceae</i>	<i>Cariceae</i>	<i>Carex</i>	<i>Carex pilulifera</i>	Alien	Weed	Perennial
<i>Cyperaceae</i>	<i>Cypereae</i>	<i>Cyperus</i>	<i>Cyperus aromaticus</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Cypereae</i>	<i>Cyperus</i>	<i>Cyperus polystachyos</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Cypereae</i>	<i>Cyperus</i>	<i>Cyperus rotundus</i>	Alien	Weed	Perennial
<i>Cyperaceae</i>	<i>Schoeneae</i>	<i>Asterochaete</i>	<i>Asterochaete nitens</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Schoeneae</i>	<i>Costularia</i>	<i>Costularia cadetii</i>	Native	Native	Perennial
<i>Cyperaceae</i>	<i>Schoeneae</i>	<i>Costularia</i>	<i>Costularia melicoides</i>	Native	Native	Perennial
<i>Juncaceae</i>	<i>Juncus</i>	<i>Juncus</i>	<i>Juncus effusus</i>	Native	Native	Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Bothriochloa</i>	<i>Bothriochloa sp.</i>	Native	Weed	Annual
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Chrysopogon</i>	<i>Chrysopogon zizanioides</i>	Alien	Cultivated	Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Dichanthium</i>	<i>Dichanthium aristatum</i>	Alien	Weed	Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Heteropogon</i>	<i>Heteropogon contortus</i>	Native	Native	Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Pennisetum</i>	<i>Pennisetum glaucum</i>	Alien	Cultivated	Annual
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Rottboellia</i>	<i>Rottboellia cochinchinensis</i>	Alien	Weed	Annual
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Saccharum</i>	<i>Saccharum spp.</i>	Alien	Cultivated	Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Sorghum</i>	<i>Sorghum arundinaceum</i>	Alien	Weed	Annual/Perennial
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Sorghum</i>	<i>Sorghum bicolor</i>	Alien	Cultivated	Annual
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Themada</i>	<i>Themada quadrivalvis</i>	Alien	Weed	Annual
<i>Poaceae</i>	<i>Andropogoneae</i>	<i>Zea</i>	<i>Zea mays</i>	Alien	Cultivated	Annual
<i>Poaceae</i>	<i>Aristideae</i>	<i>Aristida</i>	<i>Aristida adscensionis</i>	Alien	Weed	Annual
<i>Poaceae</i>	<i>Bromeae</i>	<i>Bromus</i>	<i>Bromus catharticus</i>	Alien	Weed	Annual
<i>Poaceae</i>	<i>Cynodonteae</i>	<i>Chloris</i>	<i>Chloris barbata</i>	Alien	Weed	Perennial
<i>Poaceae</i>	<i>Cynodonteae</i>	<i>Chloris</i>	<i>Chloris gayana</i>	Alien	Weed	Perennial
<i>Poaceae</i>	<i>Cynodonteae</i>	<i>Chloris</i>	<i>Chloris pycnothrix</i>	Alien	Weed	Annual

CHAPITRE 2

Poaceae	Cynodonteae	Chloris	Chloris sp.	Alien	Weed	NA
Poaceae	Cynodonteae	Cynodon	Cynodon dactylon	Native	Native	Perennial
Poaceae	Cynodonteae	Cynodon	Cynodon sp.	Alien	Weed	Perennial
Poaceae	Cynodonteae	Dactyloctenium	Dactyloctenium sp.	Native	Native	Annual
Poaceae	Cynodonteae	Eleusine	Eleusine indica	Alien	Weed	Annual
Poaceae	Cynodonteae	Tragus	Tragus mongolorum	Alien	Weed	Annual
Poaceae	Eragrostideae	Enneapogon	Enneapogon cenchroides	Alien	Weed	Annual
Poaceae	Eragrostideae	Eragrostis	Eragrostis sp.	Alien	Weed	Annual/Perennial
Poaceae	Paniceae	Cenchrus	Cenchrus clandestinus	Alien	Weed	Perennial
Poaceae	Paniceae	Cenchrus	Cenchrus echinatus	Alien	Weed	Annual
Poaceae	Paniceae	Cenchrus	Cenchrus purpureus	Alien	Cultivated	Perennial
Poaceae	Paniceae	Cenchrus	Cenchrus sp.	Alien	Weed	Perennial
Poaceae	Paniceae	Digitaria	Digitaria ciliaris	Alien	Weed	Annual
Poaceae	Paniceae	Digitaria	Digitaria radicata	Alien	Weed	Annual
Poaceae	Paniceae	Echinochloa	Echinochloa colona	Alien	Weed	Annual
Poaceae	Paniceae	Megathyrsus	Megathyrsus maximus	Alien	Weed	Perennial
Poaceae	Paniceae	Melinis	Melinis repens	Alien	Weed	Annual/Perennial
Poaceae	Paniceae	Panicum	Panicum juniperinum	Native	Native	Perennial
Poaceae	Paniceae	Setaria	Setaria pumila	Alien	Weed	Annual
Poaceae	Paniceae	Setaria	Setaria sp.	Alien	Weed	Annual
Poaceae	Paniceae	Stenotaphrum	Stenotaphrum dimidiatum	Alien	Weed	Perennial
Poaceae	Paniceae	Urochloa	Urochloa decumbens	Alien	Weed	Annual
Poaceae	Paniceae	Urochloa	Urochloa deflexa	Alien	Weed	Annual
Poaceae	Paniceae	Urochloa	Urochloa panicoides	Alien	Weed	Annual
Poaceae	Paniceae	Urochloa	Urochloa ramosa	Alien	Weed	Annual
Poaceae	Paspaleae	Paspalum	Paspalum dilatatum	Alien	Weed	Perennial
Poaceae	Paspaleae	Paspalum	Paspalum notatum	Alien	Weed	Perennial
Poaceae	Paspaleae	Paspalum	Paspalum paniculatum	Alien	Weed	Perennial
Poaceae	Paspaleae	Paspalum	Paspalum virgatum	Alien	Weed	Perennial
Poaceae	Poeae	Anthoxanthum	Anthoxanthum	Alien	Weed	Perennial

			<i>odoratum</i>			
<i>Poaceae</i>	<i>Poeae</i>	<i>Dactylis</i>	<i>Dactylis glomerata</i>	Alien	Weed	Perennial
<i>Poaceae</i>	<i>Poeae</i>	<i>Festuca</i>	<i>Festuca borbonica</i>	Native	Native	Perennial
<i>Poaceae</i>	<i>Poeae</i>	<i>Holcus</i>	<i>Holcus lanatus</i>	Alien	Weed	Perennial
<i>Poaceae</i>	<i>Zoysieae</i>	<i>Sporobolus</i>	<i>Sporobolus africanus</i>	Alien	Weed	Perennial

Supplementary Table 3. List of the other CRESS DNA viruses detected.

Type of Site	Site	Host Specie	Virus
CG	CH	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
CG	CH	<i>Digitaria ciliaris</i>	<i>Genomovirus</i>
CG	PS	<i>Chloris barbata</i>	<i>Genomovirus</i>
CG	PS	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
CG	PS	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
CG	PS	<i>Dichanthium aristatum</i>	<i>Genomovirus</i>
AE	LM	<i>Eragrostis sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Chloris gayana</i>	<i>Genomovirus</i>
AE	BP	<i>Chloris gayana</i>	<i>Genomovirus</i>
AE	BP	<i>Chloris gayana</i>	<i>Genomovirus</i>
AE	BP	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
AE	BP	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
AE	BP	<i>Cynodon dactylon</i>	<i>Genomovirus</i>
AE	BP	<i>Dactyloctenium sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Dactyloctenium sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Dactyloctenium sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Eragrostis sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Eragrostis sp.</i>	<i>Genomovirus</i>
AE	BP	<i>Stenotaphrum dimidiatum</i>	<i>Genomovirus</i>
AE	BP	<i>Sorghum arundinaceum</i>	Sorghum mastrevirus associated alphasatellite
AE	LP	<i>Zea mays</i>	<i>Genomovirus</i>
AE	CL	<i>Eragrostis sp.</i>	<i>Genomovirus</i>
AE	CL	<i>Eragrostis sp.</i>	<i>Genomovirus</i>
AE	CL	<i>Cenchrus purpureus</i>	Putative new <i>Cressdnaviricota</i> species

Supplementary Table 4. Summary of the sequenced *Mastrevirus* species in each plant pool.

Survey	Type of Site	Site	Pool Id	Host Species	Virus Species	Sequencing Procedure	Restriction Site	Accession Number
Jan 2021	AE	BP	21_REU_E0111	<i>Cenchrus echinatus</i>	MSV-F	RCA-MinION	NA	NA
Jan 2021	AE	BP	21_REU_E0111	<i>Cenchrus echinatus</i>	USV	RCA-MinION	NA	NA
Jan 2021	AE	BP	21_REU_E0111	<i>Cenchrus echinatus</i>	USV	RCA-RFLP	<i>KpnI</i>	NA
Jan 2021	AE	BP	21_REU_E0111	<i>Cenchrus echinatus</i>	USV	RCA-RFLP	<i>KpnI</i>	NA
Jan 2021	AE	BP	21_REU_E0111	<i>Cenchrus echinatus</i>	USV	RCA-RFLP	<i>KpnI</i>	NA
Apr 2021	CG	TR	21_REU_E0203	<i>Heteropogon contortus</i>	USV	RCA-MinION	NA	NA
Apr 2021	CG	CH	21_REU_E0612	<i>Urochloa panicoides</i>	USV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0116	<i>Stenotaphrum dimidiatum</i>	USV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0117	<i>Cenchrus echinatus</i>	CEAV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0117	<i>Cenchrus echinatus</i>	MSRV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0117	<i>Cenchrus echinatus</i>	MSV-B	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0117	<i>Cenchrus echinatus</i>	SSRV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0117	<i>Cenchrus echinatus</i>	USV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0119	<i>Chrysopogon zizanioides</i>	MSRV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0120	<i>Cynodon dactylon</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0122	<i>Eleusine indica</i>	EIAV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0124	<i>Megathyrsus maximus</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0127	<i>Saccharum spp.</i>	SWSV	RCA-MinION	NA	NA
Apr 2021	AE	BP	21_REU_E0128	<i>Sorghum arundinaceum</i>	SAAV-B	RCA-MinION	NA	NA
Apr 2021	AE	CL	21_REU_E0807	<i>Urochloa decumbens</i>	UDAV	RCA-MinION	NA	NA
Apr 2021	AE	CL	21_REU_E0807	<i>Urochloa decumbens</i>	UDAV	RCA-RFLP	<i>XbaI</i>	OQ451139

CHAPITRE 2

Apr 2021	AE	CL	21_REU_E0807	<i>Urochloa decumbens</i>	UDAV	RCA-RFLP	Xbal	OQ451140
Apr 2021	AE	CL	21_REU_E0807	<i>Urochloa decumbens</i>	UDAV	RCA-RFLP	Xbal	OQ451141
Apr 2021	AE	CL	21_REU_E0807	<i>Urochloa decumbens</i>	UDAV	RCA-RFLP	Xbal	OQ451142
Apr 2021	AE	CL	21_REU_E0821	<i>Urochloa panicoides</i>	MSV-B	RCA-MinION	NA	NA
Nov 2021	CG	PS	21_REU_E0715	<i>Megathyrsus maximus</i>	PanSV-C	RCA-MinION	NA	NA
Nov 2021	AE	BP	21_REU_E0138	<i>Digitaria ciliaris</i>	MSV-B	RCA-MinION	NA	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	DefMS	RCA-MinION	NA	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	DefMS	RCA-RFLP	Ncol	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	DefMS	RCA-RFLP	Ncol	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	DefMS	RCA-RFLP	Ncol	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	DefMS	RCA-RFLP	Ncol	NA
Nov 2021	AE	BP	21_REU_E0143	<i>Saccharum spp.</i>	SWSV	RCA-MinION	NA	NA
Nov 2021	AE	CL	21_REU_E0816	<i>Cenchrus purpureus</i>	CPMSV	RCA-MinION	NA	NA
Nov 2021	AE	CL	21_REU_E0816	<i>Cenchrus purpureus</i>	CPMSV	RCA-RFLP	Sal	OQ451138
Nov 2021	AE	CL	21_REU_E0819	<i>Urochloa decumbens</i>	UDAV	RCA-MinION	NA	NA
Nov 2021	AE	CL	21_REU_E0821	<i>Sporobolus africanus</i>	MSV-B	RCA-MinION	NA	NA
Nov 2021	AE	LP	21_REU_E1007	<i>Saccharum spp.</i>	SWSV	RCA-MinION	NA	NA
Nov 2021	AE	LP	21_REU_E1008	<i>Sorghum arundinaceum</i>	SAAV-A	RCA-MinION	NA	NA
Apr 2022	CG	CH	22_REU_E0607	<i>Digitaria ciliaris</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	CG	CH	22_REU_E0617	<i>Urochloa ramosa</i>	USV	RCA-MinION	NA	NA
Apr 2022	CG	PS	22_REU_E0708	<i>Megathyrsus maximus</i>	PanSV-C	RCA-MinION	NA	NA
Apr 2022	CG	PS	22_REU_E0711	<i>Urochloa deflexa</i>	PanSV-C	RCA-MinION	NA	NA
Apr 2022	CG	PS	22_REU_E0711	<i>Urochloa deflexa</i>	USV	RCA-MinION	NA	NA

CHAPITRE 2

Apr 2022	AE	BP	22_REU_E0101	<i>Cenchrus echinatus</i>	CEAV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0101	<i>Cenchrus echinatus</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0101	<i>Cenchrus echinatus</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0101	<i>Cenchrus echinatus</i>	SSRV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0101	<i>Cenchrus echinatus</i>	USV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0103	<i>Chloris gayana</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0105	<i>Cynodon dactylon</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0109	<i>Eleusine indica</i>	EIAV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0111	<i>Megathyrsus maximus</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0116	<i>Saccharum spp.</i>	SWSV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0117	<i>Sorghum arundinaceum</i>	SAAV-A	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0118	<i>Sorghum bicolor</i>	SAAV-B	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0121	<i>Urochloa panicoides</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0121	<i>Urochloa panicoides</i>	USV	RCA-MinION	NA	NA
Apr 2022	AE	BP	22_REU_E0122	<i>Zea mays</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	CL	22_REU_E0809	<i>Urochloa decumbens</i>	UDAV	RCA-MinION	NA	NA
Apr 2022	AE	LP	22_REU_E1007	<i>Eleusine indica</i>	EIAV	RCA-MinION	NA	NA
Apr 2022	AE	LP	22_REU_E1014	<i>Sorghum arundinaceum</i>	SAAV-A	RCA-MinION	NA	NA
Apr 2022	AE	LP	22_REU_E1016	<i>Urochloa panicoides</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	AE	LP	22_REU_E1017	<i>Zea mays</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1201	<i>Cenchrus echinatus</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1201	<i>Cenchrus echinatus</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1201	<i>Cenchrus echinatus</i>	PanSV-J	RCA-MinION	NA	NA

CHAPITRE 2

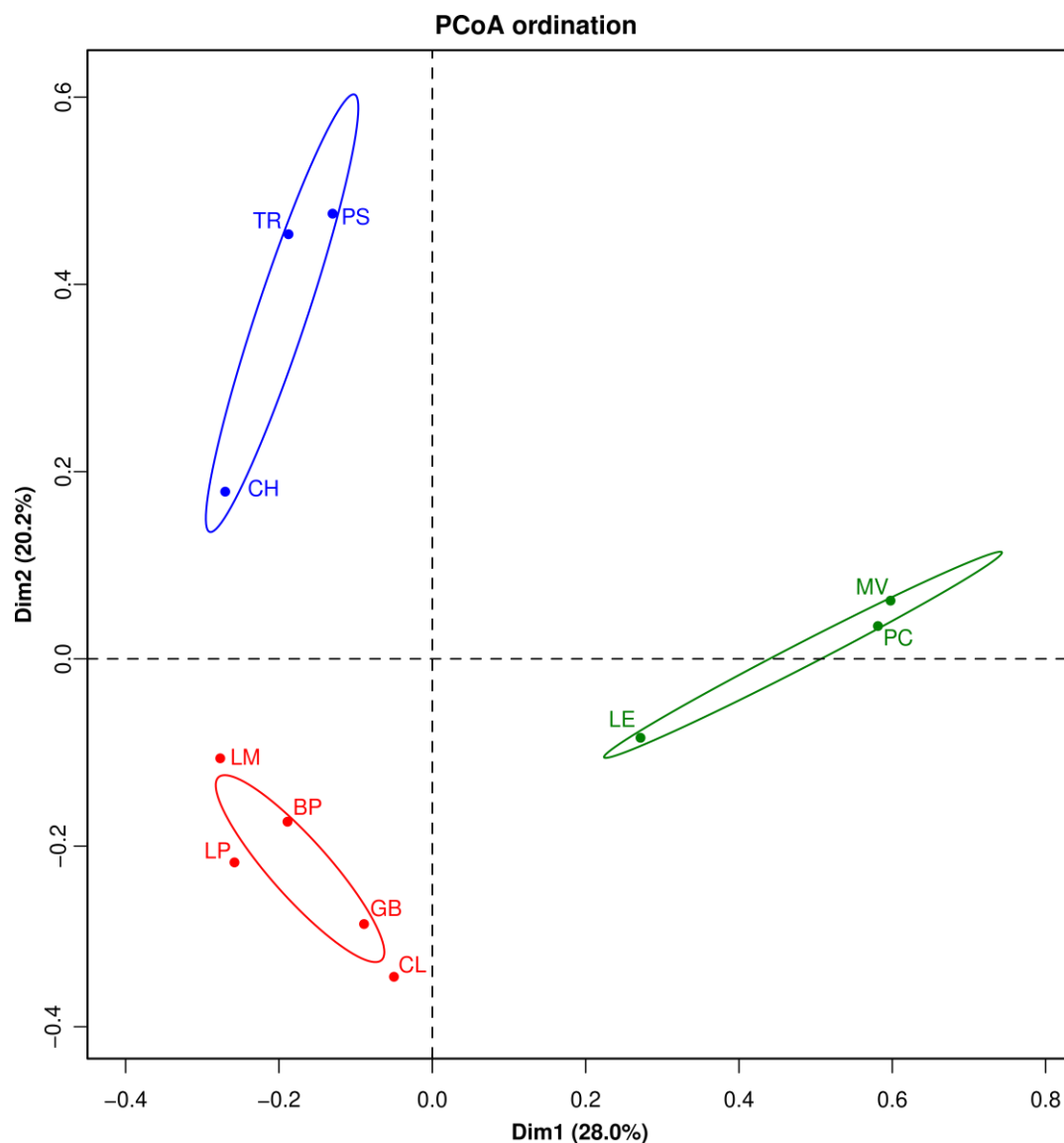
Apr 2022	AE	LM	22_REU_E1202	<i>Chloris barbata</i>	EIAV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1202	<i>Chloris barbata</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1207	<i>Eleusine indica</i>	EIAV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1207	<i>Eleusine indica</i>	MSV-B	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1209	<i>Megathyrsus maximus</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1209	<i>Megathyrsus maximus</i>	SWSV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1210	<i>Melinis repens</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1213	<i>Rottboellia cochinchinensis</i>	MSRV	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1213	<i>Rottboellia cochinchinensis</i>	PanSV-J	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1214	<i>Saccharum spp.</i>	DefMS	RCA-MinION	NA	NA
Apr 2022	AE	LM	22_REU_E1215	<i>Sorghum arundinaceum</i>	SAAV-A	RCA-MinION	NA	NA

Supplementary Table 5. List of the *Mastrevirus* species with the accession numbers of the nucleotide complete, amino acid replication associated protein (Rep) and capsid associated protein (CP) sequences.

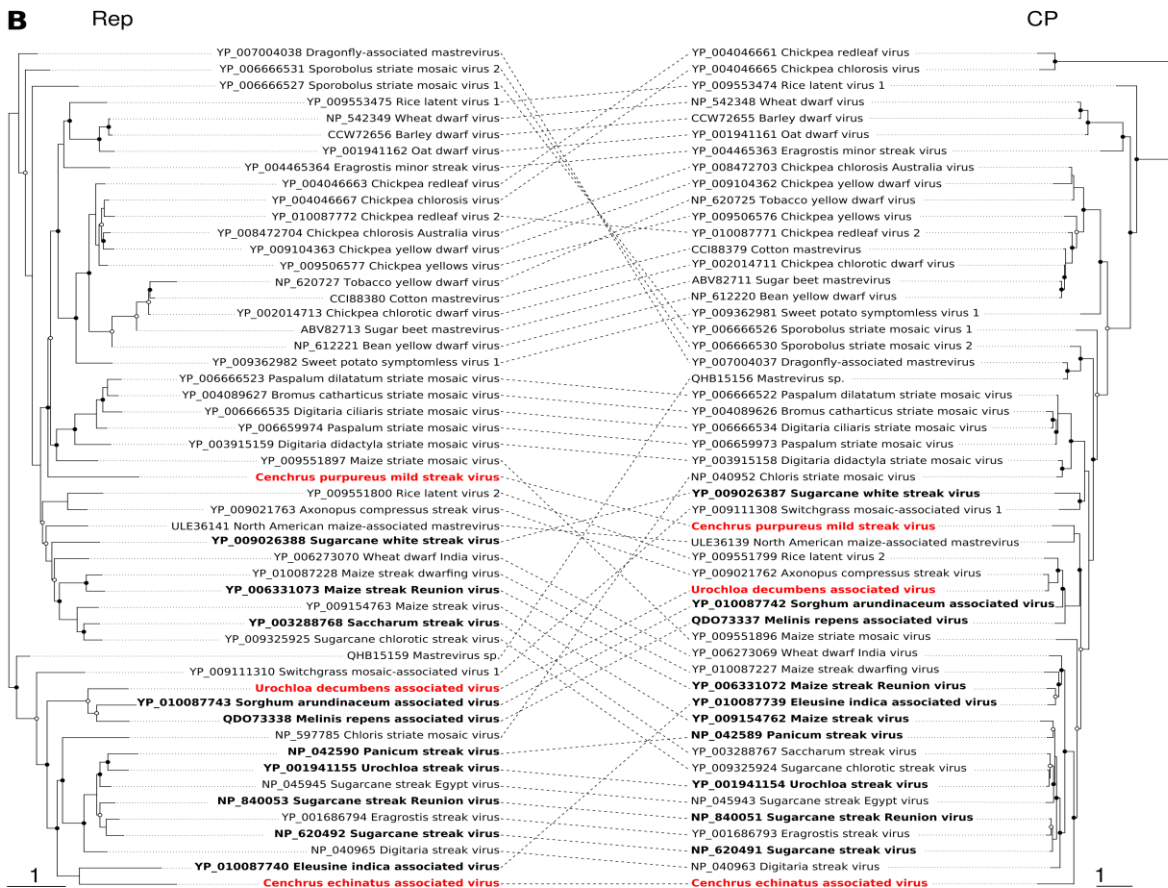
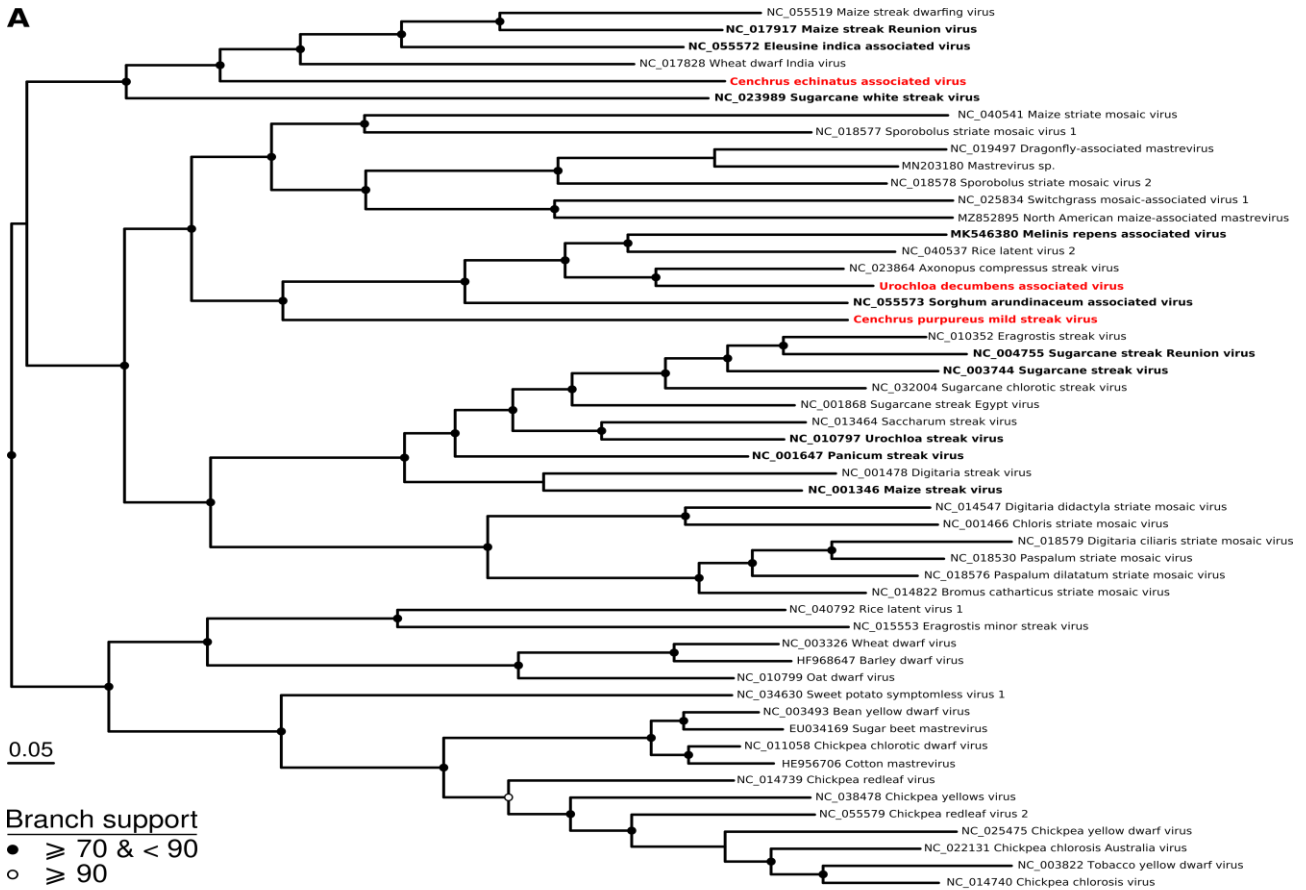
<i>Mastrevirus</i> Species	Complete	Rep	CP
North American maize-associated mastrevirus	MZ852895	ULE36141	ULE36139
Sugar beet mastrevirus	EU034169	ABV82713	ABV82711
Melinis repens associated virus	MK546380	QDO73338	QDO73337
Mastrevirus sp.	MN203180	QHB15159	QHB15156
Barley dwarf virus	HF968647	CCW72656	CCW72655
Cotton mastrevirus	HE956706	CCI88380	CCI88379
Maize streak virus	NC_001346	YP_009154763	YP_009154762
Chloris striate mosaic virus	NC_001466	NP_597785	NP_040952
Digitaria streak virus	NC_001478	NP_040965	NP_040963
Panicum streak virus	NC_001647	NP_042590	NP_042589
Sugarcane streak Egypt virus	NC_001868	NP_045945	NP_045943
Wheat dwarf virus	NC_003326	NP_542349	NP_542348
Bean yellow dwarf virus	NC_003493	NP_612221	NP_612220
Sugarcane streak virus	NC_003744	NP_620492	NP_620491
Tobacco yellow dwarf virus	NC_003822	NP_620727	NP_620725
Sugarcane streak Reunion virus	NC_004755	NP_840053	NP_840051
Eragrostis streak virus	NC_010352	YP_001686794	YP_001686793
Urochloa streak virus	NC_010797	YP_001941155	YP_001941154
Oat dwarf virus	NC_010799	YP_001941162	YP_001941161
Chickpea chlorotic dwarf virus	NC_011058	YP_002014713	YP_002014711
Saccharum streak virus	NC_013464	YP_003288768	YP_003288767
Digitaria didactyla striate mosaic virus	NC_014547	YP_003915159	YP_003915158
Chickpea redleaf virus	NC_014739	YP_004046663	YP_004046661
Chickpea chlorosis virus	NC_014740	YP_004046667	YP_004046665
Bromus catharticus striate mosaic virus	NC_014822	YP_004089627	YP_004089626
Eragrostis minor streak virus	NC_015553	YP_004465364	YP_004465363
Wheat dwarf India virus	NC_017828	YP_006273070	YP_006273069
Maize streak Reunion virus	NC_017917	YP_006331073	YP_006331072
Paspalum striate mosaic virus	NC_018530	YP_006659974	YP_006659973
Paspalum dilatatum striate mosaic virus	NC_018576	YP_006666523	YP_006666522
Sporobolus striate mosaic virus 1	NC_018577	YP_006666527	YP_006666526
Sporobolus striate mosaic virus 2	NC_018578	YP_006666531	YP_006666530
Digitaria ciliaris striate mosaic virus	NC_018579	YP_006666535	YP_006666534
Dragonfly-associated mastrevirus	NC_019497	YP_007004038	YP_007004037
Chickpea chlorosis Australia virus	NC_022131	YP_008472704	YP_008472703
Axonopus compressus streak virus	NC_023864	YP_009021763	YP_009021762
Sugarcane white streak virus	NC_023989	YP_009026388	YP_009026387
Chickpea yellow dwarf virus	NC_025475	YP_009104363	YP_009104362
Switchgrass mosaic-associated virus 1	NC_025834	YP_009111310	YP_009111308
Sugarcane chlorotic streak virus	NC_032004	YP_009325925	YP_009325924
Sweet potato symptomless virus 1	NC_034630	YP_009362982	YP_009362981
Chickpea yellows virus	NC_038478	YP_009506577	YP_009506576
Rice latent virus 2	NC_040537	YP_009551800	YP_009551799
Maize striate mosaic virus	NC_040541	YP_009551897	YP_009551896
Rice latent virus 1	NC_040792	YP_009553475	YP_009553474
Maize streak dwarfing virus	NC_055519	YP_010087228	YP_010087227

CHAPITRE 2

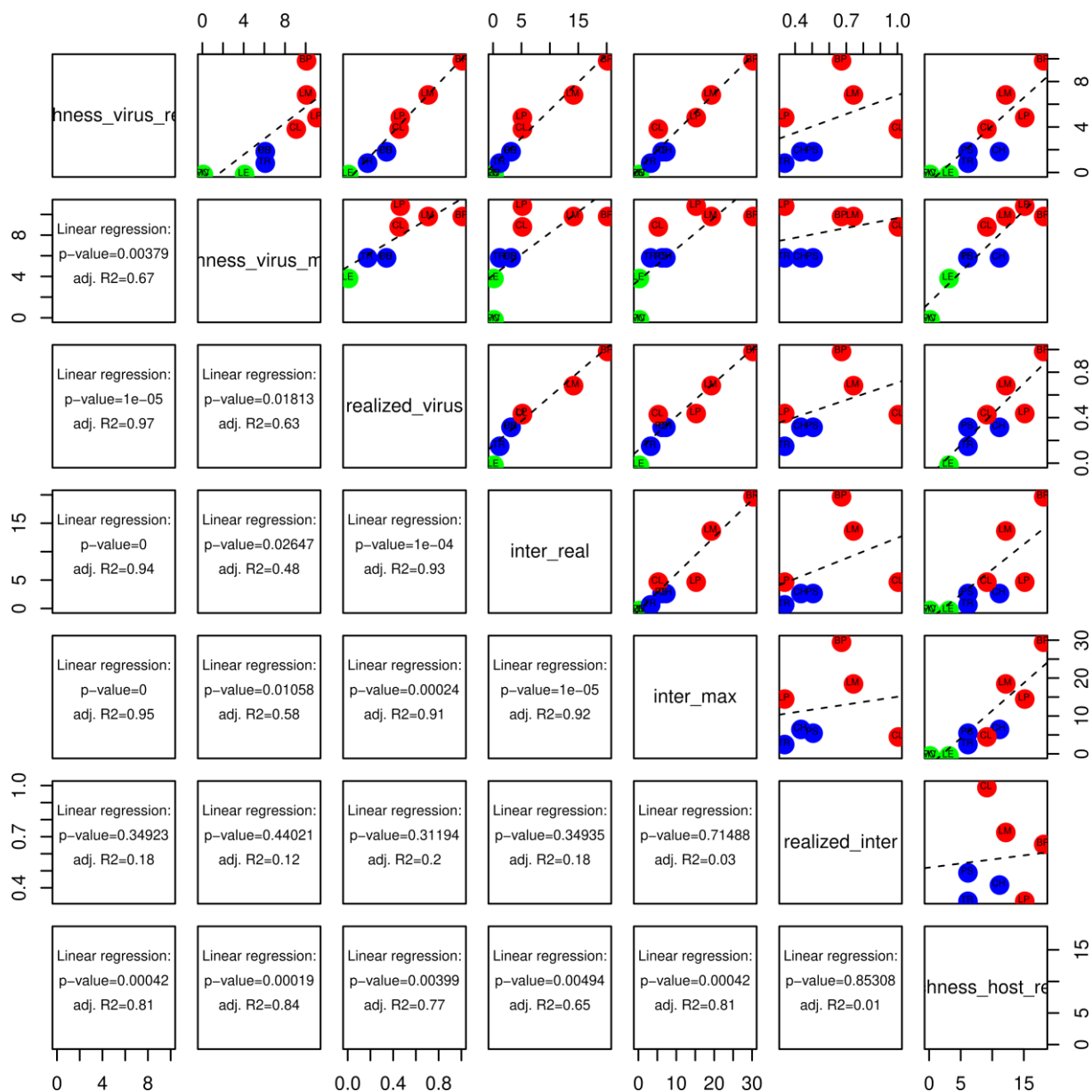
Eleusine indica associated virus	NC_055572	YP_010087740	YP_010087739
Sorghum arundinaceum associated virus	NC_055573	YP_010087743	YP_010087742
Chickpea redleaf virus 2	NC_055579	YP_010087772	YP_010087771



Supplementary Figure 1. Principal coordinate analyses (PCoA) of the diversity of Poales sampled in each site colored by site-type with AE in red, CG in blue and SG in green.



Supplementary Figure 2. (A) Unrooted maximum likelihood tree inferred from aligned complete genome (nucleotidic sequences) of the GenBank reference sequences from various *Mastrevirus* species. (B) On the left, unrooted maximum likelihood tree inferred from aligned replication associated protein (Rep) and on the right, unrooted likelihood tree inferred from aligned capsid protein (CP) amino acid sequences of reference sequences from various *Mastrevirus* species. Species detected in Reunion are in bold font. The three newly identified species, *Cenchrus echinatus associate virus*, *Urochloa decumbens associated virus* and *Cenchrus purpureus mild streak virus* are colored in red. Open and closed circles on nodes indicate bootstrap support for the branches to their left of 70-89% and $\geq 90\%$ respectively.



Supplementary Figure 3. Scatter plots of pairs of variable relative to plant and viruses richness with “richness_virus_real” for the number of viruses species, “richness_virus_max” for the maximum number of viruses species given the sampled host, “realized_virus” as the proportion of virus species observed in comparison to the maximum number of virus, “inter_real” for the number of interaction within the observed virus-host network, “inter_max” for the maximum number of interactions given the observed host and viruses, “realized_inter” as the proportion of interaction observed in comparison to the maximum interaction and “richness_host_real” for the number of host plant species. Sites are presented as colored dots (red for AE sites, blue for CG sites and green for SG sites). Linear regression are indicated with dotted lines. For each comparison, the Pearson correlation coefficient and the p-value for the linear regression slope to be different from zero is indicated.

Chapitre 3

Chapitre 3

Diversité et structure des communautés virales des plantes herbacées indigènes et exotiques de différents écosystèmes réunionnais

L'île volcanique de La Réunion, point chaud de biodiversité du SOOI (Myers *et al.*, 2000 ; Mittermeier *et al.*, 2011), présente un fort taux d'endémisme (Thébaud *et al.*, 2009). Comme de nombreuses îles océaniques, elle est particulièrement touchée par les invasions végétales. Alors que 40 % de la surface de l'île fait partie du parc national, composé essentiellement de zones sauvages protégées avec près de 14 % de plantes vasculaires strictement endémiques, 30 % de la flore réunionnaise (indigène et endémique) est en voie de disparition. La propagation des plantes exotiques envahissantes (PEEs), dans l'ensemble des écosystèmes réunionnais et l'intégralité de leurs gradients altitudinaux (Strasberg *et al.*, 2005), représente sans conteste la plus grande menace pour la flore indigène réunionnaise (Baret *et al.*, 2006 ; Cadet, 1977 ; Lavergne, 1978 ; Pyšek *et al.*, 2012 ; IUCN, 2010 ; Vilà *et al.*, 2011).

Si, précédemment, nous avons montré que la composition des communautés végétales a un impact sur les communautés de phytovirus, on peut supposer que l'introduction de PEEs dans un écosystème va impacter la structure et la composition des communautés de phytovirus. Inversement, il est également attendu que les phytovirus présents dans les milieux envahis vont influencer la dynamique des invasions végétales (Jones, 2020). Enfin, ces introductions de plantes peuvent aussi s'accompagner de nouvelles introductions virales pouvant elles-mêmes impacter la capacité d'invasion des plantes (Faillace *et al.*, 2017).

Dans le chapitre précédent, nous avons démontré l'existence à La Réunion d'une importante communauté de mastrévirus et le lien de celle-ci avec les agro-écosystèmes, des suites probables de l'introduction de plantes cultivées et d'adventices des cultures, de leurs virus et vecteurs (**Chapitre 2 – Article 3**). Nos travaux ont ainsi révélé le rôle

fondamental de l'agriculture dans la distribution de ce genre de phytovirus. Les zones cultivées étant en contact direct avec les limites du parc national (Fenouillas *et al.*, 2021), on peut supposer que les interfaces entre les zones agricoles et naturelles sont des voies d'introduction privilégiées pour les phytovirus.

L'objectif de ce chapitre a été d'explorer la diversité de l'ensemble des phytovirus présents le long d'un gradient altitudinal afin de déterminer si le patron de diversité décrit pour les mastrévirus est généralisable à l'ensemble des communautés virales ou bien dépendant des spécificités des groupes viraux et groupes de plantes étudiés. Par ailleurs, les degrés d'invasions différents des plantes selon les sites étudiés offrent la possibilité d'investiguer comment les introductions de plantes, par le biais de l'agriculture ou non, impactent les diversités virales observées. Pour cela, nous avons étendu les travaux précédents (**Chapitre 2 – Article 3**) à une plus large gamme de familles de plantes en utilisant deux protocoles généralistes de métagénomique virale basés sur des stratégies différentes d'enrichissement viral (*virion-associated nucleic acids*, VANA, François *et al.*, 2018 ; et dsRNA, Marais *et al.*, 2018), (**Chapitre 3 – Article 4**).

