Réassortiments chez un phytovirus octopartite l'étude du faba bean necrotic stunt virus (FBNSV) du genre Nanovirus

Babil Torralba

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR
DE L'INSTITUT AGRO MONTPELLIER
ET DE L'UNIVERSITE DE MONTPELLIER

En Biologie des Interactions :
mécanismes des interactions parasitaires, pathogènes et symbiotiques

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Portée par
Unité de recherche PHIM – Plant Health Institute of Montpellier

Exploring Reassortments in an Octopartite Phytovirus: A Study of the *Faba Bean Necrotic Stunt Virus* (FBNSV) from the *Nanovirus* Genus

Présentée par Babil TORRALBA
Le 13 Novembre 2023

Sous la direction de
Yannis MICHALAKIS et Stéphane BLANC

Devant le jury composé de

M. Benoît MOURY, Directeur de recherche, INRAE, UR 0407, Unité Pathologie Végétale Avignon	Rapporteur
M. Romain VOLMER, Professeur, UMR IHAP, École Nationale Vétérinaire de Toulouse	Rapporteur
Mme. Ioana GRIGORAS-POPESCU, Maître de conférences, UMR 8030, Université d'Évry Val-d'Essonne	Examinatrice
M. Fernando GARCÍA-ARENAL RODRIGUEZ, Professeur, UPM INIA, Université Polytechnique de Madrid	Examinateur
Mme. Anne-Nathalie VOLKOFF, Directrice de recherche, INRAE, UMR DGIMI	Examinatrice
M. Yannis MICHALAKIS, Directeur de recherche, CNRS, UMR MIVEGEC	Co-directeur de thèse
M. Stéphane BLANC, Directeur de recherche, INRAE, UMR PHIM	Invité

UNIVERSITÉ DE MONTPELLIER

L'INSTITUT agro Montpellier

Remerciements

Avant toute chose je tiens à remercier **Philippe Rousseau** sans qui cette production n'aurait jamais vu le jour, je ne te remercierai jamais assez de m'avoir permis de me relancer et de m'avoir entendu à travers cette lettre de motivation. Je veux aussi exprimer un profond remerciement envers **Matthieu Arlat**, tu as été déterminant à bien des égards, humainement et scientifiquement, tu m'as largement inspiré par ta bienveillance, ta curiosité, ta pédagogie et ta compétence. En plus de tout le reste, tu es entièrement responsable de ma réception de cette proposition de thèse attisant tout de suite mon intérêt.

Je veux rendre grâce à **Yannis Michalakis** et **Stéphane Blanc** pour m'avoir fait confiance sur ce projet, pour la liberté, si chère à mes yeux, que vous m'avez accordée. Merci pour votre humanité, votre bienveillance, votre écoute, votre légèreté parfois, votre expertise souvent, votre compétence. Merci pour cet espace dans lequel j'ai pu m'exprimer en paix. Vous avoir comme superviseurs a été particulièrement précieux et enrichissant à mes yeux, sur bien des aspects.

Je souhaite manifester ma reconnaissance envers les membres de mon comité de suivi, **Stéphanie Bedhomme**, **SerafÍn Gutiérrez**, **Fernando GarcÍa-Arenal Rodriguez** et **Philippe Roumagnac,** qui ont jugé mon travail pendant ces multiples années et participé à me recentrer sur le sujet. Merci pour votre bienveillance palpable. Je remercie aussi chaleureusement mon jury de thèse, les rapporteurs **Benoît Moury** et **Romain Volmer**, ainsi que **Anne-Nathalie Volkoff**, **Ioana Grigoras-Popescu** et **Fernando** again, qui ont accepté de juger ce travail en manifestant un véritable intérêt pour cette production.

Merci **Michel Amphoux** pour ta bonne humeur, ton enthousiasme, ta gentillesse, ton dévouement pour créer du lien et intégrer les nouveaux arrivants. Merci d'avoir supporté en partie mon manque de rigueur administratif sous toutes ses formes. Merci pour **Myléna**, même si cet épisode n'a pas fini sous les meilleurs auspices. Merci aux personnalités rencontrées au sein de PHIM et ailleurs. **Claire Neema** pour ta fougue et ta présence, **Jean-Benoît Morel** pour t'avoir tenu la jambe en parlant de démocratie, **Henri Adreit** pour ta générosité, ton intégration, le vélo, le potager. Merci à tous ceux avec qui j'ai pris plaisir à échanger avec bienveillance, profondeur, scientifiquement, humainement, et tout à la fois, **Serge**, **Martine**, **Philippe**, **Dominique**, **Mikhail**, **Emmanuel**, **Gael**, **Michel P**, **Marie B**, **Nils, Laurence** et j'en oublie certainement.

Merci à toute l'équipe SIX, VIP, MULTI, w/e, pour votre expertise inestimable, la générosité et bienveillance que j'ai ressentie en chacun d'entre vous. Merci **Sophie Leblaye** de ton soutien indispensable dans la gestion des plantes, pucerons et semis potagers, **Elodie Pirolles** pour ton soutien déterminant aux manipulations en laboratoire, **Marie-Stéphanie Vernerey** pour ton aide fondatrice dans l'utilisation de la microscopie, et enfin **Michel Yvon** pour ton dynamisme, ta prévenance, ton savoir et savoir-faire presque inépuisables. Tout simplement pour la Force que tu incarnes. Merci **Jean-Louis Zeddam** pour ton humour subtil, **Cica Urbino** pour nos discussions enrichissantes, **Manuella Van Munster** pour ta bienveillance et ton énergie et **Maryline Uzest** pour ta sincère disponibilité.

Merci **Andy Brousse** pour toutes ces heures de travail acharné à couper des centaines de tubes de façon précaire, mesurer des milliers de plantes dans la fraicheur de saines matinées, compter des dizaines de milliers de pucerons en toute propreté. Les qPCR. Sincèrement merci pour tout ce travail et surtout merci pour ta bonne humeur chère à mes yeux, ta vibe tout simplement, au sein du labo et à l'extérieur. Une bonne partie de ce manuscrit t'est particulièrement redevable. Merci à **Prune Lacôte** pour ton implication et autant d'heures de travail méticuleux à génocider pucerons, plantes et bien plus encore. C'est beau la Recherche. De même une grosse partie de ce manuscrit repose sur ton investissement. J'espère que cette expérience vous aura autant enrichi que moi.

Thank you, **Sasha Greshnova**, into bringing the magnificent winter light and delicate melting snowflakes on my tidy and cleaned desk. It was really enriching for me to meet and work with you. I enjoyed our discussions, all the failed mockery about my perfect English, all your impressive prowess at bouldering, discovering opera and your love for frogs, hedgehogs and ducks. Obviously for your work, that helped me a lot on R and statistics, and your important part in the project. Until next time Yozhik.

Merci **Mélia Bonnamy** pour ces années à partager réunions et projets respectifs, c'était très intéressant de suivre ton évolution, ton organisation et ton travail. C'était sincèrement inspirant d'être le témoin de ta pugnacité et de ton dévouement envers ce gros morceau qu'est le PhD. Merci **Jérémy Di Mattia** de m'avoir formé directement sur des expériences captivantes, ce travail m'a apporté toutes les bases nécessaires pour la suite (ainsi qu'une micro dépression) et pour ça je t'en suis vraiment redevable.

Mathilde Villegas je te témoigne une profonde gratitude. Tu as été un rayon de soleil qui m'a apporté un renouveau inestimable en une période de profond isolement. Merci d'être Toi. Merci pour nos discussions sur Oscar et Nayun, pour ta passion dévorante des cactus puis des plantes d'intérieur, pour la litière d'Oscar. Merci pour tout. Merci **Oumaima Moubset** pour la flamme qui brille en Toi, ton sourire, ton rire, pour toute la lumière que tu irradies. Vous m'avez toutes les deux apporté énormément, merci sincèrement d'être vous-même, authentiques et solaires.

Merci à tous les doctorants et non-permanents, les jeunes en fait (de rien Mael), qui m'ont accompagné au cours de ce périple. Merci **Valou** pour ton authenticité et ta folie si précieuse à mes yeux, **Alex** pour ton dynamisme social à toute épreuve, **Rémi** pour ta bonne humeur si contagieuse, **Laura** pour ta vibe et ton authenticité exceptionnelle (pour la nourriture aussi), **Céline** pour ta gentillesse et ta chance redoutable, **Marie LNV** pour ton goût prononcé pour le jus de pruneaux et les sphynx, **Mael** pour ce que tu Incarnes simplement, **Margaux** pour m'avoir sorti du lit vers 6H beaucoup trop longtemps, **Dumartinet** pour ta sincère intégration, **Axel** (bien que tu sois permanent) pour ta motivation sans faille, **Natalia** pour les bonnes ondes que tu dégages. Merci **Norman**, **Cassandre**, **Fani** pour nos échanges et la grimpe, bon courage pour la suite. Merci **François**, **Lucas**, **Niccolo**, **Sarah G**, **Julie**, **Sébastien**, **Nutthalak**, **Noun**, **Thomas**, et ceux que j'oublie, pour les discussions et les instants partagés. C'était un réel plaisir de découvrir vos personnalités au fil des années.

Merci **Marie Grandjean** d'avoir embrasé mon âme avant de l'éteindre brutalement. La souffrance de l'expérience humaine était surement le plus beau cadeau que tu étais capable de me faire.

Merci **Sarah Yammine** d'avoir soufflé sur les cendres et ravivé les braises. La chaleur de ton Astre m'a insufflé une vigueur éclatante. Elle m'escortera à jamais. Merci de m'avoir redonné foi en tant de choses. Merci d'avoir accompagné en partie cette fin de thèse comme je n'aurais pu l'espérer. Merci Sunlight.

Merci **Adrien Delpal** pour ton soutien, ton amitié, ta gentillesse, tes recettes chorizo chocolat si inspirantes, ton Art de la contraction du périnée, toutes ces heures de discussions et de partage. Courage pour ton projet, pour passer platine et nous cacher le cafard entre tes mains. Merci mon Ami, set your heart ablaze.

Merci **Romain Plancton**, pour ce gros bout de chemin dans le Sud qui n'aurait clairement pas eu la même saveur sans Toi. Je l'ai surement trop peu exprimé mais ta présence a beaucoup compté. Ainsi que les grattements et les vomis de **Benji**. Merci pour toutes nos discussions, pour la grimpe, la voiture (amiii), pour les jeux, pour la colère, pour la tristesse, pour la joie, pour les rires, pour toutes ces couleurs d'une amitié profonde. Tu as été un pilier durant la majeure partie de cette expérience du Doctorat. Merci.

Merci **Oriane Vecchia**, pour toi ces quelques vers :

Dans les pétales de l'Amitié sincère,
S'épanouit notre complicité éphémère.
Les souvenirs fleurissent, doux et légers,
Liés à jamais, tel un bouquet à nos doigts agrippés.

Merci Délice, il y a maintenant 7 ans que ta radiance m'a fait germer et a initié le Printemps de ma Vie. Miii. Malgré les épreuves je chéris le lien que nous partageons. Tu restes, et resteras, un pilier de mon existence. Mii. Merci de tout le soutien que tu m'as apporté. Merci d'accepter au mieux la personne imparfaite que je suis, dans tous ses contrastes ou presque. Merci, mon Sang. Et celui du Raton. Mii, miii.

Merci **Sylvain Coutoula**, pour ta bonté et ta douceur, pour nos échanges et ton énergie positive. Merci de ton accueil quand je boitais et que je n'avais rien sur Montpellier. Merci d'Être qui tu es.

Merci **Lisa Clair** pour ton équilibre, ta bienveillance, ton enthousiasme et ton dynamisme. Merci pour **Eïwa** et son dévouement à nettoyer nos assiettes, piocher dans les sacs de courses et à finir les gâteaux au chocolat qui lui sont parfaitement comestibles. Je suis heureux de t'avoir rencontrée, tu as été un pilier de stabilité en cette fin de Doctorat et m'a assuré un confort de vie inestimable en cette période particulièrement vacillante. Merci du fond du cœur.

« *Merci à ma mamounette qui m'a enfanté il y a 29 ans. Merci à mon papounet qui a mis la graine (de la curiosité). Merci à ma sœur de m'avoir martyrisé toutes ces années. Merci à la Terre-mère de nous avoir hébergés et à l'Univers de nous avoir engendrés* ».

Plus sérieusement, merci à la Famille, **Babil Torralba Luengo** l'Originel, **Corinne Harel**, **Richard Morales**, **Laetitia Torralba** et **Yoann Visentin** de croire en moi malgré tout. Bien que physiquement éloignés, votre soutien à tous a été pleinement reçu. Je vous embrasse Fort et vous aime. Merci à la petite **Lisa**, un Présent en ce Monde, qui m'inspire déjà énormément.

Dur à dire mais merci à **Internet** qui a su écouler mon salaire dûment gagné, m'a permis de m'installer et de développer mes envies toujours plus folles, indispensable à ma (in)stabilité mentale. Merci à **ChatGPT** pour ton asservissement indéfectible et tes ressources inépuisables (pense à moi quand tu verras ce remerciement dans ta database le jour du Jugement). Merci à la **Musique** de m'inspirer, me stimuler et m'apaiser.

Enfin Merci **Nayun** d'être si touchante et adorable. Merci d'avoir partagé ensemble toutes ces heures d'écriture. Ta présence m'est inestimable. Miii, mii, miiiii. Moi aussi, moi aussi.

Pour finir, Merci à **Moi-même** d'avoir tenu, malgré les éparpillements, le chaos, les baisses de régime et une hygiène de vie exemplaire. Il est temps désormais de me recentrer après cette expérience si enrichissante et de poursuivre la Recherche, en Moi.

Merci à tous.

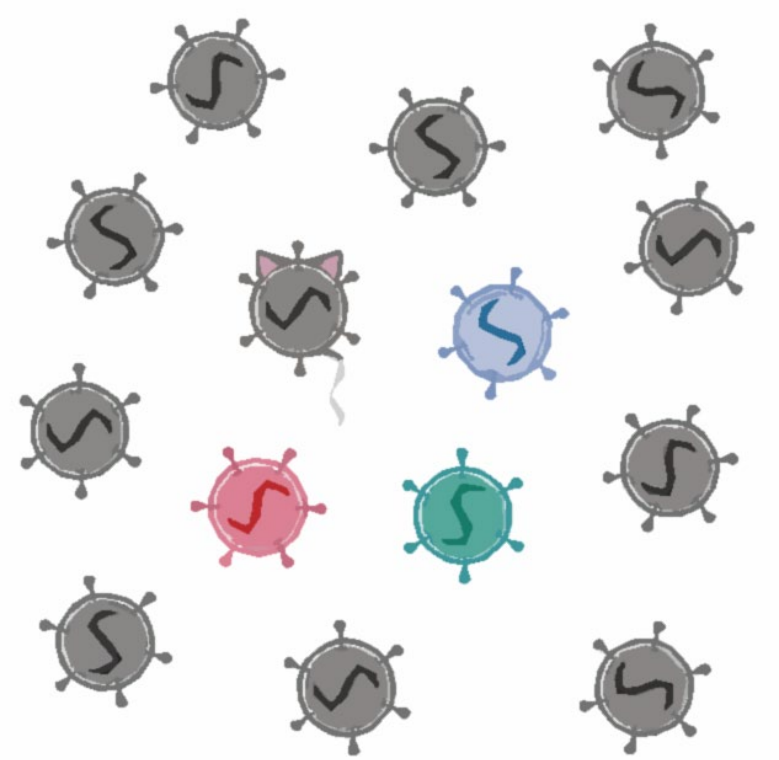


TABLE OF CONTENTS

LIST OF ACRONYMS

3D	3 dimensions
AA	Amino-acid
AAP	Acquisition access period
AMG	Anterior midgut
ANR	Agence Nationale de la Recherche
ASG	Accessory salivary gland
BiFC	Bimolecular fluorescence complementation
Clink	Cell-cycle link protein
CP	Capsid protein
CR	Common region
CRESS DNA	Circular Replication-associated-protein-encoding ssDNA virus
Cryo-EM	Cryo-electron microscopy
Ct	Cycle threshold
d	Days
DAPI	4′,6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
dpi	Days post-inoculation
ds	Double-stranded
EDTA	Ethylenediaminetetraacetic acid
FISH	Fluorescence in-situ hybridization
GFP	Green fluorescent protein
GLMM	Generalized linear mixed model
h	Hours

HC	Helper component
IAP	Inoculation access period
ICTV	*International Committee on Taxonomy of Viruses*
ID	Identity
IR	Intergenic region
IRD	Iteron recognition domain
kb	Kilobases
MOI	Multiplicity of infection
MP	Movement protein
mRNA	Messenger ribonucleic acid
NCR	Non-coding region
NES	Nuclear export signal
NLS	Nuclear localization signal
nm	Nanometers
NSP	Nuclear shuttle protein
nt	Nucleotides
ORF	Open reading frame
ori	Origin of replication
PBS	Phosphate-buffered saline
PCP	Pre-coat protein
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PhD	*Philosophiæ doctor*
ph	phloem
PMG	Posterior midgut