Ainsi, nous avons sélectionné quatre sites le long d'un transect altitudinal, présentant des degrés de gestion et de perturbation différents (plus l'altitude du site était grande, moins il était perturbé), dans lesquels nous avons collecté entre 16 et 18 pools des espèces de plantes herbacées les plus répandues dans chacun des sites, en s'efforçant d'échantillonner des espèces exotiques et indigènes. Nous avons ainsi collecté 69 pools de 50 espèces différentes appartenant à 21 familles de plantes. Comme attendu, l'analyse de la distribution des espèces prélevées révèle une plus grande similarité entre les deux sites les plus perturbés (BP et LE) et entre les deux sites les moins perturbés (MV et PC) avec une augmentation de la proportion de plantes indigènes avec l'altitude.

A partir des données de séquençage obtenues par la combinaison de deux protocoles d'enrichissement viral et du séquençage NGS, nous avons assemblé 810 séquences virales dont 134 ont été assignées à 28 genres de 18 familles de phytovirus. Au total, 45% des séquences de phytovirus obtenues semblent appartenir à des espèces virales jusque-là non décrites. La majorité des séquences obtenues a été assignée à la famille des

Genomoviridae avec l'identification de 15 génomes complets voir quasi-complets assignés à des espèces de génomovirus préalablement décrits et neuf génomes appartenant à des espèces non décrites.

Nos investigations ont également montré que les virus présents dans les agro-écosystèmes était majoritairement associés aux adventices des cultures. Globalement, les plantes indigènes et pérennes présentent une plus grande richesse virale que les espèces exotiques et annuelles. La différence de richesse virale observée entre les PEE et les espèces indigènes pourrait s'expliquer par un phénomène de relâchement de la pression parasitaire (*enemy release*), suite à l'introduction des PEE dans un nouvel environnement insulaire sans leur cortège d'agents pathogènes. Cette hypothèse est compatible avec l'introduction de graines et le fait que la majorité des phytovirus ne sont pas transmis verticalement (Blumenthal *et al.*, 2009). Elle est également compatible avec le scénario d'une introduction de plantes infectés et l'absence dans la zone d'introduction de vecteur adaptés à la transmission des virus les infectant.

Nos analyses d'assignations taxonomiques révèlent la présence de phytovirus associés à des maladies des cultures connues comme le cotton leafroll dwarf virus (CLRDV ; famille des *Solemoviridae* ; Edula *et al.*, 2023), le potato virus Y et le turnip mosaic virus (PVY ; Torrance and Talianksy, 2020 ; TMV ; Gibbs *et al.*, 2015 ; famille des *Potyviridae*) ou encore le lettuce chlorosis virus (LCV ; famille des *Closteroviridae* ; Garcian and Janssen, 2020). Comme pour les mastrévirus, ces virus ont été détectés à partir d'échantillons d'espèces végétales non-cultivées exotiques collectées dans les agro-écosystèmes.

Enfin, nos travaux ont également permis la détection du barley yellow dwarf virus (BYDV ; famille des *Tombusviridae* ; Walls *et al.*, 2019) sur flouve odorante, *Anthoxanthum odoratum* (Famille *Poaceae*) dans deux écosystèmes d'altitude, faiblement anthropisés. La flouve odorante est considérée comme une espèce de plante exotique très invasive à La Réunion (Macdonald *et al.*, 1991). Quant au BYDV, il s'agit d'un luteovirus responsable d'une des plus importantes viroses des céréales en milieu tempéré (Banks *et al.*, 1995 ; McKirdy *et al.*, 2002), n'ayant, à notre connaissance, encore jamais été décrite à La Réunion (<https://db.e-prpv.org>). L'apparente absence du BYDV sur les autres poacées présentes

dans les prairies altimontaines suggère l'absence de phénomène de saut d'hôte (*spillover*) du BYDV de la flouve odorante vers d'autres poacées. La possibilité que le BYDV confère à la flouve odorante un avantage sélectif par rapport aux autres espèces de poacées reste à explorer (Power and Mitchell, 2004). En outre, la spécificité d'habitat de la flouve odorante aux prairies altimontaines (Fenouillas *et al.*, 2021) et l'absence de détection de ce virus dans les zones agricoles suggèrent que ce virus a emprunté une autre voie d'introduction à La Réunion que celle décrite pour les autres virus identifiés dans notre étude et généralement décrits comme des virus associés aux agro-écosystèmes.

Ainsi, dans leur globalité, nos résultats concordent avec, d'une part, l'hypothèse d'introductions d'hôtes et de leurs phytovirus à travers les importants échanges de matériel végétal à La Réunion depuis le XVII^{ème} siècle avec la Compagnie française pour le commerce des Indes orientales, et d'autre part, que certaines espèces végétales ont été introduites en échappant à une partie de leurs pathogènes (*enemy release hypothesis*). La possibilité que certains virus aient pu faciliter l'invasion de certaines plantes reste à être investiguée plus finement.

Article 4

**Unraveling the enigmatic influences of phytoviruses in
plant invasions on Reunion Island**

Article 4**Unraveling the enigmatic influences of phytoviruses in plant invasions on Reunion Island**

Sélim Ben Chehida¹, Murielle Hoareau¹, Oumaima Moubset^{2,3}, Charlotte Julian^{2,3}, Laurence Blondin^{2,3}, Emmanuel Fernandez^{2,3}, Denis Filloux^{2,3}, Marie Lefebvre⁴, Hugo Fontes⁵, Christophe Lavergne⁶, Pascal Gentit⁷, Matthieu Rolland⁷, Armelle Marais⁴, Thierry Candresse⁴, Philippe Roumagnac^{2,3}, Gordon W. Harkins⁸, Arvind Varsani^{9,10}, Darren P. Martin¹¹, Jean-Michel Lett¹, Pierre Lefeuvre¹

¹CIRAD, UMR PVBMT, F-97410 St Pierre, La Réunion, France

²CIRAD, UMR PHIM, F-34090 Montpellier, France

³PHIM Plant Health Institute, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

⁴Univ. Bordeaux, INRAE, UMR BFP, 33140 Villenave d'Ornon, France

⁵Tour du Valat, Institut de recherche pour la conservation des zones humides méditerranéennes, Le Sambuc, Arles, France

⁶Conservatoire Botanique National de Mascarin, F-97436, St Leu, La Réunion, France

⁷Anses, Plant Health Laboratory, Bacteriology Virology GMO Unit, 7 rue Jean Dixmèras, 49044, Angers Cedex 01, France

⁸South African Medical Research Council Capacity Development Unit, South African National Bioinformatics Institute, University of the Western Cape, Bellville 7635, South Africa

⁹The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, USA

¹⁰Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Rondebosch, Cape Town, South Africa

¹¹Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory, South Africa

Keywords: ecology, viral metagenomics, phytovirus, exotic plants, invasion, Reunion

Abstract

Plant communities composition is one of the main driver of viral ecology, influencing both the richness and structure of viral communities. However, as perturbations of host communities through agriculture or plant invasion are frequent, it is cardinal to understand how viruses diversities and distributions are impacted with these changes. In this study, we conducted sampling across different herbaceous species in four sites, that displayed varied degree of plant invasions and human management on Reunion Island. Our aim was to explore the diversity and distribution of phytoviruses, and to analyze the impact of viral traits in relation to the invasiveness of their hosts. Out of the 69 plant pools collected from 50 plant species spanning 21 plant families, we detected 28 viral genera belonging to 18 viral families across five Baltimore classes. Whereas few clear associations of plant traits with viral traits were uncovered in our study, interestingly, we found that a greater number of native species were infected by phytoviruses, suggesting a potential 'enemy release' mechanism for exotic plants. Additionally, multiple virus species that are usually associated to crop disease were identified on uncultivated plant species from cultivated areas, implying that the patterns previously observed for the single *Mastrevirus* genus were more general and that human activities were paramount for the introduction of viruses in the agro-ecological contexts. Finally, the discovery in one of the most invasive plant of Reunion of another instance of infection by an usually crop-associated virus raise the question of the potential influence of that virus on the invasion success of its host.

Introduction

Now that it is firmly established that viruses are ubiquitous in all environments, questions have emerged on their contributions and impact on the functioning of the ecosystems. One of the aspect that had been studied is their influence on the dynamics of plant invasions (Pyšek *et al.*, 2020 ; Rua *et al.*, 2011). Indeed, viruses can influence plant invasion in several and contrasted ways (Faillace *et al.*, 2017). Among the possible scheme of interactions of viruses with plant invasion, the pathogen spillover (Power and Mitchell, 2004 ; Becker *et al.*, 2019) and the enemy release (Mitchell and Power, 2003) processes are the more conceptually simple. In the pathogen spillover hypothesis, introduction of plants along with

their infecting phytoviruses may cause the infection and reduction of the native population. In the enemy release hypothesis, the invasion is facilitated by the absence of a pathogen burden in the introduced area, contrasting with the native populations that continue to be infected by their pathogens. As vertical transmission of antagonist viruses is rare (Lovisol *et al.*, 2003 ; Roossinck *et al.*, 2010 ; Paudel and Sanfaçon, 2018), any introduction through seeds would free the plant from these viruses. Additionally, the possible absence in the introduced area of vectors adapted to the transmission of viruses that the introduced plants may harbor, would preclude plant to plant transmission (Power and Flecker, 2008). Examples of each of this effect of invasion have been documented (Agrawal *et al.*, 2005 ; Blumenthal *et al.*, 2009 ; Chun *et al.*, 2010 ; Xu *et al.*, 2023) but no general rules, if it they exist, have emerged on how viruses are most likely to influence an invasion. Understanding such processes would require obtaining comprehensive distributions of viruses within ecosystems over the time span of invasions.

In particular, in islands, native plant diversity is relatively reduced and fragile (Bellard *et al.*, 2016 ; Pyšek *et al.*, 2020), and currently, mostly threaten by the invasion of alien plant species introduced through the human activities. The absence of extensive competition between native plants render these ecosystems extremely prone to invasion of all sorts. Reunion Island is an insular ecosystem located in the southwest of the Indian Ocean, where nearly 14% of the local vascular plants are strictly endemic to the island. However, 30% of the flora (both indigenous and endemic) is currently endangered. In the particular insular context, invasive alien species (IAS) represent the greatest threat to the indigenous flora (Baret *et al.*, 2006; Cadet, 1977; Lavergne, 1978; Pyšek *et al.*, 2012; IUCN, 2010; Vilà *et al.*, 2011). They spread throughout the ecosystems of Reunion, covering the entire range of altitudinal gradients (Strasberg *et al.*, 2005). Studies by Fenouillas *et al.* (2021) have demonstrated that over 85% of the vegetation in the Reunion National Park is invaded by IAS, raising questions about the potential impact of phytoviruses on both indigenous and exotic plant communities in this island.

Recently, in a study that was focusing on the plant infecting mastreviruses (Genus *Mastrevirus*, family *Geminiviridae*), Ben Chehida *et al.* (in preparation) demonstrated that whereas the viruses were extensively found in the context of agro-ecosystems, they were

mostly absent from the more natural areas despite, for some of the sampling sites, the presence of hosts and vectors. Additionally, at odds with the pathogen release hypothesis, it also revealed that the presence of these viruses in the island most likely resulted from the step by step introduction of hosts, viruses and vectors from their native areas. Importantly, as these observations were performed on a specific group of viruses (here a group of insect transmitted ssDNA viruses) and plants (from the unique order of Poales), it was difficult to determine whether this observed patterns were general or contingent of the interplay between peculiar plant and viruses traits.

Here, in order to extend our understanding of how traits of viruses impact their distribution in environment and the interplay between these traits and environmental disturbance, we chose to generalize the previous study to all the phytoviruses that could be detected using up to date methods. Using the combination of two viral sequence enrichment approaches, the virion-associated nucleic acids (VANA ; François *et al.*, 2018) and dsRNA (Marais *et al.*, 2018) and next generation sequencing, we tested the viromes of four different ecosystems, situated all along an altitudinal transect, where sites were displaying decreasing degree of invasion and disturbance with increasing altitude. The plant species selected for the sampling were the most widespread herbaceous species of each site from the most common families between the sites. Doing so, we aimed to complete our knowledge of the viral diversity circulating in Reunion, to provide insights into the structure of viral communities and to investigate the factors driving their richness and distributions in the context of plant invasions. Finally, we explored how the distribution of different viral lineages in different landscapes can help to understand the traits of life histories of these viruses.

Materials and Methods

Selection of the sites

Four different sites representing a gradient of altitude and intensity of agricultural management were investigated. Bassin Plat (BP – 160m altitude), an agroecosystem site with intense management, is a mixed environment containing grassland, wasteland, fallow and agricultural plots (**Figure 1A ; Supplementary Table 1**). Ligne d'Equerre (LE – 1,062m altitude) is a former grazed pasture surrounded by grazed pasture and patch of native forest ;

Espace Naturel Sensible des Hauts de Mont-Vert (MV – 1,586m altitude) is a former grazed pasture enclosed within native forest ; and Piton de Caille (PC – 2,130m altitude) is a former extensive pasture located into the current Reunion National Park and is surrounded by shrublands of native species.

Sampling strategy

Sampling were performed in November 2020, after the dry and cold season of the South West of the Indian Ocean (SWIO). For each site 16 to 18 herbaceous or shrub plants were collected. The most representative plants of each sites were selected. Each plant species were collected regardless of symptoms, into pools leaf samples from 50 individuals per species, when it was possible. Pools were then dried overnight at 55°C and stored at room temperature before use. Plant life-cycle (annual or perennial), the status in Reunion (native or alien) and the invasion score (from 0 for native species to 5 for extremely invasive) were determined using the CBNM botanical database (<https://mascarine.cbnm.org/>).

Dissimilarity analyses were performed on plant species presence/absence binary data to generate a Jaccard distance matrix. Based on this distance matrix, hierarchical clusterings were performed using hclust with the “complete” algorithm and the R “vegan” package (Murtagh *et al.*, 2014). The significance of comparisons among different sites was assessed using the nonparametric statistical test analysis of similarity (ANOSIM) in the R “vegan” package (Oksanen *et al.*, 2015).

From purification of virus-like particles and nucleic acid extraction to library preparation and high-throughput sequencing

Two viral metagenomic protocols were used with the virion-associated nucleic acids (VANA ; François *et al.*, 2018) and the double-stranded RNA (dsRNA ; Marais *et al.*, 2018) protocols. The VANA protocol is based on the amplification of RNA and DNA obtained after the semi-purification of viral particles. Samples were pooled together after a molecular tagging step (Each sample corresponding to the pool of leaves of 50 different individuals of the same plant species). The samples were treated as detailed in François *et al.* (2018) before 2×150bp paired-ends sequencing on the Illumina NextSeq 500 platform. To control for proper extraction of RNA and DNA viruses and to deal with cross-sample contamination, a series of blank (sugarcane extraction from virus-free seedlings) and alien controls were used

in each batch of sequenced samples. The alien controls consisted of sugarcane sample infected with sugarcane streak Egypt virus (SSEV) or sugarcane yellow leaf virus (SCYLV), from the *Mastrevirus* and *Polerovirus* genus respectively. The percentage of reads originating from alien viruses, in comparison to the total viral reads detected per sample, was considered as potential cross-contamination between samples. This percentage was used as a threshold below which the presence of a viral read is considered to be cross-contamination background noise. In the double-stranded RNA (dsRNA) protocol, dsRNA, a nucleic form specifically associated to RNA viruses replication is targeted. The strategy was implemented as described by Marais *et al.* (2018). Equal amounts of amplification products from each pool were sent for Illumina sequencing in a multiplexed format (2×150 bp) on the NovaSeq 6000 system at the GetPlaGe platform (GenoToul INRAE Toulouse, France). As no positive control was included in the process, cross-contamination thresholds determined after VANA sequencing were used.

Sequence analyses and annotation.

The raw reads obtained after VANA strategy were demultiplexed with cutadapt 4.1. (Martin, 2011) and pairs with different barcodes or unidentifiable barcode on both reads were discarded from the analysis. Raw sequencing reads obtained from the dsRNA strategy were imported into CLC Genomics Workbench (CLC-GW), version 21.0.3, demultiplexed and trimmed on quality and length using default settings and a minimum read length of 60 nucleotides (nt). Quality control (sliding window of 30 bases with an average quality of 25 bases required) and adapter identification and removal (remove Illumina adapter sequences (with a maximum mismatch of 2 bases, a palindromic match threshold in the range of 30 bases and simple clip threshold of 10 bases) were then performed on all the reads using Trimmomatic (Bolger *et al.*, 2014). Sequences shorter than 90 nucleotides were discarded. Then, a *de novo* assembly was performed from all the reads using Megahit 1.2.9 (Li *et al.*, 2016) with the “meta-large” presets. After mapping back the reads to the contigs using minimap2 v2.17 (Li, 2018), coverage statistics were obtained using Samtools V1.11 (Danecek *et al.*, 2021). Contigs shorter than 200 nt or assembled from less than 30 reads were discarded. In order to limit intersample cross-talk associated with index-hopping (Illumina, 2017), the number of reads assigned to the same contigs from the same well ID

in different sequencing plates was compared, and only the reads from the well with the highest count were retained for further analyses.

Taxonomic assignment of sequences

Contigs taxonomic assignment was obtained using MMSeqs2 (Steinegger and Soeding, 2017) on the entire nucleotide database of viral reference sequences, and Diamond2 (Buchfink *et al.*, 2021) on the complete amino acid database of viral reference sequences (downloaded in August 2023). Then, contigs presenting with viral hits form the first screen were used in a second step of taxonomic assignment. Using BLASTn and BLASTx, contigs were searched against databases containing sequences from the studied plants or plants related to the studied plants and the complete viral database as obtained from GenBank on August 2023. Only viral contigs assigned to plant pathogenic viral families (phytoviruses) according to ICTV virus metadata resource (Siddell *et al.*, 2023), were retained for further analysis.

Phytovirus inventory and host-virus association

Based on reads mapping statistics, a global plant/virus contingency matrix was obtained and was represented as a heatmap using the R “ComplexHeatmap” package (Gu, 2016). A bipartite network was constructed from the binary plant species-viral genus interactions matrix (indicating presence or absence), using the “bipartite” package in R (Dormann *et al.*, 2008). Differences in viral richness (here approximate as the number of unique phytovirus genus identified in a sample) according to the sample characteristics (plant classification, life history, invasiveness) or virus characteristics (classification and transmission) were assessed using Kruskal-Wallis rank sum test (Hollander and Wolfe, 1973) of the “stats” R package. Tests were performed with a 95% confidence level.

Additionally, a principal coordinates analysis (PCoA) using the *cmdscale* function of the “stats” R package was obtained from the Bray-Curtis distance matrix obtained on sample viral composition data using the *vegdist* function of the “vegan” R package (Oksanen *et al.*, 2019). After selecting the optimal number of clusters using the *fviz_nbclust* function from “factoextra” R package (Kassambara and Mundt, 2020), clustering of samples based on the PCoA coordinates was performed using the *kmeans* function from “stats” R package. Then, RLQ and fourth corner analyses (Dray and Legendre, 2008) were performed using function

of the “ade4” R package (Dray and Dufour, 2007) in order to assess the significance of associations between hosts and virus characteristics. This involved linking matrices of sample variables (R), samples-virus contingency matrix (L), and virus variables (Q). The statistical significance of the co-structure between plant (R) and virus (Q) characteristics was determined by contrasting the total inertia in the RLQ analysis with the total inertia following 9999 Monte Carlo permutations of the rows in the R and Q tables (Doledec *et al.*, 1996).

Phylogenetic inference

In order to gain more knowledge for some viral groups identified in the study, phylogenetic trees were constructed for this group including sequences obtained from GenBank in August 2023. When possible, outgroup sequences were used to root the tree. Sequences were aligned using MAFFT v.7.475 (Kato and Standley, 2013) with the L-INS-i algorithm. After selecting the best fitting model using ModelFinder (Kalyaanamoorthy *et al.*, 2017), a maximum likelihood (ML) phylogenetic tree was inferred using FastTree2 (Price *et al.*, 2010) with an SH-like approximate likelihood ratio test (Guindon *et al.*, 2010) to test branch supports. Trees were edited using Figtree 1.4.4. (<http://tree.bio.ed.ac.uk/software/figtree/>) and the “ape” R package (Paradis and Schliep, 2019).

Results and discussion

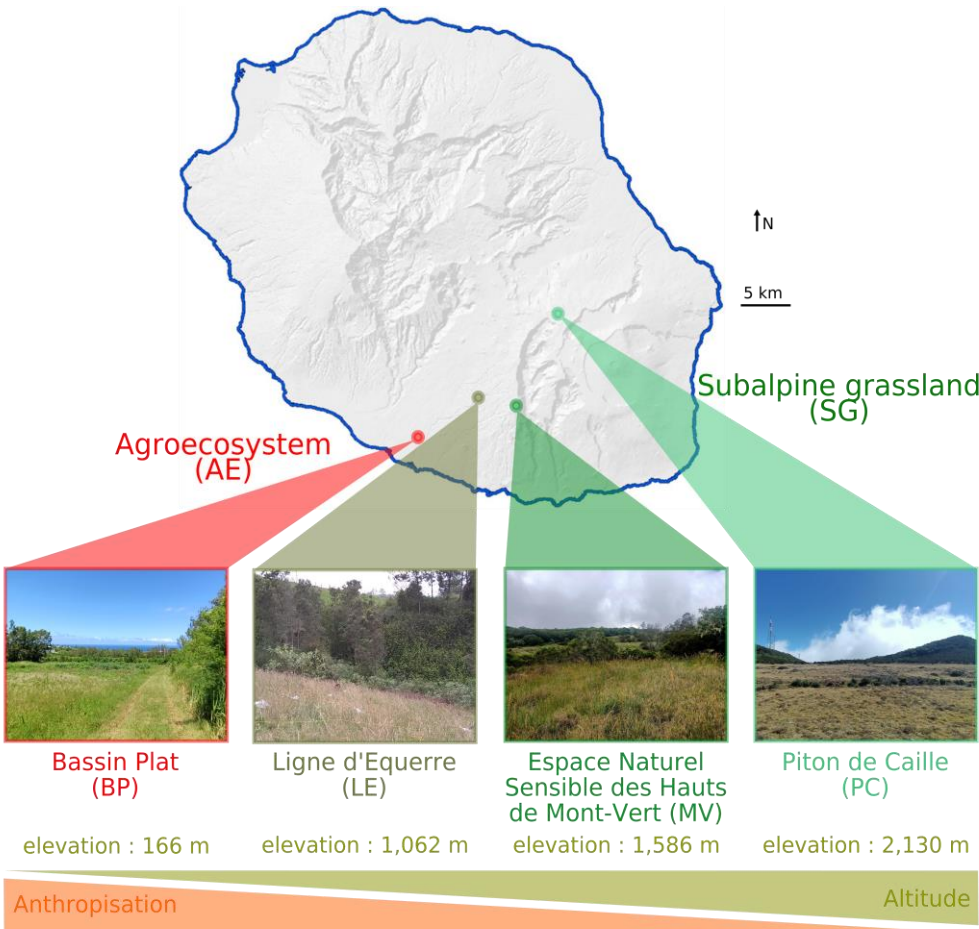
Plant diversity collected along an altitudinal gradient.

A total of 69 pools, collectively representing 3,387 individual plants, were collected from the four sites in November 2020. These pools represented 50 plant species (33 alien and 17 native; 13 annual and 37 perennial) from 21 families (**Figure 1B, Supplementary Table 1 & 2**). It is important to notice here first that all these samples, while representative of each site were not chosen at random. Due to the large diversity of herbaceous and shrubs species, a selection of species was made. Besides sampling the most frequent plants at each sites, special emphasis was put on plants present in multiple sites. Additionally, indigenous plants were specifically selected for the BP site, bearing in mind that they were mostly absent in that site. Indeed, more alien species were collected from BP (n=14), LE (n=14) and MV (n=11) than in PC (n=8). Similarly, more annual species were collected in BP

(n=10) than in LE (n=6), MV (n=1) and PC (n=1). However, it is important to acknowledge that our life cycle categorization relies on data from a plant's native habitat. Alien species may exhibit distinct life cycles in their introduced area compared to their original environment. Consequently, this could introduce a potential source of bias in our analysis.

Among the 21 host families selected in our study, *Asteraceae*, *Cyperaceae*, *Plantaginaceae* and *Poaceae* species were collected in all the sites. *Amaranthaceae*, *Apiaceae* and *Malvaceae* species were collected in BP and LE. *Polygonaceae* in LE and MV, *Juncaceae* and *Rosaceae* in MV and PC and all the eleven other plant families were collected exclusively in a single site. Therefore, ten (20%), eight (16%), seven (14%) and nine (18%) of the plant species were specific to BP, LE, MV and PC respectively (**Figure 1B**). Visually the results of hierarchical clustering analyses based on the distribution of plant species per site suggest the existence of two clusters: a first one grouping BP and LE, which are the most managed ecosystems in low altitude sites, and a second group with MV and PC, which are the least managed and in higher altitude sites. In contrast, the ANOSIM similarity analysis suggested a very significant grouping of the plant species collected for each site ($R = 0.96$, $P = 0.001$).

A.



B.

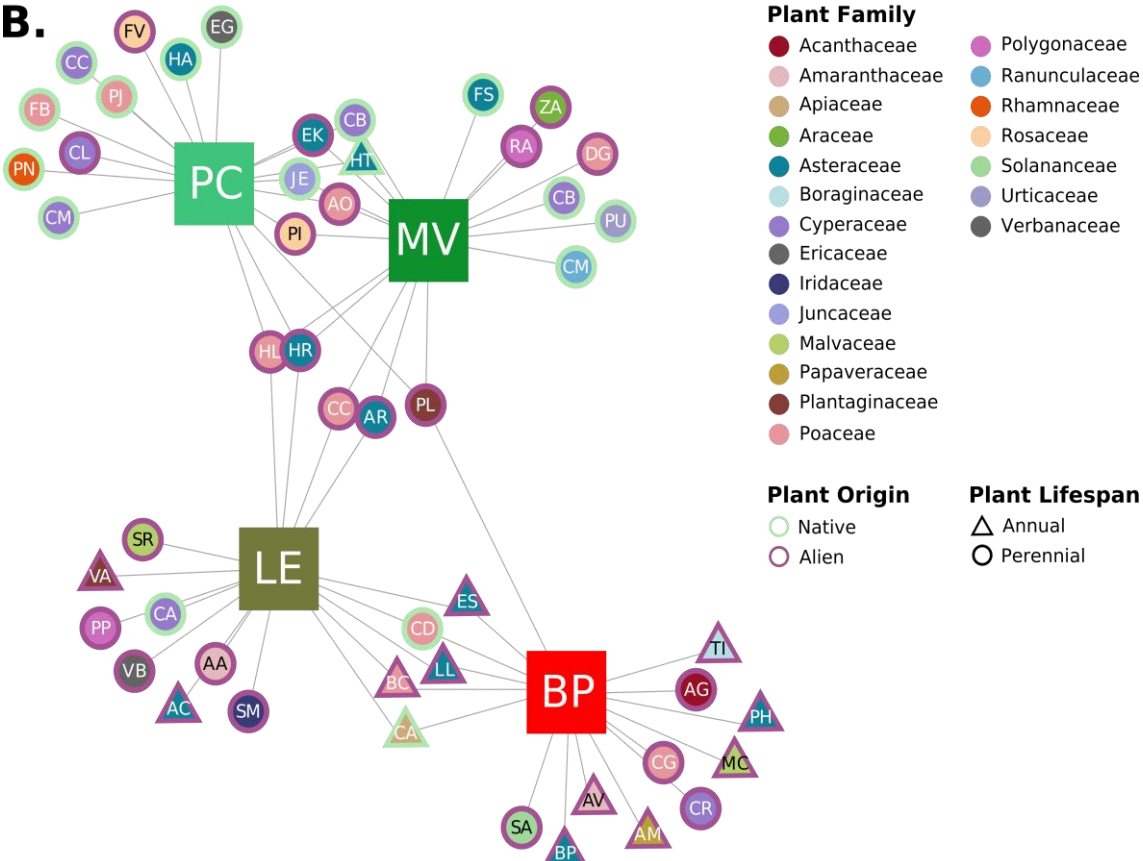


Figure 1. Summary of the herbaceous plant species sampling in Reunion. **(A)** Map of Reunion indicating the four sampling sites. Bassin Plat (BP), Ligne d'Equerre (LE), Espace Naturel Sensible des Hauts de Mont-Vert (MV) and Piton de Caille (PC). Different abiotic conditions and levels of anthropisation characterized the sampling sites. **(B)** Undirected network connecting plant species to each sampling site. Each square indicates the site and circle or triangle a plant species. Color of the circle indicates the plant family following the color scale on the right side of the figure. Native and alien plant nodes are outlined in light green and purple, respectively. Plant node shape indicates plant lifespan with triangle for annuals and circles for perennials.

Phytovirus diversity in herbaceous species from Reunion

From the raw reads obtained using the VANA (~ 71M PE reads) and dsRNA procedures (~ 62 M PE reads), and after quality control and demultiplexing, ~55 M PE reads were used for assembly (**Supplementary Table 3**). A total of 810 viral contigs, accounting for ~ 12 M reads (**Supplementary Table 3**) was obtained. After similarity search against the public viral database, 644 contigs had significant hit to viruses (**Figure 2A**). Of these 644 contigs, 134 (21%) were most similar to plant infecting viruses (**Figure 2B**). These contigs accounted for 3,595,195 reads (making 31% of the reads used for assembly).

dsRNA and VANA procedures have been characterized as predominantly retrieving similar viral families from plant samples, with only small number of viral families not being captured by one method or the other (Ma *et al.*, 2019). We thus chose to combine both procedures to obtain a more comprehensive view of viral diversity. Doing so, phytoviruses from five of the seven Baltimore classes were retrieved and their proportions (94.4% ssRNA(+), < 0.1% ssRNA(-), 0.8% dsRNA, 4.7% ssDNA and 0.1% dsDNA-RT) were in the range of previous observations (Koonin *et al.*, 2021). However, no RNA-RT phytoviruses were detected here whereas this group represents 3% of the currently described phytoviruses (Koonin *et al.*, 2021). The length of the contigs ranged from 304 to 15,052 nt with a mean of 1,209 nt. According to our initial classification, they belong to 28 genera of 18 known phytovirus families (**Supplementary Table 4**). Of these contigs, 102 contigs were related to ssRNA(+) from twelve families (3,392,417 reads) ; five contigs to ssRNA(-) from a single family (*Rhabdoviridae* family; 2,166 reads), three contigs to two dsRNA families (*Spinareoviridae* and *Amalgaviridae* ; 27,384 reads), 22 contigs to two ssDNA families (*Genomoviridae* and *Geminiviridae* ; 168,758 reads) and two contigs to a single dsDNA-RT family (*Caulimoviridae* family; 4,470 reads).

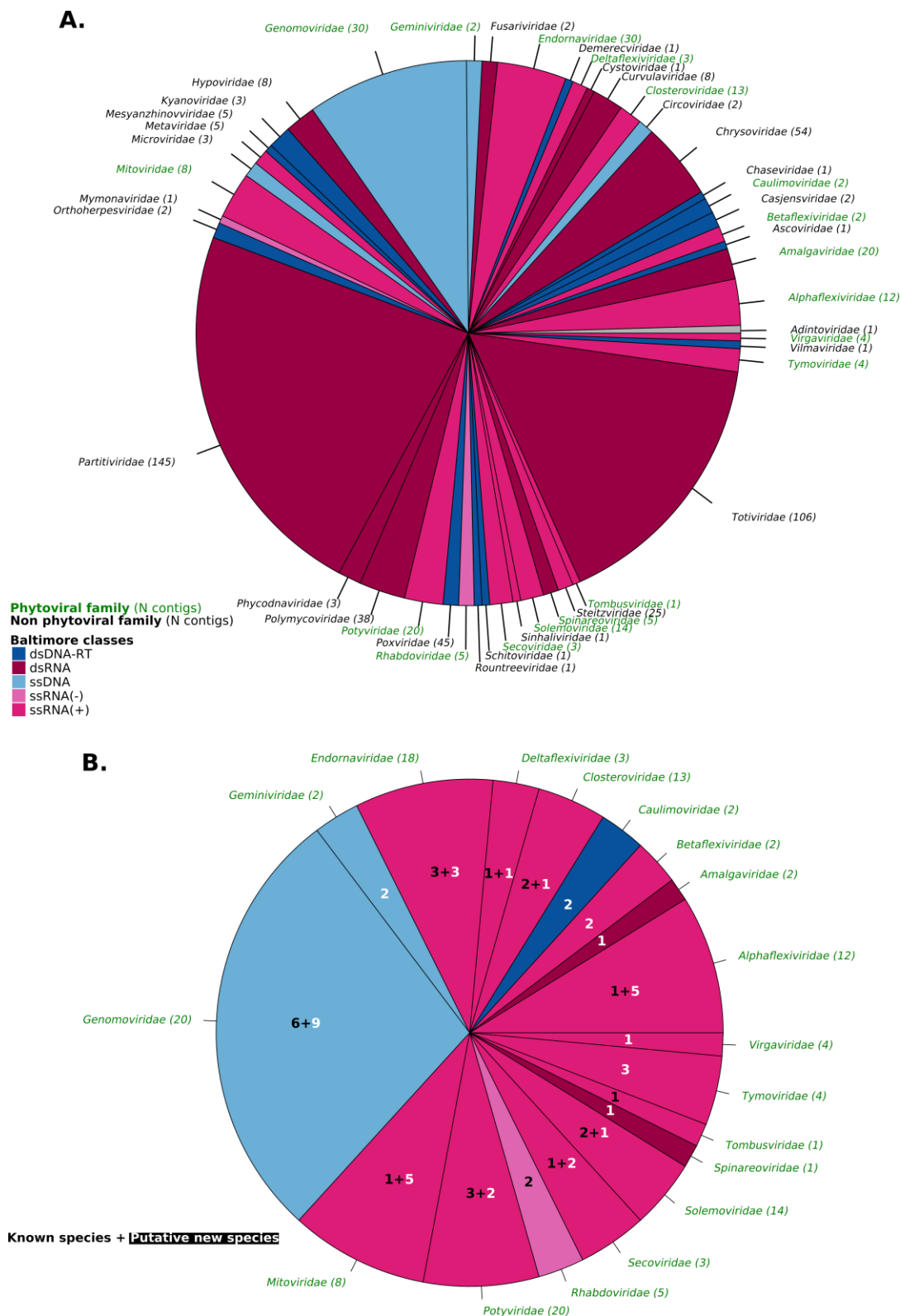


Figure 2. (A) Pie chart of the richness of viral families detected in term of OTUs. In green are indicated the viral families described as phytoviruses. **(B)** Pie chart of the richness of phytoviral families detected in term of OTUs. Number in white is the count of putative new species. The number of contigs are indicated between brackets.

Complete or near complete viral genomes of 15 phytoviruses were identified within the contigs. Four contigs were assigned to *Gemycircularvirus* species (contigs k_599761, k_701128, k_607041 and k_68885; presenting from 78.8% to 95.6% nucleotidic similarity with species previously described ; see **Supplementary Table 4** for details). And, one contig was assigned to an *Alphaendornavirus* species (k_504949; 98.7% nucleotidic similarity to Carrot alphaendornavirus 1 ; Schönegger, 2023). According to the International Committee on Taxonomy of Viruses [ICTV] species demarcation, the eleven other full genomic sequences were putative new species belonging to seven different viral families (*Geminiviridae*, *Genomoviridae*, *Mitoviridae*, *Potyviridae*, *Secoviridae*, *Solemoviridae* and *Virgaviridae*). First, the contig k_699320, initially assigned to *Geminiviridae* family, seems to be closer to *Genomoviridae* family. It presented with two putative non-overlapping open reading frames (ORFs). The first ORF encoded a putative capsid protein in the viral sense (CP, 354 aa). The second ORF encoded a putative replication-associated protein in the complementary sense (Rep, 335 aa). Predicted protein products presented with 67% and 38% amino-acid sequence identities with respectively the CP and Rep of red panda feces-associated genomovirus (UBJ26205). The genomic sequence also contained a typical geminivirus nonanucleotide motif of the replication origin (TAATATT↓AC; Roumagnac *et al.*, 2022) within the long intergenic region (IR). Hence, given the genomic architecture and the amino-acid identity of that genome to previously described genomoviruses, it may represent a member of a previously undescribed genus within the *Genomoviridae* family. Importantly, the *Cressdnaviricota* phylum (*i.e.* the phylum that comprises the *genomoviridae*), gathers a large number of divergent single stranded DNA viruses, with diverse genomic architectures (Rosario *et al.*, 2012 ; Roux *et al.*, 2013). Discovery of distantly related genomoviruses is not rare and had largely increased in frequency with the use of phi29 based metagenomic protocols (*i.e.* circomic ; Chabi-Jesus *et al.*, 2020 ; Hao *et al.*, 2021 ; Feng *et al.*, 2021). Then, two putative new species of *Genomoviridae* (contigs k_376567 and k_167758), presenting with the usual *Genomoviridae* genome architecture, a nonanucleotide at the replication origin and two ORFs (a *cp* gene in the viral sense and a *rep* gene the complementary sense), presented with pairwise identity of the complete genome and of the ORFs under the demarcation criteria of the ICTV (< 78%) similar to a *Gemycircularvirus* species (NC_038493) and a *Gemykibivirus* species (NC_024691), respectively. Phylogenetic reconstruction using the *Genomoviridae* reference sequences and the full

genomic sequences uncovered in our study (**Supplementary Figure 1**) clearly supported the placement of the sequences within the *Gemycircularvirus* genera but also indicated that these were not closely related to any of the currently described genomoviruses. It is noteworthy that during a previous survey on *Poaceae*, Ben Chehida *et al.* (in preparation) identified several genomoviruses from plants collected in Reunion (indicated as underlined tip on **Figure 3**). Thus, it's not surprising that this family was again identified. However, none of the sequence identified here were closely similar to those from the previous study.

The mitovirus full genome sequence (k_175451) presented with an RdRp associated protein with amino-acid sequence identity of 51.6% to *Dahlia pinnata* mitovirus 1 RdRP associated protein (NC_076528.1; *Duamitovirus* genera). The secovirus (k_2822470) had a putative polyprotein of the RNA 2 of a *Nepovirus* genera with 59.7% amino-acid sequence similarity to that of beet ringspot virus (NC_003694). The solemovirus genomic sequence had a polyprotein presenting with 63% amino-acid sequence similarity to that of an isolate of the subterranean clover mottle virus (NC_004346; *Sobemovirus* genera). The three *Virgaviridae* sequences (k_121208, k_678, k_688358) were representative of the the three RNA genomic component that compose the *Virgaviridae* genomes and most probably formed the genome of an unique virus. Indeed, the three contigs were similar at 84.3%, 84.3% and 90.8% each to the RNA 1 (alpha, NC_003469) gene, RNA 2 (beta, NC_003481) gene and RNA-3 (gamma, NC_003478) gene of an isolate of the barley stripe mosaic virus, respectively. Finally, the probable potyvirus complete genome (k_502700), presented a polyprotein with an amino acid identity of 36.9% with the polyprotein of the blackberry virus Y (NC_008558) from the *Brambyvirus* genus (Susaimuthu *et al.*, 2008). The polyprotein of 2,615 amino acid in length presented with six potyvirid proteins (P1, HC-Pro, P3, NI-Pro, RdRp, CP). As such and following the ICTV classification rules for the *Potyviridae* family (Inoue-Nagata *et al.*, 2022), it may represent a member of a previously undescribed *Potyviridae* genera.