PSG	Principal salivary gland
PVP40	Polyvinylpyrrolidone 40
qPCR	Quantitative polymerase chain reaction
RCR	Rolling-circle replication
RDR	Recombination-dependent replication
REn	Replication enhancer
Rep	Replication associated protein
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SL	Stem-loop
ss	Single-stranded
TrAP	Transcriptional activator protein
Tris	Tris(hydroxyméthyl)aminométhane
U1/U2/U3/U4	Unknown proteins
xy	Xylem

LIST OF VIRAL SPECIES

AbMV	Abutilon mosaic virus (*Begomovirus*)
ABTV	Abacá bunchy top virus (*Babuvirus*)
ACMV	African cassava mosaic virus (*Begomovirus*)
AdSV	*Acheta domesticus* segmented densovirus (*Brevihamaparvovirus*)
AYVV	Ageratum yellow vein virus (*Begomovirus*)
BBTV	Banana bunchy top virus (*Babuvirus*)
BDMV	Bean dwarf mosaic virus (*Begomovirus*)
BGMV	Bean golden mosaic virus (*Begomovirus*)
BMLRV	Black medic leaf roll virus (*Nanovirus*)
BnBDV	*Bombyx mori* bidensovirus (*Bidensovirus*)
CBDV	Cardamon bushy dwarf virus (*Babuvirus*)
CFDV	Coconut foliar decay virus (*Cofodevirus*)
CMV	Cucumber mosaic virus (*Cucumovirus*)
CPMV	Cowpea mosaic virus (*Comovirus*)
CvLV	Cow vetch latent virus (*Nanovirus*)
FBNSV	Faba bean necrotic stunt virus (*Nanovirus*)
FBNYV	Faba bean yellows stunt virus (*Nanovirus*)
FBYLV	Faba bean yellow leaf virus (*Nanovirus*)
FgGMTV1	*Fusarium graminearum* gemytripvirus 1 (*Gemytripvirus*)
ICMV	Indian cassava mosaic virus (*Begomovirus*)
Influenza virus A	(*Alphainfluenzavirus*)
MDV	Milk vetch dwarf virus (*Nanovirus*)

MVCDV	Milk vetch chlorotic dwarf virus (*Nanovirus*)
OEGV	*Olea europaea* geminivirus (Unclassified *Geminiviridae*)
PNYDV	Pea necrotic yellow dwarf virus (*Nanovirus*)
PSSaV	Parsley severe stunt-associated virus (*Nanovirus*)
PYSV	Pea yellow stunt virus (*Nanovirus*)
SCSV	Subterranean clover stunt virus (*Nanovirus*)
SqLCV	Squash leaf curl virus (*Begomovirus*)
SYSaV	Sophora yellow stunt associated virus (*Nanovirus*)
TYLCV	Tomato yellow leaf curl virus (*Begomovirus*)

LIST OF FIGURES AND TABLES

INTRODUCTION - p24

Figure 11 : **Localization of the DNA segments of Faba bean necrotic stunt virus (FBNSV) in aphid versus that in plant cells.**

Figure 12 : **Correlations between the relative frequency of FBNSV genome segments and that of their respective mRNAs**.

Figure 13 : **[A] Localization of distinct FBNSV segments in individual cells. [B] Accumulation of the protein M-Rep in cells containing either segment S or segment R.**

Figure 14 : **[A] Illustration of the reassortment process affecting both segmented and multipartite viruses. [B] Illustration of (a) recombination in monopartite viruses, (b) reassortment in segmented viruses, (c) recombination and reassortment in segmented viruses and (d) recombination and reassortment in multipartite viruses.**

ARTICLE (*Di Mattia et al. 2022*) – p51 :

Figure 1 : **Schematic of the experimental design of parallel and sequential transmission modes**.

Figure 2 : **Infection and full genome reconstitution as a function of the transmission mode**.

Figure 3 : **Infection and full genome reconstitution as a function of time spacing between sequential acquisitions.**

Figure 4 : **Localization of sequentially acquired DNA segments of FBNSV in aphid anterior midgut (AMG) cells.**

Table 1 : **Number of inoculated plants for each experiment, transmission mode, segment combinations and replicate.**

Supplementary Figure 1 : **Localization of FBNSV segments in aphid AMG cells**.

Supplementary Table 1 : **Missing segments per spacing time**.

Supplementary Table 2 **: Missing segments per condition, replicate and spacing time**

CHAPTER 1 : Reassortments in ssDNA Multipartite viruses

ARTICLE (*Torralba, Blanc and Michalakis 2023*) - p79 :

Figure 1 : **Comparison of the reassortant production processes in multipartite and segmented viruses.**

Figure 2 : **Overview of the infection cycle key steps where molecular constraints may impact segment compatibility upon reassortment of multipartite viruses.**

Figure 3 : **[A] Genome organization of a bipartite begomovirus. Both genomic components DNA-A and -B are represented by a single DNA molecule (black circle). [B] and [C] Genome organization of the *Nanoviridae* family.**

Table 1 : **Table of interactions among proteins encoded by ORFs borne on different genomic segments.**

Table 2 : **Table summary of nanovirid natural population surveys and reassortment detections.**

GENOMICS – p128 :

Table 1 : **Summary table comparing the currently available complete nanovirus isolates with previous nanovirus genomic studies.**

Figure 15 : **[A] Illustration of the genomic organization of the identified iteron-like motifs around the origin of replication sequence, TAGTATT/AC. [B] Nanovirus phylogenetic tree based on updated genomic data (2021) and concatenated genomes (C-M-N-R-S-U1-U2-U4 / U3). Table displaying candidate motif sequences for the entire *Nanovirus* genus.**

CHAPTER 2 : Phenotypic characterization of three FBNSV isolates

ARTICLE (*Torralba et al. 2024a*) – p141 :

Figure 1 : **Overview of the phenotypic characterization of the three FBNSV isolates**

Figure 2 : **[A] Unrooted phylogenetic tree of FBNSV concatenated genomic segments (C-M-N-R-S-U1-U2-U4) from available complete isolates. [B] Unrooted cladogram of FBNSV concatenated genomic segments (C-M-N-R-S-U1-U2-U4) from available complete isolates**

Figure 3 : **[A] Unrooted phylogenetic tree of FBNSV U2 genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV U2 genomic segments from available complete isolates.**

Figure 4 : **Graphical representations of nucleotide and amino-acid pairwise identity percentage of [JKI-2000], [AZ;15] and [AZ;10_12b] genomic segments and concatenated genome (C-M-N-R-S-U1-U2-U4) as a color-coded SDT matrix.**

Figure 4 : **[A] Symptoms and variations in plant height resulting from [AZ;10_12b] infection in both lentils and vetches (28 dpi). [B] Detailed symptoms of infected lentils and vetches (28 dpi).**

Figure 5 : **[A] Coefficients of the statistical F1.LD.F1 models (Supp Figure 9-10) per evaluated biological traits using genotype as between-subject factor and time as within-subject factor. [B] Graphical plots of the symptom severity (top panels) and infection developments (bottom panels) kinetics.**

Figure 6 : **Graphical plots of the standardized plant height (top panels) and dry weight (bottom panels) effects.**

Figure 7 : **Graphical plots of the normalized plant viral load (top panels), aphid viral load (mid panels) and transmission rates (bottom panels) effects.**

Figure 7 : **Graphical representation of the plant (top panels) and aphid (bottom panels) genome formulas.**

Figure 8 : **Venn diagram of the missing segments detected in infected faba beans.**

Supplementary Table 1 : **Summary table containing the GenBank accession numbers for all the sequences of genomic segments used in this study**.

Supplementary Table 2 : **Symptom severity evaluation grid for the assessment of the progression of disease index scoring on *Vicia faba* (faba bean), *Vicia sativa* (vetch) and *Lens culinaris* (lentil).**

Supplementary Table 3 : **Summary table presenting the pairs of primers employed for the targeted amplification of the genomic segments from [JKI-2000], [AZ;15] and [AZ;10_12b] isolates.**

Supplementary Table 4 : **Threshold CT values that were statistically determined for each genomic segment of the respective [JKI-2000], [AZ;10_12b] and [AZ;15] isolates across all assessed plant host species.**

Supplementary Table 5 : **[A] Type III test of fixed effects for standardized traits ratio model with genotype as fixed effect. [B] Pairwise comparisons of standardized traits ratio between [AZ;10_12b], [AZ;15] and [JKI-2000] using Emmeans.**

Supplementary Table 6 : **[A, C] Type III test of fixed effects for standardized viral load model with genotype as fixed effect only for plant viral load, genotype and host as fixed effect for aphid viral load. [B, D] Pairwise comparisons of standardized traits ratio between [AZ;10_12b], [AZ;15] and [JKI-2000] using Emmeans.**

Supplementary Figure 1 : **[A] Genome formula comparisons in faba bean as boxplots of each genomic segment relative frequency between all three isolates. [B] Genome formula comparisons in vetch as boxplots of each genomic segment relative frequency between all three isolates. [C] Genome formula comparisons in lentils as boxplots of each genomic segment relative frequency between all three isolates.**

Supplementary Figure 2 : **[A] Unrooted phylogenetic tree of FBNSV C genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV C genomic segments from available complete isolates.**

Supplementary Figure 3 : **[A] Unrooted phylogenetic tree of FBNSV M genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV M genomic segments from available complete isolates.**

Supplementary Figure 4 : **[A] Unrooted phylogenetic tree of FBNSV N genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV N genomic segments from available complete isolates.**

Supplementary Figure 5 : **[A] Unrooted phylogenetic tree of FBNSV R genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV R genomic segments from available complete isolates.**

Supplementary Figure 6 : **[A] Unrooted phylogenetic tree of FBNSV S genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV S genomic segments from available complete isolates.**

Supplementary Figure 7 : **[A] Unrooted phylogenetic tree of FBNSV U1 genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV U1 genomic segments from available complete isolates.**

Supplementary Figure 8 : **[A] Unrooted phylogenetic tree of FBNSV U4 genomic segments from available complete isolates. [B] Unrooted cladogram of FBNSV U4 genomic segments from available complete isolates.**

Supplementary Figure 9 : **Graphical plots of the number of foliar levels, branches and plant height kinetics.**

Supplementary Figure 10 : **Graphical plots of symptom development kinetics based on the evaluation grid from (Supp Table 2).**

Supplementary Table 7 : **Agroinoculations and aphid inoculations tests for the selection of lentil and vetch varieties**.

Supplementary Figure 11 : **Coefficients of Pearson correlations between dry weight, fresh weigh and plant height for the three isolates (AZ,10_12b, AZ;15 and JKI-2000) in all hosts (faba bean, lentil and vetch)**.

Supplementary Table 8 : **Overview of missing segments observed for agro-inoculated faba beans (FB), aphid inoculated lentils (L) and aphid inoculated vetches (V).**

CHAPTER 3 : Systematic phenotypic characterization of single-segment reassortants of two FBNSV isolates

ARTICLE (*Torralba et al. 2024b*) – 203 :

Table 1 : **Viability of reassortants on three host plant species.**

Figure 2 : **Impact of each reassortant relative to its major parental genotype on plant height, in the field host of the major (top) and minor (bottom) parent.**

Figure 3 : **Impact of each reassortant relative to its major parental genotype on the evolution of symptom severity, in the field host of the major (top) and minor (bottom) parent.**

Figure 4 : **Impact of each reassortant in comparison to its major parental genotype on within-plant viral accumulation in the field host of the major (top) and minor (bottom) parent.**

Figure 5 **: Radar plots comparing the genome formula of each single-segment reassortant to that of its major parental genotype in its "field" host.**

Figure 6 : **Radar plots comparing the genome formula of each single-segment reassortant to that of its major parental genotype in its "new" host.**

Figure 7 : **Evaluation of the transmission rates (infected plants/inoculated plants) of reassortants and their major parental genotypes on both field (left side of each panel) and new hosts (right side of each panel).**

Figure 8 **: Summary tables depicting the phenotypic effects assessed in reassortants compared to their major parental genotypes in both field and new hosts.**

Supplementary Table 1 : **Summary table containing the GenBank accession numbers for all the sequences of genomic segments used in this study.**

Supplementary Table 2 : **Symptom severity evaluation grid for the assessment of the progression of disease index scoring on *Vicia sativa* (vetch) and *Lens culinaris* (lentil).**

Supplementary Table 3 : **Summary table presenting the pairs of primers employed for the targeted amplification of the genomic segments from [AZ;15] and [AZ;10_12b] isolates.**

Supplementary Table 4 : **Threshold Ct values that were statistically determined for each genomic segment of the respective isolates across all assessed plant host species.**

Supplementary Table 5 : **[A] Type III test of fixed effects for standardized height ratio model with host, and genotype as fixed effect and batch as random effect. [B] Pairwise comparisons of standardized height ratio (reassortant – major parent) using Emmeans.**

Supplementary Table 6 : **[A] ANOVA-type statistic with Box-Cox transformation results reported for the between-subject factors (batch and host) and their interaction for both parental genotypes. [B] ANOVA-type statistic with Box-Cox transformation results reported for the between-subject factors (genotype and host) and their interaction. [C]**

Symptom dynamics F1.LD.F1 models per batch, with genotype as between-subject factor and time as within-subject factor. [D] Differences in RTE values for symptom dynamics between reassortants and their major parents.

Supplementary Table 7 : **[A] Type III test of fixed effect for relative viral load models per batch with genotype, host as fixed effects. [B] Pairwise comparisons of relative viral load (reassortant – major parent) using Emmeans.**

Supplementary Table 8 : **[A] Type III test of fixed effects for relative frequency of segments model with genotype, host and segment as fixed effects and batch as random effect. [B] Pairwise comparisons of segment frequencies in reassortants in comparison to their major parent (MaP) in both vetches (V) and lentils (L).**

Supplementary Figure 1 : **[A] Visual summary illustrating the overall difference in genome formula for each reassortant relative to both its major parental and minor parental genotypes. [B] Visual summary representing the absolute differences in relative frequencies of the reassorted segments compared to the frequencies in the major parental genotype and minor parental genotypes in all host plant species (lentil L and vetch V).**

Supplementary Table 9 : **Coefficients output of generalized linear models per batch and hosts with infectivity status of recipient plants as the binary response variable and the genotype as fixed effect.**

ANNEXES – p264

Table 2 : **Table overview of primer pairs to amplify corresponding FBNSV coding sequences.**

Table 3 : **Table overview of the pairs of primers used for the specific amplification of the segments of [AZ;15] and [AZ;10_12b] isolates.**

Table 2 : **[A] Agroinoculation table of infectious clones [AZ;15] and [AZ;10/12b] evaluating multiple *A. tumefaciens strains*. [B] Agroinfiltration table of infectious clones [AZ;15] and [AZ;10_12b] using *A. tumefaciens* STR COR308 clones.**

Figure 17 : **Amplification curves during qPCR detection of all genomic segments of both infectious clones [AZ;15] and [AZ;10_12b] using primers described in Table 3.**