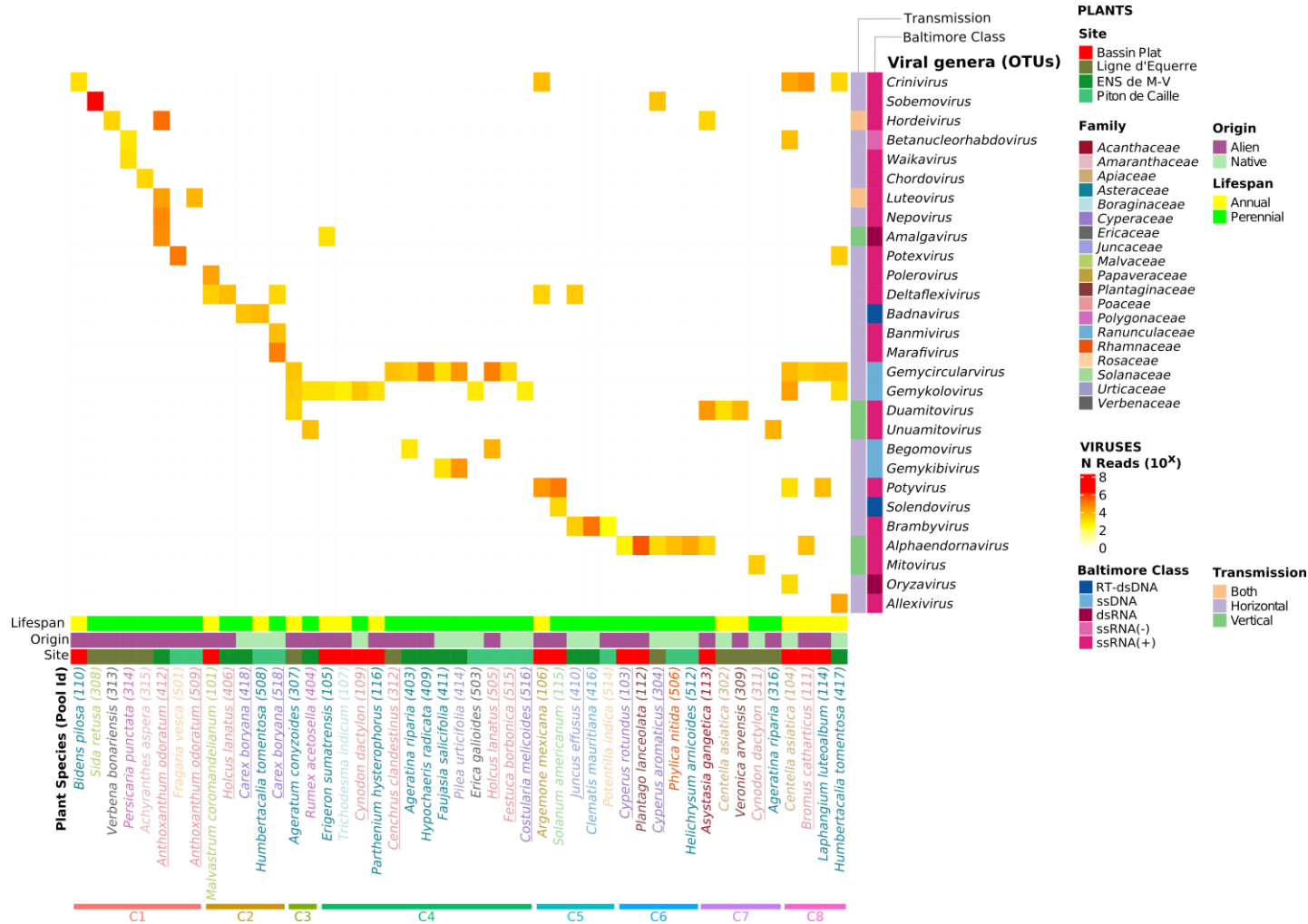


Figure 3. Heatmap of the number of reads of each viral genus (in row) detected in each plant sample (in column). Number of reads are represented by a base 10 logarithmic color scale (yellow to red, white for absence). The plant species are organized by k-means calculated clusters and colored by families. The pool ID of each sample is indicated between brackets. Sampling sites (BP, LE, MV and PC), origin (native; alien) and lifespan (annual; perennial) are indicated in the first, second and third line below the matrix. Transmission and Baltimore classes are indicated in the first and second columns respectively on the right side of the matrix for each viral genus.

The 119 other contigs were probably partial genome sequences of which 50 might represent sequences from novel species, according to the different ICTV classification rules. With the exception of the *Rhabdoviridae* and *Tombusviridae* families, new viral species were detected in all the families identified in this study (**Figure 2**).

Globally, among all the phytoviruses identified here, 45% were potential members of new species (representing 83% of the reads). It must be noticed however that this proportion is very conservative and that some viral sequences of our dataset might remain unidentified. Indeed, the use of similarity search for virus identification makes it difficult to identify viruses that are only remotely similar to the viruses that compose our databases (Roux *et al.*, 2021).

Richness patterns among the sites and plant samples

After mapping back the reads from individual samples to the contigs, it was possible to define the plant – virus infection matrix. Filters based on sequencing controls were then used in order to remove spurious assignments that were most probably associated to contamination (Mitra *et al.*, 2015; Massart *et al.*, 2017). Basically, a first filter for probable index-hopping contamination between sequencing libraries was applied on VANA read counts, removing a total of 4% of the reads. As a single library was sequenced for dsRNA, the filter wasn't used on the reads obtained from that procedure. A second filter was based on the number of reads of viruses from alien controls found in other samples. Analyses of the presence of viruses from these alien controls (SSEV and SCYLV) indicated that, per sample, up to 1.55% of the total reads could be associated to cross-sample contamination. No viral detection below this level was thus considered in the VANA and the dsRNA results. It must be noticed that the use of a global threshold is a very simplistic approach to filter cross-sample contamination. Contamination patterns most probably differ from sample types and virus taxa but the absence of high resolution on cross contamination preclude the use of more sophisticated model based methods for filtering.

Of all the 69 pools, 68% (47/69) were identified as positive for identifiable phytovirus infection (**Figure 3**). These positive pools belong to 80% (40/50) of the plant species, 87% (39/45) of the plant genus and 90% (19/21) of the plant families. The samples from the *Iridaceae* (n=1) and *Araceae* (n=1) families remained negative to any of the identified viruses. It is noteworthy that, with the exception of the samples from the *Amaranthaceae*, *Malvaceae*, *Juncaceae*, and *Plantaginaceae* families, viruses were detected in at least one sample from each families at all their collection sites (**Figure 3**). Among the four plant

families collected all along the altitudinal transect, 12/18, 5/8, 2/4 and 9/15 of the samples form, respectively, the *Asteraceae*, *Cyperaceae*, *Plantaginaceae* and *Poaceae* families were positive. Moreover, regardless of the type of plant tested, more than 60% of each were tested positive, with 62% and 82% for the alien and native species respectively, and 78% and 65% for the annual and perennial species respectively. Similarly, according to the altitudinal transect, in each site, more than 60% of the collected species were positive, with 88% in BP, 65% in LE, 65% in MV, and PC. Overall, it seems that the infection rate remains relatively consistent for each type of sampled plants.

Whereas samples from BP and MV present the highest number of viral families with twelve different families, in samples from PC we detected ten families and eight in samples from LE. Among the 18 viral families, only *Spinareoviridae* and *Tymoviridae* were found in a single site whereas nine viral families were detected in two sites, six in three sites and *Genomoviridae* species were detected in all sites. Similarly, at the genus level, 13, 11, 15 and 11 genera were detected from BP, LE, MV and PC respectively, with 57% of the genera being shared between sites. The percentage of infected plants was ranging from 61% (for MV and PC) to 87% (BP). Richness comparative analysis were then performed, considering virus assignation at the genus levels (*i.e.* in a given sample, the richness score is the number of distinct plant virus genera identified in that sample). The mean viral genera richness per sample was 1.72 with means of 2.1, 1.4, 2.1 and 1.3 for BP, LE, MV and PC respectively. Globally, 18 and 21 phytovirus genera were found in native and alien plants respectively. Mean viral richness per samples were mostly similar with a per sample mean of 1.83 genus for native plants and 1.65 genus for exotic plants. When considering the sampling sites, there was a trend of higher infection rates for native plants than alien, mean viral richness per plant were mostly similar for most sites between native and alien plants. Solely in BP, the mean viral richness of native plants was superior ($R=2.4$) to that of alien plants ($R=1.8$). Similarly, mean viral richness was higher for annual plants ($R=2.6$) than perennial plants ($R=1.8$). This tendency was conserved from site to site except for PC where richness were equivalent between both plant types ($R=1.8$ for annuals and $R=1.3$ for perennials). It is interesting to notice that this results is at odds with the global higher richness of perennials ($R=24$) than annuals ($R=14$). Plant viruses are more shared between individuals for annuals (a genus found in annuals is found in a mean of 6.2 other plants) than they are for perennials (a genus found in perennials is found in a mean of 4.9 other plants). It means that whereas

diversity of viruses is less important for annuals, annuals experiments more co-infection of a reduced set of viruses. Note that similar results were obtained at the species levels.

The viral cortege relatedness

In order to obtain a better understanding of the factors that structure viruses distribution among plants, a PCoA analysis of the viral cortege composition (at the genus level) along with k-means clustering were performed. Eight clusters of samples that comprises from two to 13 samples were identified (**Figure 4**). Most of the traits characteristic of plants and viruses were scattered among the cluster and few association were visually identifiable. It indicates that no traits had simple effect on the viral distribution between the samples. Either the structure of the community was dependent of multiple interactions between the traits or that we only obtained a glimpse of the existing viral diversity and that in the absence of the whole virome characterization, factors at play in the structuring would remain hidden. Nevertheless, it was noteworthy that cluster 1 was only composed of eight alien plant species from all the sites. Cluster 4 gather the highest number (and proportion, n=6, 46%) of native plants among every samples. These plants were clearly more infected with ssDNA viruses. Clusters 6 and 7 were mostly composed of plants infected with (+)ssRNA viruses.

In order to better understand how the characteristics of plants could structure viral diversity, we analyzed pairs of samples from the same plant species that were collected in different sites. Whereas sites should have an influence on the dynamics of viral infection (biotic and abiotic conditions could be markedly different from site to site), it is likely that plants from the same species would more likely shared more viruses from one site to the other than plants from different species. Of the 16 plant species that were gathered in multiple sites, one was negative for all the samples, eight were positive in a single sample and seven were positive for two samples. For these seven plants, two had their replicates gathered in the same cluster while the other five pairs were retrieved in distinct clusters. Globally our result indicates that whereas no factors was apparently paramount in the grouping of viruses, host relatedness, approximated here with plant of the same species was not either explaining the global clustering of the samples. Interestingly, some of the pools that were collected here were also analyzed for another study (Ben Chehida *et al.*, in preparation) that employs a NGS protocols dedicated to the detection of viruses from the *Cressdnaviricota* phylum. Of the 25 common plant pools between the two studies, 19 were negative to *Cressdnaviricota* in both study, one was positive to genomoviruses here, while two plant species were positive

to genomoviruses and geminiviruses in the Ben Chehida study. These differences in detection between the two studies may indicate that repeatability should be improved. It must be noticed however that in both studies, pools of plants were analyzed and as such, sample infections heterogeneity may explain the difference in viral detection. Nevertheless, additional experiments dealing with virus detection repeatability would help assess whether the absence of strong association between biotic and abiotic factors with viral diversity is genuine or associate to our inability to obtain a comprehensive description of the viral diversity.

Significant associations between plant traits and viral infections

Owing for the absence of obvious association between group of viruses and plants, we used a statistical method called RLQ analysis that is dedicated to the search of significant association between plant and virus traits given the pattern of association (*i.e.* the contingency matrix). The RLQ analysis produced a two-axis graph where each plant and viral variables were disposed according to their influence on each other (**Supplementary Figure 2**). Using the fourth corner method, it was then possible to identify significant correlations (significance level of 0.05) between the variables (**Figure 5**). The analysis reveals several positive association between traits. First, the viruses from the *Genomoviridae* family were involved in several of the significant associations. Indeed, they were positively correlated to alien and invasive plants and annual and perennial plants. Genomoviruses were the most prevalent and frequently found in multiple samples from all the sites and most of the plant species (eight of the 19 plant families). In particular, genomoviruses were significantly associated with two sites (BP and MV) and four plant families, including the *Poaceae*. A similar association was found for the ssDNA viruses in general. As genomoviruses accounted for most of the ssDNA viruses identified in our study, it is not surprising that similar patterns of association are found for both traits. The fact that the main virus family identified here (20% of the viral contigs) was the *Genomoviridae* raised the question of whether these are genuine phytoviruses (in the sense of plant infecting viruses) or mostly viruses infecting fungi (Krupovic *et al.*, 2016 ; Liu *et al.*, 2016 ; Varsani and Krupovic, 2021) associated to our plant samples. The presence of these viruses in association with most plant species could indicate an extremely generalist viral genus or, more probably, viruses infecting unidentified hosts associated with our plant samples.

Interestingly, horizontal transfer was associated to annuals plants. Previous studies on the ecology of mastreviruses have indicated that species within this horizontally-transmitted genus more frequently infect annual host species (Claverie *et al.*, 2023; Ben Chehida *et al.*, in preparation), with only a few perennial hosts serving as long-term pathogen reservoirs (Moore *et al.*, 2011). Here, whereas *Acanthaceae* were preferentially associated with vertical transmission (indeed the *Asystasia gangetica* pool from BP was infected with two distinct genera of vertically transmitted viruses), *Apiaceae* (*Centella asiatica* being the only species sampled here) were found preferentially associated to horizontally transmitted viruses. Besides the *Genomoviridae*, no viral family, nor virus traits were associated to the most sampled plant families (*Asteraceae* and *Poaceae*).

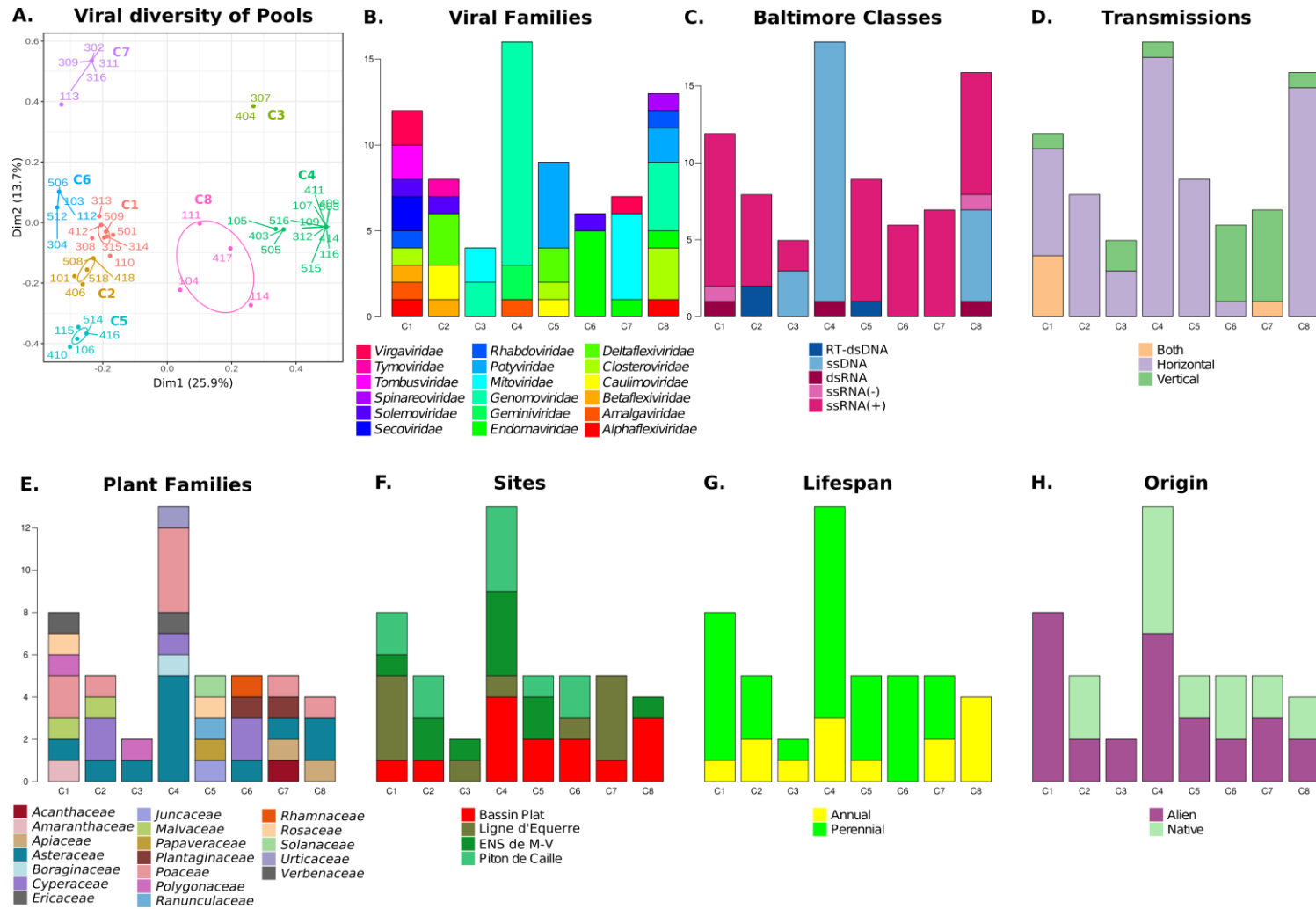


Figure 4. (A) Principal component analysis (PCoA) of the diversity of viral families in each sampled pool and colored by more probable clusters that are represented by a 95% confidence ellipse. Barplots represent for each cluster (B) the number of viral families, (C) the number of viral genera by Baltimore classes, (D) transmission types, (E) the number of pool of each plant family sampling site, (F) from each sampling site, (G) plant species lifespans and (H) plant species origin.

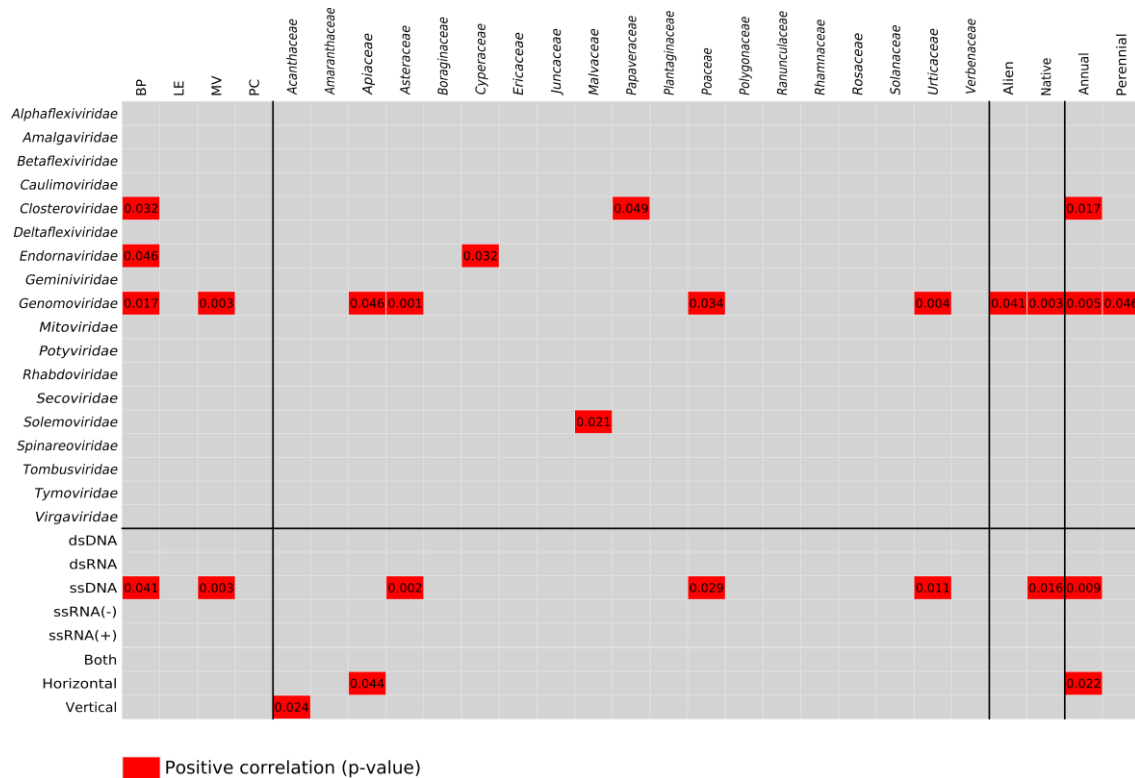


Figure 5. Correlation analysis results, with the fourth corner method, between the plant and virus characteristics. Only correspondence with p-value ≤ 0.05 are represented.

Are crop viruses involved in plant invasion?

Interestingly within the 134 distinct phytoviral contigs identified, 39 presented with relatively high similarity (85% to 98% nucleotidic identity on BLAST alignments) with well characterized plant viruses. Whereas some were described as emerging pests on plants (cotton leafroll dwarf virus, CLRDV, *Solemoviridae* family; Edula *et al.*, 2023; lettuce chlorosis virus, LCV, *Closteroviridae* family; Garcian and Janssen, 2020), long time foes for some others (turnip mosaic virus, TMV; Gibbs *et al.*, 2015) or established viruses that have been the central focus for plant pathologist all round the globe (potato virus Y, PVY; *Potyviridae* family, Torrance and Talianksy, 2020 ; barley yellow dwarf virus, BYDV, *Tombusviridae* family; Walls *et al.*, 2019). CLRDV, LCV, PVY and TMV were all detected in BP. PVY was found on *Solanum americanum*, *Laphangium luteoalbum* and *Centella asiatica*. CLRDV was detected on *Malvastrum coromandelianum*. TMV was detected on *Argemone mexicana*. LCV had the largest host range and was characterized on *Bromus catharticus*, *C. asiatica*, *A. mexicana* and *Bidens pilosa*. All these viruses are transmitted by aphids that were detected mainly in BP and LE (Ben Chehida *et al.*, in preparation). The fact that these viruses were detected on alien plants in an agro-ecosystem raised several

questions on the role of invasive plants as entry points for viruses. Moreover, it would question on how these viruses may contribute on the fitness of their host. Indeed, drought events frequently occurred in BP and examples of beneficial viral infection for drought resistance have been reported (Gonzalez *et al.*, 2021). In that sense, the detection of BYDV on *Anthoxanthum odoratum* in both MV and PC is striking. *A. odoratum* is one, if not the most invasive plant in Reunion (Macdonald *et al.*, 1991) exclusively distributed in the subalpine grasslands (Fenouillas *et al.*, 2021). The absence of detection of BYDV in other plants may indicate that viruses from the invaders are not causing the extinction of the native (Faillace *et al.*, 2017). Furthermore, the specific habitat preference of *A. odoratum* for subalpine grasslands (Fenouillas *et al.*, 2021) and the absence of virus detection in agricultural areas suggests that this virus may not have follow the agricultural connection for introduction in Reunion.

Concluding remarks

Through this large-scale sampling of various herbaceous plant species along an altitudinal transect in Reunion, putative novel viral species and genera were detected. Whereas our results may indicate that a more thorough sampling and sequencing effort is required, we nevertheless uncover some trends in plant-virus associations with (i) higher frequency of infection for native plants and (ii) horizontal transmission associated to annuals. The first results may be associated to the enemy release hypothesis (Mitchell and Power, 2003) where the alien plants were introduced in Reunion escaping to an important part of their pathogens (*e.g.* horizontally-transmitted phytoviruses) reducing their frequency of infection (Blumenthal *et al.*, 2009; Chun *et al.*, 2010). Moreover, this study also revealed the presence of multiple viruses frequently associated with crops and wild plants. Whereas probable diffusion of the viruses to other plant from the same sampling site were detected, it also indicates that cropping system could be important entry points for plant viruses. For BYDV as for the other foes usually associated with crops, future studies will certainly be required to assess the infection frequency of plants in different areas and to understand whether these crop pathogens could be beneficial to invasive alien plants beyond the context of cropping system.

Authors information

Affiliations

CIRAD, UMR PVBMT, Pôle de Protection des Plantes, 7 chemin de l'Irat, 97410 Saint Pierre, La Réunion, France

Sélim Ben Chehida, Murielle Hoarau, Jean-Michel Lett & Pierre Lefeuvre

CIRAD, UMR BGPI, TA A-54/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

Oumaima Moubset, Charlotte Julian, Laurence Blondin, Emmanuel Fernandez, Denis Filloux & Philippe Roumagnac

Université de Bordeaux, INRAE, UMR BFP, 33140 Villenave d'Ornon, France

Marie Lefebvre, Armelle Marais & Thierry Candresse

Tour du Valat, Institut de recherche pour la conservation des zones humides méditerranéennes, Le Sambuc, Arles, France

Hugo Fontes

Conservatoire Botanique National de Mascarin, F-97436, St Leu, La Réunion, France

Christophe Lavergne

Anses, Plant Health Laboratory, Bacteriology Virology GMO Unit, 7 rue Jean Dixmèras, 49044, Angers Cedex 01, France

Pascal Gentit & Matthieu Rolland

South African Medical Research Council Capacity Development Unit, South African National Bioinformatics Institute, University of the Western Cape, Bellville 7635, South Africa

Gordon W. Harkins

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

Arvind Varsani

Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Observatory 7925, South Africa

Darren P. Martin

Author contributions

S.B.C., O.M., A.M., T.C., P.R., J.M.L. and P.L. conceived and designed the experiments. S.B.C., M.H., O.M., C.J., L.B., E.F., D.F., M.L., H.F., C.L., P.G., M.R., A.M., T.C., P.R., J.M.L. and P.L. performed the experiments. S.B.C., O.M., D.F., M.L., A.M., T.C., P.R., D.P.M., J.M.L. and P.L. analysed the data. S.B.C., T.C., P.R., G.H., A.V., D.P.M., J.M.L. and P.L. wrote the paper. J.M.L. and P.L. secured funding for the project's execution.

Corresponding author

Correspondence to Pierre Lefevre

Funding

This study was funded by the European Union (ERDF, contract GURDT I2016-1731-0006632), the Conseil Régional de la Réunion and CIRAD. S.B.C. is a recipient of a PhD fellowship from CIRAD and ANR (Phytovirus project number: ANR-19-CE35-0008-02).

Data Availability Statement

Publicly available datasets were analyzed in this study.

Acknowledgments

This work was conducted on the Plant Protection Platform (3P, IBISA).

Conflicts of Interest

The authors declare no conflict of interest.

References

- Agrawal, A. A., Kotanen, P. M., Mitchell, C. E., Power, A. G., Godsoe, W., & Klironomos, J. (2005). Enemy release? An experiment with congeneric plant pairs and diverse above-and belowground enemies. *Ecology*, 86(11), 2979-2989.
- Baret, S., Rouget, M., Richardson, D. M., Lavergne, C., Egoh, B., Dupont, J., & Strasberg, D. (2006). Current distribution and potential extent of the most invasive alien plant species on La Réunion (Indian Ocean, Mascarene islands). *Austral Ecology*, 31(6), 747-758.
- Becker, D. J., Washburne, A. D., Faust, C. L., Pulliam, J. R., Mordecai, E. A., Lloyd-Smith, J. O., & Plowright, R. K. (2019). Dynamic and integrative approaches to understanding pathogen spillover. *Philosophical Transactions of the Royal Society B*, 374(1782), 20190014.
- Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, 12(2), 20150623.
- Ben Chehida, S., Bunwaree, H. D., Hoareau, M., Moubset, A., Julian, C., Blondin, L., Filloux, D., Lavergne, C., Roumagnac, P., Varsani, A., Martin, D. P., Lett, J.-M. & Lefeuvre, P. (in preparation). Reconstructing the *Mastrevirus* communities structure in Reunion: The tale of agricultural associated pathogens.
- Blumenthal, D., Mitchell, C. E., Pyšek, P., & Jarošík, V. (2009). Synergy between pathogen release and resource availability in plant invasion. *Proceedings of the National Academy of Sciences*, 106(19), 7899-7904.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.
- Buchfink, B., Reuter, K., & Drost, H. G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature methods*, 18(4), 366-368.
- Cadet, L. J. T. (1977). La végétation de l'île de La Réunion: étude phytoécologique et phytosociologique [dissertation]. *Marseille, France: Université Aix-Marseille III*.
- Chabi-Jesus, C., Najar, A., Fontenele, R. S., Kumari, S. G., Ramos-González, P. L., Freitas-Astúa, J., Kraberger, S. & Varsani, A. (2020). Viruses representing two new genomovirus species identified in citrus from Tunisia. *Archives of virology*, 165, 1225-1229.
- Chun, Y. J., Van Kleunen, M., & Dawson, W. (2010). The role of enemy release, tolerance and resistance in plant invasions: linking damage to performance. *Ecology Letters*, 13(8), 937-946.
- Claverie, S., Hoareau, M., Ben Chehida, S., Filloux, D., Varsani, A., Roumagnac, P., Martin, D. P., Lett, J. M. & Lefeuvre, P. (2023). Metagenomics reveals the structure of Mastrevirus-host interaction network within an agro-ecosystem. *Virus Evolution*, vead043.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2), giab008.
- Dolédec, S., Chessel, D., Ter Braak, C. J., & Champely, S. (1996). Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics*, 3, 143-166.
- Dray, S., Dufour, A. B., & Chessel, D. (2007). The ade4 package-II: Two-table and K-table methods. *R news*, 7(2), 47-52.
- Dray, S., & Legendre, P. (2008). Testing the species traits–environment relationships: the fourth-corner problem revisited. *Ecology*, 89(12), 3400-3412.
- Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite package: analysing ecological networks. *interaction*, 1(0.2413793), 8-11.

- Edula, S. R., Bag, S., Milner, H., Kumar, M., Suassuna, N. D., Chee, P. W., Kemerait, R.C., Hand, L. C., Snide, J. L., Srinivasan, R. & Roberts, P. M. (2023). Cotton leafroll dwarf disease: An enigmatic viral disease in cotton. *Molecular Plant Pathology*, 24(6), 513-526
- Faillace, C. A., Lorusso, N. S., & Duffy, S. (2017). Overlooking the smallest matter: viruses impact biological invasions. *Ecology Letters*, 20(4), 524-538.
- Feng, C., Feng, J., Wang, Z., Pedersen, C., Wang, X., Saleem, H., Domier, L. & Marzano, S. Y. L. (2021). Identification of the viral determinant of hypovirulence and host range in Sclerotiniaceae of a genomovirus reconstructed from the plant metagenome. *Journal of Virology*, 95(17), 10-1128.
- Fenouillas, P., Ah-Peng, C., Amy, E., Bracco, I., Dafreville, S., Gosset, M., Ingrassia, F., Lavergne, L., Lequette, B., Notter, J.-C., Pausé, J.-M., Payer, G., Payet, N., Picot, F., Pougavanon, N., Strasberg, D., Thomas, H., Triolo, J., Turquet, V. & Rouget, M. (2021). Quantifying invasion degree by alien plants species in Reunion Island. *Austral Ecology*, 46(7), 1025-1037.
- François, S., Filloux, D., Fernandez, E., Ogliastro, M., & Roumagnac, P. (2018). Viral metagenomics approaches for high-resolution screening of multiplexed arthropod and plant viral communities. *Viral Metagenomics: Methods and Protocols*, 77-95.
- Garcia, L. R., & Janssen, D. (2020). Epidemiology and control of emerging criniviruses in bean. *Virus Research*, 280, 197902.
- Gibbs, A. J., Nguyen, H. D., & Ohshima, K. (2015). The 'emergence' of turnip mosaic virus was probably a 'gene-for-quasi-gene' event. *Current Opinion in Virology*, 10, 20-26.
- González, R., Butković, A., Escaray, F. J., Martínez-Latorre, J., Melero, Í., Pérez-Parets, E., Gomez-Cadenas, A., Carrasco, P. & Elena, S. F. (2021). Plant virus evolution under strong drought conditions results in a transition from parasitism to mutualism. *Proceedings of the National Academy of Sciences*, 118(6), e2020990118.
- Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18), 2847-2849.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321.
- Hao, F., Wu, M., & Li, G. (2021). Characterization of a novel genomovirus in the phytopathogenic fungus *Botrytis cinerea*. *Virology*, 553, 111-116.
- Hill, M. O., & Smith, A. J. E. (1976). Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon*, 249-255.
- Hollander, M., & Wolfe, D. A. (1973). Kruskal-Wallis: a distribution-free test. *Nonparametric Statistical Methods*, 114-137.
- Illumina, I. (2017). Effects of index misassignment on multiplexing and downstream analysis. URL: www.illumina.com.
- IUCN (2010). IUCN Technical Evaluation. Pitons, cirques and ramparts of Reunion Island. *Technical Report*. IUCN, France
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587-589.
- Kassambara, A., & Mundt, F. (2021). Factoextra: extract and visualize the results of multivariate data analyses, R package version 1.0. 7. 2020. Preprint at.

- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.
- Koonin, E. V., Krupovic, M., & Agol, V. I. (2021). The Baltimore classification of viruses 50 years later: how does it stand in the light of virus evolution?. *Microbiology and Molecular Biology Reviews*, 85(3), e00053-21.
- Krupovic, M., Ghabrial, S. A., Jiang, D., & Varsani, A. (2016). Genomoviridae: a new family of widespread single-stranded DNA viruses. *Archives of Virology*, 161, 2633-2643.
- Lavergne, R. (1978). Les pestes végétales de l'île de La Réunion. *Info-nature*, 16, 9-60.
- Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., Yamashita, H. & Lam, T. W. (2016). MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3-11.
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094–3100.
- Lovisolo, O., Hull, R., & Rosler, O. (2003). Coevolution of viruses with hosts and vectors and possible paleontology. *Advances in Virus Research*, 62, 325-379.
- Ma, Y., Fort, T., Marais, A., Lefebvre, M., Theil, S., Vacher, C., & Candresse, T. (2021). Leaf-associated fungal and viral communities of wild plant populations differ between cultivated and natural ecosystems. *Plant-Environment Interactions*, 2(2), 87-99.
- Macdonald, I. A., Thébaud, C., Strahm, W. A., & Strasberg, D. (1991). Effects of alien plant invasions on native vegetation remnants on La Réunion (Mascarene Islands, Indian Ocean). *Environmental conservation*, 18(1), 51-61.
- Marais, A., Faure, C., Bergey, B., & Candresse, T. (2018). Viral double-stranded RNAs (dsRNAs) from plants: Alternative nucleic acid substrates for high-throughput sequencing. *Viral Metagenomics: Methods and Protocols*, 45-53.
- Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M., Reynard, J.-S., Rumbou, A., Saldarelli, P., Škorić, D., Vainio, E.J., Valkonen, J.P., 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.
- Mitchell, C. E., & Power, A. G. (2003). Release of invasive plants from fungal and viral pathogens. *Nature*, 421(6923), 625-627.
- Mitra, A., Skrzypczak, M., Ginalski, K. & Rowicka, M. (2015). Strategies for achieving high sequencing accuracy for low diversity samples and avoiding sample bleeding using Illumina platform. *PloS one*, 10(4), e0120520.
- Moore, S. M., Manore, C. A., Bokil, V. A., Borer, E. T., & Hosseini, P. R. (2011). Spatiotemporal model of barley and cereal yellow dwarf virus transmission dynamics with seasonality and plant competition. *Bulletin of mathematical biology*, 73, 2707-2730.
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?. *Journal of classification*, 31, 274-295.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P; R., O'hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. & Maintainer, H. W. (2019). Vegan: community ecology package. *Ordination Methods, Diversity Analysis and Other Functions for Community and Vegetation Ecologists*, 05–26. Version 2.5-1.
- Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.

- Paudel, D. B., & Sanfaçon, H. (2018). Exploring the diversity of mechanisms associated with plant tolerance to virus infection. *Frontiers in Plant Science*, 9, 1575.
- Power, A. G., & Mitchell, C. E. (2004). Pathogen spillover in disease epidemics. *The American Naturalist*, 164(S5), S79-S89.
- Power, A., & Flecker, A. (2008). The role of vector diversity in disease dynamics (pp. 30-47). Princeton, NJ: Princeton University Press.
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*, 5(3), e9490.
- Pyšek, P., Chytrý, M., Pergl, J., Sadlo, J., & Wild, J. (2012). Plant invasions in the Czech Republic: current state, introduction dynamics, invasive species and invaded habitats. *Preslia*, 84(3), 575-629.
- Pyšek, P., Hulme, P. E., Simberloff, D., Bacher, S., Blackburn, T. M., Carlton, J. T., Dawson, W., Essl, F., Foxcroft, L. C., Genovesi, P., Jeschke, J. M., Kühne, I., Liebhold, A. M., Mandrak, N. E., Meyerson, L. A., Pauchard, A., Pergl, J., Roy, H. E., Seebens, H., van Kleunen, M., Vilà, M., Wingfield, M. J. & Richardson, D. M. (2020). Scientists' warning on invasive alien species. *Biological Reviews*, 95(6), 1511-1534.
- Roossinck, M. J., Saha, P., Wiley, G. B., Quan, J., White, J. D., Lai, H., Chavarria, F., Shen, G. & Roe, B. A. (2010). Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*, 19, 81-88.
- Roumagnac, P., Lett, J. M., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F. M., Martin, D. P., & Varsani, A. (2022). Establishment of five new genera in the family Geminiviridae: Citlodavirus, Maldovirus, Mulcrilevirus, Opunvirus, and Topilevirus. *Archives of Virology*, 1-16.
- Roux, S., Matthijssens, J., & Dutilh, B. E. (2021). Metagenomics in virology. *Encyclopedia of Virology*, 133.
- Rúa, M. A., Pollina, E. C., Power, A. G., & Mitchell, C. E. (2011). The role of viruses in biological invasions: friend or foe?. *Current Opinion in Virology*, 1(1), 68-72.
- Schönegger, D. (2023). Analysis of the virome and reciprocal transfers of viruses between cultivated and wild carrot populations (Doctoral dissertation, Université de Bordeaux).
- Siddell, S. G., Smith, D. B., Adriaenssens, E., Alfenas-Zerbini, P., Dutilh, B. E., Garcia, M. L., Jungle, S., Krupovic, M., Kuhn, J. H., Lambert, A. J., Lefkowitz, E. J., Lobočka, M., Mushegian, A. R., Oksanen, H. M., Robertson, D. L., Ruino, L., Sabanadzovic, S., Simmonds, P., Suzuki, N., Van Doorslaer, K., Vandamme, A.-M., Varsani, A. & Zerbini, F. M. (2023). Virus taxonomy and the role of the International Committee on Taxonomy of Viruses (ICTV). *Journal of General Virology*, 104(5), 001840.
- Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11), 1026-1028.
- Strasberg, D., Rouget, M., Richardson, D. M., Baret, S., Dupont, J., & Cowling, R. M. (2005). An assessment of habitat diversity and transformation on La Réunion Island (Mascarene Islands, Indian Ocean) as a basis for identifying broad-scale conservation priorities. *Biodiversity & Conservation*, 14, 3015-3032.
- Torrance, L., & Taliaknsy, M. E. (2020). Potato Virus Y emergence and evolution from the Andes of South America to become a major destructive pathogen of potato and other solanaceous crops worldwide. *Viruses*, 12(12), 1430.
- Varsani, A., & Krupovic, M. (2021). Family Genomoviridae: 2021 taxonomy update. *Archives of Virology*, 166, 2911-2926.
- Vilà, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarošík, V., Maron, J. L., Pergl, J., Schaffner, U., Sun, Y. & Pyšek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, 14(7), 702-708.

Walls III, J., Rajotte, E., & Rosa, C. (2019). The past, present, and future of barley yellow dwarf management. *Agriculture*, 9(1), 23.

Xu, W., Li, H., Sivasithamparam, K., Tran, D. T., Jones, M. G., Chen, X., & Wylie, S. J. (2022). Spillover of a Tobamovirus from the Australian Indigenous Flora to Invasive Weeds. *Viruses*, 14(8), 1676.

Supplementary Table 1. Sampling sites and summary of sampled pools. The number in brackets refers to the number of pools detected for which phytovirome genomes were obtained.

Type of Site	Site	Lat	Lon	Alt (m)	Nov 2020
AE	BP	-21.3216	55.4892	166	16 (8)
SG	LE	-21.2916	55.5733	1,062	17 (2)
SG	MV	-21.2859	55.6030	1,586	18 (7)
SG	PC	-21.1902	55.6380	2,130	18 (6)
Total					69 (23)

Supplementary Table 2. Summary of sampled pool in each type site and site.

Date Sampling	Type of Site	Site	N Samples	Pool Id	Plant Species	Host Family	Life Cycle	Invasion Status	Invasion Score
02/11/20	AE	BP	50	20_REU_E010100	<i>Malvastrum coromandelianum</i>	Malvaceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E010200	<i>Chloris gayana</i>	Poaceae	Perennial	Alien	2
02/11/20	AE	BP	50	20_REU_E010300	<i>Cyperus rotundus</i>	Cyperaceae	Perennial	Alien	3
02/11/20	AE	BP	50	20_REU_E010400	<i>Centella asiatica</i>	Apiaceae	Annual	Native	0
02/11/20	AE	BP	50	20_REU_E010500	<i>Erigeron sumatrensis</i>	Asteraceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E010600	<i>Argemone mexicana</i>	Papaveraceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E010700	<i>Trichodesma indicum</i>	Boraginaceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E010800	<i>Amaranthus viridis</i>	Amaranthaceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E010900	<i>Cynodon dactylon</i>	Poaceae	Perennial	Native	0
02/11/20	AE	BP	50	20_REU_E011000	<i>Bidens pilosa</i>	Asteraceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E011100	<i>Bromus catharticus</i>	Poaceae	Annual	Alien	3
02/11/20	AE	BP	50	20_REU_E011200	<i>Plantago lanceolata</i>	Plantaginaceae	Perennial	Alien	3
02/11/20	AE	BP	50	20_REU_E011300	<i>Asystasia gangetica</i>	Acanthaceae	Perennial	Alien	4
02/11/20	AE	BP	50	20_REU_E011400	<i>Laphangium luteoalbum</i>	Asteraceae	Annual	Alien	2
02/11/20	AE	BP	50	20_REU_E011500	<i>Solanum americanum</i>	Solanaceae	Both	Alien	3
02/11/20	AE	BP	50	20_REU_E011600	<i>Parthenium hysterophorus</i>	Asteraceae	Annual	Alien	3
03/11/20	SG	LE	50	20_REU_E030100	<i>Laphangium luteoalbum</i>	Asteraceae	Annual	Alien	2
03/11/20	SG	LE	50	20_REU_E030200	<i>Centella asiatica</i>	Apiaceae	Annual	Native	0
03/11/20	SG	LE	50	20_REU_E030300	<i>Bromus catharticus</i>	Poaceae	Annual	Alien	3
03/11/20	SG	LE	50	20_REU_E030400	<i>Cyperus aromaticus</i>	Cyperaceae	Perennial	Native	0
03/11/20	SG	LE	50	20_REU_E030500	<i>Hypochoeris radicata</i>	Asteraceae	Perennial	Alien	5
03/11/20	SG	LE	50	20_REU_E030600	<i>Holcus lanatus</i>	Poaceae	Perennial	Alien	4
04/11/20	SG	LE	50	20_REU_E030700	<i>Ageratum conyzoides</i>	Asteraceae	Annual	Alien	3
03/11/20	SG	LE	50	20_REU_E030800	<i>Sida retusa</i>	Malvaceae	Both	Alien	2
03/11/20	SG	LE	50	20_REU_E030900	<i>Veronica arvensis</i>	Plantaginaceae	Annual	Alien	3
03/11/20	SG	LE	50	20_REU_E031000	<i>Sisyrinchium micranthum</i>	Iridaceae	Perennial	Alien	3
03/11/20	SG	LE	50	20_REU_E031100	<i>Cynodon dactylon</i>	Poaceae	Perennial	Native	0
03/11/20	SG	LE	50	20_REU_E031200	<i>Cenchrus clandestinus</i>	Poaceae	Perennial	Alien	3
03/11/20	SG	LE	50	20_REU_E031300	<i>Verbena bonariensis</i>	Verbanaceae	Perennial	Alien	3
03/11/20	SG	LE	50	20_REU_E031400	<i>Persicaria punctata</i>	Polygonaceae	Both	Alien	4
03/11/20	SG	LE	50	20_REU_E031500	<i>Achyranthes aspera</i>	Amaranthaceae	Perennial	Alien	1
03/11/20	SG	LE	50	20_REU_E031600	<i>Ageratina riparia</i>	Asteraceae	Perennial	Alien	5
03/11/20	SG	LE	50	20_REU_E031700	<i>Erigeron sumatrensis</i>	Asteraceae	Annual	Alien	3
04/11/20	SG	MV	50	20_REU_E040100	<i>Dactylis glomerata</i>	Poaceae	Perennial	Alien	3
04/11/20	SG	MV	50	20_REU_E040200	<i>Erigeron karvinskianus</i>	Asteraceae	Perennial	Alien	5
04/11/20	SG	MV	50	20_REU_E040300	<i>Ageratina riparia</i>	Asteraceae	Perennial	Alien	5
04/11/20	SG	MV	50	20_REU_E040400	<i>Rumex acetosella</i>	Polygonaceae	Perennial	Alien	4
04/11/20	SG	MV	50	20_REU_E040500	<i>Plantago lanceolata</i>	Plantaginaceae	Perennial	Alien	3
04/11/20	SG	MV	50	20_REU_E040600	<i>Holcus lanatus</i>	Poaceae	Perennial	Alien	4
04/11/20	SG	MV	50	20_REU_E040700	<i>Potentilla indica</i>	Rosaceae	Perennial	Alien	4
04/11/20	SG	MV	50	20_REU_E040800	<i>Carex balfourii</i>	Cyperaceae	Perennial	Native	0
04/11/20	SG	MV	50	20_REU_E040900	<i>Hypochoeris radicata</i>	Asteraceae	Perennial	Alien	5
04/11/20	SG	MV	50	20_REU_E041000	<i>Juncus effusus</i>	Juncaceae	Perennial	Native	0
04/11/20	SG	MV	50	20_REU_E041100	<i>Faujasia salicifolia</i>	Asteraceae	Perennial	Native	0

04/11/20	SG	MV	50	20_REU_E041200	<i>Anthoxanthum odoratum</i>	Poaceae	Perennial	Alien	5
04/11/20	SG	MV	50	20_REU_E041300	<i>Zantedeschia aethiopica</i>	Araceae	Perennial	Alien	4
04/11/20	SG	MV	50	20_REU_E041400	<i>Pilea urticifolia</i>	Urticaceae	Perennial	Native	0
04/11/20	SG	MV	50	20_REU_E041500	<i>Cenchrus clandestinus</i>	Poaceae	Perennial	Alien	3
04/11/20	SG	MV	50	20_REU_E041600	<i>Clematis mauritiana</i>	Ranunculaceae	Perennial	Native	0
04/11/20	SG	MV	50	20_REU_E041700	<i>Humbertacalia tomentosa</i>	Asteraceae	Annual	Native	0
04/11/20	SG	MV	50	20_REU_E041800	<i>Carex boryana</i>	Cyperaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E050100	<i>Fragaria vesca</i>	Rosaceae	Perennial	Alien	4
04/11/20	SG	PC	50	20_REU_E050200	<i>Juncus effusus</i>	Juncaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E050300	<i>Erica galioides</i>	Ericaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E050400	<i>Erigeron karvinskianus</i>	Asteraceae	Perennial	Alien	5
05/11/20	SG	PC	50	20_REU_E050500	<i>Holcus lanatus</i>	Poaceae	Perennial	Alien	4
05/11/20	SG	PC	50	20_REU_E050600	<i>Phyllica nitida</i>	Rhamnaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E050700	<i>Plantago lanceolata</i>	Plantaginaceae	Perennial	Alien	3
05/11/20	SG	PC	50	20_REU_E050800	<i>Humbertacalia tomentosa</i>	Asteraceae	Annual	Native	0
05/11/20	SG	PC	50	20_REU_E050900	<i>Anthoxanthum odoratum</i>	Poaceae	Perennial	Alien	5
05/11/20	SG	PC	50	20_REU_E051000	<i>Carex leporina</i>	Cyperaceae	Perennial	Alien	2
05/11/20	SG	PC	50	20_REU_E051100	<i>Hypochaeris radicata</i>	Asteraceae	Perennial	Alien	5
05/11/20	SG	PC	50	20_REU_E051200	<i>Helichrysum arnicoides</i>	Asteraceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E051300	<i>Panicum juniperinum</i>	Poaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E051400	<i>Potentilla indica</i>	Rosaceae	Perennial	Alien	4
05/11/20	SG	PC	19	20_REU_E051500	<i>Festuca borbonica</i>	Poaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E051600	<i>Costularia melicoides</i>	Cyperaceae	Perennial	Native	0
05/11/20	SG	PC	50	20_REU_E051700	<i>Costularia cadetii</i>	Cyperaceae	Perennial	Native	0
05/11/20	SG	PC	18	20_REU_E051800	<i>Carex boryana</i>	Cyperaceae	Perennial	Native	0

Supplementary Table 3. HTS sequencing statistics.

	Raw paired-end reads	Trimmed paired-end reads	Passed paired-end reads	Viral reads (20_REU)	Viral contigs (reads)	Phytoviral contigs (reads)
VANA	71,370,367	61,908,343	46,245,965	8,002,545	810	134
dsRNA	14,044,710	13,764,898	9,302,142	3,525,597	(11,528,142)	(3,595,195)

Supplementary Table 4. Summary of taxonomic classification and assignment results of viral sequences per plant species for each site.

Baltimore	Transmission	Viral Family	Viral Genus	TaxId	Contig Id	Contig Length	N Reads	Alignment Length	Percentage Identity	E Value	Pool Id	Host Family	Host Species	Site
ssDNA	Horizontal	<i>Geminiviridae</i>	<i>Begomovirus</i>	714308	k_645012	1482	171	345	34.2	2.04e-35	403	<i>Asteraceae</i>	<i>Ageratina riparia</i>	MV
ssDNA	Horizontal	<i>Geminiviridae</i>	<i>Begomovirus</i>	1046573	k_699320	2422	4125	227	34.4	2.1e-21	505	<i>Poaceae</i>	<i>Holcus lanatus</i>	PC
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1986478	k_700775	983	4219	487	71.3	1.63e-57	414	<i>Urticaceae</i>	<i>Pilea urticifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1699094	k_160144	586	2928	93	82.8	2.55e-14	414	<i>Urticaceae</i>	<i>Pilea urticifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2705538	k_250030	1654	3261	1201	70.6	2.97e-139	409	<i>Asteraceae</i>	<i>Hypochaeris radicata</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2230900	k_424954	440	1645	63	65.1	2.05e-17	312	<i>Poaceae</i>	<i>Cenchrus clandestinus</i>	LE
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2230900	k_424954	440	407	63	65.1	2.05e-17	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2230900	k_424954	440	470	63	65.1	2.05e-17	515	<i>Poaceae</i>	<i>Festuca borbonica</i>	PC
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2291614	k_534182	707	1500	266	75.6	4.93e-37	414	<i>Urticaceae</i>	<i>Pilea urticifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1367218	k_554681	1346	781	1315	80.1	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2219127	k_599761	2363	575	634	89.6	0	307	<i>Asteraceae</i>	<i>Ageratum conyzoides</i>	LE
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2219127	k_599761	2363	40442	634	89.6	0	409	<i>Asteraceae</i>	<i>Hypochaeris radicata</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2219127	k_599761	2363	239	634	89.6	0	411	<i>Asteraceae</i>	<i>Faujasia salicifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2219123	k_701128	2275	14438	432	80.1	1.18e-101	505	<i>Poaceae</i>	<i>Holcus lanatus</i>	PC
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2518602	k_62268	888	1621	211	49.8	4.29e-57	114	<i>Asteraceae</i>	<i>Laphangium luteoalbum</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1985392	k_172156	1493	7075	295	39.7	5.81e-46	414	<i>Urticaceae</i>	<i>Pilea urticifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1985404	k_376567	2386	45605	586	70.8	4.62e-69	505	<i>Poaceae</i>	<i>Holcus lanatus</i>	PC
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1985400	k_607041	2429	1837	883	95.6	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1985400	k_607041	2429	819	883	95.6	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1985400	k_155623	1183	620	302	85.4	1.25e-91	307	<i>Asteraceae</i>	<i>Ageratum conyzoides</i>	LE
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	664785	k_393682	902	587	227	53.7	3.88e-66	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	1983771	k_421722	1582	543	1564	78.8	0	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	2584384	k_68885	2220	791	1735	78.8	0	403	<i>Asteraceae</i>	<i>Ageratina riparia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykibivirus</i>	1986479	k_596495	393	255	166	48.2	3.72e-27	411	<i>Asteraceae</i>	<i>Faujasia salicifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykibivirus</i>	1516081	k_167758	2470	19284	218	49.5	9.99e-50	414	<i>Urticaceae</i>	<i>Pilea urticifolia</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2705543	k_112883	1873	11301	553	95.1	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2705543	k_112883	1873	77	553	95.1	0	107	<i>Boraginaceae</i>	<i>Trichodesma indicum</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2705543	k_112883	1873	1211	553	95.1	0	109	<i>Poaceae</i>	<i>Cynodon dactylon</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	224	116	94.0	5.46e-41	105	<i>Asteraceae</i>	<i>Erigeron sumatrensis</i>	BP

ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	30	116	94.0	5.46e-41	107	<i>Boraginaceae</i>	<i>Trichodesma indicum</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	263	116	94.0	5.46e-41	116	<i>Asteraceae</i>	<i>Parthenium hysterophorus</i>	BP
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	817	116	94.0	5.46e-41	307	<i>Asteraceae</i>	<i>Ageratum conyzoides</i>	LE
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	263	116	94.0	5.46e-41	404	<i>Polygonaceae</i>	<i>Rumex acetosella</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	346	116	94.0	5.46e-41	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	138	116	94.0	5.46e-41	503	<i>Ericaceae</i>	<i>Erica galioides</i>	PC
ssDNA	Horizontal	<i>Genomoviridae</i>	<i>Gemykolovirus</i>	2584403	k_215086	505	113	116	94.0	5.46e-41	516	<i>Cyperaceae</i>	<i>Costularia melicoides</i>	PC
dsDNA-RT	Horizontal	<i>Caulimoviridae</i>	<i>Badnavirus</i>	2056885	k_351999	555	1385	104	36.5	5.58e-12	418	<i>Cyperaceae</i>	<i>Carex boryana</i>	MV
dsDNA-RT	Horizontal	<i>Caulimoviridae</i>	<i>Badnavirus</i>	2056885	k_351999	555	2600	104	36.5	5.58e-12	508	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	PC
dsDNA-RT	Horizontal	<i>Caulimoviridae</i>	<i>Solendovirus</i>	107324	k_550545	337	485	76	53.9	1.17e-09	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(-)	Horizontal	<i>Rhabdoviridae</i>	<i>Betanucleorhabdovirus</i>	2813288	k_144399	316	186	172	78.5	6.75e-31	314	<i>Polygonaceae</i>	<i>Persicaria punctata</i>	LE
ssRNA(-)	Horizontal	<i>Rhabdoviridae</i>	<i>Betanucleorhabdovirus</i>	11307	k_329634	330	341	328	80.5	3.43e-79	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(-)	Horizontal	<i>Rhabdoviridae</i>	<i>Betanucleorhabdovirus</i>	11307	k_256373	356	676	329	81.8	3.98e-85	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(-)	Horizontal	<i>Rhabdoviridae</i>	<i>Betanucleorhabdovirus</i>	11307	k_382783	496	281	489	78.5	1.74e-110	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(-)	Horizontal	<i>Rhabdoviridae</i>	<i>Betanucleorhabdovirus</i>	11307	k_550558	595	682	589	75.4	1.99e-104	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	1985968	k_680764	391	308	121	42.1	5.43e-17	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	1043184	k_332154	595	601	201	49.8	1.95e-49	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	1043184	k_33092	601	1597	126	59.5	2.25e-40	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	742819	k_11120	358	291	100	48.0	9.1e-15	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Allexivirus</i>	2016426	k_550934	591	5391	68	64.7	1.58e-17	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	447171	k_20109	508	796	152	48.7	3.96e-29	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_257588	2134	9574	1799	83.4	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_472649	1207	2132	1063	86.5	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_517795	1994	19546	1985	86.2	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_630273	765	346	752	79.8	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_192722	2216	16014	2216	81.2	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	12187	k_517686	3794	31431	3781	86.0	0	501	<i>Rosaceae</i>	<i>Fragaria vesca</i>	PC
ssRNA(+)	Horizontal	<i>Betaflexiviridae</i>	<i>Banmivirus</i>	148879	k_353423	304	2060	103	61.2	7.52e-25	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Horizontal	<i>Betaflexiviridae</i>	<i>Chordovirus</i>	1425362	k_2886	330	507	95	44.2	3.21e-12	315	<i>Amaranthaceae</i>	<i>Achyranthes aspera</i>	LE
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	558690	k_428396	834	88	69	84.1	8.16e-10	110	<i>Asteraceae</i>	<i>Bidens pilosa</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	558690	k_428396	834	7808	69	84.1	8.16e-10	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_141614	822	442	817	96.5	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_26613	506	1308	502	93.8	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_5814	415	342	394	92.1	8.49e-158	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP

ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_268717	953	632	953	97.4	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_84353	619	262	422	90.8	1.39e-163	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_110556	1757	1337	1757	95.3	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_321094	1667	1629	1642	88.2	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_628992	825	392	733	90.5	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_296538	715	625	690	96.5	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_34158	874	896	876	87.7	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_268717	953	588	953	97.4	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_26613	506	1076	502	93.8	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_34158	874	443	876	87.7	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_110556	1757	199	1757	95.3	0	110	<i>Asteraceae</i>	<i>Bidens pilosa</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_141614	822	2123	817	96.5	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_321094	1667	1818	1642	88.2	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_26613	506	2683	502	93.8	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_110556	1757	2979	1757	95.3	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_34158	874	3778	876	87.7	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	642478	k_566269	441	1409	276	91.7	3.39e-106	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Horizontal	<i>Closteroviridae</i>	<i>Crinivirus</i>	67754	k_481877	362	472	119	56.3	5.44e-22	417	<i>Asteraceae</i>	<i>Humbertacalia tomentosa</i>	MV
ssRNA(+)	Horizontal	<i>Deltaflexiviridae</i>	<i>Deltaflexivirus</i>	1788309	k_597170	780	804	148	41.9	1.65e-22	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Deltaflexiviridae</i>	<i>Deltaflexivirus</i>	1788309	k_597170	780	581	148	41.9	1.65e-22	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Deltaflexiviridae</i>	<i>Deltaflexivirus</i>	1788309	k_597170	780	431	148	41.9	1.65e-22	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Horizontal	<i>Deltaflexiviridae</i>	<i>Deltaflexivirus</i>	1719047	k_441622	408	2191	327	87.8	4.05e-111	406	<i>Poaceae</i>	<i>Holcus lanatus</i>	MV
ssRNA(+)	Horizontal	<i>Deltaflexiviridae</i>	<i>Deltaflexivirus</i>	1719047	k_455952	979	886	368	84.0	6.46e-107	410	<i>Juncaceae</i>	<i>Juncus effusus</i>	MV
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_107893	325	30	99	45.5	3.24e-19	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_208794	1094	595	247	74.9	9.46e-36	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_193616	2438	1739	732	49.5	0	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_27074	1356	528	457	34.4	7.92e-54	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_184957	1283	702	408	40.2	6.45e-73	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1776177	k_107905	317	47	104	45.2	7.37e-20	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1777015	k_558648	666	79	209	84.2	1.16e-95	103	<i>Cyperaceae</i>	<i>Cyperus rotundus</i>	BP
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1468454	k_176450	1449	596	1305	71.9	3.15e-176	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1468454	k_190583	1151	734	1109	72.7	3.46e-168	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1468454	k_92653	504	629	151	38.4	5.98e-14	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1468454	k_650261	409	489	289	70.9	1.02e-23	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC

ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1468454	k_210209	961	362	324	42.6	8.79e-62	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	1131702	k_468163	1310	817	107	39.3	1.56e-14	512	<i>Asteraceae</i>	<i>Helichrysum amnicoides</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	310749	k_180801	380	610	105	75.2	1.23e-09	113	<i>Acanthaceae</i>	<i>Asystasia gangetica</i>	BP
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	310749	k_370149	488	361	139	82.0	1.61e-28	304	<i>Cyperaceae</i>	<i>Cyperus aromaticus</i>	LE
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	310749	k_370149	488	305	139	82.0	1.61e-28	506	<i>Rhamnaceae</i>	<i>Phylica nitida</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	310749	k_415313	492	668	127	76.4	2.12e-14	506	<i>Rhamnaceae</i>	<i>Phylica nitida</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	310749	k_181609	600	854	65	84.6	2.01e-09	506	<i>Rhamnaceae</i>	<i>Phylica nitida</i>	PC
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	3051685	k_504949	15052	1663	14304	98.7	0	111	<i>Poaceae</i>	<i>Bromus catharticus</i>	BP
ssRNA(+)	Vertical	<i>Endornaviridae</i>	<i>Alphaendornavirus</i>	3051685	k_504949	15052	430230	14304	98.7	0	112	<i>Plantaginaceae</i>	<i>Plantago lanceolata</i>	BP
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Duamitovirus</i>	2080450	k_334083	578	608	137	47.4	1.1e-22	307	<i>Asteraceae</i>	<i>Ageratum conyzoides</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Duamitovirus</i>	2080450	k_334083	578	3254	137	47.4	1.1e-22	309	<i>Plantaginaceae</i>	<i>Veronica arvensis</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Duamitovirus</i>	1629666	k_530067	437	116	434	96.1	0	302	<i>Apiaceae</i>	<i>Centella asiatica</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Duamitovirus</i>	1629666	k_265870	480	162	475	91.6	0	302	<i>Apiaceae</i>	<i>Centella asiatica</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Duamitovirus</i>	2080459	k_175451	2849	16061	749	51.9	0	113	<i>Acanthaceae</i>	<i>Asystasia gangetica</i>	BP
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Mitovirus</i>	2052563	k_413185	650	291	212	55.2	2.18e-59	311	<i>Poaceae</i>	<i>Cynodon dactylon</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Mitovirus</i>	2052563	k_252992	727	471	245	52.7	1.18e-59	311	<i>Poaceae</i>	<i>Cynodon dactylon</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Unuamitovirus</i>	1564031	k_661638	662	3671	207	34.8	2.1e-10	316	<i>Asteraceae</i>	<i>Ageratina riparia</i>	LE
ssRNA(+)	Vertical	<i>Mitoviridae</i>	<i>Unuamitovirus</i>	1133728	k_264399	347	1770	115	66.1	5.27e-37	404	<i>Polygonaceae</i>	<i>Rumex acetosella</i>	MV
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Brambyvirus</i>	323118	k_502700	9606	817	2718	37.7	0	410	<i>Juncaceae</i>	<i>Juncus effusus</i>	MV
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Brambyvirus</i>	323118	k_502700	9606	102048	2718	37.7	0	416	<i>Ranunculaceae</i>	<i>Clematis mauritiana</i>	MV
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Brambyvirus</i>	323118	k_502700	9606	40	2718	37.7	0	514	<i>Rosaceae</i>	<i>Potentilla indica</i>	PC
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12196	k_77593	428	1223	387	80.4	8.25e-95	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12196	k_478055	533	600	506	84.2	7.01e-154	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	2734556	k_208629	332	921	326	81.0	8.12e-81	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_365973	1703	283	1298	83.7	0	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_676000	575	2038	479	87.9	4.79e-169	114	<i>Asteraceae</i>	<i>Laphangium luteoalbum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_365973	1703	110	1298	83.7	0	114	<i>Asteraceae</i>	<i>Laphangium luteoalbum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_608697	456	396	415	75.2	6.8e-71	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_623339	882	21952	881	87.7	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_614832	675	701	671	88.8	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_483	2474	7492	2469	84.5	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_676000	575	2504	479	87.9	4.79e-169	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_265772	1814	15647	1646	86.8	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_271316	630	2219	629	76.6	3.34e-127	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP

ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_96417	562	6481	127	91.3	2.14e-40	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_591933	968	424	938	87.8	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_572007	514	807	479	88.7	1.58e-174	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12216	k_365973	1703	19337	1298	83.7	0	115	<i>Solanaceae</i>	<i>Solanum americanum</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12230	k_467430	425	669	411	93.4	4.51e-174	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12230	k_476306	1934	12071	1934	94.5	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12230	k_571737	500	1025	499	94.2	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12230	k_561021	347	606	282	94.0	1.04e-117	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Potyviridae</i>	<i>Potyvirus</i>	12230	k_569512	368	490	368	98.9	0	106	<i>Papaveraceae</i>	<i>Argemone mexicana</i>	BP
ssRNA(+)	Horizontal	<i>Secoviridae</i>	<i>Nepovirus</i>	191547	k_104465	1401	12561	1405	74.6	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
ssRNA(+)	Horizontal	<i>Secoviridae</i>	<i>Nepovirus</i>	2822470	k_687623	4503	22759	1257	58.3	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
ssRNA(+)	Horizontal	<i>Secoviridae</i>	<i>Waikavirus</i>	2419797	k_144568	426	302	103	45.6	1.01e-18	314	<i>Polygonaceae</i>	<i>Persicaria punctata</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_24674	422	2565	420	88.3	4.2e-149	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_163952	347	719	347	91.9	2.44e-138	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_419109	344	228	327	93.9	2.41e-138	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_129189	411	3660	411	90.0	4.35e-155	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_560050	402	590	401	92.5	2.51e-164	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_564342	599	1674	599	91.0	0	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Polerovirus</i>	312295	k_254343	448	335	445	89.7	5.45e-167	101	<i>Malvaceae</i>	<i>Malvastrum coromandelianum</i>	BP
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12470	k_465756	543	607	132	76.5	4.54e-17	304	<i>Cyperaceae</i>	<i>Cyperus aromaticus</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12470	k_75194	368	47505	101	50.5	9.3e-17	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_236717	1335	751	149	58.4	1.5e-36	304	<i>Cyperaceae</i>	<i>Cyperus aromaticus</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_380776	899	5464	182	54.4	2.39e-42	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_75391	469	456982	57	54.4	7.21e-10	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_7348	4247	1695518	1373	70.0	9.41e-157	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_398694	380	3704	117	62.4	5.85e-31	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Horizontal	<i>Solemoviridae</i>	<i>Sobemovirus</i>	12472	k_236717	1335	109603	149	58.4	1.5e-36	308	<i>Malvaceae</i>	<i>Sida retusa</i>	LE
ssRNA(+)	Both	<i>Tombusviridae</i>	<i>Luteovirus</i>	2169985	k_429211	1751	10358	1734	91.6	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
ssRNA(+)	Both	<i>Tombusviridae</i>	<i>Luteovirus</i>	2169985	k_429211	1751	3122	1734	91.6	0	509	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	PC
ssRNA(+)	Horizontal	<i>Tymoviridae</i>	<i>Marafivirus</i>	2169998	k_521437	878	22646	559	72.5	6.17e-75	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Horizontal	<i>Tymoviridae</i>	<i>Marafivirus</i>	56879	k_189043	576	10740	576	71.4	2.05e-72	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Horizontal	<i>Tymoviridae</i>	<i>Marafivirus</i>	56879	k_538080	921	28530	766	72.2	1.74e-107	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Horizontal	<i>Tymoviridae</i>	<i>Marafivirus</i>	1778580	k_75097	333	2774	313	70.0	4.52e-27	518	<i>Cyperaceae</i>	<i>Carex boryana</i>	PC
ssRNA(+)	Both	<i>Virgaviridae</i>	<i>Hordeivirus</i>	12327	k_499164	434	494	90	84.4	1.02e-17	113	<i>Acanthaceae</i>	<i>Asystasia gangetica</i>	BP

ssRNA(+)	Both	<i>Virgaviridae</i>	<i>Hordeivirus</i>	12327	k_121208	3630	541	3414	78.4	0	313	<i>Verbenaceae</i>	<i>Verbena bonariensis</i>	LE
ssRNA(+)	Both	<i>Virgaviridae</i>	<i>Hordeivirus</i>	12327	k_688358	2570	37844	2345	83.3	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
ssRNA(+)	Both	<i>Virgaviridae</i>	<i>Hordeivirus</i>	12327	k_121208	3630	56874	3414	78.4	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
ssRNA(+)	Both	<i>Virgaviridae</i>	<i>Hordeivirus</i>	12327	k_678	3037	54808	1784	72.5	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
dsRNA	Vertical	<i>Amalgaviridae</i>	<i>Amalgavirus</i>	1985164	k_447453	691	253	122	77.9	5.87e-17	105	<i>Asteraceae</i>	<i>Erigeron sumatrensis</i>	BP
dsRNA	Vertical	<i>Amalgaviridae</i>	<i>Amalgavirus</i>	1985164	k_559702	2176	26761	709	62.9	0	412	<i>Poaceae</i>	<i>Anthoxanthum odoratum</i>	MV
dsRNA	Horizontal	<i>Spinareoviridae</i>	<i>Oryzavirus</i>	42475	k_215568	459	370	153	38.6	3.13e-21	104	<i>Apiaceae</i>	<i>Centella asiatica</i>	BP

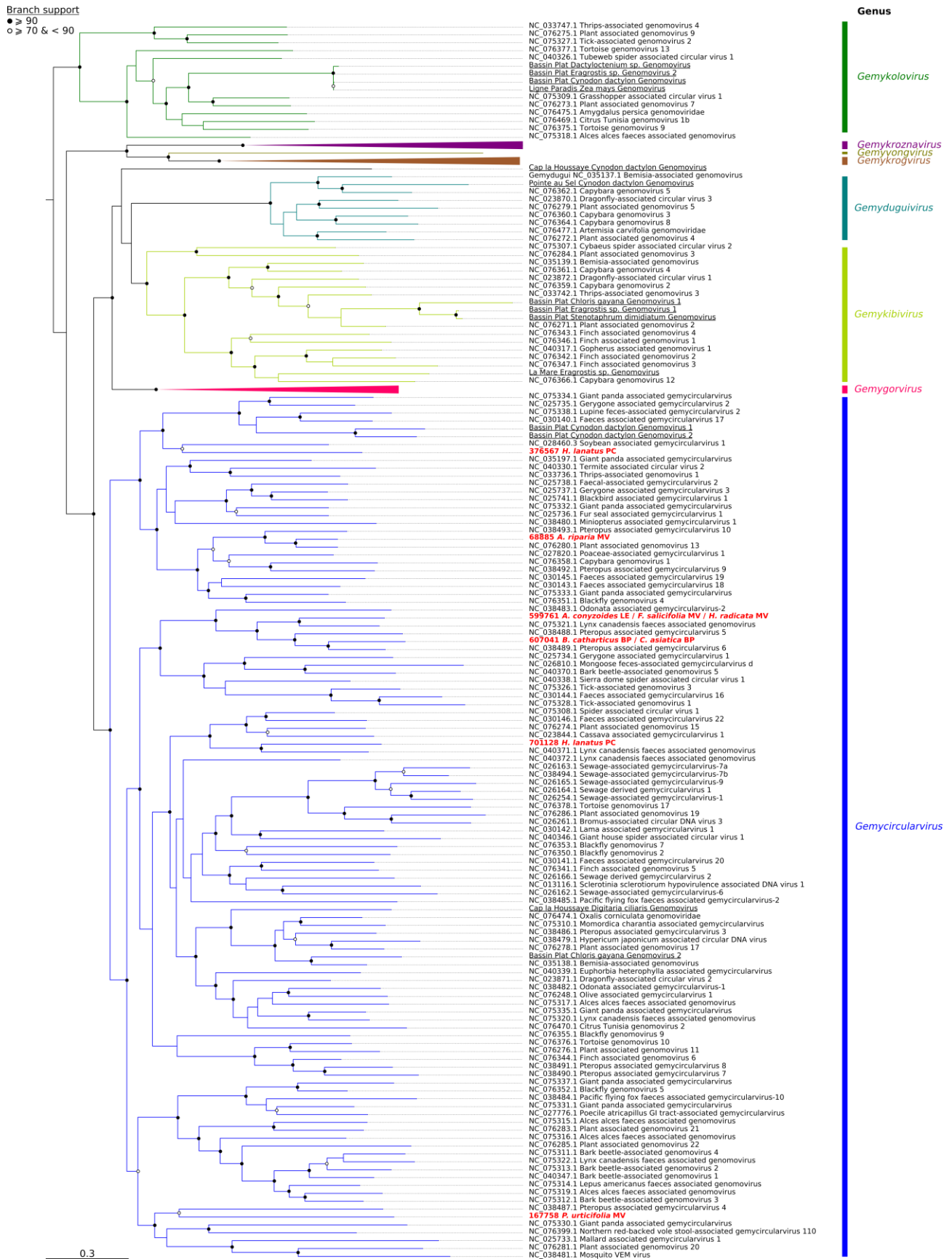
Supplementary Table 5. List of the viral species with the accession numbers of the nucleotide complete sequences, and the outgroups used for the phylogenetic analysis.