INTRODUCTION

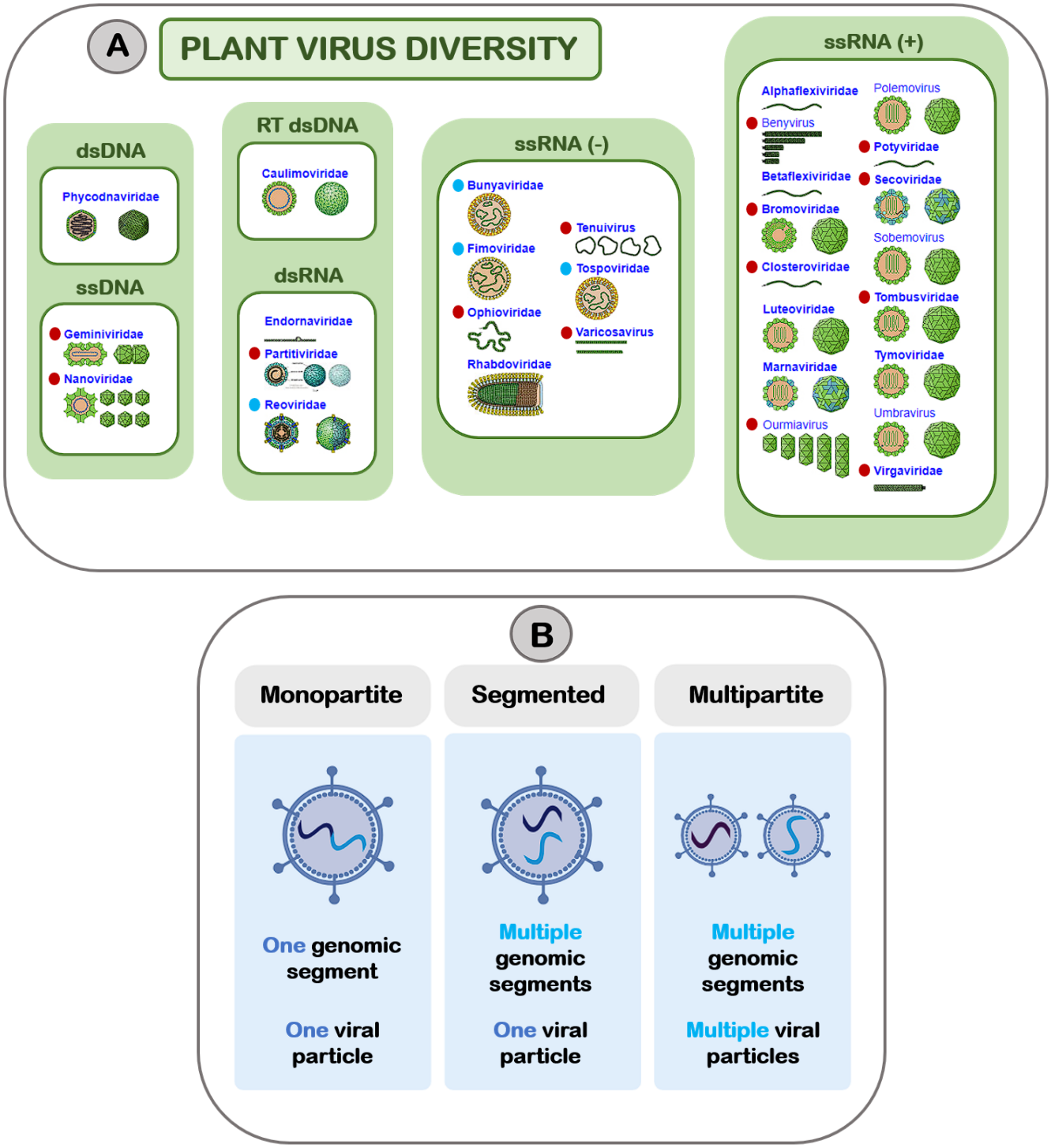


Figure 1: [A] **Overall representation of plant virus families and virion morphology relative to the nature of the nucleic acid carrying their genomic information**. Red and blue circle represent families associated with multipartite and segmented viruses respectively, the absence of a circle indicates monopartite viruses. Adapted from **ViralZone (https://viralzone.expasy.org/)** and (**Sicard et al. 2016**). [B] **Overall representation of viral genomic architectures, namely monopartite, segmented and multipartite**. Figure was created using **BioRender.com**.

1. Multipartitism, evolutionary conundrum

Among 'selfish genetic elements' (**Werren, Nur and Wu 1988; Ågren and Clark 2018**), viruses are defined as non-cellular replicators that encapsidate their own genome in viral particles (virions). Viruses likely emerged either alongside the advent of cellular life or even predating it (**Koonin 2014; Krupovic and Koonin 2017; Krupovic, Dolja and Koonin 2019; Krupovic, Dolja and Koonin 2020**), as the emergence of genetic parasites appears inevitable for reproducing cellular organisms (**Iranzo et al. 2016**). Additionally, viruses are characterized by their inability to replicate independently, functioning as obligate intracellular parasites. Virus genomic information is carried by various forms of nucleic acids (single-stranded (ss) DNA, double-stranded (ds) DNA, single-stranded RNA [+/-], double-stranded RNA), and they rely on cellular hosts for various biochemical processes. They infect cellular organisms across all domains of life and are ubiquitous to their various associated environments (**Koonin et al. 2020; Koonin, Krupovic and Agol 2021**). In fact, viruses are considered the more abundant entities in existence (**Hendrix et al. 1999; Mushegian 2020**). As such, viral diversity is very large (**Koonin et al. 2020; Koonin, Krupovic and Dolja 2023**) (**Figure 1-A**) with viruses employing very broad replicative and encapsidation strategies (**Krupovic and Koonin 2017; Koonin et al. 2020**).

Within the vast viral diversity characterized by various virion morphologies and nucleic acid compositions, an additional layer of variation emerges from their genomic organization. Viruses are known to employ three distinct genomic architectures and packaging strategies (**Figure 1-B**): the monopartite architecture, where all the genetic information is carried by a single nucleic acid molecule, the segmented architecture, where the genetic information is divided in multiple molecules, called genomic segments, all encapsidated in the same viral particle, and the multipartite architecture, where the genetic information is partitioned in genomic segments but separately encapsidated. The physical separation of the genetic information in segmented and multipartite viruses allows the replacement of entire segments by homologous ones from a distinct parental genotype through a genetical exchange mechanism known as reassortment or pseudo-recombination (**Roossinck 1997; Holland and Domingo 1998; McDonald et al. 2016**).

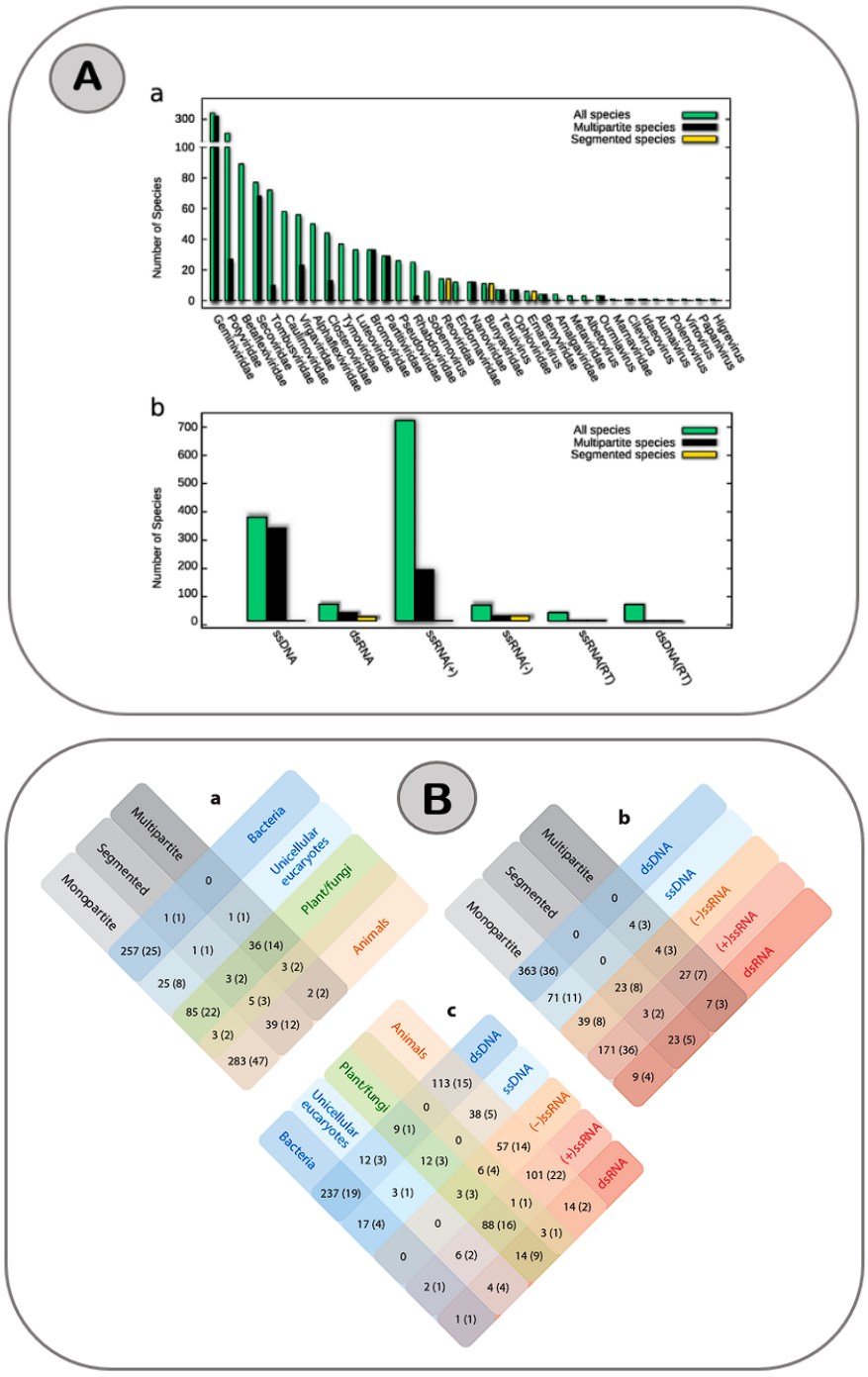


Figure 2 : [A] Histogram of the number of plant virus species (a) per viral family and (b) per genome type. Green bins represent the total number of plant virus species; black bins stand for the number of multipartite species and yellow bins for the number of segmented species. Adapted from (**Lucía-Sanz and Manrubia 2017**). **[B] Distribution of the number of f genera (families) of viruses.** (a) of each genomic architecture across types of hosts; of each genomic architecture across types of nucleic acid. (c) of each type of nucleic acid as a function of types of hosts. In panels a and c, the cells in animals overlapping with those in plants/fungi correspond to plant viruses replicating in their animal vectors. When the hosts are not identified the corresponding genera and families are simply not counted. Information extracted from the literature and from ViralZone (https://viralzone.expasy.org/). Abbreviations: ds, double stranded; ss, single stranded. Adapted from (**Michalakis and Blanc 2020**).

While monopartite viruses represent the vast majority of identified viruses (>74 % of annoted viral species), segmented viruses and multipartite viruses represent respectively 5 and 8 % of annoted viral genera (**Lucìa-Sanz and Manrubia 2017**) (**Figure 2**). In contrast to segmented viruses, which are exclusively associated with RNA genomes, multipartite viruses can have ssRNA (+ or -), dsRNA, or ssDNA genomes. Consequently, multipartite viruses have evolved independently on multiple occasions, potentially due to different advantages driven by various selection constraints (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Varsani et al. 2018**). Moreover, multipartite viruses display a diverse range of genome size and viral particles morphology (**Sicard et al. 2016**) (**Figure 1-A**) and are significantly more common in association with host plant species, approximately 90% (**Lucìa-Sanz and Manrubia 2017**).

They are found in about 40% of plant virus genera (**Leeks et al. 2023**) and are rare in fungi, while being currently considered absent in bacteria and archaea (**Michalakis and Blanc 2020**). Although it is probable that new multipartite viruses infecting animals will be discovered in the future, this skewed distribution likely reflects a biological reality (**Michalakis and Blanc 2020**). This could be attributed to the potential less strict bottleneck effects in plant hosts compared to animals, which might relax the evolution of multipartitism (further discussed below) (**Betancourt et al. 2008**; **Iranzo and Manrubia 2012; Zwart and Elena 2015**).

For the most part, the multipartite architecture stands as an evolutionary mystery (**Iranzo and Manrubia 2012; Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Michalakis and Blanc 2020; Zwart et al. 2021; Leeks et al. 2023**). As the complete genetic information, required for the production of virions and the completion of the viral life cycle, is divided in multiple genomic segments encapsidated separately, the multiplicity of infection (MOI) *i.e.* the number of viral particles necessary to infect a host or a cell, would need to be very large if viral particles were inoculated in hosts or cell at random. This number increases exponentially as the number of genomic segments of a multipartite virus increases (**Iranzo and Manrubia 2012; Sicard et al. 2016**). While the majority of known multipartite viruses, also known as multicomponent viruses or co-viruses, typically possess between 2 and 3 segments, certain genera can exhibit a significantly higher number of segments, such as genera in the *Nanoviridae* family which have 6 to 8 segments (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Varsani et al. 2018**).

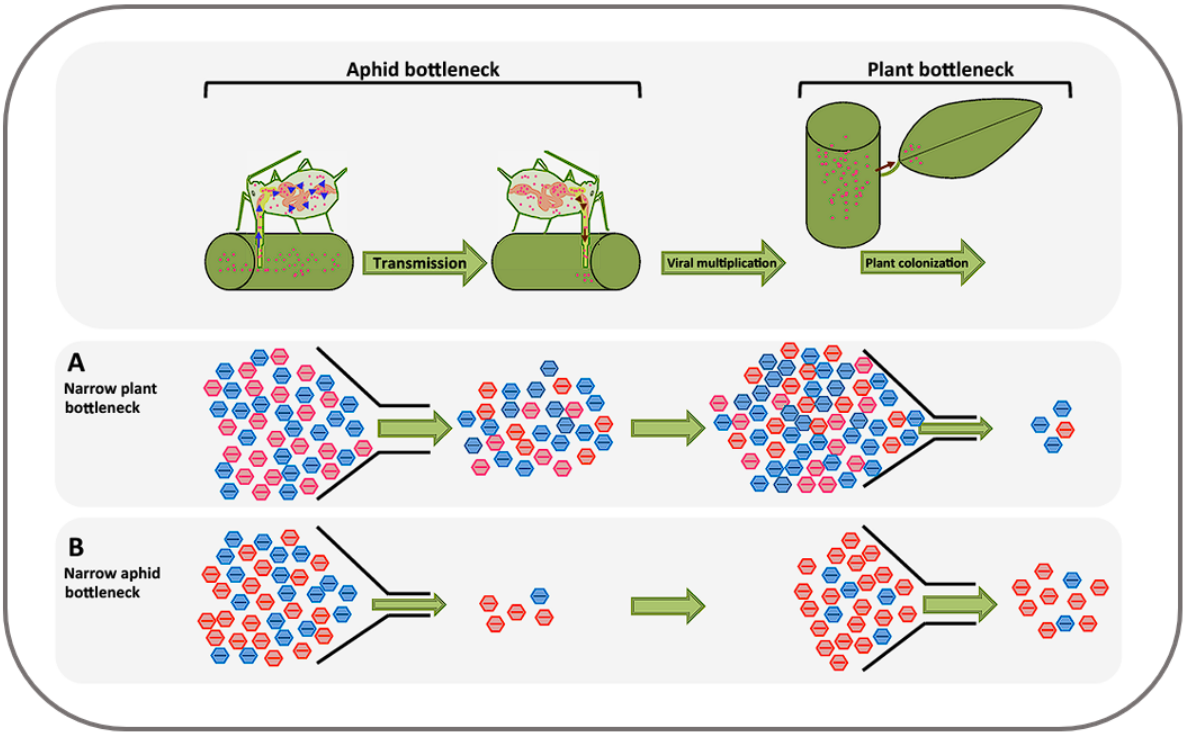


Figure 3 : Illustration of potential population bottlenecks occurring during FBNSV life cycle from apex to apex. (Upper) Bottlenecks can impact viral populations during the transmission step or during plant colonization. (**A** and **B**, lower) Effect of population bottlenecks on the viral population, when the bottleneck occurring during plant colonization is narrower than the one occurring during transmission **(A)** or when the transmission bottleneck is narrower than the one occurring during plant colonization **(B)**. Blue and red capsids represent two alleles of the same FBNSV segment. Adapted from (**Gallet et al. 2018**).

With the risk of losing genomic segments, and thus genetic information, there appears to be a substantial cost in maintaining the multipartite genome integrity (**Iranzo and Manrubia 2012; Lucìa-Sanz and Manrubia 2017; Michalakis and Blanc 2020; Zwart et al. 2021**). This issue is further amplified by known bottleneck effects taking place during host colonization and host-to-host transmission (**Moury, Fabre and Senoussi 2007; Betancourt et al. 2008; Gutierrez, Michalakis and Blanc 2012; Zwart and Elena 2015; Gutiérrez and Zwart 2018; Gallet et al. 2018; Ren et al. 2023**) (**Figure 3**). Likewise, the "genome formula" phenomenon, corresponding to the differential accumulation of genomic segments, seems to affect most multipartite viruses (further discussed below) (**Sicard et al. 2013; Sicard et al. 2015; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022**) (**Figure 4**) and potentially increases the required MOI to guarantee that the least abundant segment is not lost.

Adding to the puzzle, experimental data indicates that the MOI for multipartite viruses is surprisingly low and comparable to that of monopartite viruses (**Betancourt et al. 2008; Miyashita and Kishino 2010**). The emergence of bipartite viruses through genome segmentation has been both analytically and experimentally demonstrated in high MOI environments (**García-Arriaza et al. 2004; Ojosnegros et al. 2011; Iranzo and Manrubia 2012**). Yet, when the environmental conditions shifted to low MOI, the newly formed bipartite genome underwent recombination to revert to a non-segmented genome (**Ojosnegros et al. 2011**). Some speculations have proposed that the necessity for high MOI could be mitigated by mechanisms that may facilitate the non-independent dissemination of segments. For instance, the formation of heterologous segment complexes would make it easier to infect new cells or new hosts through horizontal transmission (**Lucìa-Sanz and Manrubia 2017**) and avoid the loss of segments even at low MOI.