Viral Family	Viral Genus	Viral species	Accession Number
<i>Geminiviridae</i>	<i>Mastrevirus</i>	Maize streak virus	NC_001346.2
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Alces alces faeces associated genomovirus MP111	NC_075316.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Alces alces faeces associated genomovirus MP157	NC_075315.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Alces alces faeces associated genomovirus MP43	NC_075319.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Alces alces faeces associated genomovirus MP84	NC_075317.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bark beetle-associated genomovirus 1	NC_040347.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bark beetle-associated genomovirus 2	NC_075313.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bark beetle-associated genomovirus 3	NC_075312.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bark beetle-associated genomovirus 4	NC_075311.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bark beetle-associated genomovirus 5	NC_040370.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bemisia-associated genomovirus AdO	NC_035138.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackbird associated gemycircularvirus 1	NC_025741.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackfly genomovirus 2	NC_076350.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackfly genomovirus 4	NC_076351.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackfly genomovirus 5	NC_076352.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackfly genomovirus 7	NC_076353.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Blackfly genomovirus 9	NC_076355.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Bromus-associated circular DNA virus 3	NC_026261.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Capybara genomovirus 1	NC_076358.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Cassava associated gemycircularvirus 1	NC_023844.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Citrus Tunisia genomovirus 2	NC_076470.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Dragonfly-associated circular virus 2	NC_023871.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Euphorbia heterophylla associated gemycircularvirus	NC_040339.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faecal-associated gemycircularvirus 2	NC_025738.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 16	NC_030144.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 17	NC_030140.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 18	NC_030143.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 19	NC_030145.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 20	NC_030141.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Faeces associated gemycircularvirus 22	NC_030146.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Finch associated genomovirus 5	NC_076341.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Finch associated genomovirus 6	NC_076344.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Fur seal associated gemycircularvirus 1	NC_025736.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Gerygone associated gemycircularvirus 1	NC_025734.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Gerygone associated gemycircularvirus 2	NC_025735.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Gerygone associated gemycircularvirus 3	NC_025737.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant house spider associated circular virus 1	NC_040346.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075333.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_035197.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075332.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075334.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075335.1

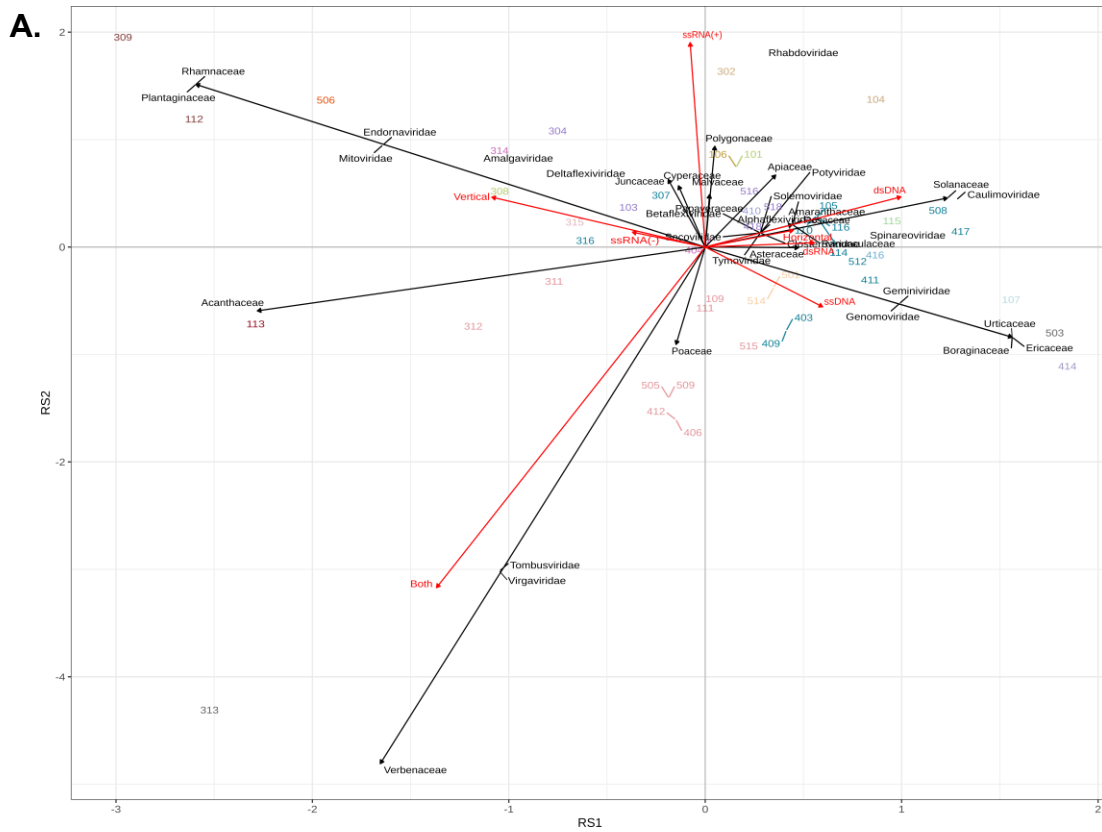
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075337.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075331.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Giant panda associated gemycircularvirus	NC_075330.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Hypericum japonicum associated circular DNA virus	NC_038479.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lama associated gemycircularvirus 1	NC_030142.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lepus americanus faeces associated genomovirus SHP9	NC_075314.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lupine feces-associated gemycircularvirus 2	NC_075338.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lynx canadensis faeces associated genomovirus CL1 148	NC_040371.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lynx canadensis faeces associated genomovirus CL1 46	NC_075322.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lynx canadensis faeces associated genomovirus CL1 48	NC_040372.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lynx canadensis faeces associated genomovirus CL1 71	NC_075321.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Lynx canadensis faeces associated genomovirus CL5 48	NC_075320.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Mallard associated gemycircularvirus 1	NC_025733.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Miniopterus associated gemycircularvirus 1	NC_038480.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Momordica charantia associated gemycircularvirus	NC_075310.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Mongoose feces-associated gemycircularvirus d	NC_026810.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Mosquito VEM virus SDBVL G	NC_038481.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Northern red-backed vole stool-associated gemycircularvirus 110	NC_076399.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Odonata associated gemycircularvirus-1	NC_038482.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Odonata associated gemycircularvirus-2	NC_038483.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Olive associated gemycircularvirus 1	NC_076248.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Oxalis corniculata genomoviridae	NC_076474.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pacific flying fox faeces associated gemycircularvirus-10	NC_038484.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pacific flying fox faeces associated gemycircularvirus-2	NC_038485.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 11	NC_076276.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 13	NC_076280.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 15	NC_076274.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 17	NC_076278.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 19	NC_076286.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 20	NC_076281.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 21	NC_076283.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Plant associated genomovirus 22	NC_076285.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Poaceae-associated gemycircularvirus 1	NC_027820.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Poecile atricapillus GI tract-associated gemycircularvirus	NC_027776.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 10	NC_038493.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 3	NC_038486.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 4	NC_038487.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 5	NC_038488.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 6	NC_038489.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 7	NC_038490.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 8	NC_038491.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Pteropus associated gemycircularvirus 9	NC_038492.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sclerotinia sclerotiorum hypovirulence associated DNA virus 1	NC_013116.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage derived gemycircularvirus 1	NC_026164.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage derived gemycircularvirus 2	NC_026166.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage-associated gemycircularvirus-1	NC_026254.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage-associated gemycircularvirus-6	NC_026162.1

<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage-associated gemycircularvirus-7a	NC_026163.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage-associated gemycircularvirus-7b	NC_038494.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sewage-associated gemycircularvirus-9	NC_026165.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Sierra dome spider associated circular virus 1	NC_040338.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Soybean associated gemycircularvirus 1	NC_028460.3
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Spider associated circular virus 1	NC_075308.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Termite associated circular virus 2	NC_040330.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Thrips-associated genomovirus 1	NC_033736.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Tick-associated genomovirus 1	NC_075328.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Tick-associated genomovirus 3	NC_075326.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Tortoise genomovirus 10	NC_076376.1
<i>Genomoviridae</i>	<i>Gemycircularvirus</i>	Tortoise genomovirus 17	NC_076378.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Artemisia carvifolia genomoviridae	NC_076477.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Bemisia-associated genomovirus AdDF	NC_035137.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Capybara genomovirus 3	NC_076360.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Capybara genomovirus 5	NC_076362.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Capybara genomovirus 8	NC_076364.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Dragonfly-associated circular virus 3	NC_023870.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Plant associated genomovirus 4	NC_076272.1
<i>Genomoviridae</i>	<i>Gemyduguivirus</i>	Plant associated genomovirus 5	NC_076279.1
<i>Genomoviridae</i>	<i>Gemygorvirus</i>	Capybara genomovirus 6	NC_076363.1
<i>Genomoviridae</i>	<i>Gemygorvirus</i>	Human genital-associated circular DNA virus-1	NC_026817.1
<i>Genomoviridae</i>	<i>Gemygorvirus</i>	Plant associated genomovirus 1	NC_076270.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Bemisia-associated genomovirus	NC_035139.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Capybara genomovirus 12	NC_076366.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Capybara genomovirus 2	NC_076359.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Capybara genomovirus 4	NC_076361.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Chicken genomovirus mg4 1218	NC_076434.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Chicken genomovirus mg7 73	NC_076436.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Chicken genomovirus mg8 401	NC_076438.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Cybaeus spider associated circular virus 2	NC_075307.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Dragonfly-associated circular virus 1	NC_023872.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Finch associated genomovirus 1	NC_076346.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Finch associated genomovirus 2	NC_076342.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Finch associated genomovirus 3	NC_076347.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Finch associated genomovirus 4	NC_076343.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Gopherus associated genomovirus 1	NC_040317.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Plant associated genomovirus 2	NC_076271.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Plant associated genomovirus 3	NC_076284.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Rhinolophus associated gemykibivirus 1	NC_038499.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Rhinolophus associated gemykibivirus 2	NC_038500.1
<i>Genomoviridae</i>	<i>Gemykibivirus</i>	Thrips-associated genomovirus 3	NC_033742.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Alces alces faeces associated genomovirus MP68	NC_075318.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Amygdalus persica genomoviridae	NC_076475.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Citrus Tunisia genomovirus 1b	NC_076469.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Grasshopper associated circular virus 1	NC_075309.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Plant associated genomovirus 7	NC_076273.1

<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Plant associated genomovirus 9	NC_076275.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Thrips-associated genomovirus 4	NC_033747.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Tick-associated genomovirus 2	NC_075327.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Tortoise genomovirus 13	NC_076377.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Tortoise genomovirus 9	NC_076375.1
<i>Genomoviridae</i>	<i>Gemykolovirus</i>	Tubeweb spider associated circular virus 1	NC_040326.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Apis mellifera genomovirus 2	NC_076287.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Caribou associated gemykrogvirus 1	NC_024909.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Chicken genomovirus mg4 1165	NC_076432.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Chicken genomovirus mg4 1173	NC_076433.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Chicken genomovirus mg4 1247	NC_076435.1
<i>Genomoviridae</i>	<i>Gemykrogvirus</i>	Chicken genomovirus mg7 78	NC_076437.1
<i>Genomoviridae</i>	<i>Gemykroznavirus</i>	Capybara genomovirus 10	NC_076365.1
<i>Genomoviridae</i>	<i>Gemykroznavirus</i>	Finch associated genomovirus 8	NC_076345.1
<i>Genomoviridae</i>	<i>Gemykroznavirus</i>	Plant associated genomovirus 23	NC_076277.1
<i>Genomoviridae</i>	<i>Gemykroznavirus</i>	Plant associated genomovirus 24	NC_076282.1
<i>Genomoviridae</i>	<i>Gemykroznavirus</i>	Zizania latifolia genomoviridae	NC_076476.1
<i>Genomoviridae</i>	<i>Gemyvongvirus</i>	Human associated gemyvongvirus 1	NC_028459.1



Supplementary Figure 1. Phylogenetic tree of the GenBank reference sequences of *Geminoviridae*, rooted on one *Geminiviridae* sequence. Sequences determined in this study are indicated in bold font. Open and closed circles at nodes indicate bootstrap support for the branches to their left of 70-89% and ≥90% respectively.



Supplementary Figure 2. Representation of the two principal axes of the RLQ analysis. **(A)** Correlation between plant families with viral family life history traits. **(B)** Correlation between plant sampling site, origin and life cycle with viral family life history traits.

Discussion générale

Discussion générale

Les études portant sur la diversité des phytovirus à l'échelle des écosystèmes ont été à l'origine d'avancées majeures dans le domaine de l'écologie virale (Roossinck and Garcia-Arenal, 2015 ; Sullivan *et al.*, 2017 ; O'Malley, 2016). En particulier, l'obtention de données concernant l'influence de la composition des communautés d'hôtes et de l'ensemble des interactions réalisées au sein des pathosystèmes étudiés a permis d'explorer de nouvelles pistes de compréhension de l'écologie des phytovirus à l'échelle des écosystèmes (Malmstrom and Alexander, 2016 ; Bernardo *et al.*, 2018 ; Susi *et al.*, 2019 ; Valverde *et al.*, 2020 ; McLeish *et al.*, 2022 ; Holmes, 2022).

Cependant, la réalisation d'une approche intégrative portant sur l'ensemble des facteurs pouvant influencer le fonctionnement des communautés de phytovirus à l'échelle de différents écosystèmes reste une entreprise complexe. C'est pourquoi, dans nos travaux, nous nous sommes focalisés dans un premier temps sur les virus à ADNsb, récemment étudiés à l'échelle d'un agro-écosystème de La Réunion (Claverie *et al.*, 2023). Ainsi, nous avons mis au point un protocole de métagénomique virale dédié à l'étude des molécules d'ADN circulaire de petite taille, tels que les espèces de la famille *Geminiviridae* et leurs molécules satellites, permettant d'obtenir des résultats de séquençage relativement fiables en peu de temps (**Chapitre 1**). Cette approche a été utilisée pour la détection de nouveaux membres du phylum *Cressdnaviricota*, mais aussi pour la description de la richesse et de la structure des communautés des espèces du genre *Mastrevirus* dans des pools de plantes de l'ordre des Poales, échantillonnés dans dix sites de trois types d'écosystèmes de La Réunion (**Chapitre 2**). Enfin, à partir de deux protocoles de métagénomique virale plus holistiques que celui précédemment utilisé, nous avons tenté d'approfondir nos connaissances sur les potentiels impacts des phytovirus sur la dynamique des plantes exotiques envahissantes à La Réunion, auxquelles les écosystèmes naturels de l'île sont fortement confrontés et susceptibles (**Chapitre 3**).

Les résultats prometteurs des TGS pour l'étude du phytovirome

Le génome des virus ne comportant pas de régions conservées spécifiques pouvant être utilisées comme marqueurs cibles (*e.g.* marqueurs ribosomiques), des méthodes spécifiques basées sur la présence de siRNA, d'ARN double brin ou encore sur la semi-purification de particules virales ont été développées (Kreuze *et al.*, 2009 ; François *et al.*, 2018 ; Marais *et al.*, 2018). Dans le cas particulier des virus présentant des génomes à ADN circulaire, les protocoles les plus utilisés sont ceux reposant sur l'amplification en cercle

roulant (RCA ; Inoue-Nagata *et al.*, 2004). Combinés à des méthodes de séquençage à haut-débit, l'utilisation de ces protocoles dit de circomics (Wyant *et al.*, 2012) ont ouvert de nouvelles voies d'études sur la diversité de ces virus. En combinant l'approche RCA à l'utilisation de la technologie nanopore sur la plateforme MinION, nous avons pu mettre au point une alternative crédible au séquençage Sanger pour la caractérisation de virus appartenant au phylum *Cressdnaviricota* (recombinants, défectifs, variants, molécules satellites) (**Chapitre 1 & 2**). L'application de ce protocole nous a permis non seulement de retrouver la diversité virale déjà connue et décrite à La Réunion mais aussi de mettre à jour de nouvelles espèces virales jamais décrites. Ce protocole, dont le coût de mise en oeuvre se rapproche sensiblement des méthodes les moins onéreuses de séquençage, permet la détection et le diagnostic de plusieurs espèces, voire de souches, de *Cressdnaviricota* en une seule analyse et ouvre la voie vers de plus amples utilisations.

Bien que l'accessibilité actuelle du séquençage nanopore soit encore limitée par la lourdeur des étapes qui précèdent (biologie moléculaire pour la préparation des librairies) et succèdent au séquençage (analyses bioinformatiques; Schalamun *et al.*, 2018), des protocoles simplifiés ont été développés permettant de passer de l'échantillon de plante à l'obtention des séquences virales directement sur le terrain (Boykin *et al.*, 2019). Si cette voie de développement est poursuivie, on peut s'attendre à une véritable démocratisation de cet outil pour des applications spécifiques, telles que le diagnostic de maladies végétales virales *in natura* en un temps réduit.

Toutefois, la qualité des séquences obtenues pour l'instant à partir de cette approche reste sous optimale (Boykin *et al.*, 2019), avec en particulier des erreurs de lectures retrouvées dans les assemblages sur les portions d'homopolymères des génomes (**Chapitre 1 - Article 1**). Cette limitation rend l'utilisation de la RCA-RFLP et du séquençage Sanger obligatoire pour l'obtention d'une séquence génomique précise. Les améliorations constantes des protocoles et les évolutions permanentes de la technologie de séquençage nanopore devraient toutefois permettre à court terme de contourner les difficultés actuelles associées au manque de précision des séquences obtenues. En effet, (1) le développement de nouvelles générations de cassettes et kits de séquençage plus fiables (Sun *et al.*, 2022; Cheng *et al.*, 2023), (2) la disponibilité d'algorithmes de traitement du signal (*basecalling*) plus performants (Samarkoon *et al.*, 2023), et surtout (3) le séquençage de longs multimères de génomes viraux obtenus après RCA (Filloux *et al.*, 2023) devraient permettre d'obtenir

des séquences de meilleure qualité avec *in fine* des portions d'homopolymères fidèles au génome d'origine. Ce type de protocole devra toutefois être testé et validé selon les mêmes méthodologies que celles que nous avons adoptées dans le **chapitre 1** de cette thèse.

Aujourd'hui, l'utilisation du séquençage nanopore avec le dispositif de paillasse MinION a dépassé le cadre de la recherche fondamentale pour des applications en diagnostic de maladies humaines directement en milieu clinique (Avershina *et al.*, 2022). Ce fut notamment le cas pour l'étude de l'épidémie récente du SARS-CoV-2 (Zhang *et al.*, 2021), l'exploration du microbiome intestinal complet (Matsuo *et al.*, 2021 ; Cha *et al.*, 2023) et l'épigénétique (Goldsmith *et al.*, 2021 ; Javaran *et al.*, 2021). Il est vraisemblable que les améliorations à venir permettront de banaliser son utilisation et que son adoption et vulgarisation permettront le développement et le déploiement de nombreux protocoles généralistes ou dédiés à des virus particuliers.

L'exploration de la diversité des African streak virus de La Réunion révèle la coexistence de trajectoires évolutives potentiellement opposées

L'exploration de la diversité des African streak virus (AfSVs) infectant les *Poaceae* collectées dans différents écosystèmes à La Réunion nous a permis d'élargir les inventaires précédemment réalisés (Krabberger *et al.*, 2017 ; Claverie *et al.*, 2023 ; **Chapitre 2**) avec la description de sept des huit espèces d'AfSV précédemment décrites à La Réunion, d'une espèce décrite jusqu'à lors uniquement sur le continent africain, mais aussi la caractérisation de trois nouvelles espèces de *Mastrevirus* inconnues.

Surtout, nos travaux ont permis de mettre en évidence deux génomes de virus aux trajectoires diamétralement opposées. Alors que le CPAV est un virus dont le génome présente une ossature d'alphasatellite qui a acquis et domestiqué des gènes *cp* et *mp* de mastrevirus, le DefMS est une forme défective d'un mastrevirus, encore inconnu sous sa forme complète, présentant la délétion d'un ORF codant pour la protéine associée à la réplication. La détection à plusieurs reprises de ces deux virus dans plusieurs échantillons respectivement de *Cenchrus purpureus* et de canne à sucre pour le CPAV et le DefMS, semble indiquer que ces virus sont circulants et font partie intégrante des populations virales associées aux poacées à La Réunion.

Ces deux découvertes virales soulignent la modularité des virus du phylum des *Cressdnaviricota*, pour lesquels de nombreuses organisations génomiques distinctes ont pu être recensées (Rosario *et al.*, 2012 ; Dayaram *et al.*, 2016). La propension de ces virus à évoluer par délétion ou recombinaison est intimement liée aux mécanismes de réplication, favorisant l'apparition de génomes défectifs et leur éventuel sauvetage par recombinaison (Michalakis and Blanc, 2020). La détection de cressdnavirus défectifs est en effet fréquente. Ainsi des génomes défectifs de cressdnavirus ont été détectés dans les échantillons de luzernes infectés par les souches de *Capulavirus* (**Chapitre 1 – Article 1**), dans la canne à sucre avec le DefMS (**Chapitre 2 - Article 3**) ou encore pour des génomovirus issus de plantes collectées le long du gradient altitudinal (**Chapitre 3 – Article 4**). Si ces génomes défectifs présentent des régions intergéniques compatibles (itérons et site d'initiation à la réplication) avec la protéine Rep (Stanley and Townsend, 1985), ils pourraient participer à l'infection. Alors que les mécanismes aboutissant à la production des virus défectifs sont bien identifiés, le rôle de ces derniers reste sujet à débats. Au-delà de faciliter l'évolution par recombinaison (homologue ou hétérologue), la présence de défectifs pourrait permettre un autre niveau de régulation de la transcription (Bonnamy *et al.*, 2023) et même impacter la dynamique de leurs virus-assistants en interférant dans les processus de réplication de ces derniers (Patil and Dasgupta, 2006).

De manière intéressante, l'évolution en catégorie discrète de la taille des génomes souligne un autre aspect de la biologie de ces virus. En effet, alors qu'un composant complet présente une taille d'environ 2.7kb, les satellites présentent des tailles d'environ 1.3kb. Les travaux de Frischmuth *et al.* (2001) ont mis en évidence que la taille du génome chez les géminivirus avait un impact sur la taille et la géométrie de la capsid. Si les génomes viraux complets d'environ 2.7kb sont encapsidés dans des particules icosaédriques associées en doublet, les satellites le seraient, par trans-encapsidation avec le virus assistant, dans une particule icosaédrique simple (Frischmuth *et al.*, 2001 ; Saunders *et al.*, 2020). Les contraintes associées à la géométrie des particules et les nécessaires interactions ADN-CP semblent contraindre la taille des molécules d'acides nucléiques pouvant former des virions stables et viables (Hesketh *et al.*, 2018). Au regard des tailles de génomes observées pour le CPAV (2734 nt) et le DefMS (1505 nt), il sera intéressant de confirmer la géométrie des virions par microscopie électronique à transmission et la relation entre la taille des molécules à ADNs_b et la géométrie des particules virales.

Association de molécules à ADNsb : Avantage évolutif du multipartisme

La description de la coexistence d'une importante diversité de molécules à ADNsb circulaire dans une plante suggère que ces associations et plus largement la stratégie du multipartisme présentent des avantages. Les approches théoriques ont notamment prédit que les avantages de tels systèmes biologiques seraient liés à une plus grande stabilité des particules virales, une réplication plus rapide ou encore une recombinaison accrue pour cet ensemble de segments génomiques (Sicard *et al.*, 2019). Par conséquent, la stratégie du multipartisme pourrait être essentielle pour relâcher la contrainte des limites de taille des génomes encapsidés en offrant de nouvelles possibilités d'évolution et d'adaptation.

Dans ce sens, les exemples d'interactions entre composants génomiques, leur influence sur la pathogénicité, et leur rôle dans le contournement des mécanismes de résistance de l'hôte sont multiples (Gnanasekaran and Chakraborty, 2018 ; Lucia-Sanz *et al.*, 2022). En effet, les interactions entre molécules de familles différentes peuvent avoir des répercussions sur la pathogénicité des virus, allant de la synergie à l'antagonisme (Power, 2008 ; Syller, 2012 ; Mascia and Gallitelli, 2016). Par exemple, l'association non stricte entre les ADN-A et -B du pepper yellow vein Mali virus (PepYVMLV, genre *Begomovirus* ; Ouattara *et al.*, 2022) sont synergiques aboutissant à une augmentation de la virulence, de la charge virale et de la capacité de transmission. Dans le cas des molécules satellites, différentes situations ont été caractérisées. Alors que les deltasatellites semblent réduire l'accumulation des virus assistants et atténuer les symptômes dus à l'infection, les betasatellites semblent induire une augmentation de l'abondance de leur virus assistants et des symptômes induits (Gnanasekaran and Chakraborty, 2018). Les alphasatellites peuvent soit induire une sévérité accrue des symptômes associés à leurs virus assistants, comme par exemple lors de co-infection avec des begomovirus (Paprotko *et al.*, 2010 ; Mar *et al.*, 2017 ; Nogueira *et al.*, 2021) ou mastrévirus (Kumar *et al.*, 2014), ou soit entraîner une réduction des symptômes (Bridson *et al.*, 2004 ; Mar *et al.*, 2017).

Dans le cas du DefMS, des ORFs complets codants potentiellement pour des protéines fonctionnelles de capsid et de mouvement de mastrévirus ont été caractérisés. L'absence d'ORF codant pour une protéine associée à la réplication nécessite que cette molécule défective s'associe avec un virus assistant compatible. La co-détection du DefMS et du SWSV dans un pool de canne à sucre suggère que le SWSV peut assister la réplication du DefMS. Comprendre les avantages et les inconvénients d'une telle association, stricte pour la molécule défective et vraisemblablement non stricte pour le virus assistant, nécessite des investigations supplémentaires: premièrement sur le terrain pour évaluer la prévalence de

cette association et le gain de *fitness* qu'elle apporte au virus assistant, et deuxièmement en conditions contrôlées, pour évaluer les traits biologiques (virulence, abondance et capacité de transmission) des deux partenaires en infections simples et mixtes.

Concernant le CPAV, nos investigations moléculaires suggèrent que nous sommes en présence d'un alphasatellite qui grâce à l'acquisition et la domestication d'un gène Rep de mastrévirus a réussi à s'extraire de la dépendance du multipartisme et de l'association avec un virus assistant. En effet, nos travaux en conditions contrôlées avec l'utilisation de clones agroinfectieux ont permis la détection sur les feuilles systémiques de *C. purpureus* de l'ADN viral du CPAV en inoculation simple. Ces résultats suggèrent que le CPAV est autonome dans sa capacité à se répliquer et à contaminer les étages foliaires de la plante agroinoculée. Toutefois, l'observation de la diminution de la charge virale du CPAV, en dessous du seuil de détection de la PCR, après plusieurs semaines et en cas d'infection simple, suggère que l'assistance du CPMSV reste nécessaire pour maintenir une charge virale du CPAV détectable. A ce stade de nos investigations, nous ne sommes pas en mesure de dire si au bout de quelques semaines les plantes de *C. purpureus* ont réussi à juguler définitivement la contamination virale ou à maintenir la charge virale en dessous d'un certain seuil. Parmi les mécanismes de résistance potentiellement incriminés, on peut citer le mécanisme d'extinction de l'ARN viral (ARN interférence ou *silencing*) (Moon and Park, 2016). L'activation de ce mécanisme de résistance conduit à la dégradation spécifique des ARN messagers viraux. Ceci a notamment été montré pour les bégomovirus du manioc qui ont la capacité d'induire le silencing post-transcriptionnel (PTGS) chez les plantes infectées et la production d'ARN interférents (ARNi) (Chellappan *et al.*, 2004). L'activation du PTGS et l'accumulation des ARNi a notamment été corrélé avec la réduction de la virulence et des charges virales. L'évaluation de la dynamique de production des ARNi au cours de l'infection virale permettrait d'étudier l'hypothèse de l'activation du PTGS par le CPAV et/ou le CPMSV en cas d'infection simple ou mixte et le rôle potentiel d'inactivation du PTGS par le CPMSV. Par ailleurs, l'évaluation précise de l'intensité des symptômes de stries et de chloroses le long des nervures lors d'expérimentations en conditions contrôlées avec les clones agroinfectieux permettra d'évaluer la synergie entre ces deux virus et les effets sur la pathogénicité.

Relations complexes entre la richesse de plantes hôtes et la diversité virale : cas des poacées et des AfSVs

Notre étude sur la diversité des espèces d'AfSVs et les réseaux d'interactions plantes hôtes-virus (**Chapitre 2 - Article 3**) a permis de cartographier la richesse spécifique des hôtes et de leurs virus et d'identifier les principaux facteurs associés à leurs variations.

Tout d'abord, l'analyse de la relation entre la richesse des hôtes et des virus suggère une corrélation positive. Compatible avec l'hypothèse d'effet d'amplification (Keesing *et al.*, 2006), ce résultat était attendu dans le cas de l'existence d'une association forte entre les espèces végétales associées aux milieux cultivés (cultures et adventices) et les AfSVs, mais aussi au regard des spécificités d'hôtes propres à chacun de ces virus. En effet, plus on analyse de plantes hôtes différentes, plus on a de chance de trouver chacun des virus pouvant les infecter. A l'inverse, le cas particulier du site LE où aucun AfSVs n'a été détecté pourrait être expliqué par le phénomène de dilution (Keesing *et al.*, 2006) et la faible présence d'hôtes (trois espèces d'hôtes potentiels) dans l'ensemble de la flore présente sur ce site. Ainsi, si un effet d'amplification est décrit pour les AfSVs, il est uniquement associé aux hôtes potentiels de ces virus, et non à l'ensemble de la flore comme le propose la théorie (Keesing *et al.*, 2006). En effet, notre résultat est basé sur la richesse spécifique simple sans la prise compte de l'ensemble de la flore. Par conséquent, il n'est pas directement comparable à d'autres études démontrant l'effet d'amplification (Keesing *et al.*, 2006 ; Hechinger and Lafferty, 2005 ; Lafferty, 2012 ; Kamiya *et al.*, 2014 ; Halliday *et al.*, 2020). Afin de mieux comprendre comment la richesse en hôtes influence la fréquence des infections à AfSVs, il nous semble primordial de s'intéresser à la fréquence des infections en fonction de la composition floristique globale des milieux.

L'exploitation des connaissances du réseau d'interaction plante/virus et notamment la comparaison de la richesse de virus réalisée (*i.e.* nombre d'espèces virales détectées) par rapport à la richesse de virus réalisable (*i.e.* nombre maximal d'espèces virales attendu au regard des espèces d'hôtes présentes) a permis des observations originales. En effet, la prise en compte de l'ensemble des espèces virales pouvant infecter les hôtes a révélé que les proportions de richesses virales réalisées pour chacun des sites ne sont pas constantes mais positivement corrélées aux richesses en hôtes. En d'autres termes, la proportion d'AfSVs détectée est plus élevée pour des communautés riches en hôtes en comparaison de communautés pauvres en hôtes. Cette relation, nommée "les riches seront plus riches" ("*rich get richer*" ; Edward and Rohwer, 2005), a été modélisée dans le cadre d'études théoriques portant sur la diversité virale. Ainsi, dans le cas d'une communauté riche en

hôtes, le résultat de l'infection par des pathogènes ne peut se solder que par une amplification du nombre de pathogènes infectant ces hôtes. Dans cette modélisation, Edward et Rohwer (2005) propose qu'une loi de puissance (forte richesse en hôte) engendre forcément une autre loi de puissance (forte richesse en pathogènes). La forte richesse d'hôtes de mastrevirus dans les agro-écosystèmes pourrait ainsi faciliter la forte richesse des AfSVs.

Un autre mécanisme, cette fois-ci associé à la synergie inter-spécifique (*i.e.* la réalisation d'une gamme d'hôtes plus large en co-infection qu'en infection simple), pourrait faciliter l'observation d'une plus forte richesse virale associée aux fortes richesses d'hôtes par la co-occurrence de plusieurs espèces virales sur un même site. Cependant, aucune corrélation significative n'a été détectée entre la proportion d'interactions réalisées (*i.e.* non plus le nombre de virus, mais la somme des infections réalisées entre virus et plante divisé par le total des interactions possibles) et la richesse virale ou la richesse en hôtes. Par conséquent, toutes les relations réalisables ne sont pas réalisées et l'absence de certaines interactions semble aller à l'encontre de l'hypothèse d'une synergie virale à l'origine de nos résultats d'amplification des interactions hôtes-virus.

Malgré tous ces résultats, il demeure complexe d'établir des règles générales capables d'expliquer la relation entre la composition (en termes de richesse) des communautés végétales et virales. En effet, les relations identifiées peuvent être influencées ou restreintes par un ensemble de facteurs tels que les échelles spatio-temporelles choisies pour l'étude (McLeish *et al.*, 2017 ; Martinez, 2018 ; Valverde *et al.*, 2020), ou encore les voies de transmission.

L'importance du vecteur dans l'écologie virale : la grande inconnue du triptyque plante/virus/vecteur

L'ensemble de nos études, focalisées sur les virus et leurs hôtes, n'ont pas pris en compte l'impact des vecteurs sur la richesse observée des virus et des hôtes et les réseaux d'interactions réalisés. Or, la gamme d'hôte d'un phytovirus transmis par insecte vecteur ne pourra jamais être plus large que celle de ce dernier (Power and Flecker, 2003).

Les mastrévirus sont décrits comme des virus transmis horizontalement par des cicadelles de la famille des Cicadellidae. A ce jour, aucune transmission des mastrévirus par la graine n'a été documentée. Par conséquent, la distribution et la gamme de plantes hôtes des

mastrévirus semblent strictement limitées par celles de leurs insectes vecteurs (Power and Flecker, 2003). Cependant, que cela concerne la répartition des vecteurs, leurs préférences d'hôtes (McLeish *et al.*, 2018) ou les éventuelles manipulations qu'ils subissent de la part des virus (Wang *et al.*, 2020), la transmission reste la plus grande inconnue dans le triptyque plante/virus/vecteur. A ce jour, seules quelques études soulignent des préférences alimentaires de *Cicadulina mbila* (van Rensburg, 1992 ; Reynaud *et al.*, 1988 ; Mesfin *et al.*, 1995) ou la manipulation de l'insecte par le MSV (Oluwafemi and Jackai, 2013).

Nos analyses concernant les interactions hôte-virus suggèrent que la richesse virale est plus importante dans les milieux les plus riches en hôtes (**Chapitre 2 - Article 3**). Ce résultat majeur pourrait être influencé par la dynamique des populations d'insectes vecteurs. En effet, le site BP (agro-écosystème) semble présenter une population de *Cicadulina mbila* plus abondante que le site TR (prairie littorale) où une richesse virale moindre a été décrite. Dans le cas particulier du site LE où aucun AfSVs n'a été détecté, un nombre très important d'individus de *C. mbila* a été collecté. D'après Rodelo-Urrego *et al.* (2013), la distribution spatiale des virus devrait refléter les mouvements écologiques des vecteurs. Or, *C. mbila* est décrite comme le principal vecteur des AfSVs à La Réunion. Pour expliquer l'absence des AfSVs sur le site LE, on peut proposer trois scénarios explicatifs : (i) soit les AfSVs ne sont pas présents sur le site, ce qui peut être dû à l'existence d'une barrière écologique empêchant l'arrivée de vecteurs virulifères sur le site (cependant, il faut noter que le MSV a été décrit à La Réunion sur des plants de maïs cultivés à plus de 1000 m d'altitude (Reynaud, 1988)); (ii) soit les plantes sur lesquelles les insectes se nourrissent préférentiellement à LE ne sont pas des hôtes d'AfSVs; (iii) soit les populations vectrices présentent des cycles de vie saisonniers entraînant une réduction de la dispersion des phytovirus à certains moments de l'année. Par ailleurs, les populations de *C. mbila* pourraient présenter des différences génétiques et biologiques avec des préférences alimentaires contrastées et des capacités de transmission variables selon les virus, les hôtes et les saisons. Le pathosystème largement documenté à ce sujet est celui des bégomovirus et de leur vecteur, l'aleurode *Bemisia tabaci* (Hogenhout *et al.*, 2008). *B. tabaci* est un complexe d'espèces, anciennement nommés biotypes, qui présentent des espèces invasives (anciennement biotypes B et Q) et particulièrement polyphages, pouvant s'alimenter sur plus de 1000 espèces végétales différentes (Fiallo-Olivé and Navas-Castillo, 2023). Capables de transmettre jusqu'à 500 espèces virales distinctes (Navas-Castillo *et al.*, 2011), elles ont été associées à l'émergence d'un grand nombre de bégomovirus (TYLCV), mais aussi de crinivirus (beet yellow virus ; Wisler *et al.*, 1998) et ipomovirus

(cucumber vein yellowing virus ; Adkins *et al.*, 2010), impactant lourdement les productions agricoles dans les régions tropicales, subtropicales et tempérées. Leur diffusion à l'échelle mondiale a été causée par les échanges commerciaux et facilitée par leur caractère polyphage. Malgré la préférence alimentaire des espèces invasives de *B. tabaci* pour les espèces de plantes cultivées, leur polyphagie permet à ces aleurodes invasives de s'adapter et de s'alimenter sur les plantes non cultivées pendant les périodes d'inter-saisons culturaux (Navas-Castillo *et al.*, 2011 ; de Barro *et al.*, 2011). D'autre part, dans le cas du TYLCV, des manipulations directes et indirectes du vecteur ont été constatées (Whitfield and Rotenberg, 2015). En effet, le virus semble impacter directement le choix d'alimentation des aleurodes virulifères vers des plantes saines. De plus, il a été observé que le temps d'alimentation des individus infectés est plus long que celui des individus sains (Liu *et al.*, 2013 ; Wang *et al.*, 2020). En parallèle, les plantes hôtes infectées par le TYLCV présentent une capacité d'attraction plus importante pour les aleurodes non virulifères favorisant ainsi l'acquisition virale (Liu *et al.*, 2013). En conséquence, l'ensemble de ces mécanismes permettent au TYLCV de faciliter sa transmission horizontale et d'accroître sa diffusion.

Au vu de l'importance de la transmission dans les réseaux d'interactions hôte-virus et en complément de notre étude de la diversité des mastrévirus à l'échelle de divers écosystèmes, il nous apparaît essentiel d'élargir notre travail en intégrant les insectes vecteurs. Premièrement, les insectes phytophages ou entomophages collectés sur un site ou dans un écosystème particulier peuvent être utilisés comme "collecteur" de la diversité virale présente. Ainsi des études de métagénomiques, appelées VEM pour *vector enabled metagenomics* (métagénomique par l'intermédiaire du vecteur) ont déjà prouvé leur efficacité pour identifier la présence de bégomovirus (Ng *et al.*, 2011; Rosario *et al.*, 2015) et de mastrévirus (Fontenele *et al.*, 2018) dans un agro-écosystème à partir des insectes vecteurs connus comme l'aleurode *B. tabaci* ou la cicadelle *Dalbulus maidis* respectivement. De plus, l'utilisation de ce protocole VEM sur des insectes prédateurs en l'occurrence des libellules a permis de détecter la présence d'une nouvelle espèce de mastrévirus associée à un alphasatellite dans un agro-écosystème à Puerto Rico (Rosario *et al.*, 2013). Deuxièmement, cette stratégie indirecte basée sur les insectes vecteurs connus collecté sur un site devrait permettre d'évaluer la fréquence des insectes virulifères et potentiellement infectieux *in natura* et *in fine* la probabilité pour un insecte vecteur de croiser la route d'une plante virosée que ce soit dans un agro-écosystème ou des écosystèmes non cultivées. La réalisation d'une telle étude sur plusieurs années permettrait de déterminer la dynamique d'accès aux plantes virosées par les insectes vecteurs au cours des saisons et le rôle

potentiel de ces vecteurs sur la persistance des communautés d'AfSVs dans les écosystèmes. Troisièmement, la capacité et l'efficacité de transmission des AfSVs par *C. mbila* pourra être évaluée en conditions contrôlées non seulement en infection simple mais également en infections multiples. En effet, la comparaison des modalités d'infection simple ou multiple permettra d'évaluer finement les spécificités d'hôtes (gammes d'hôtes des mastrévirus décrits à La Réunion) et de vection (efficacité de transmission par vecteur) et les conséquences des interactions inter-spécifiques entre virus sur ces paramètres biologiques (existence de synergie inter-spécifique). Dans ce cadre là, les études récentes de Armand *et al.* (2023) ont montré l'existence d'interactions asymétriques lors de la co-infection du BYDV (famille *Luteoviridae*) et du WDV (famille *Geminiviridae*) sur le taux d'infection, l'accumulation virale et l'efficacité de transmission par vecteur. Ces travaux permettront également d'aborder les questions concernant la qualité d'hôte et de source virale pour un insecte vecteur et l'influence de leurs variations sur la dynamique écologique des virus. L'ensemble de ces investigations nécessitera cependant de contourner nos difficultés expérimentales actuelles concernant l'utilisation de clones agroinfectieux et notamment les faibles taux d'inoculation.

Richesse et structure des communautés virales associées aux milieux agricoles : sur la piste de l'origine des AfSVs à La Réunion

Nos résultats relatifs à la distribution des AfSVs à La Réunion soulignent l'absence d'une communauté locale de ces virus en dehors des agro-écosystèmes et ayant coévolué avec les Poales natives. Nos observations suggèrent plutôt que les agro-écosystèmes offrent les conditions environnementales les plus favorables à la survie et à la propagation de la communauté formée par les AfSV, leurs hôtes et leurs vecteurs. Par ailleurs, nos résultats semblent indiquer que les interactions hôtes-virus détectées à l'échelle des agro-écosystèmes ne résultent pas d'une communauté stable. En effet, la structure des réseaux d'interactions observée présente une configuration emboîtée (*nestedness*) et non modulaire. Si ces caractéristiques d'interactions restent toujours difficiles à interpréter, l'absence de modularité semble indiquer une instabilité du réseau (Grilli *et al.*, 2016 ; Dormann *et al.*, 2017 ; Valverde *et al.*, 2020). En effet, la majorité des interactions réalisées concerne des espèces annuelles sauvages ou cultivées dont la disparition saisonnière pourrait entraîner une instabilité du réseau d'interaction hôte-virus. Cette instabilité du réseau semble correspondre à un marqueur typique des agro-écosystèmes, qui contrairement aux prairies non gérées des zones littorales et d'altitudes, sont fortement soumis aux perturbations anthropiques (espèces exotiques introduites, récoltes régulières,

dés herbages, etc.). D'ailleurs, il est intéressant de noter que pour le site BP, les travaux de Claverie *et al.*, (2023) entre 2014 et 2019 n'avaient pas révélé la présence des CEAV, SAAV-B et USV (**Chapitre 2 - Article 3**). Les analyses réalisées *a posteriori* sur les échantillons collectés lors de cette étude ont permis de confirmer l'absence de ces virus pendant cette période (2014-2019). La description nouvelle de ces virus sur le site BP à partir de 2020 renforce l'hypothèse émise précédemment sur l'instabilité des populations de mastrevirus présentes dans les agro-écosystèmes. Il se pose alors la question de l'origine et de l'historique d'introduction de chacune des espèces virales sur leurs sites de détection.

A ce titre, il est intéressant de noter que la comparaison des gammes d'hôtes réalisées par les AfSVs sur le continent africain et dans les îles du SOOI révèle une large concordance. Cette similarité de la gamme d'hôte suggère l'absence de phénomène majeur de saut d'hôte des AfSVs entre les poacées hôtes décrites sur ces îles et celles décrites sur le continent. Elle suggère également la reconstitution, introduction après introduction, des communautés virales présentant des interactions plante/virus similaires à celles de leurs aires d'origine.

L'étude de la diversité génétique des génomes, en utilisant des modèles d'inférences phylogéographiques dédiées (Kühnert *et al.*, 2011 ; Picard *et al.*, 2017 ; Dellicour *et al.*, 2021), a permis d'identifier les sources (Pande *et al.*, 2017) ou les voies de diffusion des virus à l'échelle mondiale (Talbi *et al.*, 2010). Dans le cas des géminivirus, ces avancées ont permis de mieux comprendre l'histoire évolutive et les principales routes de dissémination du TYLCV (Lefeuvre *et al.*, 2019) et du MSV (Monjane *et al.*, 2011) par exemple et d'établir les aires d'origine de ces virus introduits à La Réunion. Les connaissances phylogéographiques pourraient nous renseigner sur les migrations réalisées par les espèces d'AfSVs depuis le continent africain vers les îles du SOOI, mais également vers les autres régions du monde (*e.g.* la détection du MSRV en Chine ; Chen *et al.*, 2015). Pour des raisons de disponibilités de génomes complets et de puissance d'analyse, il apparaît judicieux d'axer dans un premier temps ces études sur les espèces les plus répandues d'AfSVs tels que le MSV pour lequel onze souches ont été décrites infectant plus de 80 espèces de *Poaceae* et présentant une distribution géographique couvrant 13 pays africains et les îles du SOOI et le PanSV présentant dix souches différentes détectées dans sept pays d'Afrique (Varsani *et al.*, 2009 ; Marwal *et al.*, 2019 ; Emeraghi *et al.*, 2021). Ces analyses phylogéographiques permettront d'actualiser les études précédentes (Monjane *et al.*, 2011) et peut-être d'étendre nos connaissances sur les voies de diffusion proposées pour ces virus.

Espèces végétales invasives à La Réunion : quel rôle pour les phytovirus ?

Les échanges de matériel végétal dans le cadre des activités agricoles, dont l'intention première est l'introduction de plantes cultivées, se sont soldés et continuent de se solder par l'introduction d'un cortège non désiré de plantes adventices des cultures, d'insectes ravageurs et, comme dans le cas des mastrévirus, de microorganismes phytopathogènes (Wilson *et al.*, 2016 ; Pyšek *et al.*, 2017). Les agro-écosystèmes sont souvent considérés comme la porte d'entrée des espèces exotiques (Catford *et al.*, 2011; O'Donnell *et al.*, 2012 ; Shrestha and Shrestha, 2019). Si ces populations exotiques se maintiennent sans intervention humaine, se dispersent et entraînent des répercussions sur la biodiversité, la santé, l'agriculture ou le fonctionnement des écosystèmes, on parle alors d'invasion (Simberloff *et al.*, 2013).

Les pathogènes, et dans notre cas les phytovirus, peuvent impacter les dynamiques invasives de leurs hôtes exotiques selon quatre modèles proposés par Faillace *et al.* (2017). Les deux modèles les plus décrits dans le cas des phytovirus sont le *pathogen spillover*, mécanisme par lequel la transmission d'un phytovirus depuis une espèce hôte primaire vers une espèce hôte secondaire peut réduire les populations d'hôtes secondaires au profit des hôtes primaires (Power and Mitchell, 2004) ; et le *pathogen release*, décrivant l'introduction d'espèces hôtes exotiques sans leurs agents pathogènes, leur permettant de proliférer et d'envahir les niches des espèces hôtes indigènes toujours contraintes par leurs pathogènes (Mitchell and Power, 2003).

Face à la distribution des mastrévirus principalement dans les agro-écosystèmes et au vu de l'importance de l'introduction de plantes exotiques dans l'historique de ces milieux, nous avons tenté de déterminer si le statut d'indigénat des plantes pouvait avoir un impact sur les patrons de diversité des autres familles de phytovirus pouvant être détectées par des méthodes génériques de *viromics* (*i.e.* métagénomique virale; Roux *et al.*, 2017). Notre analyse de la richesse phytovirale révèle qu'elle est plus importante pour les hôtes indigènes que pour les plantes exotiques envahissantes (PEEs). Ce résultat suggère que les PEEs ont potentiellement été introduites sans leur cortège d'agents pathogènes viraux, compatible avec le modèle du *pathogen release* favorable à la dynamique invasive des PEEs. L'introduction des PEEs à La Réunion par l'intermédiaire de graines, adossées ou non à l'absence des principaux insectes vecteurs, représente le principal élément explicatif à cette perte de virus à transmission horizontale. Par ailleurs, nous avons observé dans nos études une richesse phytovirale plus importante pour les espèces de plantes pérennes que

pour les espèces annuelles. Ce résultat semble souligner le rôle majeur des plantes pérennes comme réservoirs viraux, permettant aux phytovirus de persister d'une saison à l'autre dans les écosystèmes, contrairement aux plantes annuelles qui disparaissent régulièrement de la communauté végétale .

Parmi les autres espèces virales décrites, on constate plusieurs phytovirus généralement associés aux maladies des cultures. Ainsi, le cotton leafroll dwarf virus (CLDV ; famille des *Solemoviridae* ; Edula *et al.*, 2023), le potato virus Y et le turnip mosaic virus (PVY ; Torrance and Talianksy, 2020 ; TMV ; Gibbs *et al.*, 2015 ; famille des *Potyviridae*) ou encore le lettuce chlorosis virus (LCV ; famille des *Closteroviridae* ; Garcian and Janssen, 2020) ont tous été trouvés en association avec des PEEs annuelles échantillonnées sur le site agricole de BP. Ces phytovirus, n'ayant pas été détectés sur les espèces herbacées indigènes, ont potentiellement été introduits avec des adventices des cultures dans les agro-écosystèmes réunionnais auxquelles ils sont restés inféodés. De plus, ces quatre espèces sont transmises horizontalement par des pucerons (famille Aphididae), dont plusieurs espèces ont été identifiées après l'échantillonnage d'hémiptères réalisé à BP (**Chapitre 2 - Article 3**). Ainsi, à l'instar des AfSVs, ces quatre espèces de phytovirus sont associées à des PEEs annuelles distribuées en milieu agricole, et dont les vecteurs sont potentiellement présents en nombre. Par conséquent, le cas des AfSVs et leurs associations étroites avec les agro-écosystèmes ne semble pas être une exception et suggère que d'autres phytovirus ont suivi la même voie. Cependant, nos investigations ne permettent pas d'apporter d'éléments de réponse à la question de l'influence potentielle de ces virus sur les dynamiques de leurs hôtes et notamment les PEEs.

La détection d'un autre phytovirus dans le contexte des prairies altimontaines et son rôle potentiel dans le succès d'invasion de sa plante hôte nous a ouvert de nouvelles perspectives. En effet, dans cette étude nous avons détecté le BYDV (*Luteovirus pashordei*, famille *Tombusviridae*) sur les deux pools de flouve odorante (*Anthoxanthum odoratum*, famille des *Poaceae*) collectés dans les sites altimontains MV et PC. D'une part, le BYDV est responsable d'une des plus importantes viroses des céréales en milieu tempéré (Banks *et al.*, 1995 ; McKirdy *et al.*, 2002) et d'autre part la flouve odorante représente une des espèces de *Poaceae* les plus invasives à La Réunion (McDonald *et al.*, 1991). L'absence du BYDV sur les autres poacées présentes dans les prairies altimontaines suggère l'absence de saut d'hôte (*spillover*) du BYDV de la flouve odorante vers d'autres poacées. La possibilité que le BYDV confère à la flouve odorante un avantage sélectif par rapport aux

autres espèces de poacées pourrait alors être envisagée mais reste à être explorée (Power and Mitchell, 2004). En outre, la spécificité d'habitat de la flouve odorante aux prairies altimontaines (Fenuillas et al., 2021) et l'absence de détection de ce virus dans les zones agricoles suggèrent que ce virus a emprunté une autre voie d'introduction à La Réunion que celle décrite pour les autres virus associés aux agro-écosystèmes. Ainsi, la possibilité d'une introduction volontaire d'une poaceae originaire du continent européen sur l'île de La Réunion afin de servir de plantes fourragères dans la seconde moitié du XXème siècle a été proposée pour la flouve odorante (Cadet, 1974 ; McDonald *et al.*, 1991).

Concernant le BYDV, une étude récente a détecté la présence de ce virus sur les îles Kerguelen, sur trois espèces de *Poaceae* indigènes et deux espèces de PEEs (Svanella-Dumas *et al.*, 2013). La présence de plusieurs souches, la distribution géographique sur l'ensemble de l'archipel, la large gamme de poacées infectées, et la prévalence plus forte décrite pour les espèces indigènes que les PEEs sont des éléments indicatifs de l'installation du BYDV sur l'archipel des Kerguelen depuis plusieurs décennies. Une situation qui contraste avec celle du BYDV et de la flouve odorante à La Réunion et qui renforce le questionnement sur le rôle que pourrait jouer le BYDV dans la dynamique invasive de cette PEE.

Dans l'ensemble, cette étude de la richesse et de la structure des communautés virales et des facteurs pouvant les influencer révèlent l'absence d'une tendance unique chez les phytovirus. Nos travaux soulignent en effet l'influence globale des richesses en hôte et virus dans la structuration des communautés. De plus, si dans certaines conditions, ces virus sont des agents pathogènes aux conséquences négatives voire létales, dans d'autres conditions, ils semblent représenter un atout pour la survie et la prolifération de leur hôte. Les interactions réalisées par les virus avec leurs hôtes ne peuvent être considérées que dans le contexte global des environnements biotiques et abiotiques dans lesquels elles se réalisent. Face à la complexité de tels systèmes biologiques et au vu des enjeux, il apparaît essentiel d'approfondir notre compréhension des patrons de distribution et de diversité des phytovirus pour pouvoir anticiper les impacts des changements globaux imposés par l'agriculture et l'expansion continue des activités anthropiques sur les milieux naturels.

Références

Références

- Adams, M. J., Antoniw, J. F., & Kreuze, J. (2009). Virgaviridae: a new family of rod-shaped plant viruses. *Archives of Virology*, *154*, 1967-1972.
- Adams, I. P., Braidwood, L. A., Stomeo, F., Phiri, N., Uwumukiza, B., Feyissa, B., Mahuku, G., Wangi, A., Smith, J., Mumford, R. & Boonham, N. (2017). Characterising maize viruses associated with maize lethal necrosis symptoms in sub saharan africa. *BioRxiv*, 161489.
- Agindotan, B. O., Domier, L. L., & Bradley, C. A. (2015). Detection and characterization of the first North American mastrevirus in switchgrass. *Archives of Virology*, *160*(5), 1313-1317.
- Adkins, S., Webb, S. E., Roberts, P. D., Kousik, C. S., Stansly, P. A., Bruton, B. D., Achor, D., Muchovej, R. M. & Baker, C. A. (2010). A review of Ipomoviruses and watermelon vine decline in Florida. *Bemisia: Bionomics and Management of a Global Pest*, 333-337.
- Aiewsakun, P., & Katzourakis, A. (2016). Time-dependent rate phenomenon in viruses. *Journal of Virology*, *90*(16), 7184-7195.
- Al Rwahnih, M., Daubert, S., Golino, D., & Rowhani, A. (2009). Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology*, *387*(2), 395-401.
- Alcaide, C., Rabadán, M. P., Moreno-Perez, M. G., & Gómez, P. (2020). Implications of mixed viral infections on plant disease ecology and evolution. *Advances in virus research*, *106*, 145-169.
- 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.
- Alexander, H. M., Steets, J. A., & Ali, A. (2020). Distribution of asclepias asymptomatic virus and exploration of possible effects on the wild plant host, asclepias viridis. *Plant Health Progress*, *21*(1), 54-59.
- Alexander, H. M., & Holt, R. D. (1998). The interaction between plant competition and disease. *Perspectives in Plant Ecology, Evolution and Systematics*, *1*(2), 206-220.
- Ali, A., Li, H., Schneider, W. L., Sherman, D. J., Gray, S., Smith, D., & Roossinck, M. J. (2006). Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. *Journal of Virology*, *80*(17), 8345-8350.
- Amin, I., Hussain, K., Akbergenov, R., Yadav, J. S., Qazi, J., Mansoor, S., Hohn, T., Fauquet, C. M. & Briddon, R. W. (2011). Suppressors of RNA silencing encoded by the components of the cotton leaf curl begomovirus-betasatellite complex. *Molecular Plant-Microbe Interactions*, *24*(8), 973-983.
- Anderson, J. A., Ghabrial, S. A., & Taylor, N. L. (1991). Natural incidence of peanut stunt virus infection in hybrid populations of *Trifolium ambiguum* × *T. repens*. *Plant Disease*, *75*(2), 156-159.
- Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends of Ecology and Evolution* *19*: 535–544
- Armand, T., Souquet, M., Korn, L., Gauthier, K., & Jacquot, E. Asymmetric interactions between barley yellow dwarf virus-PAV and wheat dwarf virus in wheat. *Frontiers in Plant Science*, *14*, 1194622.
- Ashby, B., Gupta, S., & Buckling, A. (2014). Effects of epistasis on infectivity range during host-parasite coevolution. *Evolution*, *68*(10), 2972-2982.
- Astier, S., Albouy, J., Maury, Y., & Lecoq, H. (2001). Principles of plant virology: genome, pathogenicity, virus ecology. Institut National de la Recherche Agronomique.
- Avershina, E., Frye, S. A., Ali, J., Taxt, A. M., & Ahmad, R. (2022). Ultrafast and cost-effective pathogen identification and resistance gene detection in a clinical setting using Nanopore Flongle sequencing. *Frontiers in Microbiology*, *13*, 822402.

REFERENCES

- Baret, S., Rouget, M., Richardson, D. M., Lavergne, C., Egoh, B., Dupont, J., & Strasberg, D. (2006). Current distribution and potential extent of the most invasive alien plant species on La Réunion (Indian Ocean, Mascarene islands). *Austral Ecology*, *31*(6), 747-758.
- Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, *458*(7241), 1018-1020.
- Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriological Reviews*, *35*(3), 235-241.
- Banks, P. M., Davidson, J. L., Bariana, H., & Larkin, P. J. (1995). Effects of barley yellow dwarf virus on the yield of winter wheat. *Australian Journal of Agricultural Research*, *46*(5), 935-946.
- Bass, D., Stentiford, G. D., Wang, H. C., Koskella, B., & Tyler, C. R. (2019). The pathobiome in animal and plant diseases. *Trends in Ecology & Evolution*, *34*(11), 996-1008.
- Bedhomme, S., Hillung, J., & Elena, S. F. (2015). Emerging viruses: why they are not jacks of all trades?. *Current Opinion in Virology*, *10*, 1-6.
- Bengyella, L., Waikhom, S. D., Allie, F., & Rey, C. (2015). Virus tolerance and recovery from viral induced-symptoms in plants are associated with transcriptome reprogramming. *Plant Molecular Biology*, *89*, 243-252.
- 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. (2018). Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME Journal*, *12*(1), 173-184.
- Betancourt, M., Fereres, A., Fraile, A., & García-Arenal, F. (2008). Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. *Journal of Virology*, *82*(24), 12416-12421.
- Bhattacharyya, D., & Chakraborty, S. (2018). Chloroplast: the Trojan horse in plant-virus interaction. *Molecular Plant Pathology*, *19*(2), 504-518.
- Bigarré, L., Salah, M., Granier, M., Frutos, R., Thouvenel, J. C., & Peterschmitt, M. (1999). Nucleotide sequence evidence for three distinct sugarcane streak mastreviruses. *Archives of Virology*, *144*(12), 2331-2344.
- Bilgin, D. D., Zavala, J. A., Zhu, J. I. N., Clough, S. J., Ort, D. R., & DeLUCIA, E. H. (2010). Biotic stress globally downregulates photosynthesis genes. *Plant, Cell & Environment*, *33*(10), 1597-1613.
- Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazariants, D., Si-Ammour, A., Park, H. S., Vazquez, F., Robertson, D., Meins, F. Jr., Hohn, T. & Pooggin, M. M. (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research*, *34*(21), 6233-6246.
- Blumenthal, D., Mitchell, C. E., Pyšek, P., & Jarošík, V. (2009). Synergy between pathogen release and resource availability in plant invasion. *Proceedings of the National Academy of Sciences*, *106*(19), 7899-7904.
- Bonnamy, M., Blanc, S., & Michalakis, Y. (2023). Replication mechanisms of circular ssDNA plant viruses and their potential implication in viral gene expression regulation. *Mbio*, e01692-23.
- Borer, E. T., Hosseini, P. R., Seabloom, E. W., & Dobson, A. P. (2007). Pathogen-induced reversal of native dominance in a grassland community. *Proceedings of the National Academy of Sciences*, *104*(13), 5473-5478.
- Bosque-Pérez, N. A. (2000). Eight decades of maize streak virus research. *Virus Research*, *71*(1-2), 107-121.
- Böttcher, B., Unseld, S., Ceulemans, H., Russell, R. B., & Jeske, H. (2004). Geminata structures of African cassava mosaic virus. *Journal of Virology*, *78*(13), 6758-6765.
- Bousalem, M., Douzery, E. J. P., & Fargette, D. (2000). High genetic diversity, distant phylogenetic relationships and intraspecies recombination events among natural populations of Yam mosaic virus: a contribution to understanding potyvirus evolution. *Journal of General Virology*, *81*(1), 243-255.

- Böttcher, B., Unseld, S., Ceulemans, H., Russell, R. B., & Jeske, H. (2004). Geminata structures of African cassava mosaic virus. *Journal of Virology*, *78*(13), 6758-6765.
- Boykin, L. M., Sseruwagi, P., Alicai, T., Ateka, E., Mohammed, I. U., Stanton, J. A. L., Kayuki, C., Mark, D., Fute, T., Erasto, J., Bachwenkizi, H., Muga, B., Mumo, N., Mwangi, J., Abidrabo, P., Okao-Okuja, G., Omuut, G., Akol, J., Apio, H. B., Osingada, F., Kehoe, M. A., Eccles, D., Savill, A., Lamb, S., Kinene, T., Rawle, C. B., Muralidhar, A., Mayall, K., Tairo, F. & Ndunguru, J. (2019). Tree lab: Portable genomics for early detection of plant viruses and pests in sub-saharan africa. *Genes*, *10*(9), 632.
- Bragard, C., Caciagli, P., Lemaire, O., Lopez-Moya, J. J., MacFarlane, S., Peters, D., Susi, P. & Torrance, L. (2013). Status and prospects of plant virus control through interference with vector transmission. *Annual Review of Phytopathology*, *51*, 177-201.
- Breitbart, M., & Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus?. *Trends in Microbiology*, *13*(6), 278-284.
- 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*(22), 14250-14255.
- Briddon, R. W., Lunness, P., Chamberlin, L. C. L., Pinner, M. S., Brundish, H., & Markham, P. G. (1992). The nucleotide sequence of an infectious insect-transmissible clone of the geminivirus Panicum streak virus. *Journal of General Virology*, *73*(5), 1041-1047.
- 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 β complexes. *Virology*, *324*(2), 462-474.
- Briddon, R. W., & Stanley, J. (2006). Subviral agents associated with plant single-stranded DNA viruses. *Virology*, *344*(1), 198-210.
- Bronzato Badial, A., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., & King, J. (2018). Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. *Plant Disease*, *102*(8), 1648-1652.
- Butkovic, A., Dolja, V. V., Koonin, E. V., & Krupovic, M. (2023). Plant virus movement proteins originated from jelly-roll capsid proteins. *PLoS Biology*, *21*(6), e3002157.
- Cadet, L. J. T. (1977). La végétation de l'île de La Réunion: étude phytoécologique et phytosociologique [dissertation]. Marseille, France: Université Aix-Marseille III.
- Campbell, C.S. (2020). Economic and ecological importance. *Britannica*, Chicago, IL. <https://www.britannica.com/plant/Poaceae/Economic-and-ecological-importance>.
- 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*(7), e102945.
- Carlsson-Granér, U., & Thrall, P. H. (2015). Host resistance and pathogen infectivity in host populations with varying connectivity. *Evolution*, *69*(4), 926-938.
- Carr, J. P., Donnelly, R., Tungadi, T., Murphy, A. M., Jiang, S., Bravo-Cazar, A., Yoon, J.-Y., Cunniffe, N. J., Glover, B. J. & Gilligan, C. A. (2018). Viral manipulation of plant stress responses and host interactions with insects. *Advances in Virus Research*, *102*, 177-197.
- Casteel, C. L., Yang, C., Nanduri, A. C., De Jong, H. N., Whitham, S. A., & Jander, G. (2014). The Nla-Pro protein of Turnip mosaic virus improves growth and reproduction of the aphid vector, *Myzus persicae* (green peach aphid). *The Plant Journal*, *77*(4), 653-663.
- Catford, J. A., Vesk, P. A., White, M. D., & Wintle, B. A. (2011). Hotspots of plant invasion predicted by propagule pressure and ecosystem characteristics. *Diversity and Distributions*, *17*(6), 1099-1110.

REFERENCES

- Cha, T., Kim, H. H., Keum, J., Kwak, M. J., Park, J. Y., Hoh, J. K., Kim, C.-R., Jeon, B.-H. & Park, H. K. (2023). Gut microbiome profiling of neonates using Nanopore MinION and Illumina MiSeq sequencing. *Frontiers in Microbiology*, *14*, 1148466.
- Chalupowicz, L., Dombrovsky, A., Gaba, V., Luria, N., Reuven, M., Beerman, A., Lachman, O., Dror, O., Nissan, G. & Manulis-Sasson, S. (2019). Diagnosis of plant diseases using the Nanopore sequencing platform. *Plant Pathology*, *68*(2), 229-238.
- Chare, E. R., & Holmes, E. C. (2006). A phylogenetic survey of recombination frequency in plant RNA viruses. *Archives of Virology*, *151*, 933-946.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195-205.
- Chatzivassiliou, E. K., Boubourakas, I., Drossos, E., Eleftherohorinos, I., Jenser, G., Peters, D., & Katis, N. I. (2001). Weeds in greenhouses and tobacco fields are differentially infected by Tomato spotted wilt virus and infested by its vector species. *Plant Disease*, *85*(1), 40-46.
- Chellappan, P., Vanitharani, R., & Fauquet, C. M. (2004). Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. *Journal of Virology*, *78*(14), 7465-7477.
- Chen, S., Huang, Q., Wu, L., & Qian, Y. (2015). Identification and characterization of a maize-associated mastrevirus in China by deep sequencing small RNA populations. *Virology Journal*, *12*, 1-9.
- Cheng, B., Hu, K., Song, Z., An, R., & Liang, X. (2023). Nanopore sequencing of short dsDNA after elongation by combination of ligation and PEAR. *Bulletin of the Chemical Society of Japan*, *96*(8), 785-792.
- Choua, M., Heath, M. R., Speirs, D. C., & Bonachela, J. A. (2020). The effect of viral plasticity on the persistence of host-virus systems. *Journal of Theoretical Biology*, *498*, 110263.
- Christian, M. L., & Willis, W. G. (1993). Survival of wheat streak mosaic virus in grass hosts in Kansas for wheat harvest to fall wheat emergence. *Plant Disease*, *77*(3), 239-242.
- Civitello, D. J., Cohen, J., Fatima, H., Halstead, N. T., Liriano, J., McMahon, T. A., Ortega, C. N., Sauer, E. L., Sehgal, T., Young, S. & Rohr, J. R. (2015). Biodiversity inhibits parasites: broad evidence for the dilution effect. *Proceedings of the National Academy of Sciences*, *112*(28), 8667-8671.
- Clark, M. F., & Adams, A. N. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology*, *34*(3), 475-483.
- Claverie, S., Ouattara, A., Hoareau, M., Filloux, D., Varsani, A., Roumagnac, P., Lett, J. M. & Lefevre, P. (2019). Exploring the diversity of Poaceae-infecting mastreviruses on Reunion Island using a viral metagenomics-based approach. *Scientific Reports*, *9*(1), 1-11.
- Claverie, S., Varsani, A., Hoareau, M., Filloux, D., Roumagnac, P., Martin, D. P., Lefevre, P. & Lett, J. M. (2020). Sorghum mastrevirus-associated alphasatellites: new geminalphasatellites associated with an African streak mastrevirus infecting wild Poaceae plants on Reunion Island. *Archives of Virology*, *165*, 1925-1928.
- Claverie, S., Hoareau, M., Ben Chehida, S., Filloux, D., Varsani, A., Roumagnac, P., Martin, D. P., Lett, J. M. & Lefevre, P. (2023). Metagenomics reveals the structure of Mastrevirus-host interaction network within an agro-ecosystem. *Virus Evolution*, *vead043*.
- Colvin, J., Omongo, C. A., Govindappa, M. R., Stevenson, P. C., Maruthi, M. N., Gibson, G., Seal, S. E. & Muniyappa, V. (2006). Host-plant viral infection effects on arthropod-vector population growth, development and behaviour: management and epidemiological implications. *Advances in Virus Research*, *67*, 419-452.
- Cooper, I., & Jones, R. A. (2006). Wild plants and viruses: under-investigated ecosystems. *Advances in Virus Research*, *67*, 1-47.
- Cramer, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* *391*: 288-291.

- Crespo-Bellido, A., & Duffy, S. (2023). The how of counter-defense: viral evolution to combat host immunity. *Current Opinion in Microbiology*, 74, 102320.
- Csorba, T., Kontra, L., & Burgyán, J. (2015). Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology*, 479, 85-103.
- Dáder, B., Then, C., Berthelot, E., Ducouso, M., Ng, J. C., & Drucker, M. (2017). Insect transmission of plant viruses: Multilayered interactions optimize viral propagation. *Insect Science*, 24(6), 929-946.
- 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.
- de Alba, A. E. M., Elvira-Matelot, E., & Vaucheret, H. (2013). Gene silencing in plants: a diversity of pathways. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1829(12), 1300-1308.
- De Barro, P. J. (2011). *Bemisia tabaci*, the capacity to invade. In *The Whitefly, Bemisia tabaci (Homoptera: Aleyrodidae) Interaction with Geminivirus-Infected Host Plants: Bemisia tabaci, Host Plants and Geminiviruses* (pp. 181-204). Dordrecht: Springer Netherlands.
- 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(15), 4075-4081.
- Del Blanco, I. A., Hegarty, J., Gallagher, L. W., Falk, B. W., Brown-Guedira, G., Pellerin, E., & Dubcovsky, J. (2014). Mapping of QTL for tolerance to Cereal yellow dwarf virus in two-rowed spring barley. *Crop Science*, 54(4), 1468-1475.
- Dellicour, S., Rose, R., Faria, N. R., Vieira, L. F. P., Bourhy, H., Gilbert, M., Lemey, P. & Pybus, O. G. (2017). Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Molecular Biology and Evolution*, 34(10), 2563-2571.
- Delwart, E. (2012). Animal virus discovery: improving animal health, understanding zoonoses, and opportunities for vaccine development. *Current Opinion in Virology*, 2(3), 344-352.
- Dietrich, C., & Maiss, E. (2003). Fluorescent labelling reveals spatial separation of potyvirus populations in mixed infected *Nicotiana benthamiana* plants. *Journal of General Virology*, 84(10), 2871-2876.
- Dolja, V. V., & Koonin, E. V. (2011). Common origins and host-dependent diversity of plant and animal viromes. *Current Opinion in Virology*, 1(5), 322-331.
- Dolja, V. V., & Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Research*, 244, 36-52.
- Dolja, V. V., Krupovic, M., & Koonin, E. V. (2020). Deep roots and splendid boughs of the global plant virome. *Annual Review of Phytopathology*, 58, 23-53.
- Domingo, E. J. J. H., & Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annual Review of Microbiology*, 51(1), 151-178.
- Domingo-Calap, P., Cuevas, J. M., & Sanjuan, R. (2009). The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genetics*, 5(11), e1000742.
- Donaire, L., Wang, Y., Gonzalez-Ibeas, D., Mayer, K. F., Aranda, M. A., & Llave, C. (2009). Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology*, 392(2), 203-214.
- Dormann, C. F., Fründ, J., & Schaefer, H. M. (2017). Identifying causes of patterns in ecological networks: opportunities and limitations. *Annual Review of Ecology, Evolution, and Systematics*, 48, 559-584.

REFERENCES

- Duffy, S., Shackelton, L. A., & Holmes, E. C. (2008). Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, 9(4), 267-276.
- Duffy, S., Turner, P. E., & Burch, C. L. (2006). Pleiotropic costs of niche expansion in the RNA bacteriophage $\Phi 6$. *Genetics*, 172(2), 751-757.
- Dunn, A. M., Torchin, M. E., Hatcher, M. J., Kotanen, P. M., Blumenthal, D. M., Byers, J. E., Coon, C. A. C., Frankel, V. M., Holt, R. D., Hufbauer, R. A., Kanarek, A. R., Schierenbeck, K. A., Wolfe, L. M. & Perkins, S. E. (2012). Indirect effects of parasites in invasions. *Functional Ecology*, 26(6), 1262-1274.
- Edula, S. R., Bag, S., Milner, H., Kumar, M., Suassuna, N. D., Chee, P. W., Kemerait, R.C., Hand, L. C., Snide, J. L., Srinivasan, R. & Roberts, P. M. (2023). Cotton leafroll dwarf disease: An enigmatic viral disease in cotton. *Molecular Plant Pathology*.
- Edwards, R. A., & Rohwer, F. (2005). Viral metagenomics. *Nature Reviews Microbiology*, 3(6), 504-510.
- Elena, S. F., & Lenski, R. E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4(6), 457-469.
- Elena, S. F., Agudelo-Romero, P., & Lalić, J. (2009). The evolution of viruses in multi-host fitness landscapes. *The Open Virology Journal*, 3, 1.
- Elena, S. F., Fraile, A., & García-Arenal, F. (2014). Evolution and emergence of plant viruses. *Advances in Virus Research*, 88, 161-191.
- Emeraghi, M., Achigan-Dako, E. G., Nwaoguala, C. N., & Oselebe, H. (2021). Maize streak virus research in Africa: an end or a crossroad. *Theoretical and Applied Genetics*, 134(12), 3785-3803.
- Faillace, C. A., Lorusso, N. S., & Duffy, S. (2017). Overlooking the smallest matter: viruses impact biological invasions. *Ecology Letters*, 20(4), 524-538.
- Fargette, D., Konate, G., Fauquet, C., Muller, E., Peterschmitt, M., & Thresh, J. M. (2006). Molecular ecology and emergence of tropical plant viruses. *Annual Review of Phytopathology*, 44, 235-260.
- Faust, C. L., Dobson, A. P., Gottdenker, N., Bloomfield, L. S., McCallum, H. I., Gillespie, T. R., Diuk-Wasser, M. & Plowright, R. K. (2017). Null expectations for disease dynamics in shrinking habitat: dilution or amplification?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1722), 20160173.
- Fereres, A., & Moreno, A. (2009). Behavioural aspects influencing plant virus transmission by homopteran insects. *Virus Research*, 141(2), 158-168.
- Fenouillas, P., Ah-Peng, C., Amy, E., Bracco, I., Dafreville, S., Gosset, M., Ingrassia, F., Lavergne, L., Lequette, B., Notter, J.-C., Pausé, J.-M., Payer, G., Payet, N., Picot, F., Pougavanon, N., Strasberg, D., Thomas, H., Triolo, J., Turquet, V. & Rouget, M. (2021). Quantifying invasion degree by alien plants species in Reunion Island. *Austral Ecology*, 46(7), 1025-1037.
- Fiallo-Olivé, E., & Navas-Castillo, J. (2023). Begomoviruses: what is the secret (s) of their success?. *Trends in Plant Science*.
- Filloux, D., Fernandez, E., Loire, E., Claude, L., Galzi, S., Candresse, T., Winter, S., Jeeva, M. L., Makesh Kumar, T., Martin, D. P. & Roumagnac, P. (2018). Nanopore-based detection and characterization of yam viruses. *Scientific Reports*, 8(1), 17879.
- Filloux, D., Fernandez, E., Bonnamy, M., Ben Chehida, S., Moubset, O., Hoareau, M., Julian, C., Blondin, Lett, J.-M., Lefeuvre, P. & Roumagnac, P. (2023). Nanopore sequencing of long tandem copies produced by rolling circle amplification is a credible alternative to recover complete genomes of circular ssDNA plant viruses. In : 19e Rencontres de virologie végétale (RVV2023) Rencontres de virologie végétale (RVV2023). 19, Aussois, France, 15 Janvier 2023/19 Janvier 2023.
- Fontenele, R. S., Alves-Freitas, D. M., Silva, P. I., Foresti, J., Silva, P. R., Godinho, M. T., Varsani, A. & Ribeiro, S. G. (2018). Discovery of the first maize-infecting mastrevirus in the Americas using a vector-enabled metagenomics approach. *Archives of Virology*, 163, 263-267.

REFERENCES

- Foresman, B. J., Oliver, R. E., Jackson, E. W., Chao, S., Arruda, M. P., & Kolb, F. L. (2016). Genome-wide association mapping of barley yellow dwarf virus tolerance in spring oat (*Avena sativa* L.). *PLoS One*, *11*(5), e0155376.
- Fraenkel-Conrat, H. (1986). Tobacco Mosaic Virus The History of Tobacco Mosaic Virus and the Evolution of Molecular Biology. In *The Plant Viruses: The Rod-Shaped Plant Viruses* (pp. 5-17). Boston, MA: Springer US.
- Fraile, A., & García-Arenal, F. (2010). The coevolution of plants and viruses: resistance and pathogenicity. *Advances in Virus Research*, *76*, 1-32.
- Fraile, A., & García-Arenal, F. (2016). Environment and evolution modulate plant virus pathogenesis. *Current Opinion in Virology*, *17*, 50-56.
- Fraile, A., Alonso-Prados, J. L., Aranda, M. A., Bernal, J. J., Malpica, J. M., & Garcia-Arenal, F. (1997). Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. *Journal of Virology*, *71*(2), 934-940.
- Fraile, A., McLeish, M. J., Pagán, I., González-Jara, P., Piñero, D., & García-Arenal, F. (2017). Environmental heterogeneity and the evolution of plant-virus interactions: viruses in wild pepper populations. *Virus Research*, *241*, 68-76.
- François, S., Filloux, D., Fernandez, E., Ogliastro, M., & Roumagnac, P. (2018). Viral metagenomics approaches for high-resolution screening of multiplexed arthropod and plant viral communities. *Viral Metagenomics: Methods and Protocols*, 77-95.
- Fraser, R. S. S. (1992). The genetics of plant-virus interactions: implications for plant breeding. *Euphytica*, *63*, 175-185.
- French, R. K., & Holmes, E. C. (2020). An ecosystems perspective on virus evolution and emergence. *Trends in Microbiology*, *28*(3), 165-175.
- French, R., & Stenger, D. C. (2003). Evolution of Wheat streak mosaic virus: dynamics of population growth within plants may explain limited variation. *Annual Review of Phytopathology*, *41*(1), 199-214.
- Frischmuth, T., Ringel, M., & Kocher, C. (2001). The size of encapsidated single-stranded DNA determines the multiplicity of African cassava mosaic virus particles. *Journal of General Virology*, *82*(3), 673-676.
- Froissart, R., Roze, D., Uzest, M., Galibert, L., Blanc, S., & Michalakakis, Y. (2005). Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. *PLoS Biology*, *3*(3), e89.
- Fuller, C. (1901). Mealic variegation. *Mealic Variegation.*, 17-19.
- Futuyma, D. J., & Moreno, G. (1988). The evolution of ecological specialization. *Annual Review of Ecology and Systematics*, *19*(1), 207-233.
- Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P., & Charnay, P. (1979). Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. *Nature*, *281*(5733), 646-650.
- García-Arenal, F., & Fraile, A. (2013). Trade-offs in host range evolution of plant viruses. *Plant Pathology*, *62*, 2-9.
- García-Arenal, F., & Zerbin, F. M. (2019). Life on the edge: geminiviruses at the interface between crops and wild plant hosts. *Annual Review of Virology*, *6*, 411-433.
- García-Arenal, F., Fraile, A., & Malpica, J. M. (2003). Variation and evolution of plant virus populations. *International Microbiology*, *6*, 225-232.
- García-Díaz, M., & Bebenek, K. (2007). Multiple functions of DNA polymerases. *Critical Reviews in Plant Sciences*, *26*(2), 105-122.
- García, L. R., & Janssen, D. (2020). Epidemiology and control of emerging criniviruses in bean. *Virus Research*, *280*, 197902.

- Garrett, K. A., Dendy, S. P., Frank, E. E., Rouse, M. N., & Travers, S. E. (2006). Climate change effects on plant disease: genomes to ecosystems. *Annual Review of Phytopathology*, 44, 489-509.
- Ghoshal, B., & Sanfaçon, H. (2015). Symptom recovery in virus-infected plants: revisiting the role of RNA silencing mechanisms. *Virology*, 479, 167-179.
- Gibbs, A. J. (1983). Virus ecology—"struggle" of the genes. *Physiological Plant Ecology III: Responses to the Chemical and Biological Environment*, 537-558.
- Gibbs, A. J., Nguyen, H. D., & Ohshima, K. (2015). The 'emergence' of turnip mosaic virus was probably a 'gene-for-quasi-gene' event. *Current Opinion in Virology*, 10, 20-26.
- Gilabert, A., Gauffre, B., Parisey, N., Le Gallic, J. F., Lhomme, P., Bretagnolle, V., Dedryver, C.-A., Baudry, J. & Plantegenest, M. (2017). Influence of the surrounding landscape on the colonization rate of cereal aphids and phytovirus transmission in autumn. *Journal of Pest Science*, 90, 447-457.
- Gnanasekaran, P., & Chakraborty, S. (2018). Biology of viral satellites and their role in pathogenesis. *Current Opinion in Virology*, 33, 96-105.
- Goldsmith, C., Cohen, D., Dubois, A., Martinez, M. G., Petitjean, K., Corlu, A., Testoni, B., Hernandez-Vargas, H. & Chemin, I. (2021). Cas9-targeted nanopore sequencing reveals epigenetic heterogeneity after de novo assembly of native full-length hepatitis B virus genomes. *Microbial Genomics*, 7(5).
- Gray, S. M., & Banerjee, N. (1999). Mechanisms of arthropod transmission of plant and animal viruses. *Microbiology and Molecular Biology Reviews*, 63(1), 128-148.
- Grilli, J., Rogers, T., & Allesina, S. (2016). Modularity and stability in ecological communities. *Nature Communications*, 7(1), 12031.
- Gu, H., Zhang, C., & Ghabrial, S. A. (2007). Novel naturally occurring Bean pod mottle virus reassortants with mixed heterologous RNA1 genomes. *Phytopathology*, 97(1), 79-86.
- Guadie, D., Tesfaye, K., Knierim, D., Winter, S., & Abraham, A. (2019). Molecular analysis of maize (*Zea mays* L.)-infecting mastreviruses in Ethiopia reveals marked diversity of virus genomes and a novel species. *Virus Genes*, 55, 339-345.
- Gurarie, D., & Seto, E. Y. (2009). Connectivity sustains disease transmission in environments with low potential for endemicity: modelling schistosomiasis with hydrologic and social connectivities. *Journal of the Royal Society Interface*, 6(35), 495-508.
- Guyot, V., Rajeswaran, R., Chu, H. C., Karthikeyan, C., Laboureau, N., Galzi, S., Mukwa, L. F. T., Krupovic, M., Kumar, P. L., Iskra-Caruana, M.-L. & Pooggin, M. M. (2022). A newly emerging alphasatellite affects banana bunchy top virus replication, transcription, siRNA production and transmission by aphids. *PLoS Pathogens*, 18(4), e1010448.
- Haas, S. E., Hooten, M. B., Rizzo, D. M., & Meentemeyer, R. K. (2011). Forest species diversity reduces disease risk in a generalist plant pathogen invasion. *Ecology Letters*, 14(11), 1108-1116.
- Halliday, F. W., Heckman, R. W., Wilfahrt, P. A., & Mitchell, C. E. (2017). A multivariate test of disease risk reveals conditions leading to disease amplification. *Proceedings of the Royal Society B: Biological Sciences*, 284(1865), 20171340.
- Halliday, F. W., Rohr, J. R., & Laine, A. L. (2020). Biodiversity loss underlies the dilution effect of biodiversity. *Ecology Letters*, 23(11), 1611-1622.
- Hamelin, F. M., Hilker, F. M., Sun, T. A., Jeger, M. J., Hajimorad, M. R., Allen, L. J., & Prendeville, H. R. (2017). The evolution of parasitic and mutualistic plant-virus symbioses through transmission-virulence trade-offs. *Virus Research*, 241, 77-87.
- Hamer, G. L., Chaves, L. F., Anderson, T. K., Kitron, U. D., Brawn, J. D., Ruiz, M. O., Loss, S. R., Walker, E. D. & Goldberg, T. L. (2011). Fine-scale variation in vector host use and force of infection drive localized patterns of West Nile virus transmission. *PloS One*, 6(8), e23767.