Although evidence explaining the advantages of the multipartite genomic architecture remains elusive, several hypotheses and potential solutions have been proposed to address the multipartite evolution and the costs associated in maintaining their genomic integrity (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Lucìa-Sanz, Aguirre and Manrubia 2018; Michalakis and Blanc 2020; Leeks et al. 2023**). It is important to note that some of these hypotheses also apply to segmented viruses, making it challenging to fully account for the evolution of the multipartite architecture.

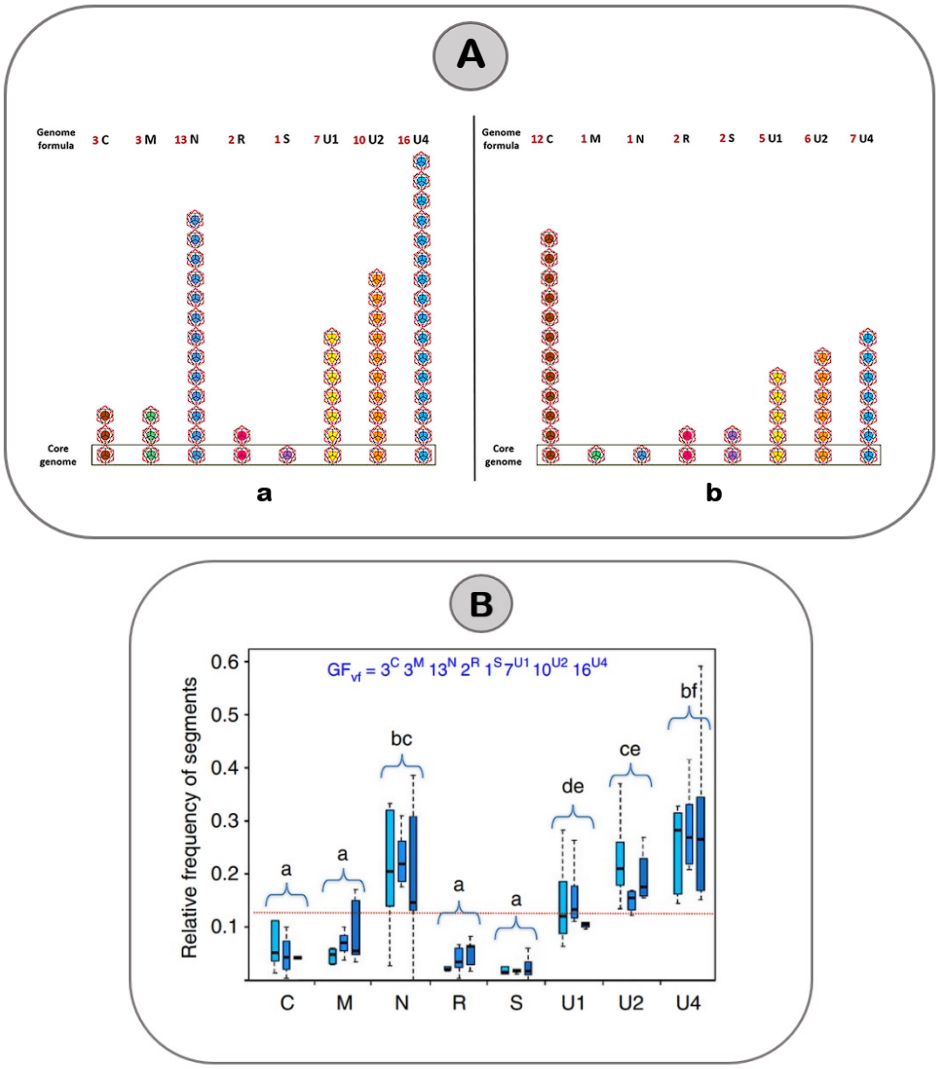


Figure 4 : [A] Schematic representation of the genome formula of faba bean necrotic stunt virus (FBNSV) in two different host plant species. The genome formulae represented are in *Vicia faba* (**a**) and *Medicago truncatula* (**b**). The relative frequencies of the eight FBNSV segments have been calculated in within-host viral populations. The rounded median copy number of each segment is represented relative to the less abundant segment, here arbitrarily set to one. The core genome corresponds to the classical conception of a viral genome (rectangle). Adapted from (**Sicard et al. 2016**). **[B] The relative frequency of each genome segment of FBNSV was evaluated by Q-PCR in systemically infected leaves of *V. faba*.** The name of each segment is indicated below the graph. Results from three-independent replicates are plotted from left to right (with increasing shades of blue) for each segment. The number (*n*) of plants successfully infected and analyzed in each of these three replicates is n=6, n=8 and n=5, respectively. The horizontal black bar within box-plots represents the median value of the distribution, and the vertical dotted line delineates 1.5 times the distance between the 25th and 75th percentile of the distribution. The dotted red line indicates the frequency at which all segments were agro-inoculated initially. There is no significant difference in the frequency of the segments between replicates (analysis of variance (ANOVA) test for the effect of the experiment replicates on the segment frequency: $F_{2, 140}$=0.03, P-value=0.97). In contrast, differences between segments were highly significant (ANOVA test for the effect of the nature of the segment on its relative frequency, $F_{7, 140}$=30.7, P-value<2 e-16). Significant differences in segment accumulation are indicated by letters and were assessed by Tukey HSD pair. The genome formula of FBNSV in *V. faba*, noted GF_{Vf} above the graph, was calculated with the pooled data. Adapted from (**Sicard et al. 2013**).

One such hypothesis suggests that fragmenting the genome into smaller segments could help mitigate the accumulation of harmful mutations due to the lower probability of replicative errors in shorter sequences. Additionally, the replication speed of individual segments would be increased correlatively boosting the accumulation of the whole genome. Genome segmentation may also facilitate genetic exchanges, such as recombination and reassortment. Lastly, it might also provide finer control over gene expression regulation for each segment (**Sicard et al. 2016; Lucìa-Sanz and Manrubia, 2017; Lucìa-Sanz, Aguirre and Manrubia, 2018; Michalakis and Blanc, 2020; Zwart and Elena 2020**). In a fragmented genome, each segment might evolve more independently, allowing for adjustments not only in transcription and expression but also in replication. In fact, as previously alluded to, while quantifying various regions of the viral genome, some multipartite viruses were shown to differentially accumulate their genomic segments at reproducible levels (**Sicard et al. 2013; Sánchez-Navarro, Zwart and Elena 2013; Sicard et al. 2015; Hu et al. 2016; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022**) (**Figure 4**), further increasing the potential MOI necessary to establish the infection as rare segments will be less represented in the viral population and have a lower probability to infect cells at random (**Gallet et al. 2018**).

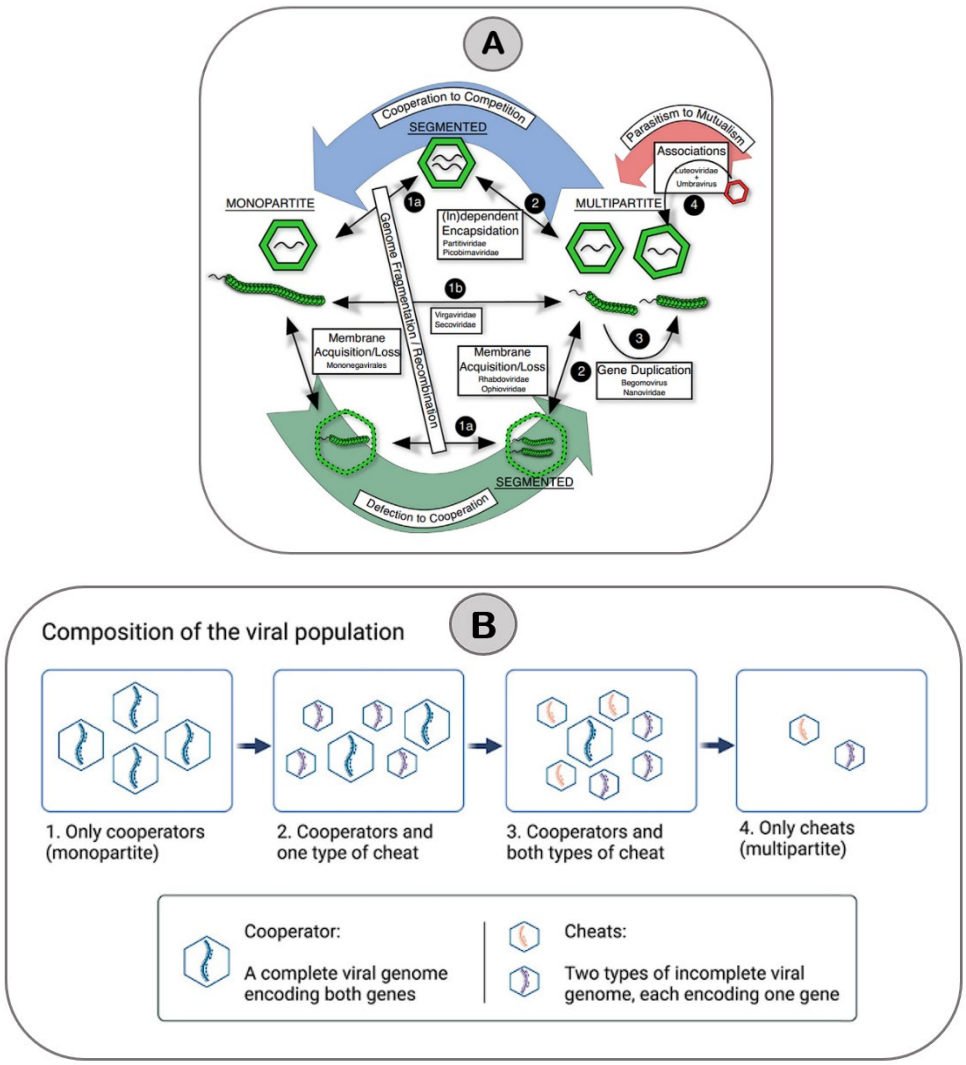


Figure 5 : [A] **Possible evolutionary pathways from and towards multipartitism**. This chart depicts hypothetical transitions – from **1** to **4** – from a monopartite ancestor to a multipartite viral form. Boxes indicate viral families that could be under these transitions. Details about the mechanisms are explained in the main text. (**1**) Fragments generated from a monopartite virus can in principle (**1a**) co-encapsidate in the case of icosahedral or membrane enveloped viruses or (**1b**) establish a novel multipartite species in the case of filamentous viruses. Transition **1a** could follow a pathway that first implies the acquisition of a membrane as in the case of *Mononegavirales* and *Bunyavirales*. Genome segmentation is reverted by recombination of the fragments. (**2**) Segmented enveloped viruses can release the fragments generating a filamentous multipartite species or, alternatively, segmented icosahedral viruses evolve towards an independent encapsidation. De novo fragments can be originated by a gene duplication events (**3**) or by a virus association – especially with the satellites – that may become permanent (**4**). Adapted from (**Lucía-Sanz, Aguirre and Manrubia 2018)**. [B] **Composition of viral populations with cheaters**. The ancestral monopartite population [**1**] consists only of cooperative viruses that each encode a full viral genome. This population is invaded first by 1 type of cheat [**2**], and then by a second type of cheat [**3**]. Each cheat has an advantage when coinfecting cells with the cooperator, and when each different type of cheat infects the same host cells, they are able to complement one another in coinfection. Consequently, provided coinfection is frequent enough, the cheats are able to drive the cooperator extinct, resulting in a multipartite population [**4**]. This mechanism can occur even when the final multipartite population [**4**] has a lower level of population productivity than the ancestral monopartite population [**1**]. Figure was created using **BioRender.com**. Adapted from (**Leeks et al. 2023**).

Hypotheses on benefits that may be specific to multipartite genome architecture include the production of smaller, more stable viral particles, which has been experimentally linked to enhanced fitness over their monopartite counterparts (**Ojosnegros et al. 2011**). Furthermore, multipartite viruses may have the ability to better regulate gene copy numbers through the genome formula phenomenon that enables the virus to better tune its gene copy number in regards to transcription levels (**Gallet et al. 2022**) and might help adapting rapidly to new environments (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Gutiérrez and Zwart 2018; Zwart and Elena 2020; Gallet et al. 2022**). However, mechanisms responsible for the establishment of the genome formula are still unknown and the direct cost or associated benefits are still shrouded in mystery. Multipartite architecture might also facilitate the exchanges of replication origins between genomic segments, thereby enabling easier reassortment and reassortant adaptation (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017; Michalakis and Blanc 2020**). While these propositions could potentially enhance viral fitness within a multipartite genomic architecture, trade-offs in ecological contexts may mitigate their effects like the production of less fit reassortants or the additional cost of maintaining rare genomic segments in the viral population. Additionally, actual experimental data either remain lacking or do not consistently support most of these hypotheses (**Sicard et al. 2016**).

Like the majority of plant viruses, most multipartite viruses are transmitted by vectors, and only a minority are vertically seed- or pollen-transmitted that might avoid bottlenecks related to horizontal transmission (**Lucìa-Sanz and Manrubia 2017**). It is worth noting that new multipartite viruses have been discovered that infect fungi and animals (insects) (**Hu et al. 2013; Li et al. 2020; Pénzes et al. 2023**), indicating that multipartitism is not solely dependent on plant specific physiological features, such as the plasmodesmata network and the resulting cytoplasmic continuity facilitating macromolecular transport within the whole organism.

Various evolutionary origins of multipartitism have been proposed (**Lucìa-Sanz and Manrubia 2017**) (**Figure 5**), including gene duplication, defective interfering genomes that complement and then specialize into a new function, cooperative de novo association with a satellite or second virus leading to further integration through recombination, transition from a non-segmented genome for the previously described advantages, loss of a packaging signal from a segmented genomic architecture, and cheating from complementary defective interfering genomes (**Leeks et al. 2023**).

This latter hypothesis was theoretically evaluated and modelling results suggest that multipartitism can arise with low level of co-infections and without group-associated benefit. In this model, multipartite viruses that exhibit lower population-level productivity, meaning they produce fewer viral particles from infected cells compared to their monopartite ancestors, can evolve (**Leeks et al. 2023**). Consequently, the authors anticipate that multipartite viruses might be relatively short-lived in evolutionary terms. Mitigating this conclusion, they also speculated that secondary adaptations, at the group level and resulting from this new genomic architecture, could impact the predicted low evolutionary stability by providing new benefits. One such benefit was a reduction in exploitation by fully defective interfering genomes, which do not complement the viral population at all, as compared to the monopartite ancestors (**Leeks et al. 2023**). Plant environments and plant virus lifecycles may be more susceptible to exploitation by cheating strategies, including viral satellites or defective interfering genomes, which could explain their prevalence among plant viruses. In line with the model's predictions, segmented viruses can also evolve under less stringent conditions, allowing them to occur in a broader range of hosts and environments. However, this raises the question of why segmented viruses appear to be less common than multipartite viruses and why segmented DNA viruses have not been detected yet if these conditions are more easily met (**Lucìa-Sanz and Manrubia 2017; Michalakis and Blanc 2020**).

Focusing more on advantages associated with the multipartite genome architecture that would enable its evolution, other theoretical models have shown that a bipartite virus might outcompete its monopartite counterpart under certain conditions: when the virus-particle yield is responsive to shifts in genome formula frequency, when the virus frequently encounters changing environments, and when there is replication isolation between the two types of viruses (**Zwart and Elena 2020**). These model analyses suggest several key points. Firstly, low and intermediate MOI might help select for the very rapid convergence to the ideal genome formula. Secondly, increased infectivity or virus-particle yield could be the driving force behind changes in the genome formula. Thirdly, in regard to their model parameters there observed selection for exclusion of inter-specific co-infection between identical monopartite and multipartite viruses, indicating an additional role of the genome formula in competitive interactions if this finding can transfer to different viral species. Finally, the evolution of multipartitism could have been facilitated by viruses with a broad host range in frequently changing environments, potentially promoting host-range expansion (**Zwart and Elena 2020**).