REFERENCES

- Handel, A., Lebarbenchon, C., Stallknecht, D., & Rohani, P. (2014). Trade-offs between and within scales: environmental persistence and within-host fitness of avian influenza viruses. *Proceedings of the Royal Society B: Biological Sciences*, 281(1787), 20133051.
- Harkins, G. W., Martin, D. P., Duffy, S., Monjane, A. L., Shepherd, D. N., Windram, O. P., Owor, B. E., Donaldson, L., van Antwerpen, T., Sayed, R. A., Flett, B., Ramusi, M., Rybicki, E. P., Peterschmitt, M. & Varsani, A. (2009). Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. *The Journal of General Virology*, 90(Pt 12), 3066.
- Harrison, B. D., Zhou, X., Otim-Nape, G. W., Liu, Y., & Robinson, D. J. (1997). Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Annals of Applied Biology*, 131(3), 437-448.
- Hasiów-Jaroszewska, B., Boezen, D., & Zwart, M. P. (2021). Metagenomic studies of viruses in weeds and wild plants: a powerful approach to characterise variable virus communities. *Viruses*, 13(10), 1939.
- Hechinger, R. F., & Lafferty, K. D. (2005). Host diversity begets parasite diversity: bird final hosts and trematodes in snail intermediate hosts. *Proceedings of the Royal Society B: Biological Sciences*, 272(1567), 1059-1066.
- Heil, M., & Bostock, R. M. (2002). Induced systemic resistance (ISR) against pathogens in the context of induced plant defences. *Annals of Botany*, 89(5), 503-512.
- Hendrix, R. W. (2002). Bacteriophages: evolution of the majority. *Theoretical Population Biology*, 61(4), 471-480.
- Hernández-Zepeda, C., Varsani, A., & Brown, J. K. (2013). Intergeneric recombination between a new, spinach-infecting curtovirus and a new geminivirus belonging to the genus Becurtovirus: first New World exemplar. *Archives of Virology*, 158, 2245-2254.
- Hesketh, E. L., Saunders, K., Fisher, C., Potze, J., Stanley, J., Lomonosoff, G. P., & Ranson, N. A. (2018). The 3.3 Å structure of a plant geminivirus using cryo-EM. *Nature Communications*, 9(1), 2369.
- Hily, J. M., Poulicard, N., Mora, M. Á., Pagán, I., & García-Arenal, F. (2016). Environment and host genotype determine the outcome of a plant–virus interaction: from antagonism to mutualism. *New Phytologist*, 209(2), 812-822.
- Hipp, K., Grimm, C., Jeske, H., & Böttcher, B. (2017). Near-atomic resolution structure of a plant geminivirus determined by electron cryomicroscopy. *Structure*, 25(8), 1303-1309.
- Hogenhout, S. A., Ammar, E. D., Whitfield, A. E., & Redinbaugh, M. G. (2008). Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology*, 46, 327-359.
- Holmes, E. C. (2003). Error thresholds and the constraints to RNA virus evolution. *Trends in microbiology*, 11(12), 543-546.
- 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(3), 208-217.
- Huggett, D. A. J., Leather, S. R., & Walters, K. F. A. (1999). Suitability of the biomass crop *Miscanthus sinensis* as a host for the aphids *Rhopalosiphum padi* (L.) and *Rhopalosiphum maidis* (F.), and its susceptibility to the plant luteovirus Barley Yellow Dwarf Virus. *Agricultural and Forest Entomology*, 1(2), 143-149.
- Hughes, F. L., Rybicki, E. P., & von Wechmar, M. B. (1992). Genome typing of southern African subgroup 1 geminiviruses. *Journal of General Virology*, 73(5), 1031-1040.
- Hull, R. (2014). Replication of plant viruses. *Plant Virology*, 341-421.
- ICTV (2020). The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. *Nature Microbiology*, 2020, vol. 5, no 5, p. 668-674.
- Ingwell, L. L., Eigenbrode, S. D., & Bosque-Pérez, N. A. (2012). Plant viruses alter insect behavior to enhance their spread. *Scientific Reports*, 2(1), 578.

REFERENCES

- 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 ϕ 29 DNA polymerase. *Journal of Virological Methods*, *116*(2), 209-211.
- Inouye, T., & Osaki, T. (1980). The first record in the literature of the possible plant virus disease that appeared in "Manyoshu", a Japanese classic anthology, as far back as the time of the 8th century. *Japanese Journal of Phytopathology*, *46*(1), 49-50.
- IUCN (2010). IUCN Technical Evaluation. Pitons, cirques and ramparts of Reunion Island. *Technical Report*. IUCN, France
- Ivanovski, D. (1892). Über die Mosaikkrankheit der Tabakspflanze, St. Petersburg. *Academy Imperial Science Bulletin*, *35*, 67-70.
- This is the first report that extracts from infected leaves were still infectious after filtration through a Chamberland filter-candle.**
- Javaran, V. J., Moffett, P., Lemoyne, P., Xu, D., Adkar-Purushothama, C. R., & Fall, M. L. (2021). Grapevine virology in the third-generation sequencing era: from virus detection to viral epitranscriptomics. *Plants*, *10*(11), 2355.
- Jeger, M. J., Seal, S. E., & Van den Bosch, F. (2006). Evolutionary epidemiology of plant virus disease. *Advances in Virus Research*, *67*, 163-203.
- 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., Lütgemeier, M., & Preiß, W. (2001). DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *The EMBO Journal*, *20*(21), 6158-6167.
- Jeske, H. (2018). Barcoding of plant viruses with circular single-stranded DNA based on rolling circle amplification. *Viruses*, *10*(9), 469.
- Jin, H., Domier, L. L., Kolb, F. L., & Brown, C. M. (1998). Identification of quantitative loci for tolerance to barley yellow dwarf virus in oat. *Phytopathology*, *88*(5), 410-415.
- Jiu, M., Zhou, X. P., Tong, L., Xu, J., Yang, X., Wan, F. H., & Liu, S. S. (2007). Vector-virus mutualism accelerates population increase of an invasive whitefly. *PLoS One*, *2*(1), e182.
- Johnson, P. T., Preston, D. L., Hoverman, J. T., & LaFonte, B. E. (2013). Host and parasite diversity jointly control disease risk in complex communities. *Proceedings of the National Academy of Sciences*, *110*(42), 16916-16921.
- Johnson, P. T., Wood, C. L., Joseph, M. B., Preston, D. L., Haas, S. E., & Springer, Y. P. (2016). Habitat heterogeneity drives the host-diversity-begets-parasite-diversity relationship: evidence from experimental and field studies. *Ecology Letters*, *19*(7), 752-761.
- Jones, R. A. (2013). Virus diseases of perennial pasture legumes in Australia: incidences, losses, epidemiology, and management. *Crop and Pasture Science*, *64*(3), 199-215.
- Jones, R. A. (2009). Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, *141*(2), 113-130.
- Jones, R. A. C. (2014). Plant virus ecology and epidemiology: historical perspectives, recent progress and future prospects. *Annals of Applied Biology*, *164*(3), 320-347.
- Jones, R. A. (2018). Plant and insect viruses in managed and natural environments: Novel and neglected transmission pathways. *Advances in Virus Research*, *101*, 149-187.
- Jones, R. A. (2020). Disease pandemics and major epidemics arising from new encounters between indigenous viruses and introduced crops. *Viruses*, *12*(12), 1388.

REFERENCES

- Kamiya, T., O'Dwyer, K., Nakagawa, S., & Poulin, R. (2014). Host diversity drives parasite diversity: meta-analytical insights into patterns and causal mechanisms. *Ecography*, *37*(7), 689-697.
- Kanakala, S., & Kuria, P. (2018). Chickpea chlorotic dwarf virus: an emerging monopartite dicot infecting mastrevirus. *Viruses*, *11*(1), 5.
- Kawecki, T. J. (1994). Accumulation of deleterious mutations and the evolutionary cost of being a generalist. *The American Naturalist*, *144*(5), 833-838.
- Kazlauskas, D., Varsani, A., Koonin, E. V., & Krupovic, M. (2019). Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nature Communications*, *10*(1), 3425.
- Keesing, F., Holt, R. D., & Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters*, *9*(4), 485-498.
- 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*(7324), 647-652.
- Dolja, V. V., & Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Research*, *244*, 36-52.
- 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*(10), 2763-2766.
- Koonin, E. V., Dolja, V. V., & Krupovic, M. (2015). Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*, *479*, 2-25.
- Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., Zerbini, F. M. & Kuhn, J. H. (2020). Global organization and proposed megataxonomy of the virus world. *Microbiology and Molecular Biology Reviews*, *84*(2), e00061-19.
- Koonin, E. V., Krupovic, M., & Agol, V. I. (2021). The Baltimore classification of viruses 50 years later: how does it stand in the light of virus evolution?. *Microbiology and Molecular Biology Reviews*, *85*(3), e00053-21.
- 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*(1-2), 282-291.
- Kraberger, S., Saumtally, S., Pande, D., Khoodoo, M. H., Dhayan, S., Dookun-Saumtally, A., Shepherd, D. N., Hartnady, P., Atkinson, R., Lakay, F. M., Hanson, B., Redhi, D., Monjane, A. L., Windram, O. P., Walters, M., Oluwafemi, S., Lett, J. M., 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.
- 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), 1-7.
- Kreuzer, K. N., Saunders, M., Weislo, L. J., & Kreuzer, H. W. (1995). Recombination-dependent DNA replication stimulated by double-strand breaks in bacteriophage T4. *Journal of Bacteriology*, *177*(23), 6844-6853.
- Krupovic, M., Ravantti, J. J., & Bamford, D. H. (2009). Geminiviruses: a tale of a plasmid becoming a virus. *BMC Evolutionary Biology*, *9*, 1-11.
- Krupovic, M., & Koonin, E. V. (2017). Multiple origins of viral capsid proteins from cellular ancestors. *Proceedings of the National Academy of Sciences*, *114*(12), E2401-E2410.
- Krupovic, M., Cvirkaite-Krupovic, V., Iranzo, J., Prangishvili, D., & Koonin, E. V. (2018). Viruses of archaea: structural, functional, environmental and evolutionary genomics. *Virus Research*, *244*, 181-193.

REFERENCES

- Krupovic, M., Varsani, A., Kazlauskas, D., Breitbart, M., Delwart, E., Rosario, K., Yutin, N., Wolf, Y. I., Harrach, B., Zerbini, F. M., Dolja, V. V., Kuhn, J. H. & Koonin, E. V. (2020). Cressdnaviricota: a virus phylum unifying seven families of rep-encoding viruses with single-stranded, circular DNA genomes. *Journal of Virology*, *94*(12), 10-1128.
- Kühnert, D., Wu, C. H., & Drummond, A. J. (2011). Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. *Infection, Genetics and Evolution*, *11*(8), 1825-1841.
- 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*(12), 7093-7104.
- Kumar, P., Cowan, G. H., Squires, J. N., Hackett, C. A., Tobin, A. K., Torrance, L., & Roberts, A. G. (2022). Phloem connectivity and transport are not involved in mature plant resistance (MPR) to Potato Virus Y in different potato cultivars, and MPR does not protect tubers from recombinant strains of the virus. *Journal of Plant Physiology*, *275*, 153729.
- Kvarnheden, A., Lindblad, M., Lindsten, K., & Valkonen, J. P. T. (2002). Genetic diversity of Wheat dwarf virus. *Archives of Virology*, *147*, 205-216.
- Lacroix, C., Jolles, A., Seabloom, E. W., Power, A. G., Mitchell, C. E., & Borer, E. T. (2014). Non-random biodiversity loss underlies predictable increases in viral disease prevalence. *Journal of the Royal Society Interface*, *11*(92), 20130947.
- Lafferty, K. D., & Kuris, A. M. (2012). Ecological consequences of manipulative parasites. *Host Manipulation by Parasites*, 158, 168.
- Lai, M. M. (1992). RNA recombination in animal and plant viruses. *Microbiological Reviews*, *56*(1), 61-79.
- Laine, A. L. (2006). Evolution of host resistance: looking for coevolutionary hotspots at small spatial scales. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1584), 267-273.
- Laliberté, J. F., & Sanfaçon, H. (2010). Cellular remodeling during plant virus infection. *Annual Review of Phytopathology*, *48*, 69-91.
- Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences*, *82*(20), 6955-6959.
- Lappe, R. R., Elmore, M. G., Lozier, Z. R., Jander, G., Miller, W. A., & Whitham, S. A. (2022). Metagenomic identification of novel viruses of maize and teosinte in North America. *BMC Genomics*, *23*(1), 767.
- Lavergne, R. (1978). Les pestes végétales de l'île de La Réunion. *Info-Nature*, *16*, 9-60.
- Lawry, R., Martin, D. P., Shepherd, D. N., van Antwerpen, T., & Varsani, A. (2009). A novel sugarcane-infecting mastrevirus from South Africa. *Archives of Virology*, *154*, 1699-1703.
- Lebouvier, M., Laparie, M., Hulle, M., Marais, A., Cozic, Y., Lalouette, L., Vernon, P., Candresse, T., Frenot, Y., & Renault, D. (2011). The significance of the sub-Antarctic Kerguelen Islands for the assessment of the vulnerability of native communities to climate change, alien insect invasions and plant viruses. *Biological Invasions*, *13*, 1195-1208.
- Lefevre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Roumagnac, P., & Varsani, A. (2019). Evolution and ecology of plant viruses. *Nature Reviews Microbiology*, *17*(10), 632-644.
- Leiva, A. M., Siriwan, W., Lopez-Alvarez, D., Barrantes, I., Hemniam, N., Saokham, K., & Cuellar, W. J. (2020). Nanopore-based complete genome sequence of a Sri Lankan cassava mosaic virus (Geminivirus) strain from Thailand. *Microbiology Resource Announcements*, *9*(6), 10-1128.
- Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., & Smith, D. B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research*, *46*(D1), D708-D717.

REFERENCES

- Lesnaw, J. A., & Ghabrial, S. A. (2000). Tulip breaking: past, present, and future. *Plant Disease*, 84(10), 1052-1060.
- Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, 15(3), 237-240.
- Li, H., & Roossinck, M. J. (2004). Genetic bottlenecks reduce population variation in an experimental RNA virus population. *Journal of Virology*, 78(19), 10582-10587.
- Li, J. X., Liu, S. S., & Gu, Q. S. (2016). Transmission efficiency of Cucumber green mottle mosaic virus via seeds, soil, pruning and irrigation water. *Journal of Phytopathology*, 164(5), 300-309.
- Linnaeus, C. (1735). *Systema naturae sive regna tria naturae systematice proposita per classes, ordines, genera et species*. Apud T. Haak.
- Liu, Y., Cai, J., Li, D., Qin, B., & Tian, B. (1998). Chinese tomato yellow leaf curl virus—a new species of geminivirus. *Science in China Series C: Life Sciences*, 41, 337-343.
- Liu, B., Preisser, E. L., Chu, D., Pan, H., Xie, W., Wang, S., Wu, Q., Zhou, X. & Zhang, Y. (2013). Multiple forms of vector manipulation by a plant-infecting virus: Bemisia tabaci and tomato yellow leaf curl virus. *Journal of Virology*, 87(9), 4929-4937.
- Liu, X., Chen, L., Liu, M., García-Guzmán, G., Gilbert, G. S., & Zhou, S. (2020). Dilution effect of plant diversity on infectious diseases: latitudinal trend and biological context dependence. *Oikos*, 129(4), 457-465.
- López-González, S., Gómez-Mena, C., Sánchez, F., Schuetz, M., Samuels, A. L., & Ponz, F. (2021). The Effects of Turnip Mosaic Virus Infections on the Deposition of Secondary Cell Walls and Developmental Defects in Arabidopsis Plants Are Virus-Strain Specific. *Frontiers in Plant Science*, 12, 741050.
- Lovisolo, O., Hull, R., & Rosler, O. (2003). Coevolution of viruses with hosts and vectors and possible paleontology. *Advances in Virus Research*, 62, 325-379.
- Lozano, G., Trenado, H. P., Fiallo-Olivé, E., Chirinos, D., Geraud-Pouey, F., 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, 162.
- Lucia-Sanz, A., Aguirre, J., Fraile, A., Garcia-Arenal, F., & Manrubia, S. (2022). Defective subviral particles modify ecological equilibria and enhance viral coexistence. *Frontiers in Virology*, 2, 929851.
- Lythgoe, K. A., Gardner, A., Pybus, O. G., & Grove, J. (2017). Short-sighted virus evolution and a germline hypothesis for chronic viral infections. *Trends in Microbiology*, 25(5), 336-348.
- Ma, Y., Marais, A., Lefebvre, M., Faure, C., & Candresse, T. (2020). Metagenomic analysis of virome cross-talk between cultivated Solanum lycopersicum and wild Solanum nigrum. *Virology*, 540, 38-44.
- Ma, Y., Svanella-Dumas, L., Julian, C., Galzi, S., Fernandez, E., Yvon, M., Pirolles, E., Lefebvre M., Filloux, D., Roumagnac, P. & Candresse, T. (2021). Genome characterization and diversity of trifolium virus 1: identification of a novel legume-infecting capulavirus. *Archives of Virology*, 166, 2573-2578.
- Maachi, A., Donaire, L., Hernando, Y., & Aranda, M. A. (2022). Genetic differentiation and migration fluxes of viruses from melon crops and crop edge weeds. *Journal of Virology*, 96(16), e00421-22.
- Macdonald, I. A., Thébaud, C., Strahm, W. A., & Strasberg, D. (1991). Effects of alien plant invasions on native vegetation remnants on La Réunion (Mascarene Islands, Indian Ocean). *Environmental Conservation*, 18(1), 51-61.
- Maclot, F., Candresse, T., Filloux, D., Malmstrom, C. M., Roumagnac, P., Van der Vlugt, R., & Massart, S. (2020). Illuminating an ecological blackbox: using high throughput sequencing to characterize the plant virome across scales. *Frontiers in Microbiology*, 11, 578064.

REFERENCES

- Maclot, F., Candresse, T., Filloux, D., Rott, P., Malmstrom, C., van der Vlugt, R., Massart, S. & Roumagnac, P. (2021). From boots on the ground to nucleotides in the sequencer: a century of advances in the study of the plant virus ecology. *Virologie*, 25(1), 29-42.
- Maclot, F., Debue, V., Malmstrom, C. M., Filloux, D., Roumagnac, P., Eck, M., Tamisier, L., Blouin, A. G., Candresse, T. & Massart, S. (2023). Long-Term Anthropogenic Management and Associated Loss of Plant Diversity Deeply Impact Virome Richness and Composition of Poaceae Communities. *Microbiology Spectrum*, 11(2), e04850-22.
- Malmstrom, C. M., & Alexander, H. M. (2016). Effects of crop viruses on wild plants. *Current Opinion in Virology*, 19, 30-36.
- Malmstrom, C. M., McCullough, A. J., Johnson, H. A., Newton, L. A., & Borer, E. T. (2005a). Invasive annual grasses indirectly increase virus incidence in California native perennial bunchgrasses. *Oecologia*, 145, 153-164.
- Malmstrom, C. M., Hughes, C. C., Newton, L. A., & Stoner, C. J. (2005b). Virus infection in remnant native bunchgrasses from invaded California grasslands. *New Phytologist*, 168(1), 217-230.
- Malmstrom, C. M., Melcher, U., & Bosque-Perez, N. A. (2011). The expanding field of plant virus ecology: historical foundations, knowledge gaps, and research directions. *Virus Research*, 159(2), 84-94.
- Manrubia, S. C., Domingo, E., & Lázaro, E. (2010). Pathways to extinction: beyond the error threshold. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1548), 1943-1952.
- Mar, T. B., Xavier, C. A. D., Lima, A. T. M., Nogueira, A. M., Silva, J. C. F., Ramos-Sobrinho, R., Lau, D. & Zerbini, F. M. (2017). Genetic variability and population structure of the New World begomovirus Euphorbia yellow mosaic virus. *Journal of General Virology*, 98(6), 1537-1551.
- Marais, A., Faure, C., Bergey, B., & Candresse, T. (2018). Viral double-stranded RNAs (dsRNAs) from plants: Alternative nucleic acid substrates for high-throughput sequencing. *Viral Metagenomics: Methods and Protocols*, 45-53.
- Maris, P. C., Joosten, N. N., Goldbach, R. W., & Peters, D. (2003). Restricted spread of tomato spotted wilt virus in thrips-resistant pepper. *Phytopathology*, 93(10), 1223-1227.
- Marsberg, A., Kemler, M., Jami, F., Nagel, J. H., Postma-Smidt, A., Naidoo, S., Wingfield, M. J., Crous, P. W., Spatafora, J. W., Hesse, C. N., Robbertse, B. & Slippers, B. (2017). Botryosphaeria dothidea: a latent pathogen of global importance to woody plant health. *Molecular Plant Pathology*, 18(4), 477-488.
- Marshall, N., Priyamvada, L., Ende, Z., Steel, J., & Lowen, A. C. (2013). Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathogens*, 9(6), e1003421.
- Martin, D. P., Biagini, P., Lefevre, P., Golden, M., Roumagnac, P., & Varsani, A. (2011). Recombination in eukaryotic single stranded DNA viruses. *Viruses*, 3, 1699-1738.
- Martinez, M. E. (2018). The calendar of epidemics: Seasonal cycles of infectious diseases. *PLoS Pathogens*, 14(11), e1007327.
- Marwal, A., Verma, R. K., Mishra, M., Kumar, R., & Gaur, R. K. (2019). Mastreviruses in the African World: harbouring both monocot and dicot species. *Geminiviruses: Impact, Challenges and Approaches*, 85-102.
- Mascia, T., & Gallitelli, D. (2016). Synergies and antagonisms in virus interactions. *Plant Science*, 252, 176-192.
- Maskell, L. C., Raybould, A. F., Cooper, J. I., Edwards, M. L., & Gray, A. J. (1999). Effects of turnip mosaic virus and turnip yellow mosaic virus on the survival, growth and reproduction of wild cabbage (*Brassica oleracea*). *Annals of Applied Biology*, 135(1), 401-407.
- Matsuo, Y., Komiya, S., Yasumizu, Y., Yasuoka, Y., Mizushima, K., Takagi, T., Kryukov, K., Fukuda, A., Morimoto, Y., Naito, Y., Okada, H., Bono, H., Nakagawa, S. & Hirota, K. (2021). Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. *BMC Microbiology*, 21, 1-13.

REFERENCES

- Mauck, K., Bosque-Pérez, N. A., Eigenbrode, S. D., De Moraes, C. M., & Mescher, M. C. (2012). Transmission mechanisms shape pathogen effects on host–vector interactions: evidence from plant viruses. *Functional Ecology*, 26(5), 1162-1175.
- Mauck, K. E., De Moraes, C. M., & Mescher, M. C. (2014). Biochemical and physiological mechanisms underlying effects of Cucumber mosaic virus on host-plant traits that mediate transmission by aphid vectors. *Plant, Cell & Environment*, 37(6), 1427-1439.
- Maule, A. J., Hull, R., & Donson, J. (1983). The application of spot hybridization to the detection of DNA and RNA viruses in plant tissues. *Journal of Virological Methods*, 6(4), 215-224.
- McKenzie, R. I. H., Burnett, P. A., Gill, C. C., Comeau, A., & Brown, P. D. (1985). Inheritance of tolerance to barley yellow dwarf virus in oats. *Euphytica*, 34, 681-687.
- McKirby, S. J., Jones, R. A. C., & Nutter Jr, F. W. (2002). Quantification of yield losses caused by barley yellow dwarf virus in wheat and oats. *Plant Disease*, 86(7), 769-773.
- McLeish, M., Sacristán, S., Fraile, A., & García-Arenal, F. (2017). Scale dependencies and generalism in host use shape virus prevalence. *Proceedings of the Royal Society B: Biological Sciences*, 284(1869), 20172066.
- McLeish, M. J., Fraile, A., & García-Arenal, F. (2018). Ecological complexity in plant virus host range evolution. *Advances in Virus Research*, 101, 293-339.
- McLeish, M. J., Fraile, A., & García-Arenal, F. (2021). Population genomics of plant viruses: the ecology and evolution of virus emergence. *Phytopathology*, 111(1), 32-39.
- McLeish, M. J., Zamfir, A. D., Babalola, B. M., Peláez, A., Fraile, A., & García-Arenal, F. (2022). Metagenomics show high spatiotemporal virus diversity and ecological compartmentalisation: Virus infections of melon, Cucumis melo, crops, and adjacent wild communities. *Virus Evolution*, 8(2), veac095.
- McNaughton, A. L., Roberts, H. E., Bonsall, D., de Cesare, M., Mokaya, J., Lumley, S. F., Golubchik, T., Piazza, P., Martin, J. B., de Lara, C., Brown, A., Ansari, M. A., Bowden, R., Barnes, E. & Matthews, P. C. (2019). Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). *Scientific Reports*, 9(1), 7081.
- Mehle, N., Gutierrez-Aguirre, I., Kutnjak, D., & Ravnikar, M. (2018). Water-mediated transmission of plant, animal, and human viruses. *Advances in Virus Research*, 101, 85-128.
- Melcher, U., Muthukumar, V., Wiley, G. B., Min, B. E., Palmer, M. W., Verchot-Lubicz, J., Ali, A., Nelson, R. S., Roe, B. A., Thapa, V. & Pierce, M. L. (2008). Evidence for novel viruses by analysis of nucleic acids in virus-like particle fractions from *Ambrosia psilostachya*. *Journal of Virological Methods*, 152(1-2), 49-55.
- Mesfin, T., Den Hollander, J., & Markham, P. G. (1995). Feeding activities of *Cicadulina mbila* (Hemiptera: Cicadellidae) on different host-plants. *Bulletin of Entomological Research*, 85(3), 387-396.
- Michalakis, Y., & Blanc, S. (2020). The curious strategy of multipartite viruses. *Annual Review of Virology*, 7, 203-218.
- Mink, G. I. (1993). Pollen and seed-transmitted viruses and viroids. *Annual Review of Phytopathology*, 31(1), 375-402.
- Miras, M., Sempere, R. N., Kraft, J. J., Miller, W. A., Aranda, M. A., & Truniger, V. (2014). Interfamilial recombination between viruses led to acquisition of a novel translation-enhancing RNA element that allows resistance breaking. *New Phytologist*, 202(1), 233-246.
- Mitchell, C. E., & Power, A. G. (2003). Release of invasive plants from fungal and viral pathogens. *Nature*, 421(6923), 625-627.
- Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: the critical role of hotspots. In *Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas* (pp. 3-22). Berlin, Heidelberg: Springer Berlin Heidelberg.

REFERENCES

- Monci, F., Sánchez-Campos, S., Navas-Castillo, J., & Moriones, E. (2002). A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Virology*, *303*(2), 317-326.
- Monjane, A. L., Harkins, G. W., Martin, D. P., Lemey, P., Lefevre, P., Shepherd, D. N., Oluwafemi, S., Simuyandi, M., Zinga, I., Komba, E. K., Lakoutene, D. P., Mandakombo, N., Mboukoulida, J., Semballa, S., Tagne, A., Tiendrébéogo, F., Erdmann, J. B., van Antwerpen, T., Owaor, B. E., Flett, B., Ramusi, M., Windram, O. P., Syed, R., Lett, J.-M., Briddon, R. W., Markham, P. G., Rybicki, E. P. & Varsani, A. (2011). Reconstructing the history of maize streak virus strain a dispersal to reveal diversification hot spots and its origin in southern Africa. *Journal of Virology*, *85*(18), 9623-9636.
- Montes, N., Alonso-Blanco, C., & García-Arenal, F. (2019). Cucumber mosaic virus infection as a potential selective pressure on *Arabidopsis thaliana* populations. *PLoS Pathogens*, *15*(5), e1007810.
- Moon, J. Y., & Park, J. M. (2016). Cross-talk in viral defense signaling in plants. *Frontiers in Microbiology*, *7*, 2068.
- Moonan, F., Molina, J., & Mirkov, T. E. (2000). Sugarcane yellow leaf virus: an emerging virus that has evolved by recombination between luteoviral and poleroviral ancestors. *Virology*, *269*(1), 156-171.
- Moriones, E., & Navas-Castillo, J. (2000). Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. *Virus Research*, *71*(1-2), 123-134.
- 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*, *98*(15), 8306-8311.
- Moubset, O., François, S., Maclot, F., Palanga, E., Julian, C., Claude, L., Fernandez, E., Rott, P., Daugrois, J.-H., Antoine-Lorquin, A., Bernardo, P., Blouin, A. G., Temple, C., Kraberger, S., Fontenele, R. S., Harkins, G. W., Ma, Y., Marais, A., Candresse, T., Ben Chehida, S., Lefevre, P., Lett, J.-M., Varsani, A., Massart, S., Ogliastro, M., Martin, D. P., Filloux, D. & Roumagnac, P. (2022). Virion-associated nucleic acid-based metagenomics: a decade of advances in molecular characterization of plant viruses. *Phytopathology*, *112*(11), 2253-2272.
- Moury, B., & Simon, V. (2011). dN/dS-based methods detect positive selection linked to trade-offs between different fitness traits in the coat protein of potato virus Y. *Molecular Biology and Evolution*, *28*(9), 2707-2717.
- Moury, B., Ayme, V., Souche, S., Jacquemond, M., Johansen, E., Jacquot, E., & Kerlan, C. (2004). Potato virus Y diversity and evolution: which role for the host plants?. In *EAPR Virology 2004, Abstracts of the 12th European Association for Potato Research Virology Section Meeting, Rennes, France, 13-19 June 2004* (pp. 21-22). European Association for Potato Research.
- Moury, B., Fabre, F., Hébrard, E., & Froissart, R. (2017). Determinants of host species range in plant viruses. *Journal of General Virology*, *98*(4), 862-873.
- Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., Rivera-Bustamante, R., Malathi, V. G., Briddon, R. B. & 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.
- 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*(24), 7259-7265.
- Mushegian, A. R. (2020). Are there 10³¹ virus particles on earth, or more, or fewer?. *Journal of bacteriology*, *202*(9), 10-1128.
- Muthukumar, V., Melcher, U., Pierce, M., Wiley, G. B., Roe, B. A., Palmer, M. W., Thapa, V., Ali, A & Ding, T. (2009). Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. *Virus Research*, *141*(2), 169-173.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853-858.

REFERENCES

- Naito, F. Y., Melo, F. L., Fonseca, M. E. N., Santos, C. A., Chanes, C. R., Ribeiro, B. M., Gilbertson, R. L., Boiteux, L. S., Pereira-Carvalho & de Cássia Pereira-Carvalho, R. (2019). Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. *Archives of Virology*, *164*, 1907-1910.
- Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annual Review of Phytopathology*, *49*, 219-248.
- Nawaz-ul-Rehman, M. S., Nahid, N., Mansoor, S., Briddon, R. W., & Fauquet, C. M. (2010). Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology*, *405*(2), 300-308.
- Ng, J. C., & Falk, B. W. (2006). Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annual Review of Phytopathology*, *44*, 183-212.
- 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*(4), e19050.
- Nogueira, A. M., Nascimento, M. B., Barbosa, T. M., Quadros, A. F., Gomes, J. P. A., Orílio, A. F., Barros, D. R. & Zerbini, F. M. (2021). The association between New World alphasatellites and bipartite begomoviruses: effects on infection and vector transmission. *Pathogens*, *10*(10), 1244.
- O'Donnell, J., Gallagher, R. V., Wilson, P. D., Downey, P. O., Hughes, L., & Leishman, M. R. (2012). Invasion hotspots for non-native plants in Australia under current and future climates. *Global Change Biology*, *18*(2), 617-629.
- O'Malley, M. A. (2016). The ecological virus. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *59*, 71-79.
- O'Mara, F. P. (2012). The role of grasslands in food security and climate change. *Annals of Botany*, *110*(6), 1263-1270.
- Oluwafemi, S., Bruce, T. J., Pickett, J. A., Ton, J., & Birkett, M. A. (2011). Behavioral responses of the leafhopper, *Cicadulina storeyi* China, a major vector of maize streak virus, to volatile cues from intact and leafhopper-damaged maize. *Journal of Chemical Ecology*, *37*, 40-48.
- Oluwafemi, S., & Jackai, L. E. N. (2013). Stylet penetration behaviours of four *Cicadulina* leafhoppers on healthy and maize streak virus infected maize seedlings. *African Crop Science Journal*, *21*(2), 161-172.
- Oluwafemi, S., Kraberger, S., Shepherd, D.N., Martin, D. P. & Varsani, A. (2014). A high degree of African streak virus diversity within Nigerian maize fields includes a new mastrevirus from *Axonopus compressus*. *Archives of Virology*, *159*, 2765-2770.
- Ostfeld, R. S., & Keesing, F. (2000). Biodiversity and disease risk: the case of Lyme disease. *Conservation Biology*, *14*(3), 722-728.
- Ouattara, A., Tiendrébéogo, F., Becker, N., Urbino, C., Thébaud, G., Hoareau, M., Allibert, A., Chiroleu, F., Vernerey, M.-S., Traoré, E. V., Barro, N., Traoré, O., Lefeuvre & Lett, J. M. (2022). Synergy between an emerging monopartite begomovirus and a DNA-B component. *Scientific Reports*, *12*(1), 695.
- 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*(11), 3154-3165.
- Pagán, I., & García-Arenal, F. (2018). Tolerance to plant pathogens: theory and experimental evidence. *International Journal of Molecular Sciences*, *19*(3), 810.
- Pagán, I., Alonso-Blanco, C., & García-Arenal, F. (2008). Host responses in life-history traits and tolerance to virus infection in *Arabidopsis thaliana*. *PLoS Pathogens*, *4*(8), e1000124.
- Pande, D., Kraberger, S., Lefeuvre, P., Lett, J. M., Shepherd, D. N., Varsani, A., & Martin, D. P. (2012). A novel maize-infecting mastrevirus from La Reunion Island. *Archives of Virology*, *157*, 1617-1621.

REFERENCES

- Pande, D., Madzokere, E., Hartnady, P., Kraberger, S., Hadfield, J., Rosario, K., Jäschke, A., Monjane, A. L., Owor, B. E., Dida, M. M., Shepherd, D. N., Martin, D. P., Varsani, A. & Harkins, G. W. (2017). The role of Kenya in the trans-African spread of maize streak virus strain A. *Virus Research*, 232, 69-76.
- 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(2), 148-157.
- Parratt, S. R., Numminen, E., & Laine, A. L. (2016). Infectious disease dynamics in heterogeneous landscapes. *Annual Review of Ecology, Evolution, and Systematics*, 47, 283-306.
- Patil, B. L., & Dasgupta, I. (2006). Defective interfering DNAs of plant viruses. *Critical Reviews in Plant Sciences*, 25(1), 47-64.
- Paudel, D. B., & Sanfaçon, H. (2018). Exploring the diversity of mechanisms associated with plant tolerance to virus infection. *Frontiers in Plant Science*, 9, 1575.
- 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(1781), 20133374.
- Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., & González-Candelas, F. (2015). Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. *Infection, Genetics and Evolution*, 30, 296-307.
- Picard, C., Dallot, S., Bruncker, K., Berthier, K., Roumagnac, P., Soubeyrand, S., Jacquot, E. & Thébaud, G. (2017). Exploiting genetic information to trace plant virus dispersal in landscapes. *Annual Review of Phytopathology*, 55, 139-160.
- Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'connell, C., E. Wong, Russel, L., Zern, J., Aquino, T. & Tsomondo, T. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, 84(1), 1-20.
- Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S., & Fauquet, C. M. (2001). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology*, 82(3), 655-665.
- Pointeau, S., Jaguenet, E., Couty, A., Dubois, F., Rambaud, C., & Ameline, A. (2014). Differential performance and behavior of the corn leaf aphid, *Rhopalosiphum maidis*, on three species of the biomass crop miscanthus. *Industrial Crops and Products*, 54, 135-141.
- Pooggin, M. M. (2016). Role of small RNAs in virus-host interaction. *Plant-virus interactions: molecular biology, intra-and intercellular transport*, 161-189.
- Pooggin, M. M. (2018). Small RNA-omics for plant virus identification, virome reconstruction, and antiviral defense characterization. *Frontiers in Microbiology*, 9, 2779.
- Power, A. G., & Flecker, A. S. (2003). Virus specificity in disease systems: are species redundant. *The importance of species: perspectives on expendability and triage*. Princeton University Press, Princeton, 330-346.
- Power, A. G., & Mitchell, C. E. (2004). Pathogen spillover in disease epidemics. *The American Naturalist*, 164(S5), S79-S89.
- Power, A. G. (2008). Community ecology of plant viruses. In *Plant Virus Evolution* (pp. 15-26). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Preiss, W., & Jeske, H. (2003). Multitasking in replication is common among geminiviruses. *Journal of Virology*, 77(5), 2972-2980.
- Prendeville, H. R., Ye, X., Jack Morris, T., & Pilson, D. (2012). Virus infections in wild plant populations are both frequent and often unapparent. *American Journal of Botany*, 99(6), 1033-1042.