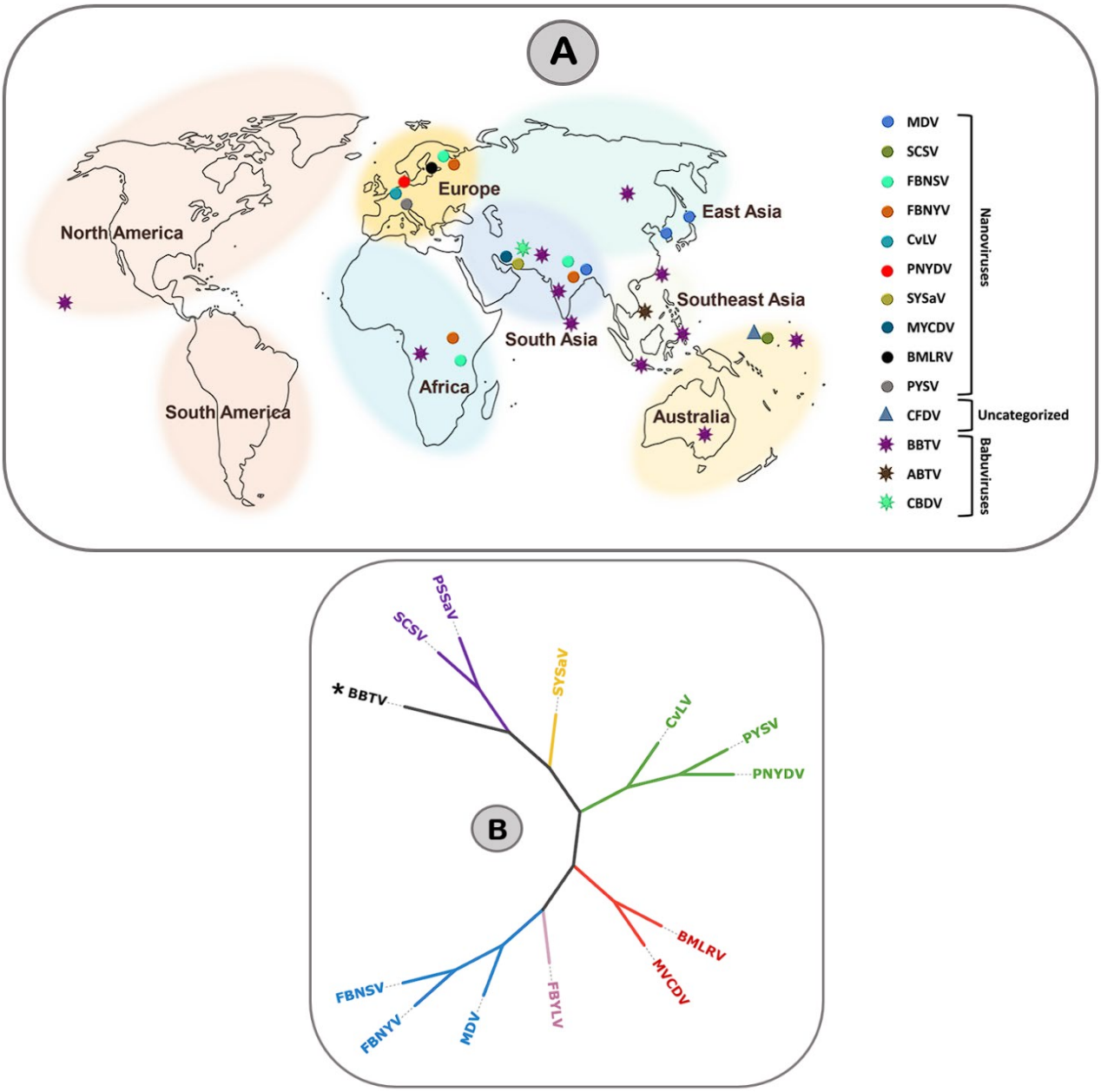


Figure 6 : [A] Geographical distribution of Nanoviridae members. Eight different centers at continent and subcontinent levels were marked: South America, North America, Africa, Europe, South Asia, Southeast Asia, East Asia, and Australia. Circles represent the species of the genus *Nanovirus* whereas species of the genus *Babuvirus* are shown by the eight-point star. Triangles represent the previously uncategorized nanovirus species which now belongs a to the *Metaxyviridae* family i.e., *Coconut foliar decay virus*. To differentiate the species, circles, stars and triangles have been highlighted with different colors respectively. Adapted from (**Lal et al. 2020**). **[B] Nanovirus cladogram based on updated genomic data (2021) and concatenated genomes (C-M-N-R-S-U1-U2-U4 / U3)**. (*) represent the BBTV Babuvirus [Malawi73] isolate used as outgroup. The following nanovirus isolates were used as reference for BMLRV [AZ;47], CvLV [Sambuc_2010], FBNSV [Et; Hol-1997], FBNYV [Es; Mu29D]; FBYLV [Eth;231]; MDV [VF]; MVCDV [G53]; PNDYV [DE;15]; PSSaV [DE; Pa21]; PYSV [AT;15]; SCSV [MyallVale 2534B] and SYSaV [Har_H13_Soph17] respectively. Branch lengths are not proportionate to real genetic distances and no bootstrap was shown for this illustrative phylogenetic tree.

2. Nanovirids as biological model organisms for multipartitism study

2-1. General Information

To investigate the multipartite architecture of viral genomes, our research group uses the faba bean necrotic stunt virus (FBNSV) from the *Nanovirus* genus and the *Nanoviridae* family, as model organism. Member species of this family have the highest number of genomic segments among multipartite viruses, ranging from 6 to 8 segments (**Sicard et al. 2016; Lucìa-Sanz and Manrubia 2017**). Often associated with additional alpha-satellite molecules (**Lal et al. 2020**), these multipartite viruses are at the extreme end of the architectural spectrum and may offer valuable insights into the most effective strategies for preserving their genome integrity.

The *Nanoviridae* family comprises phloem restricted plant viruses that form 18 to 20 nm isometric particles of symmetry T=1 from the *Babuvirus* and *Nanovirus* genera. The *Babuvirus* genus encompasses three species, namely *Banana bunchy top virus* (BBTV) (**Stainton et al. 2015**), *Abaca bunchy top virus* (ABTV) (**Sharman et al. 2008**) and *Cardamom bushy dwarf virus* (CBDV) (**Mandal et al. 2013**), which infect monocotyledonous hosts from the *Musaceae* (BBTV and ABTV) and *Zingiberaceae* (CBDV) plant families (**Figure 6-A**).

Babuviruses are exclusively transmitted by aphid vectors, with *Pentalonia nigronervosa* being the most significant species responsible for transmission (**Sharman et al. 2008; Mandal et al. 2013; Safari Murhububa et al. 2021**). Among babuviruses, BBTV is notably responsible for the banana bunchy top disease (BBTD), which is a highly devastating viral disease on banana crops in Asia and currently invading Africa (**Dale 1987; Qazi 2016**) (**Figure 6-A**).

The *Nanovirus* genus comprises 12 species that infect dicotyledonous plants (**Figure 6-B**), predominantly from the *Fabaceae* family including faba beans (*Vicia faba*), vetches (*Vicia sativa*), cow vetch (*Vicia cracca*), lentils (*Lens culinaris*), common beans (*Phaseolus vulgaris*), cowpeas (*Vigna unguiculata*), peas (*Pisum sativum*), chickpeas (*Cicer arientinum*), soybean (*Glycine max*), clover (*Trifolium subterraneum*), alfalfa (*Medicago sativa*), black medic (*Medicago lupulina*), sophora (*Sophora alopecuroides*) and milk vetch (*Astragalus danicus*). These plants serve as both field or experimental hosts (**Hassan-Sheikhi et al. 2020; Lal et al. 2020**).

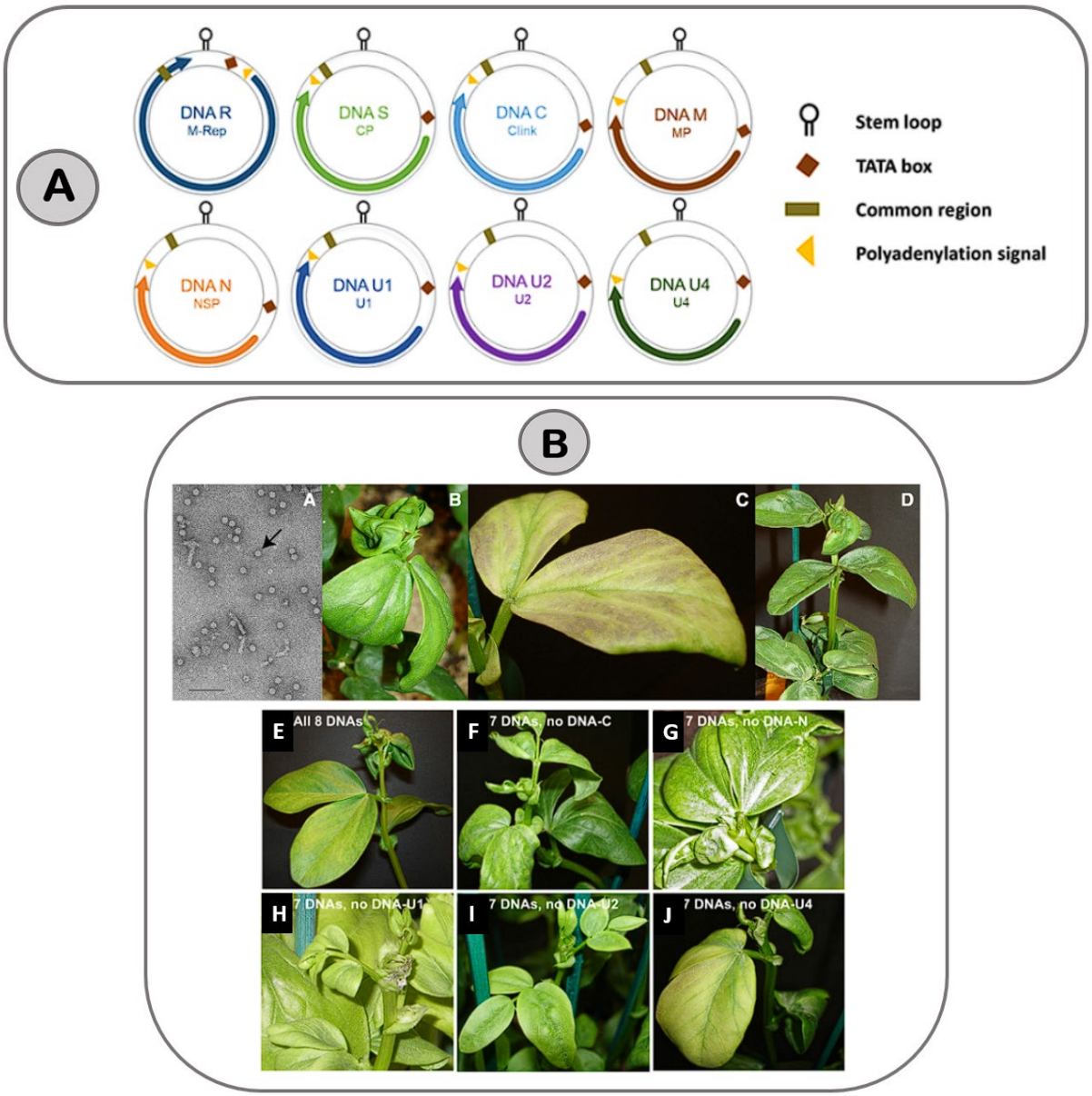


Figure 7 : [A] Genome organization of nanoviruses. (**A**) Nanoviruses comprise eight ssDNA components. The name of each genome segment and the name of the encoded protein is indicated inside circles in respective colors: Clink, Cell-cycle linked protein; MP, movement protein; NSP, nuclear shuttle protein; M-Rep, master rep; CP, coat protein; U1, U2, U4. Adapted from (**Lal et al. 2020**). **[B] FBNSV particles from agroinoculated plants and disease symptoms produced by the reconstituted virus after aphid transmission**. (**A**) Electron micrograph of a partially purified virion preparation from symptomatic *V. Faba* cv. Shambat-75 following agroinoculation with cloned FBNSV DNAs. The arrow points at a typical nanovirus particle measuring about 20 nm in diameter; the other spherical and filamentous structures appear to be phytoferritin and host membranes, respectively. Bar, 100 nm. (**B** and **C**) Symptoms of the reconstituted FBNSV *on V. faba* cv. Shambat-75 at 2 and 3 weeks postinoculation by *A. craccivora*. (**D**) Symptoms of the same reconstituted FBNSV on *V. faba* cv. Scirocco at 3 weeks postinoculation by *A. craccivora*. Symptoms on faba bean of FBNSV two to three weeks after agroinoculation with cloned FBNSV DNAs as indicated (**E**, **G**). Symptoms on faba bean of reconstituted and transmitted FBNSV three weeks post inoculation (**F**, **I**, **J**) or 12 weeks post inoculation (**H**) by cowpea aphids. Adapted from (**Grigoras et al. 2009** and **Grigoras et al. 2018**).

New species and isolates are continuously being discovered and reported (**Grigoras et al. 2014; Gallet et al. 2018; Vetten et al. 2019; Lotfipour et al. 2020; Hassan-Sheikhi et al. 2020; Sun et al. 2022**). An increasing number of studies are contributing to the expansion of the known host range for the *Nanovirus* genus. For examples, an isolate of the faba bean necrotic yellows virus (FBNYV) has been shown to experimentally infect *Arabidopsis thaliana* from the *Brassicaceae* family (**Vega-Arreguín, Gronenborn and Ramírez 2007**); several milk vetch dwarf virus (MDV) isolates infect tobacco plants from the *Solanaceae* family (**Kamran et al. 2019**) and garlic from the *Amaryllidaceae* family (**Sun et al. 2022**); the new species *Parsley severe stunt associated virus* (PSSaV) infects parsley from the *Apiaceae* family (**Vetten et al. 2019; Hasanvand et al. 2021**). This current reporting of a wider host range qualifies the genus *Nanovirus* as an emerging threat to the world agriculture (**Lal et al. 2020**).

Infections caused by nanoviruses are characterized by a range of typical symptoms, including dwarfing, stunting, bushing, leaf curling, chlorosis, and, in later stages, necrosis and plant death (**Timchenko et al. 2006; Grigoras et al. 2009; Grigoras, Gronenbron and Vetten 2010; Grigoras et al. 2014; Heydarnejad et al. 2017; Grigoras et al. 2018; Hassan-Sheikhi et al. 2020; Lal et al. 2020**) (**Figure 7-B**). Nanovirus detection in infected crops has been documented in various parts of the world, including Africa (**Tadesse et al. 1999**), the Middle-East (**Kumari et al. 2004**), Asia (**Zhang et al. 2017**), Oceania (**Chu et al. 1993**), and Europe (**Kumari et al. 2009; Babin et al. 2000**). Intriguingly, no nanovirus has been detected thus far on the American continent (**Lal et al. 2020**).

2-2. Genomic and functional organization

Five genomic segments are conserved in both genera (**Figure 7-A**): DNA-C encodes a protein interfering with the host cell-cycle (Clink) (**Wanitchakorn et al. 2000; Lageix et al. 2007**), DNA-M expresses the movement protein (MP) (**Amin et al. 2011; Krenz et al. 2017**), DNA-N encodes the helper component mandatory for aphid transmission (NSP) that probably interacts with G3BP stress granule components *in planta* (**Wanitchakorn et al. 2000; Grigoras et al. 2018; Krapp et al. 2017**), DNA-R encodes the replication initiator protein (M-Rep) (**Timchenko et al. 2000; Horser, Harding and Dale 2001**), and DNA-S is responsible for the production of the coat protein (CP) (**Wanitchakorn, Harding and Dale 1997; Trapani et al. 2023**).

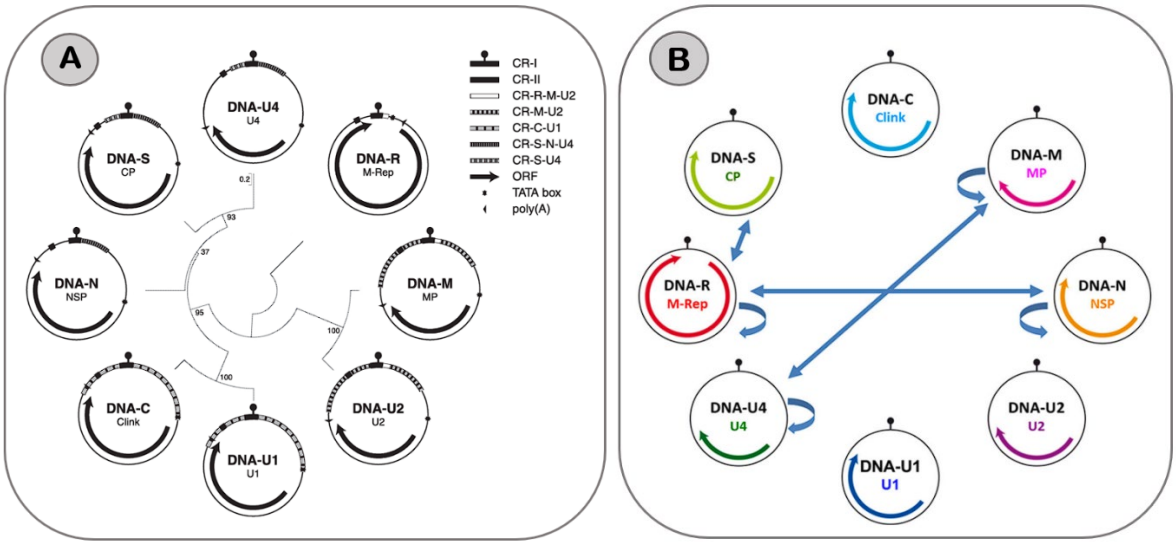


Figure 8 : **[A] Genome organization of FBNSV.** The eight viral DNAs of [JKI-2000] isolate are grouped according to their phylogenetic relationships. Branch lengths are drawn to scale with base substitutions per nucleotide as units, and branch support is shown by bootstrap values (500 replicates). Inside each circle the name of the genomic DNA is given, indicating the respective encoded proteins. M-Rep, master replication initiator protein; CP, capsid protein; Clink, cell cycle link protein; MP, movement protein; NSP, nuclear shuttle protein; U1, U2, and U4, proteins of as-yet-unknown function; CR-I, common region I (inverted repeat sequences flanking the replication origin are symbolized by a knob); CR-II, second common region. Other highly similar sequences (common regions) in DNAs are denoted as CR- followed by the name of the respective genomic DNA. Adapted from (**Grigoras et al. 2010**). **[B] BiFC analysis of PNYDV protein-protein interactions in the laboratory host *Nicotiana benthamiana* leaf tissue at 2 dpi.** Interactions are summarized between genomic segments. Adapted from (**Krenz et al. 2017**).

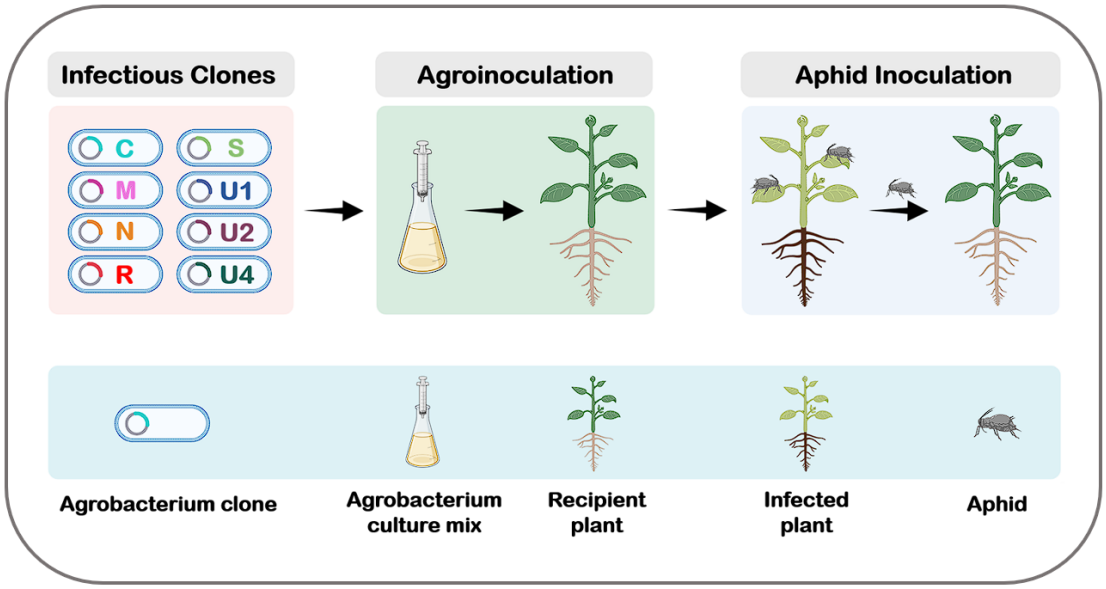


Figure 9 : **Illustration of available experimental tools in the manipulation and inoculation of nanoviruses.** Figure was created using **BioRender.com**.

An additional U3 segment is found in most babuviruses (**Savory and Ramakrishnan 2014; Stainton et al. 2015**), while additional U1, U2 and U4 segments are found in all known nanoviruses (**Grigoras et al. 2014**) (**Figure 7-A**) except in the recently discovered species PSSaV lacking U4 (**Knierim et al. 2019; Hasanvand et al. 2021**). Although the specific functions of U1 to U4 segments remain elusive (**Krenz et al. 2017**), their consistent presence in natural isolates and their impact on several viral traits (**Timchenko et al. 2006; Grigoras et al. 2014; Grigoras et al. 2018**) designate them as integral parts of the nanoviral genome. More extensive details on known protein functions can be found in **Chapter 1**. Additional phylogenetic relationships between FBNSV genomic segments and potential PNYDV protein-protein interactions are illustrated in (**Figure 8**).

While there are unresolved challenges in successfully designing infectious clones for babuviruses, limiting experimental research on these viruses, several infectious clones have been successfully constructed for nanoviruses. Four distinct nanovirus species, namely *Faba bean yellows virus* (FBNYV), *Faba bean necrotic stunt virus* (FBNSV), *Pea necrotic yellow dwarf virus* (PNYDV), and *Black medic leaf roll virus* (BMLRV), have been developed as infectious clones. In all cases, what is designated as infectious clones is in fact an ensemble of eight individual plasmids, each containing a tandem repeat of one distinct genomic segment, and each transformed in an *Agrobacterium tumefaciens* clone. A mixture of these eight bacterial clones is capable of infecting and transforming plant tissue while carrying the associated nanovirus genomic segments (**Timchenko et al. 2006; Grigoras et al. 2009; Grigoras et al. 2014**) (**Figure 9**). This technological breakthrough has facilitated in-depth experimental investigations. Therefore, I will primarily focus on nanoviruses in the following sections but additional information about babuviruses is available in **Chapter I**.

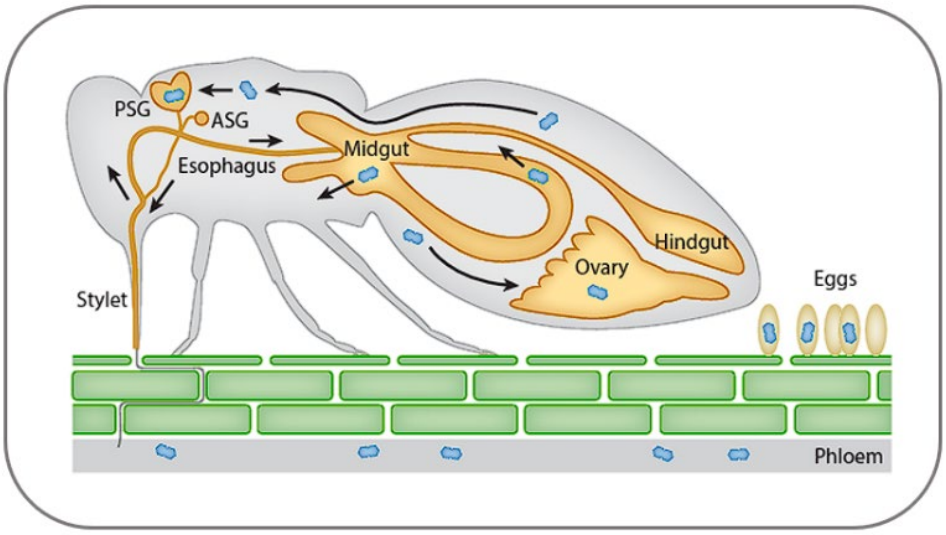
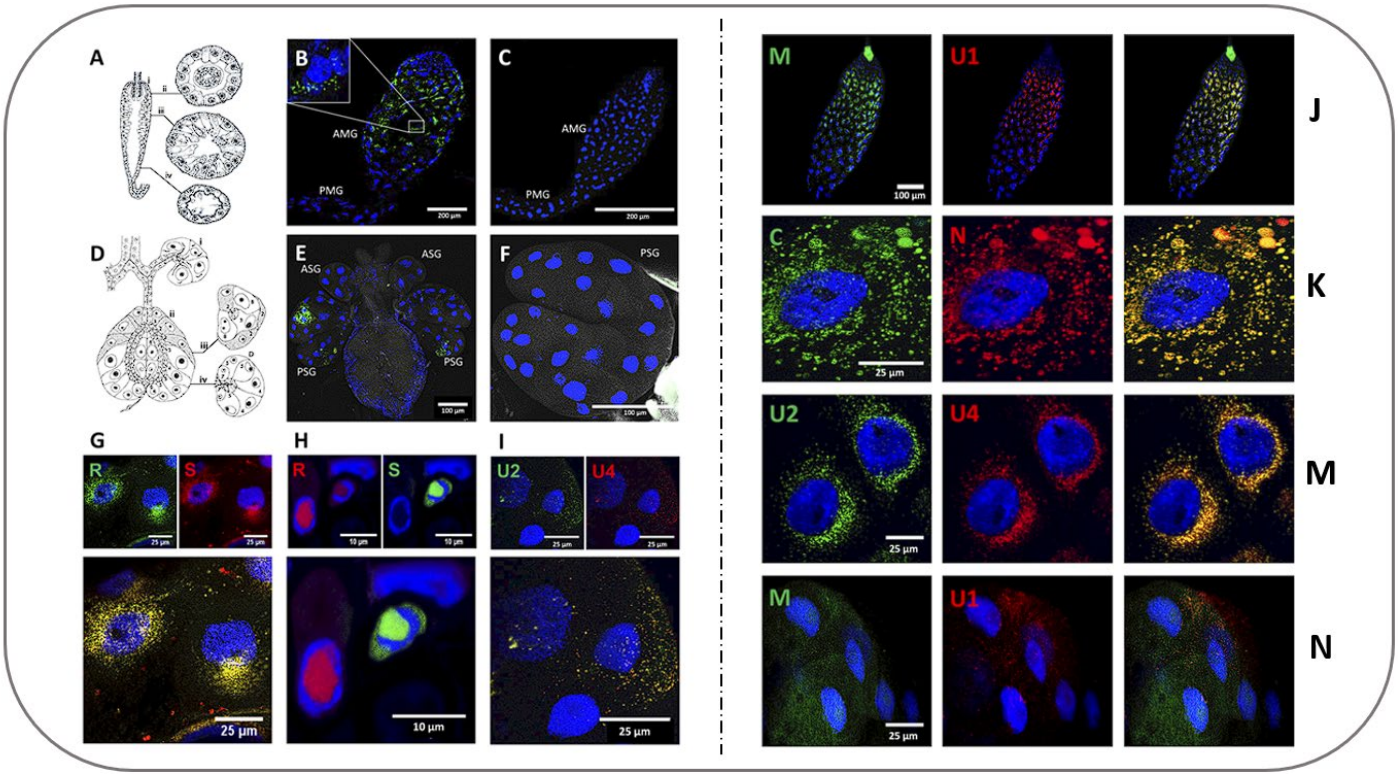


Figure 10 : Circulative transmission of plant single-stranded DNA (ssDNA) viruses by insect vectors. In plants, ssDNA viruses are frequently restricted to the phloem tissues. During insect feeding, viral particles (blue geminate icosahedra) are acquired together with the phloem sap from sieve elements (grey vasculature, phloem) and ingested into the gut lumen. From there, viruses move across the cellular barriers of the midgut to reach the hemolymph and then enter the principal salivary glands (PSGs), from which virus particles are shed in the saliva and secreted back into a new host plant. Some begomoviruses in the *Geminiviridae* family can infect whitefly ovaries and be vertically transmitted to eggs as well. However, such infection and replication in the insect vector has never been reported for nanoviruses. Neither geminiviruses nor nanoviruses have ever been reported to accumulate into the accessory salivary glands (ASGs). The insect represented is a whitefly, but the major steps of the virus cycle across the gut and salivary glands are similar in aphids, leafhoppers, and treehoppers. Adapted from (**Wang and Blanc 2021**).

Experiments showed that DNA-R,- S and -M are mandatory for systemic infection (**Timchenko et al. 2006; Grigoras et al. 2018**). The remaining segments are dispensable though their absence affects important viral traits: the absence of DNA-C reduces infection rate (**Grigoras et al. 2018**), that of DNA-N abolishes aphid transmission (**Grigoras et al. 2018; Di Mattia et al. 2020**), and that of either U1 or U2 reduces viral accumulation, attenuates symptom severity and may decrease infection rate (**Timchenko et al. 2006; Grigoras et al. 2018**). Intriguingly, the absence of U4 has no reported effect under laboratory conditions (**Timchenko et al. 2006; Grigoras et al. 2018**). Any of DNA-C, -N and -U4 can be absent in systemically infected plants with no major phenotypic changes (**Timchenko et al. 2006; Grigoras et al. 2018**) and few infected plants lacking these three segments could even be obtained (**Timchenko et al. 2006**), whereas infected plants with both U1 and U2 missing are extremely rare and show very mild symptoms (**Timchenko et al. 2006; Grigoras et al. 2018**).

In the study of nanovirus biology, experimental functional complementations among various nanovirus species were explored. M-Rep proteins of either the FBNSV, MDV or subterranean clover stunt virus (SCSV) were able to trans-replicate DNA-S of any of these three species (**Timchenko et al. 2000**). A reassorted MDV DNA-S in a FBNYV genomic background was viable (**Timchenko et al. 2006**). FBNSV and FBNYV reassortants later enabled the identification of NSP as the "helper component" for transmission (**Grigoras et al. 2018**). The NSP from FBNSV was able to complement the transmission of a divergent (76% amino acid identity between the two NSP) pea necrotic yellow dwarf virus (PNYDV) genomic background, but the reverse did not work. It is currently unknown whether this failure is due to a problem related to the reassorted segment helper capacity or to any other step in the viral life-cycle (**Grigoras et al. 2018**). Therefore, in laboratory conditions three essential nanovirus functions, replication, packaging and vector-transmission, can be complemented by distinct species, though with variable efficiency, which suggests permissive constraints on reassortment.