REFERENCES

- Pyšek, P., Jarošík, V., Hulme, P. E., Pergl, J., Hejda, M., Schaffner, U., & Vilà, M. (2012). A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. *Global Change Biology*, *18*(5), 1725-1737.
- Pyšek, P., Pergl, J., Essl, F., Lenzner, B., Dawson, W., Kreft, H., Weigelt, P., inter, M., Kartesz, J., Nichino, M., Antonova, L. A., Barcelona, J. F., Cabezas, F., J., Cardenas, D., Cardenas-Toro, J., Castano, N., Chacon, E., Chatelain, C., Dullinger, S., Ebel, A. L., Figueiredo, E., Fuentes, N., Genovesi, P., Groom, Q. J., Henderson, L., Inderjit, Kupriyanov, A., Masciadri, S., Maurel, N., Meerman, J., Morozova, O., Moser, D. Nickrent, D., Nowak, P. M., Pagad, S., Patzelt, A., Pelser, P. B., Seebens, H., Shu, W.-s., Thomas, J., Velayos, M., Weber, E., Wiereinga, J. J., Baptiste, M. P. & Kleunen, M. V. (2017). Naturalized alien flora of the world. *Preslia*, *89*(3), 203-274.
- Pyšek, P., Hulme, P. E., Simberloff, D., Bacher, S., Blackburn, T. M., Carlton, J. T., Dawson, W., Essl, F., Foxcroft, L. C., Genovesi, P., Jeschke, J. M., Kühne, I., Liebhold, A. M., Mandrak, N. E., Meyerson, L. A., Pauchard, A., Pergl, J., Roy, H. E., Seebens, H., van Kleunen, M., Vilà, M., Wingfield, M. J. & Richardson, D. M. (2020). Scientists' warning on invasive alien species. *Biological Reviews*, *95*(6), 1511-1534.
- Qiu, W., & Moyer, J. W. (1999). Tomato spotted wilt tospovirus adapts to the TSWV N gene-derived resistance by genome reassortment. *Phytopathology*, *89*(7), 575-582.
- Råberg, L., Graham, A. L., & Read, A. F. (2009). Decomposing health: tolerance and resistance to parasites in animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1513), 37-49.
- Remold, S. K. (2002). Unapparent virus infection and host fitness in three weedy grass species. *Journal of Ecology*, *96*7-977.
- Remold, S. (2012). Understanding specialism when the jack of all trades can be the master of all. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4861-4869.
- Ren, R., Zheng, L., Han, J., Perdoncini Carvalho, C., Miyashita, S., Zhang, D., & Qu, F. (2023). Intracellular bottlenecks permits no more than three tomato yellow leaf curl virus genomes to initiate replication in a single cell. *PLoS Pathogens*, *19*(5), e1011365.
- 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): Approches biologique, génétique et épidémiologique de la relation vecteur-virus-plante (Doctoral dissertation).
- Richardson, D. M., Allsopp, N., D'Antonio, C. M., Milton, S. J., & Rejmánek, M. (2000). Plant invasions—the role of mutualisms. *Biological Reviews*, *75*(1), 65-93.
- Riedel, C., Habekuß, A., Schliephake, E., Niks, R., Broer, I., & Ordon, F. (2011). Pyramiding of Ryd2 and Ryd3 conferring tolerance to a German isolate of Barley yellow dwarf virus-PAV (BYDV-PAV-ASL-1) leads to quantitative resistance against this isolate. *Theoretical and applied genetics*, *123*(1), 69-76.
- Rieux, A., Campos, P., Duvermy, A., Scussel, S., Martin, D., Gaudeul, M., Lefeuvre, P., Becker, N. & Lett, J. M. (2021). Contribution of historical herbarium small RNAs to the reconstruction of a cassava mosaic geminivirus evolutionary history. *Scientific Reports*, *11*(1), 21280.
- Rodelo-Urrego, M., Pagán, I., González-Jara, P., Betancourt, M., Moreno-Letelier, A., Ayllón, M. A., Fraile, A., Pinero, D. & García-Arenal, F. (2013). Landscape heterogeneity shapes host-parasite interactions and results in apparent plant–virus codivergence. *Molecular Ecology*, *22*(8), 2325-2340.
- Rodriguez-Valera, F., Martin-Cuadrado, A. B., Rodriguez-Brito, B., Pasic, L., Thingstad, T. F., Rohwer, F., & Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nature Precedings*, 1-1.
- Rohwer, F., & Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature*, *459*(7244), 207-212.
- 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., Macedo, M. A., Maliano, M. R., Soto-Aguilar, M., Souza, J. O., Briddon, R. W., Kenyon, L., Rivera Bustamante, R. F., Zerbini, F. M., Adkins S., Legg, J., P., Kvarnheden, A., Wintermantel, W. M., Sudarshana, M. R.,

REFERENCES

- Peterschmitt, M., Lapidot, M., Martin, D. P., Moriones, E., Inoue-Nagata, A. K. & Gilbertson, R. L. (2018). World management of geminiviruses. *Annual Review of Phytopathology*, *56*, 637-677.
- Roossinck, M. J. (2010). Lifestyles of plant viruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1548), 1899-1905.
- Roossinck, M. J., Saha, P., Wiley, G. B., Quan, J., White, J. D., Lai, H., Chavarria, F., Shen, G. & Roe, B. A. (2010). Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*, *19*, 81-88.
- Roossinck, M. J. (2011). The big unknown: plant virus biodiversity. *Current Opinion in Virology*, *1*(1), 63-67.
- Roossinck, M. J. (2012a). Persistent plant viruses: molecular hitchhikers or epigenetic elements?. In *Viruses: Essential Agents of Life* (pp. 177-186). Dordrecht: Springer Netherlands.
- Roossinck, M. J. (2012b). Plant virus metagenomics: biodiversity and ecology. *Annual Review of Genetics*, *46*, 359-369.
- Roossinck, M. J. (2015). Plants, viruses and the environment: Ecology and mutualism. *Virology*, *479*, 271-277.
- 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*(6), 716-727.
- Roossinck, M. J., & Bazán, E. R. (2017). Symbiosis: viruses as intimate partners. *Annual Review of Virology*, *4*, 123-139.
- Rosario, K., Duffy, S., & Breitbart, M. (2012). A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Archives of Virology*, *157*, 1851-1871.
- Rosario, K., Schenck, R. O., Harbeitner, R. C., Lawler, S. N., & Breitbart, M. (2015). Novel circular single-stranded DNA viruses identified in marine invertebrates reveal high sequence diversity and consistent predicted intrinsic disorder patterns within putative structural proteins. *Frontiers in Microbiology*, *6*, 696.
- Roumagnac, P., Lett, J. M., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F. M., Martin, D. P., & Varsani, A. (2022). Establishment of five new genera in the family Geminiviridae: Citlodavirus, Maldovirus, Mulcrilevirus, Opunvirus, and Topilevirus. *Archives of Virology*, 1-16.
- 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*(1), 2700.
- Roux, S., Emerson, J. B., Eloë-Fadrosch, E. A., & Sullivan, M. B. (2017). Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. *PeerJ*, *5*, e3817.
- Rúa, M. A., Pollina, E. C., Power, A. G., & Mitchell, C. E. (2011). The role of viruses in biological invasions: friend or foe?. *Current Opinion in Virology*, *1*(1), 68-72.
- Sacristán, S., Malpica, J. M., Fraile, A., & García-Arenal, F. (2003). Estimation of population bottlenecks during systemic movement of Tobacco mosaic virus in tobacco plants. *Journal of Virology*, *77*(18), 9906-9911.
- Sacristán, S., Díaz, M., Fraile, A., & García-Arenal, F. (2011). Contact transmission of Tobacco mosaic virus: a quantitative analysis of parameters relevant for virus evolution. *Journal of Virology*, *85*(10), 4974-4981.
- Samarakoon, H., Ferguson, J. M., Gamaarachchi, H., & Deveson, I. W. (2023). Accelerated nanopore basecalling with SLOW5 data format. *Bioinformatics*, *39*(6), btad352.
- Sanjuán, R., Moya, A., & Elena, S. F. (2004). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proceedings of the National Academy of Sciences*, *101*(22), 8396-8401.

REFERENCES

- Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, 73, 4433-4448.
- 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(4), 695-710.
- Saunders, D. A., Hobbs, R. J., & Margules, C. R. (1991). Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, 5(1), 18-32.
- 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(1), 142-152.
- Saunders, K., Bedford, I. D., Yahara, T., & Stanley, J. (2003). The earliest recorded plant virus disease. *Nature*, 422(6934), 831-831.
- Saunders, K., Richardson, J., Lawson, D. M., & Lomonosoff, G. P. (2020). Requirements for the packaging of geminivirus circular single-stranded DNA: effect of DNA length and coat protein sequence. *Viruses*, 12(11), 1235.
- Schalamun, M., Kainer, D., Beavan, E., Nagar, R., Eccles, D., Rathjen, J. P., Lanfear, R. & Schwessinger, B. (2018). A comprehensive toolkit to enable MinION sequencing in any laboratory. *BioRxiv*, 289579.
- Schalk, H. J., Matzeit, V., Schiller, B., Schell, J., & Gronenborn, B. (1989). Wheat dwarf virus, a geminivirus of graminaceous plants needs splicing for replication. *The EMBO Journal*, 8(2), 359-364.
- Seifert, K. A. (2009). Progress towards DNA barcoding of fungi. *Molecular Ecology Resources*, 9, 83-89.
- Shapiro, L. R., Salvaudon, L., Mauck, K. E., Pulido, H., De Moraes, C. M., Stephenson, A. G., & Mescher, M. C. (2013). Disease interactions in a shared host plant: effects of pre-existing viral infection on cucurbit plant defense responses and resistance to bacterial wilt disease. *PLoS One*, 8(10), e77393.
- Shepherd, D. N., Varsani, A., Windram, O. P., Lefeuvre, P., Monjane, A. L., Owor, B. E., & Martin, D. P. (2008). Novel sugarcane streak and sugarcane streak Reunion mastreviruses from southern Africa and La Reunion. *Archives of Virology*, 153, 605-609.
- Shepherd, D. N., Martin, D. P., Van der Walt, E., Dent, K., Varsani, A., & Rybicki, E. P. (2010). Maize streak virus: an old and complex 'emerging' pathogen. *Molecular Plant Pathology*, 11(1), 1-12.
- Sherwood, S. G. (1997). Little things mean a lot: Working with Central American farmers to address the mystery of plant disease. *Agriculture and Human Values*, 14(2), 181-189.
- Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., Qin, X. C., Li, J., Cao, J. P., Eden, J. S., Buchmann, J., Wand, W., Xu, J., Holmes, E. C. & Zhang, Y. Z. (2016). Redefining the invertebrate RNA virosphere. *Nature*, 540(7634), 539-543.
- Shrestha, B. B., & Shrestha, K. K. (2021). Invasions of alien plant species in Nepal: Patterns and process. *Invasive Alien Species: Observations and Issues from Around the World*, 2, 168-183.
- Sicard, A., Pirolles, E., Gallet, R., Vernerey, M. S., Yvon, M., Urbino, C., Peterschmitt, M., Gutierrez, S., Michalakis, Y. & Blanc, S. (2019). A multicellular way of life for a multipartite virus. *Elife*, 8, e43599.
- Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., Galil, B., Garcia-Berthou, E., Pascal, M., Pyšek, P., Sousa, R., Tabacchi, E. & Vilà, M. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*, 28(1), 58-66.
- Simmonds, P. (2015). Methods for virus classification and the challenge of incorporating metagenomic sequence data. *Journal of General Virology*, 96(6), 1193-1206.
- Simmonds, P., Adams, M. J., Benkő, M., Breitbart, M., Brister, J. R., Carstens, E. B., Davison, A. J., Delwart, E., Gorbalenya, A. E., Harrach, B., Hull, R., King, A. M. Q., Koonin, E. V., Krupovic, M., Kuhn, J. H., Lefkowitz, E. J., Nibert, M. L., Orton, R., Roossinck, M. J., Sabanadzovic, S., Sullivan, M. B., Suttle, C. A., Tesh, R. B.,

REFERENCES

- van der Vlugt, R. A., Varsani, A. & Zerbini, F. M. (2017). Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology*, 15(3), 161-168.
- Simmonds, P., Aiewsakun, P., & Katzourakis, A. (2019). Reply to 'evolutionary stasis of viruses?'. *Nature Reviews Microbiology*, 17(5), 329-330.
- Simmonds, P., Adriaenssens, E. M., Zerbini, F. M., Abrescia, N. G., Aiewsakun, P., Alfenas-Zerbini, P., Bao, Y., Barylski, J., Drosten, C., Duffy, S., Duprex, W. P., Dutilh, B. E., Elena, S. F., Garcia, M. L., Junglen, S., Katzourakis, A., Koonin, E. V., Krupovic, M., Kuhn, J. H., Lambert, A. J., Lefkowitz, E. J., Lobočka, M., Lood, C., Mahony, J., Meier-Kolthoff, J. P., Musheigna, A. R., Oksanen, H. M., Poranen, M. M., Reyes-Munoz, A., Robertson, D. L., Roux, S., Rubino, L., Sabanadzovic, S., Siddell, S., Skern, T., Smith, D. B., Sullivan, M. B., Suzuki, N., Turner, D., Van Doorslaer, L., Vandamme, A.-M., Varsani, A., & Vasilakis, N. (2023). Four principles to establish a universal virus taxonomy. *PLoS Biology*, 21(2), e3001922.
- Simon, H. Y., Siddle, K. J., Park, D. J., & Sabeti, P. C. (2019). Benchmarking metagenomics tools for taxonomic classification. *Cell*, 178(4), 779-794.
- Singh, R. P., Burnett, P. A., Albarran, M., & Rajaram, S. (1993). Bdv1: a gene for tolerance to barley yellow dwarf virus in bread wheats. *Crop Science*, 33(2), 231-234.
- Smith, O., Clapham, A., Rose, P., Liu, Y., Wang, J., & Allaby, R. G. (2014). A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. *Scientific Reports*, 4(1), 4003.
- Song, S., Sun, P., Ju, M., Zhang, X., Yang, X., Zhang, L., & Li, Z. (2022). Complete genomic characterization of a reassortment isolate of Sophora yellow stunt virus from *Sophora alopecuroides* in China. *Journal of Plant Pathology*, 104(2), 781-786.
- Stackebrandt, E., & Goebel, B. M. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology*, 44(4), 846-849.
- Stainton, D., Kraberger, S., Walters, M., Wiltshire, E. J., Rosario, K., Halafihi, M. I., Lolohea, S., Katoa, I., Faitua, T. H., Aholelei, W., Taufu, L., Thomas, J. E., Collings, D. A., Martin, D. P. & Varsani, A. (2012). Evidence of inter-component recombination, intra-component recombination and reassortment in banana bunchy top virus. *Journal of General Virology*, 93(5), 1103-1119.
- Stanley, J., & Townsend, R. (1985). Characterisation of DNA forms associated with cassava latent virus infection. *Nucleic Acids Research*, 13(7), 2189-2206.
- Stinchcombe, J. R., & Hoekstra, H. E. (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, 100(2), 158-170.
- Storey, H. H. (1928). Transmission studies of maize streak disease 1. *Annals of Applied Biology*, 15(1), 1-25.
- Strasberg, D., Rouget, M., Richardson, D. M., Baret, S., Dupont, J., & Cowling, R. M. (2005). An assessment of habitat diversity and transformation on La Réunion Island (Mascarene Islands, Indian Ocean) as a basis for identifying broad-scale conservation priorities. *Biodiversity & Conservation*, 14, 3015-3032.
- Stukenbrock, E. H., & McDonald, B. A. (2008). The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, 46, 75-100.
- Sullivan, M. B., Weitz, J. S., & Wilhelm, S. (2017). Viral ecology comes of age. *Environmental Microbiology Reports*, 9(1), 33-35.
- Sun, K., Liu, Y., Zhou, X., Yin, C., Zhang, P., Yang, Q., Mao, L., Shentu, X. & Yu, X. (2022). Nanopore sequencing technology and its application in plant virus diagnostics. *Frontiers in Microbiology*, 13, 939666.
- Susi, H., Filloux, D., Frilander, M. J., Roumagnac, P., & Laine, A. L. (2019). Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ*, 7, e6140.
- Susi, H., & Laine, A. L. (2021). Agricultural land use disrupts biodiversity mediation of virus infections in wild plant populations. *New Phytologist*, 230(6), 2447-2458.

- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nature Reviews Microbiology*, 5(10), 801-812.
- Suttle, C. A. (2013). Viruses: unlocking the greatest biodiversity on Earth. *Genome*, 56(10), 542-544.
- Svanella-Dumas, L., Candresse, T., Hulle, M., & Marais, A. (2013). Distribution of Barley yellow dwarf virus-PAV in the sub-Antarctic Kerguelen Islands and characterization of two new Luteovirus species. *Plos One*, 8(6), e67231.
- Syller, J. (2012). Facilitative and antagonistic interactions between plant viruses in mixed infections. *Molecular Plant Pathology*, 13(2), 204-216.
- Talbi, C., Lemey, P., Suchard, M. A., Abdelatif, E., Elharrak, M., Jalal, N., Faouzi, A., Echevarria, J. E., Moron, S. V., Rambaut, A., Campiz, N., Tatem, A. J., Holmes, E. C. & Bourhy, H. (2010). Phylodynamics and human-mediated dispersal of a zoonotic virus. *PLoS Pathogens*, 6(10), e1001166.
- Thébaud, C., Warren, B. H., Strasberg, D., & Cheke, A. (2009). Mascarene islands, biology. *Atoll Research Bulletin*, 127, 1-216.
- Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M., & Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature*, 423(6935), 74-77.
- Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Harimalala, M. A., De Bruyn, A., Villemot, J., Traoré, V. S., 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(1), 1-7.
- Tisza, M. J., Pastrana, D. V., Welch, N. L., Stewart, B., Peretti, A., Starrett, G. J., Pang, Y. Y. S., Krishnamurthy, S.R. Pesavento, P. A., McDermott, D. H., Murphy, P. M., Withed, J. L., Miller, B., Brenchley, J., Rosshart, S. P., Rehmann, B., Doorbar, J., Ta'ala, B. A., Pletnikova, O., Troncoso, J. C., Resnick, S. M., Bolduc, B., Sullivan, M. B., Varsani, A., Segall, A. M. & Buck, C. B. (2020). Discovery of several thousand highly diverse circular DNA viruses. *Elife*, 9, e51971.
- Tollenaere, C., Susi, H., & Laine, A. L. (2016). Evolutionary and epidemiological implications of multiple infection in plants. *Trends in Plant Science*, 21(1), 80-90.
- Torrance, L., & Taliaksky, M. E. (2020). Potato Virus Y emergence and evolution from the Andes of South America to become a major destructive pathogen of potato and other solanaceous crops worldwide. *Viruses*, 12(12), 1430.
- Tromas, N., Zwart, M. P., Poulain, M., & Elena, S. F. (2014). Estimation of the in vivo recombination rate for a plant RNA virus. *Journal of General Virology*, 95(3), 724-732.
- Tylianakis, J. M., & Morris, R. J. (2017). Ecological networks across environmental gradients. *Annual Review of Ecology, Evolution, and Systematics*, 48, 25-48.
- Uzest, M., Gargani, D., Drucker, M., Hébrard, E., Garzo, E., Candresse, T., Fereres, A. & Blanc, S. (2007). A protein key to plant virus transmission at the tip of the insect vector stylet. *Proceedings of the National Academy of Sciences*, 104(46), 17959-17964.
- Valverde, S., Vidiella, B., Montañez, R., Fraile, A., Sacristán, S., & García-Arenal, F. (2020). Coexistence of nestedness and modularity in host–pathogen infection networks. *Nature Ecology & Evolution*, 4(4), 568-577.
- Van Rensburg, G. D. J. (1992). Laboratory observations on the biology of *Cicadulina mbila* (Naudé)(Homoptera: Cicadellidae), a vector of maize streak disease. 2. The effect of selected host plants. *Phytophylactica*, 14(3), 109-112.
- Varma, A., & Malathi, V. G. (2003). Emerging geminivirus problems: a serious threat to crop production. *Annals of Applied Biology*, 142(2), 145-164.
- 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).

REFERENCES

- Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *The Journal of General Virology*, 89(Pt 9), 2063.
- Varsani, A., Monjane, A. L., Donaldson, L., Oluwafemi, S., Zinga, I., Komba, E. K., Plakoutene, D., Mandakombo, N., Mboukoulida, J., Semballa, S., Briddon, R. W., Markham, T. G., Lett, J.-M., Lefeuvre, P., Rybicki, E. P. & Martin, D. P. (2009). Comparative analysis of Panicum streak virus and Maize streak virus diversity, recombination patterns and phylogeography. *Virology Journal*, 6, 1-11.
- Varsani, A., Roumagnac, P., Fuchs, M., Navas-Castillo, J., Moriones, E., Idris, A., Briddon, R. W., Rivera-Bustamante, R., 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., Lefeuvre, P., Roumagnac, P., & Martin, D. (2018). Notes on recombination and reassortment in multipartite/segmented viruses. *Current opinion in virology*, 33, 156-166.
- Varsani, A., Martin, D. P., Randles, J. W., Vetten, H. J., Thomas, J. E., Fiallo-Olivé, E., Navas-Castillo, J., Lett, J.-M., Zerbini, F. M., Roumagnac, P. & Gronenborn, B. (2021). Taxonomy update for the family Alphsatellitidae: new subfamily, genera, and species. *Archives of Virology*, 166, 3503-3511.
- Viana, M., Mancy, R., Biek, R., Cleaveland, S., Cross, P. C., Lloyd-Smith, J. O., & Haydon, D. T. (2014). Assembling evidence for identifying reservoirs of infection. *Trends in Ecology & Evolution*, 29(5), 270-279.
- Viehweger, A., Krautwurst, S., Lamkiewicz, K., Madhugiri, R., Ziebuhr, J., Hölzer, M., & Marz, M. (2019). Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Research*, 29(9), 1545-1554.
- Vigne, E., Marmonier, A., & Fuchs, M. (2008). Multiple interspecies recombination events within RNA2 of Grapevine fanleaf virus and Arabis mosaic virus. *Archives of Virology*, 153(9), 1771-1776.
- Vilà, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarošík, V., Maron, J. L., Pergl, J., Schaffner, U., Sun, Y. & Pyšek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, 14(7), 702-708.
- Walls III, J., Rajotte, E., & Rosa, C. (2019). The past, present, and future of barley yellow dwarf management. *Agriculture*, 9(1), 23.
- Wang, Y., Gaba, V., Yang, J., Palukaitis, P., & Gal-On, A. (2002). Characterization of synergy between Cucumber mosaic virus and potyviruses in cucurbit hosts. *Phytopathology*, 92(1), 51-58.
- Wang, X. R., Wang, C., Ban, F. X., Ghanim, M., Pan, L. L., Qian, L. X., Liu, L. Q., Wang, X. W. & Liu, S. S. (2020). Apoptosis in a whitefly vector activated by a begomovirus enhances viral transmission. *Msystems*, 5(5), 10-1128.
- Wasik, B. R., & Turner, P. E. (2013). On the biological success of viruses. *Annual Review of Microbiology*, 67, 519-541.
- Webb, M. D. (1987). Species recognition in Cicadulina leafhoppers (Hemiptera: Cicadellidae), vectors of pathogens of Gramineae. *Bulletin of Entomological Research*, 77(4), 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(5), 729-742.
- Whitfield, A. E., & Rotenberg, D. (2015). Disruption of insect transmission of plant viruses. *Current Opinion in Insect Science*, 8, 79-87.
- Whitlock, M. C. (1996). The red queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. *The American Naturalist*, 148, S65-S77.
- Wilson, C. E., Castro, K. L., Thurston, G. B., & Sissons, A. (2016). Pathway risk analysis of weed seeds in imported grain: A Canadian perspective. *NeoBiota*, 30, 49-74.

REFERENCES

- Wisler, G. C., Duffus, J. E., Liu, H. Y., & Li, R. H. (1998). Ecology and epidemiology of whitefly-transmitted closteroviruses. *Plant Disease*, 82(3), 270-280.
- Woolhouse, M. E., Taylor, L. H., & Haydon, D. T. (2001). Population biology of multihost pathogens. *Science*, 292(5519), 1109-1112.
- Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, 32(4), 569-577.
- Woolhouse, M. E., & Gowtage-Sequeria, S. (2005). Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases*, 11(12), 1842.
- Wren, J. D., Roossinck, M. J., Nelson, R. S., Scheets, K., Palmer, M. W., & Melcher, U. (2006). Plant virus biodiversity and ecology. *PLoS biology*, 4(3), e80.
- Wright, E. A., Heckel, T., Groenendijk, J., Davies, J. W., & Boulton, M. I. (1997). Splicing features in maize streak virus virion-and complementary-sense gene expression. *The Plant Journal*, 12(6), 1285-1297.
- Wyant, P. S., Strohmeier, S., Schäfer, B., Krenz, B., Assunção, I. P., de Andrade Lima, G. S., & Jeske, H. (2012). Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. *Virology*, 427(2), 151-157.
- Xavier, C. A., & Whitfield, A. E. (2023). Plant virology. *Current Biology*, 33(11), R478-R484.
- Xu, P., Chen, F., Mannas, J. P., Feldman, T., Sumner, L. W., & Roossinck, M. J. (2008). Virus infection improves drought tolerance. *New Phytologist*, 180(4), 911-921.
- Xu, W., Li, H., Sivasithamparam, K., Tran, D. T., Jones, M. G., Chen, X., & Wylie, S. J. (2022). Spillover of a Tobamovirus from the Australian Indigenous Flora to Invasive Weeds. *Viruses*, 14(8), 1676.
- Yahaya, A., Dangora, D. B., Alegbejo, M. D., Kumar, P. L., & Alabi, O. J. (2017). Identification and molecular characterization of a novel sugarcane streak mastrevirus and an isolate of the A-strain of maize streak virus from sugarcane in Nigeria. *Archives of Virology*, 162, 597-602.
- Yang, S., & Hua, J. (2004). A haplotype-specific Resistance gene regulated by BONZAI1 mediates temperature-dependent growth control in Arabidopsis. *The Plant Cell*, 16(4), 1060-1071.
- Yang, H., Tang, M., & Gross, T. (2015). Large epidemic thresholds emerge in heterogeneous networks of heterogeneous nodes. *Scientific Reports*, 5(1), 13122.
- Yi, H., & Richards, E. J. (2007). A cluster of disease resistance genes in Arabidopsis is coordinately regulated by transcriptional activation and RNA silencing. *The Plant Cell*, 19(9), 2929-2939.
- Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A. & ICTV Report Consortium. (2017). ICTV virus taxonomy profile: Geminiviridae. *Journal of General Virology*, 98(2), 131-133.
- Zerbini, F. M., Siddell, S. G., Mushegian, A. R., Walker, P. J., Lefkowitz, E. J., Adriaenssens, E. M., Alfenas-Zerbini, P., Dutilh, B. E., Garica, M. L., Junglen, S., Krupovic, M., Kuhn, J. H., Lambert, A. J., Lobočka, M., Oksanen, H. M., Robertson, D. L., Rubino, L., Sabanadzovic, S., Simmonds, P., Suzuki, N., Van Doorslaer, K., Vandamme, A.-M. & Varsani, A. (2022). Differentiating between viruses and virus species by writing their names correctly. *Archives of Virology*, 167(4), 1231-1234.
- Zhang, X. S., Holt, J., & Colvin, J. (2000). Mathematical models of host plant infection by helper-dependent virus complexes: why are helper viruses always avirulent?. *Phytopathology*, 90(1), 85-93.
- 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(2), 471-477.
- Zhang, Y. Z., Chen, Y. M., Wang, W., Qin, X. C., & Holmes, E. C. (2019). Expanding the RNA virosphere by unbiased metagenomics. *Annual Review of Virology*, 6, 119-139.

REFERENCES

- Zhang, Z., Wang, X., Wei, X., Zheng, S. W., Lenhart, B. J., Xu, P., Li, J., Pan, J., Albrecht, H. & Liu, C. (2021). Multiplex quantitative detection of SARS-CoV-2 specific IgG and IgM antibodies based on DNA-assisted nanopore sensing. *Biosensors and Bioelectronics*, *181*, 113134.
- Zhao, W., Yang, P., Kang, L., & Cui, F. (2016). Different pathogenicities of Rice stripe virus from the insect vector and from viruliferous plants. *New Phytologist*, *210*(1), 196-207.
- Zhao, L., Zhang, H., Tian, W., & Xu, X. (2018). Identifying compartments in ecological networks based on energy channels. *Ecology and Evolution*, *8*(1), 309-318.
- 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*(8), 2101-2111.
- Zhou, X. (2013). Advances in understanding begomovirus satellites. *Annual Review of Phytopathology*, *51*, 357-381.
- Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. D., Liu, M. Q., Chen, Y., Shen, R., Wang, X., Zheng, X. S., Zhao, K, Chen, Q. J., Deng, F., Liu, L. L., Yan, B., Zhan, F. X., Wang, Y. Y., Xiao, G.G. & Shi, Z. L. (2020). Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *BioRxiv*, 2020-01.
- Zwart, M. P., & Elena, S. F. (2015). Matters of size: genetic bottlenecks in virus infection and their potential impact on evolution. *Annual Review of Virology*, *2*, 161-179.

Evolution and ecology of phytoviruses at the scale of Réunion Island ecosystems: impact of vegetation landscape and anthropization

The interactions carried out by phytoviruses with their hosts, other phytoviruses, and potentially their vectors are complex, ranging from mutualism to antagonism. To determine the factors explaining such variations in the spectrum of interactions performed by phytoviruses, it is necessary to integrate these communities into the broader context of ecosystems, thus involving a greater number of entities in these networks. Consequently, to determine the environmental factors involved and their influence on the ecological dynamics (*i.e.* interactions, evolution, and adaptation) of viral populations, it is essential to obtain a holistic view of the diversity and structure of viral communities (*i.e.* species assemblages) at the ecosystem level. However, natural environments are now strongly subject to anthropogenic disturbances, which are an integral part of the environmental factors impacting the ecological dynamics of ecosystems. For example, the simplification of environments through agriculture or the introduction of invasive alien species into new environments alters the composition and dynamics of plant communities and phytoviruses. Thus, despite the apparent difficulty of analyzing all factors influencing the richness and structure of viral communities, we have chosen to approach these questions through the lens of hosts at the scale of the insular ecosystems of Réunion.

To accurately assess these processes, we initially focused the study on the diversity of mastreviruses (*Geminiviridae* family) previously detected in Réunion, at the scale of all known host plants within the Poales order. This required the development and evaluation of a metagenomic approach based on third-generation high-throughput sequencing, dedicated to the detection of circular DNA viruses through the combined use of amplification methods (RCA) and long-read sequencing on MinION, and adapted for the multiplexing of numerous samples. Comparing the results to those obtained through a gold standard Sanger cloning and sequencing method, validated the quality of the assembled viral genomes and the utility of this protocol for exploring and characterizing circular DNA virus genomes.

The use of the RCA-MinION protocol on 273 pools of 61 Poales species sampled from three differently anthropized ecosystems (agroecosystems, coastal, and subalpine grasslands) enabled the identification of 92 genomes of circular ssDNA viruses. Among them, the genome of a chimeric virus resulting from recombination between a geminivirus and an alphasatellite stood out. Beyond, once again, demonstrating the high variability and strong evolutionary potential of viruses in the phylum *Cressdnaviricota*, the characterization of this virus genome and the study of its biology shed light on the role of multipartism and inter-species viral synergy in host adaptation. Additionally, a defective genome, an alphasatellite, and twelve species of the genus *Mastrevirus*, whether newly described or not, were molecularly characterized. Mastreviruses were predominantly identified in exotic annual *Poaceae* sampled in agro-ecosystems. The overall analysis of the plant-virus interaction network revealed a positive correlation between host and mastrevirus richness, consistent with an amplification effect. With the absence of an endemic mastrevirus community in Réunion, we highlighted an association between mastreviruses and alien plants linked to agricultural environments, suggesting that the origin of these viruses on the island is intimately linked to the introduction of infected plant material from the African continent.

Finally, we extended these analyses to all phytoviruses described on native and alien uncultivated herbaceous species for the subset of sites showing the strongest contrasts in their degrees of plant invasion. Through the combination of two broad-spectrum metagenomic approaches, we detected 28 genera from 18 families of phytoviruses, among which 45% appear to belong to previously undescribed viral species. Overall, native plants exhibit a greater viral richness than alien ones, consistent with the pathogen release hypothesis during introduction events. Consistently with mastreviruses, some viruses known to cause economically significant diseases were only detected on alien hosts collected in agro-ecosystems, suggesting that agroecosystems serve as preferred entry points for many phytoviruses in Réunion. Interestingly, the systematic detection of the barley yellow dwarf virus (BYDV; family *Tombusviridae*) on sweet vernal grass, *Anthoxanthum odoratum* (Family *Poaceae*), in the subalpine ecosystems of Réunion, and its absence in cultivated ecosystems, could suggest a host-virus mutualism phenomenon and the potential, if not essential, role of this virus in the successful invasion of this alien plant in these ecosystems.

The study of ecological factors influencing the richness and structure of viral communities in Réunion has led to a better understanding of these communities and plant-virus interactions according to the level of anthropogenic disturbance in these insular ecosystems. Our findings highlight the diversity and complexity of virus-host interactions depending on the environment and within a symbiotic continuum ranging from mutualism to antagonism.

Évolution et écologie des phytovirus à l'échelle des écosystèmes réunionnais : impact du paysage végétal et de l'anthropisation

Les interactions réalisées par les phytovirus avec leurs hôtes, les autres phytovirus, et éventuellement leurs vecteurs sont complexes, variant du mutualisme à l'antagonisme. Afin de déterminer les facteurs expliquant de telles variations dans le spectre des interactions réalisées par les phytovirus, il est nécessaire d'intégrer ces communautés dans le contexte plus global des écosystèmes, faisant ainsi intervenir un nombre plus conséquent d'entités dans ces réseaux. Par conséquent, afin de déterminer les facteurs environnementaux impliqués et l'influence qu'ils ont sur les dynamiques écologiques (*i.e.* interactions, évolution et adaptation) des populations virales, il apparaît indispensable d'obtenir une vision holistique de la diversité et de la structure des communautés (*i.e.* assemblage des espèces) virales à l'échelle des écosystèmes. Cependant, les milieux naturels sont aujourd'hui fortement soumis aux perturbations anthropiques, faisant partie intégrante des facteurs environnementaux impactant la dynamique écologique des écosystèmes. Par exemple, la simplification des milieux par l'agriculture ou l'introduction d'espèces exotiques envahissantes dans de nouveaux environnements, modifient la composition et la dynamique des communautés végétales et par extension les phytovirus qu'elles hébergent. Ainsi, malgré la difficulté apparente d'analyser l'intégralité des facteurs influençant la richesse et la structure des communautés virales, nous avons choisi d'aborder ces questions par le prisme des hôtes à l'échelle d'écosystèmes insulaires de La Réunion.

Afin d'évaluer avec précision ces processus, nous avons focalisé, dans un premier temps l'étude à la diversité des mastrévirus (famille *Geminiviridae*) précédemment détectés à La Réunion, à l'échelle de l'ensemble des plantes hôtes de l'ordre des Poales. Cela a nécessité l'élaboration et l'évaluation d'une approche métagénomique basée sur le séquençage haut débit de troisième génération, dédiée à la détection de virus à ADN circulaire par l'utilisation combinée de méthodes d'amplification (RCA) et de séquençage *long reads* sur MinION, et adaptée au multiplexage de nombreux échantillons. La comparaison des résultats à ceux obtenus par une méthode classique de clonage-séquençage Sanger a permis de valider la qualité des génomes viraux assemblés et l'utilisation de ce protocole pour l'exploration et la caractérisation des génomes de virus à ADN circulaire.

L'utilisation du protocole RCA-MinION sur 273 pools de 61 espèces de Poales échantillonnées sur trois types d'écosystèmes différemment anthropisés (agro-écosystèmes, prairies littorales et altimontaines) a permis d'identifier 92 génomes de virus à ADN circulaire. Parmi eux, le génome d'un virus chimérique issu de la recombinaison entre un geminivirus et un alphasatellite s'est distingué. Au-delà de démontrer une nouvelle fois la grande variabilité et le fort potentiel d'évolution des virus du phylum *Cressdnaviricota*, la caractérisation du génome de ce virus et l'étude de sa biologie nous éclairent sur le rôle du multipartisme et de la synergie virale inter-espèces dans l'adaptation à l'hôte. Par ailleurs, douze espèces du genre *Mastrevirus*, nouvellement décrites ou non, un génome défectif et un alphasatellite ont été caractérisés moléculairement. Les mastrévirus ont été identifiés en majorité sur des *Poaceae* annuelles exotiques échantillonnées dans les agro-écosystèmes. L'analyse globale du réseau d'interactions plantes-virus a révélé une corrélation positive entre la richesse en hôtes et en mastrévirus, compatible avec un effet d'amplification. Avec l'absence d'une communauté virale résidente à La Réunion, nous avons mis en lumière une association entre les mastrévirus et les plantes exotiques inféodées aux milieux agricoles, suggérant que l'origine de ces virus sur l'île est intimement liée à l'introduction de matériel végétal infecté depuis le continent Africain.

Enfin, nous avons étendu ces analyses à l'ensemble des phytovirus décrits sur des espèces herbacées non cultivées indigènes et exotiques pour la fraction des sites présentant les contrastes les plus forts dans leurs degrés d'invasion végétale. Grâce à la combinaison de deux approches métagénomiques généralistes, nous avons détecté 28 genres de 18 familles de phytovirus parmi lesquelles, 45% semblent appartenir à des espèces virales jamais décrites. Globalement, les plantes indigènes présentent une plus grande richesse virale que les exotiques, compatible avec l'hypothèse du relâchement de la pression parasitaire lors des événements d'introduction. De plus, certains virus connus comme responsables de maladies d'importance économique ont été détectés sur des hôtes exotiques collectés dans les agro-écosystèmes. Comme pour les mastrévirus, ces descriptions semblent confirmer que les agro-écosystèmes représentent les portes d'entrée privilégiées de nombreux phytovirus à La Réunion. Enfin, la détection du barley yellow dwarf virus (BYDV ; famille des *Tombusviridae*) sur flouve odorante, *Anthoxanthum odoratum* (Famille *Poaceae*), dans les écosystèmes altimontains de La Réunion, et son absence dans les agro-écosystèmes, pourraient suggérer un phénomène de mutualisme hôte-virus et le rôle potentiel, voire essentiel, de ce virus dans le succès de l'invasion de cette plante exotique dans ces écosystèmes.

L'étude des facteurs écologiques influençant la richesse et la structure des communautés virales dans un environnement insulaire a permis une meilleure description de ces communautés et des interactions plantes-virus selon le niveau de perturbation anthropique des écosystèmes réunionnais. L'ensemble de nos résultats soulignent la diversité et la complexité des interactions virus-hôte selon l'environnement et dans un continuum symbiotique entre mutualisme et antagonisme.