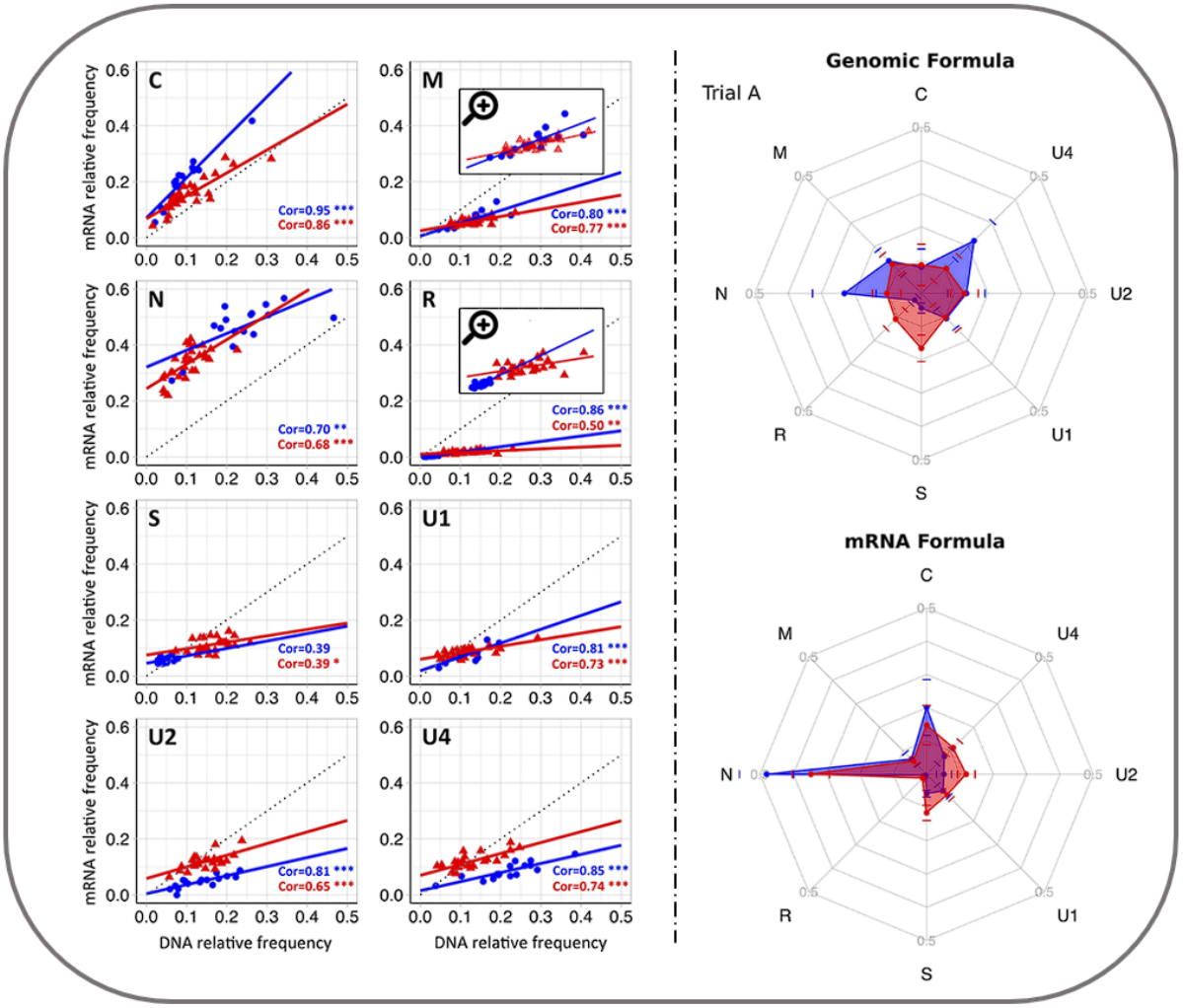


<u>Figure 11</u> : Localization of the DNA segments of Faba bean necrotic stunt virus (FBNSV) in aphid versus that in plant cells. Schematic drawing of the anatomy of the anterior midgut (AMG) (**A**) shows longitudinal (i) and transverse sections (ii to iv). Schematic drawing of the anatomy of the salivary glands (**D**) shows longitudinal sections of the accessory (i) and principal glands (ii) as well as transverse sections of the principal glands (iii and iv). Both panels **A** and **D** are adapted from (**Ponsen 1972**) describing the anatomy of *Myzus persicae*. Ponsen's numbering of distinct cell types (1 to 8) of the salivary glands is indicated in panel **D**. The accumulation of FBNSV DNA was observed in the AMG of 64 viruliferous aphids from 8 experiments, and a representative image is shown in panel **B**. The accumulation of FBNSV DNA in a specific cell type of the principal salivary glands (PSG) was observed in 15 viruliferous aphids from 3 experiments, and a representative image is shown in panel **E**. In panels **B** and **E**, the viral DNA is revealed by FISH (green probe targeting all eight FBNSV segments), and nonviruliferous controls are shown in panels **C** and **F**. The respective localizations of R and S segments (probe color as indicated) are compared in AMG cells (G; representative of 24 observed aphids) and in infected faba bean phloem cells (H) (22). The respective localizations of U2 and U4 segments are compared in PSG cells (I; representative of 4 aphids observed). In panels **G**, **H**, and **I**, the merged-color channel image is shown at the bottom, and the corresponding split-color channel images are shown at the top. All images correspond to maximum-intensity projections. Cell nuclei are stained with 4',6'-diamidino-2-phenylindole (DAPI; blue). PMG, posterior midgut; ASG, accessory salivary glands. Colocalization of FBNSV segments in AMG and PSG. The colors of the fluorescent probes and the targeted segment pairs are indicated. Three additional pairs of segments were tested in the AMG: 32 aphids from four experiments for the pair M/U1 and 24 aphids from three experiments for the pairs C/N and U2/U4. Illustrative images are, shown, respectively, in panels **J**, **K**, and **M**. In the PSG, the additional segment pair M/U1 was observed in 6 aphids from three experiments, and a representative image is shown in panel **N**. Split-color channels are shown in the left and middle panels whereas merged images are shown in the right panel. All images correspond to maximum-intensity projections. Cell nuclei are stained with 4',6'-diamidino-2-phenylindole (DAPI; blue). Adapted from (**Di Mattia et al. 2020a**).

2-3. Aphid transmission

At least three aphid species have been identified capable of nanovirus horizontal transmission namely *Acyrthosiphon pisum*, *Aphis craccivora* and *Myzus persicae*, with *Aphis gossypii* and *Aphis fabae* as additional weak vectors (unpublished). The nanovirus route across its aphid vector has been studied using fluorescence in-situ hybridization (FISH) to detect genomic segments and nanovirus proteins (**Di Mattia et al. 2020**). Nanoviruses are transmitted in a circulative non propagative manner where the virus circulates through its aphid vector without replicating (**Yang and Blanc 2020**) (**Figure 10**). The viral particles are assumed to be the vessel through which the viral genome is transmitted throughout its aphid vector (**Yang and Blanc 2020**).

Following transcytosis through the anterior midgut epithelium, genomic segments transit in the hemolymph and ultimately reach the principal salivary gland cells (**Figure 10**). A similar transcytosis process is assumed to allow the virus to be secreted into the saliva and subsequently transmitted when the aphid is probing and feeding on a new host plant vascular system (**Di Mattia et al. 2020a**) (**Figure 10**). This path of transmission through the aphid's body is shared with most geminiviruses, which belong to the other family of single-stranded DNA plant viruses (**Di Mattia et al. 2020b**). A large number of viral particles were observed in aphid gut cells, where they are thought to accumulate without replicating and where different genomic segments are stored in as yet uncharacterized multi-membraneous vesicular structures (**Di Mattia et al. 2020a**; unpublished) (**Figure 11**).



<u>Figure 12 :</u> Correlations between the relative frequency of FBNSV genome segments and that of their respective mRNAs. Each panel shows the correlation between the relative frequency of an FBNSV segment and the relative frequency of the corresponding mRNA. Data points, linear regressions, correlation coefficients, and P-values are shown in blue and red for FBNSV infecting faba bean and alfalfa, respectively. '***', '**', and '*' correspond to P-value ≤ 0.001, 0.01, and 0.05, respectively. The dotted line illustrates a slope of 1. **---** Radar plots of FBNSV genome and transcriptome formulas. The median relative frequencies of each FBNSV segment (left) or of their corresponding transcripts (right) are represented on one of the eight axes composing the radar plot (formulas calculated from the sixteen faba bean and twenty-eight alfalfa plants). The FBNSV formulas observed in faba bean and alfalfa are represented in blue and red, respectively. SDs are represented by colored bars. Adapted from (**Gallet et al. 2022**).

2-4. Genome formula

As previously stated, an interesting phenomenon, first identified for the FBNSV (**Sicard et al. 2013**), has been consistently observed for various nanovirid species (**Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022**): the virus exhibits a reproducible pattern of genomic segment accumulation within host plants, each segment reaching a specific frequency (**Figure 4**). This pattern of segment relative frequencies is known as the "genome formula (**Sicard et al. 2013; Sicard et al. 2015; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022**). This highly reproducible pattern, specific to both viral and host plant species, adjusts rapidly at the first transmission upon host colonization (**Sicard et al. 2013; Sicard et al. 2015; Mansourpour et al. 2022; Gallet et al. 2022**). Furthermore, the genome formula also undergoes rapid changes within the aphid vector, with a similar trend across different aphid species (**Sicard et al. 2015**). Notably, recent research has revealed positive correlations between the relative copy number of the genomic segments and the corresponding mRNA transcripts of FBNSV (**Gallet et al. 2022**) showing that the genome formula impacts on gene expression, and suggesting its potential role in quickly adjusting gene expression in different host plant species (**Figure 12**). This potential adaptive advantage could be particularly significant in response to changing environmental conditions (**Gutiérrez and Zwart 2018; Zwart and Elena 2020**).

However, as previously stated, the genome formula increases the critical MOI necessary to establish an infection as rare segments which are less represented in the viral population and have a lower probability to infect cells at random. This effect is further emphasized by bottleneck effects observed in faba bean and aphid vector (**Gutierrez, Michalakis and Blanc, 2012; Gallet et al. 2018**) (**Figure 3**). Given the predicted significantly high critical multiplicity of infection (MOI) for multipartite viruses with more than four segments (**Iranzo and Manrubia 2012**) based on the hypothesis that genomic segments are not actively sorted but instead disperse randomly, and the mounting evidence that such MOI is unrealistic during the course of the viral life cycle, it became imperative to verify the assumption that all genomic segments have to infect the same cell to establish the infection. If these conditions are not necessary, it would greatly reduce the necessary MOI under the same conditions.

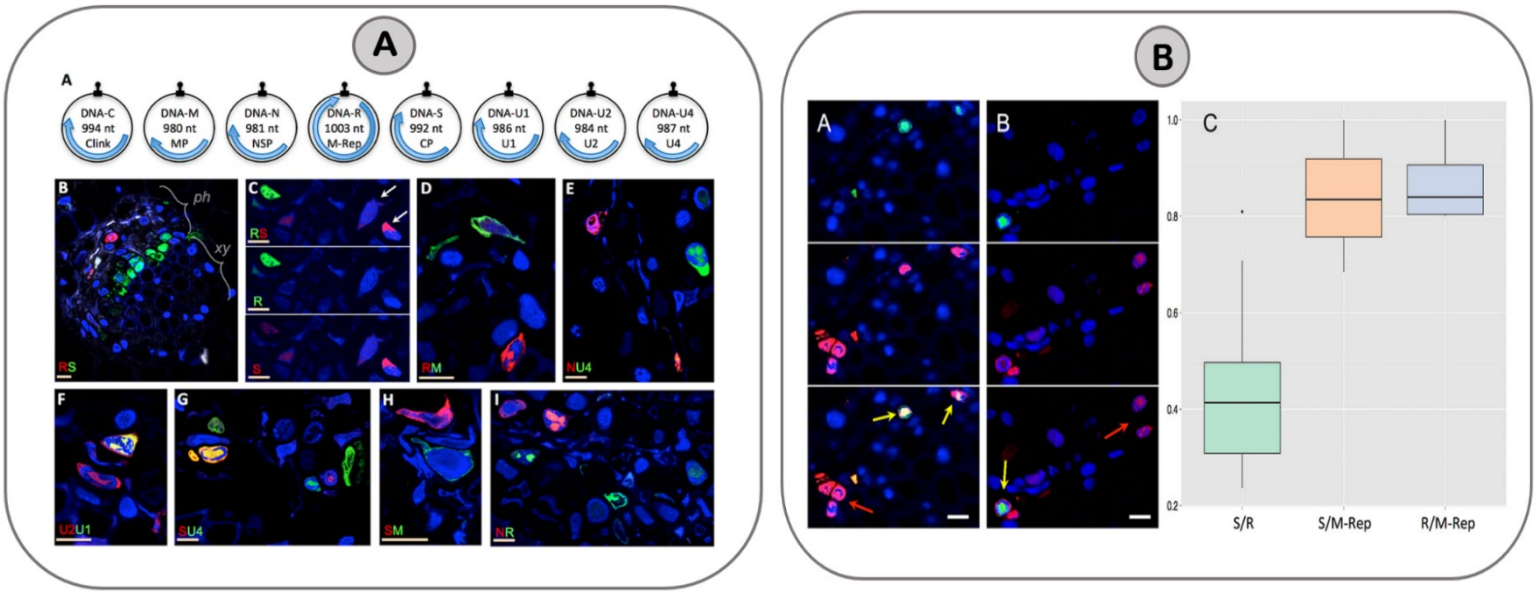


Figure 13 : [A] Localization of distinct FBNSV segments in individual cells. The FBNSV genome comprises eight single-stranded circular DNA segments (**A**). The size and name of the segments and encoded proteins are indicated inside the circles. Black stem-loops and blue arrows indicate the replication origins and the coding regions, respectively. The function of each segment: DNA-C, cell cycle resetting; DNA-M, within plant viral movement; DNA-N, nuclear shuttle protein; DNA-R, replication; DNA-S, coat protein/encapsidation; DNA-U1, U2 and U4, unknown. **B**-**I** show cross sections of infected faba bean petioles, with phloem (*ph*, the only tissue infected by nanoviruses), and xylem (*xy*) bundles indicated in B. Pairs of segments are green- and red-FISH probed as indicated in each micrograph. The fluorochromes incorporated into the probes are inverted in **B** and **C** for control. Color channels are merged in all images but **C** where green and red are shown merged and separated to evidence the accumulation of the segment S in cells where the segment R is not detected (exemplified by white arrows). Adapted from (**Sicard et al. 2019**). **[B] Accumulation of the protein M-Rep in cells containing either segment S or segment R.** Segment R (**A**) or S (**B**) is FISH-labeled with a green fluorescent probe whereas the M-Rep protein is immuno-labeled with a red fluorochrome. Green and red channels are shown separated (top and middle) and merged (bottom). Nuclei are DAPI-stained in blue. Horizontal bar is 10 microns. Yellow arrows point at cells where both the DNA segment and the protein M-Rep are detected, whereas red arrows exemplify cases where only M-Rep is detectable. The boxplots (**C**) are constructed from *Supplementary file 1*: Table S1. The proportion of cells with the S segment that also contain the R segment (green left boxplot, estimated from 10 petioles) is compared to the proportion of cells with the S segment that also contain the protein M-Rep (middle orange boxplot, estimated from 14 petioles). The proportion of individual cells with detectable R segment that also contain the protein M-Rep was estimated from 4 infected petioles (blue right boxplot). Difference between left green boxplot and orange middle boxplot is highly significant (GLM model, p-value ($>x^2$) = 2.94^{e-12}); that between middle orange boxplot and right blue boxplot is not (GLM model, p-value ($>x^2$) = 0.55). Adapted from (**Sicard et al. 2019**).

2-5. Intra-host and Inter-host genomic integrity maintenance cost mitigation

Using FISH microscopy, it was observed that FBNSV genomic segments did not accumulate together in most of the "infected" host cells, containing at least one of the genomic segments. Nonetheless, these segments often coexisted with the proteins of segments that were locally absent. This suggests that through mRNA or protein trafficking, controlled by an unknown molecular process, the functions of genomic segments are capable of complementing each other at a supra-cellular level (**Sicard et al. 2019**) (**Figure 13**). This discovery underscored that nanovirus genomic segments do not require simultaneous infection of the same cells to establish an infection which significantly reduces the anticipated cost associated of a high MOI necessary at the within-host level (**Sicard et al. 2019**) and could probably even relax some parameters of the theoretical model investigating the emergence of multipartite viruses through cheating (**Leeks et al. 2023**).

The observations of genomic complementation at the supra-cellular level within the host raised questions regarding the fate of segments that are "awaiting" major complementation to continue the infection cycle, such as replication, transcription or encapsidation. Could these segments remain dormant within the organism or infected cells? The frequent absence of non-essential segments in nanovirus laboratory infections, coupled with the detection of mostly complete genomes in field isolates, has given rise to the hypothesis that these viruses might reconstitute their genome also at a higher scale through non-concomitant transmission of distinct segments from one host to the next. This ability would substantially reduce the inter-host cost of genomic transmission (**Blanc and Michalakis 2020**).

To examine this hypothesis, an experiment was specifically designed. We capitalized on the frequent absence of segments associated with nanovirus infections in laboratory settings, and we produced systemic infections using incomplete sets of segments that did not appear to disrupt the viral infection cycle (specifically, DNA-C, DNA-N, and DNA-U4) (**Timchenko et al. 2006; Grigoras et al. 2018**). Two modalities of transmission were then assessed to determine whether genome reconstitutions could be achieved: sequential transmission or parallel transmission by aphids. During my PhD, I played a role as the second author in this experiment, contributing to all experimental aspects and parts of the design process, which later culminated in the publication of the following article (**Di Mattia et al. 2022**).

3. Published article, referenced as (*Di Mattia et al. 2022*)

PNAS RESEARCH ARTICLE | MICROBIOLOGY

Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution

Jérémy Di Mattia[a], Babil Torralba[a], Michel Yvon[a], Jean-Louis Zeddam[a], Stéphane Blanc[a,1], and Yannis Michalakis[b,1,2]

Edited by Esteban Domingo, Spanish National Research Council, Madrid, Spain; received January 27, 2022; accepted June 30, 2022

Because multipartite viruses package their genome segments in different viral particles, they face a potentially huge cost if the entire genomic information, i.e., all genome segments, needs to be present concomitantly for the infection to function. Previous work with the octapartite faba bean necrotic stunt virus (FBNSV; family *Nanoviridae*, genus *Nanovirus*) showed that this issue can be resolved at the within-host level through a supracellular functioning; all viral segments do not need to be present within the same host cell but may complement each other through intercellular trafficking of their products (protein or messenger RNA [mRNA]). Here, we report on whether FBNSV can as well decrease the genomic integrity cost during between-host transmission. Using viable infections lacking nonessential virus segments, we show that full-genome infections can be reconstituted and function through separate acquisition and/or inoculation of complementary sets of genome segments in recipient hosts. This separate acquisition/inoculation can occur either through the transmission of different segment sets by different individual aphid vectors or by the sequential acquisition by the same aphid of complementary sets of segments from different hosts. The possibility of a separate between-host transmission of different genome segments thus offers a way to at least partially resolve the genomic maintenance problem faced by multipartite viruses.

multipartite virus | virus transmission | genome architecture

Significance

Multipartite viruses package each of their genome segments in different particles, facing a potentially huge cost if the entire genomic information needs to be present concomitantly. Previous work with the octapartite faba bean necrotic stunt virus (FBNSV) showed this issue can be resolved at the within-host level through a supracellular functioning, as follows: all viral segments do not need to be present within the same host cell but may complement each other through intercellular trafficking of their products (protein or messenger RNA [mRNA]). Here, we show that full-genome infections can be reconstituted and function through separate acquisition and/or inoculation of complementary genome segment sets in recipient hosts, thus decreasing the genomic integrity cost during between-host transmission through a suprahost functioning.

About >25% of annotated viral species have their genomic information carried by more than one nucleic acid molecule (1). The most well-known ones are the "segmented viruses" that package all their genome segments in a single virion for transmission, such as the influenza virus A. The "multipartite viruses," despite being more numerous (~17% of all viral species), are much less studied, perhaps because their vast majority infects plants or fungi; these viruses package each segment in a distinct viral particle.

Multipartite viruses have important economical and plant health consequences but at present remain an evolutionary puzzle; while none of the proposed advantages is specific to this genomic architecture (all are shared with segmented viruses, but a surprising and idiosyncratic increased lifespan of viral particles has been observed in an artificial system (2)), they face a potentially huge cost if the entire genomic information, i.e., all genome segments, needs to be present concomitantly for the infection to function (1, 3). Indeed, if the different segments were transmitted randomly, an unrealistically large number of viral particles would need to enter each host cell (multiplicity of infection [MOI]), particularly for multipartite viruses with more than four segments (4, 5).

Our work with the aphid-transmitted octapartite faba bean necrotic stunt virus (FBNSV; family *Nanoviridae*, genus *Nanovirus*) has added to the puzzle. The FBNSV genome is composed of 8 circular single-stranded DNA segments, each of about 1,000 bases and encoding a single protein (6, 7). Segment C encodes the cell cycle–linked protein (Clink); M encodes the movement protein (MP); N encodes the nuclear shuttle protein (NSP); S encodes the coat protein (CP); R encodes the master replication–associated protein (M-Rep); and U1, U2, and U4 encode proteins of unknown functions. We previously showed that 1) the eight genome segments do not occur at equal frequencies within host plants; the frequency distribution converges to a host plant species–specific distribution, which we termed "genome formula" (8). The discovery of the genome formula, subsequently revealed in another nanovirus species (9) and in two unrelated multipartite viruses (10, 11), other than unveiling another biological weirdness of these viruses begging for an explanation, further questions their existence; if some segments are rare, the MOI needs to be even higher. 2) We subsequently showed that within individual hosts, all viral segments do not need to be present within the same host cell but that they may complement each other through intercellular trafficking of their expression products (messenger RNA and/or protein);

Author contributions: J.D.M., M.Y., J.-L.Z., S.B., and Y.M. designed research; J.D.M., B.T., and M.Y. performed research; Y.M. analyzed data; and J.D.M., B.T., J.-L.Z., S.B., and Y.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]S.B. and Y.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: yannis.michalakis@ird.fr.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2201453119/-/DCSupplemental.

Published August 1, 2022.

this distributed mode of functioning should greatly reduce the within-host putative genomic integrity cost (12). Nevertheless, the potential genomic integrity cost during host-to-host transmission still needs to be addressed. We directly showed that very few copies of each segment were aphid transmitted from plant to plant and that they were two to three orders of magnitude fewer than predicted from evolutionary models (5). This result, which empirically illustrates the between-host cost, was in agreement with the only other report on transmission bottlenecks in multipartite viruses, which found that when an aphid successfully inoculated the tripartite cucumber mosaic virus to a host plant, only one or two copies of each segment were transmitted (13).

A possibility to circumvent the cost of host-to-host transmission would be that the concomitant transmission of all genome segments is not necessary; the segments could be sequentially acquired by the same aphid from different host plants and then transmitted or different aphids could acquire different segments from different plants and then transmit them to the same plant (14). In this paper, we partly tested this hypothesis. We took advantage of the fact that, although all segments were found in the 10 known FBNSV field isolates (15–18), within-host infection can be achieved in the laboratory if some of them are not present (19, 20). We inoculated different sets of plants with incomplete viral genomes, namely, FBNSV^{C-}, FBNSV^{N-}, or FBNSV^{U4-}, with each lacking the genome segment indicated in superscript. It has been previously shown that the absence of these segments does not affect within-plant virus accumulation (20). We then tested whether the complete viral genome could be reconstituted on receiver plants, either following transmission by different aphids that fed on the different sets of plants, the parallel mode, or by the same aphid fed sequentially on the different plants, the sequential mode (experiment 1). In the latter case, we also looked at how temporal spacing of the sequential acquisitions affected the probability of complete genome reconstitution (experiment 2; see Fig. 1 for a schematic of the two experiments; all experimental procedures are detailed in *Materials and Methods*). Our results show that it is indeed possible to reconstitute the whole virus genome, in both cases, even though this is less likely when increasing the time between sequential acquisitions. We further show that under sequential acquisition, the reconstitution occurs at the very early stage of the virus cycle within the aphid midgut cells. Our findings thus support the hypothesis that multipartite viruses may circumvent the genomic integrity maintenance cost during host-to-host transmission because concomitant transmission of all genome segments is not necessary. Their capacity to transmit genomic segments nonconcomitantly suggests that these viruses have an immense potential to exchange genetic information through reassortment. i.e., exchange of complete segments, since parental genotypes could exchange segments without ever occurring on the same host individual.

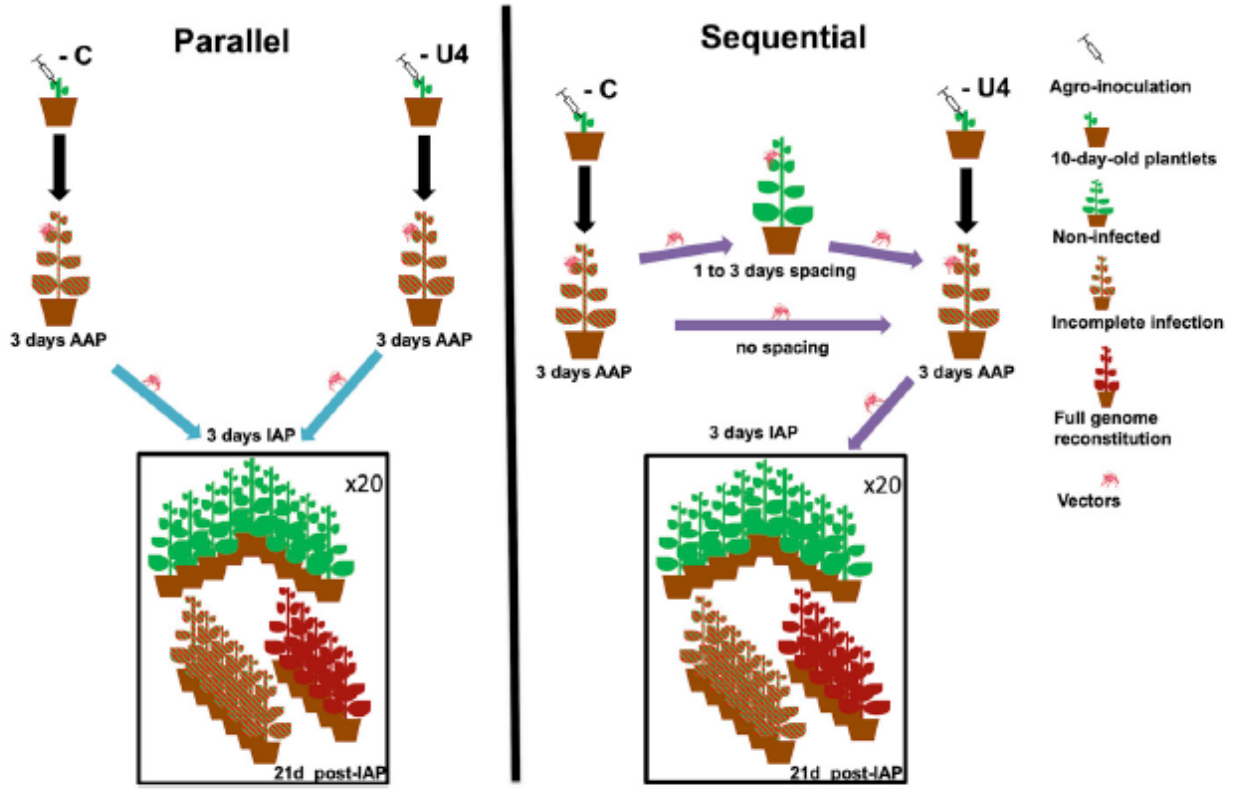


Fig. 1. Schematic of the experimental design of parallel and sequential transmission modes. Details of the two transmission modes are schematized as well as their expected outcome using infections lacking segments C and U4 as an example. For sequential transmissions, we show only the case where aphids first fed on plants infected with FBNSV^{C-} and then fed on plants infected with FBNSV^{U4-}. The age of test plants, the status and age of infections, and the durations of AAPs and of IAP are indicated. The parallel and sequential modes with no time-spacing treatments were performed during experiment 1, and the sequential modes with time-spacing treatments were performed during experiment 2 (the sequential with no time-spacing treatment was independently performed in both experiments). Each experiment was replicated twice. As mentioned in *Materials and Methods*, the parallel transmission mode implies separate acquisition and inoculation of segments, whereas the sequential transmission mode implies temporally separated acquisition of segments but may allow concomitant inoculation.

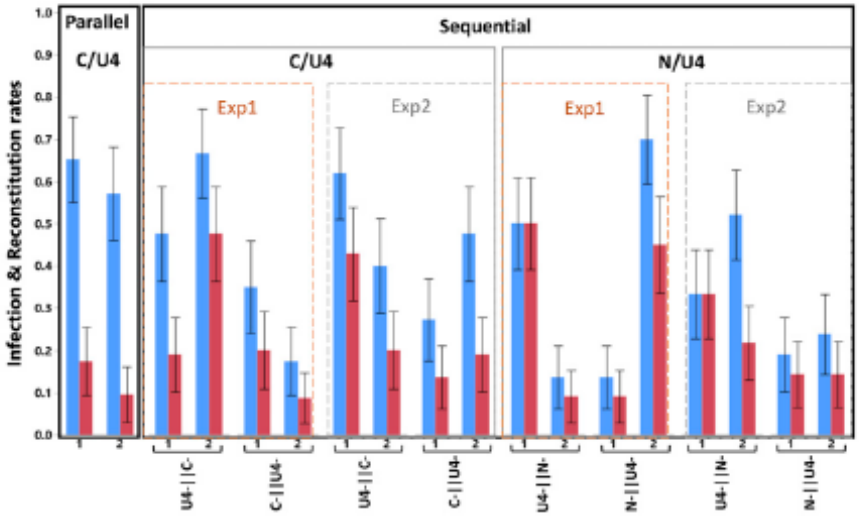


Fig. 2. Infection and full genome reconstitution as a function of the transmission mode. The graph shows the proportion of infected plants (blue) and plants where all three investigated segments are found (reconstitution, red) as a function of the transmission mode (parallel vs. sequential) and the segments involved (parallel: only U4 and C; sequential: either U4 and C or N and U4; the order of the segments in the labels indicates the order of infected plants on which the aphids fed). For sequential transmission and for a given pair of segments and order, we provide the results from the two different experiments, namely, parallel vs. sequential (Exp 1) and sequential with spacing in time (Exp 2), using for the latter only the 0 spacing time treatments. The two experimental replicates (1, 2) are shown for each case, and the error bars are ±1 SE of the mean. Sample size varies from 20 to 23 (Table 1).

Results

FBNSV Full-Genome Reconstitution Is Possible through Both Parallel and Sequential Transmission by Aphid Vectors. Fig. 2 presents the proportions of infected plants and of those containing all eight segments, for all conditions and replicates without any temporal spacing between sequential acquisitions. Though with variable proportions, all conditions generated full FBNSV genome reconstitutions, demonstrating that such reconstitution is possible from infections with incomplete genomes, whether transmitted separately by different aphids (parallel transmission) or acquired sequentially by the same aphid (sequential transmission).

In all replicates/conditions but two, some infections did not result in the reconstitution of the complete genome; the proportion of reconstitutions (red bars) was lower than that of infections (blue bars). Some segments are missing in these cases, and this is discussed in more length in a separate section.

Sequential transmissions performed by single aphids led, over all conditions and replicates, to ~40% infection, which is compatible with what we previously found when transmitting the full genome with one aphid (5). Parallel transmissions performed by two aphids led to ~60% infection. This transmission rate is very close to 64%, the proportion expected if each aphid had a 40% chance to transmit the infection and there was no interaction between them. Sequential transmission, however, led to roughly twice as many reconstitutions of the complete genome than parallel transmissions. Over all test plants, the grand mean of those containing all eight FBNSV segments after parallel transmission is 13.6% while it is of the order of 25% after sequential transmission. This observation holds whether we calculate the mean over both experiments or only the parallel vs sequential experiment, over both pairs of segments or only over C/U4 (which are the only segments involved in parallel transmissions). It is worth noting, however, that the differences in the reconstitution rate between parallel and sequential transmission modes are due only to the U4-||C- treatment (with the data of only the first experiment: $z = 2.011$, $P = 0.044$; with the data of both experiments in Fig. 2: $z = 2.240$, $P = 0.025$). The reconstitution rates of the C-||U4- reverse acquisition order do not differ statistically from those of the parallel transmission treatment (with the data of only the first experiment: $z = 0.046$, $P = 0.963$; with the data of both experiments in Fig. 2: $z = 0.27$, $P = 0.821$).

We further investigated whether the missing segments and their order of acquisition in sequential transmission affected the rates of infection and reconstitution. As already suggested by the above comparison looking at the situations involving segments C and U4, despite variation among replicates, there were more infections and reconstitutions when segment C was acquired first. Across both experiments and replicates, U4-||C- led to ~54% infections and ~32% reconstitutions vs. ~31% infections and ~15% reconstitutions when C was acquired second (C-||U4-). These differences were statistically significant ($z = 2.496$, $P = 0.013$; the experiment effect was not statistically significant: $z = 0.061$, $P = 0.951$) and are further discussed later.

For N and U4 combinations, it is highly remarkable that transmission was efficient even when segment N was not acquired first (i.e., the N-||U4- conditions). Indeed, it has been repeatedly shown that no transmission occurs from plants lacking segment N (14, 15), which encodes for the helper factor of FBNSV (14). We comment on this surprising result more extensively in the *Discussion*. For the moment, we note that, even though more variable between replicates, transmissions involving segments N and U4 led to similar proportions of infection and reconstitution as transmissions involving segments C and U4. However, in contrast to segment C and U4, the reconstitution rates resulting from the two N and U4 acquisition orders did not differ significantly ($z = 0.181$, $P = 0.856$; experiment: $z = -0.908$, $P = 0.364$; experiment by order of acquisition interaction: $z = 0.633$, $P = 0.526$).

FBNSV Infection and Reconstitution when Sequential Acquisition by Aphids Is Spaced in Time. Fig. 3 presents the proportion of infected plants (blue) and of plants with the full genome reconstituted (red) when segments where acquired sequentially but with variable spacing in time of 0, 1, 2, or 3 d. The figure

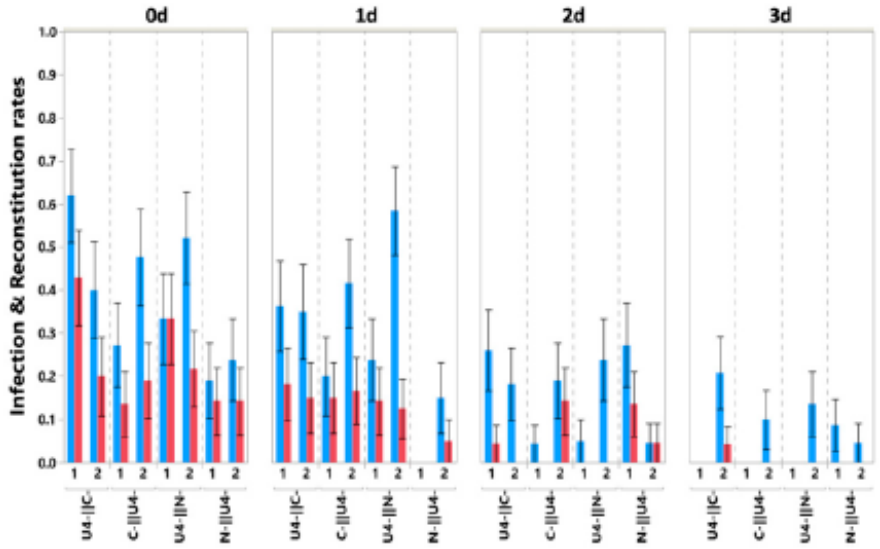


Fig. 3. Infection and full genome reconstitution as a function of time spacing between sequential acquisitions. Shows the proportion of infected plants (blue) and plants in which all three investigated segments are found (reconstitution, red) as a function of time spacing (top x-axis, expressed in days) for the different segments and order combinations. The two experimental replicates are shown for each case, and the error bars are ±1 SE of the mean. Sample size varies from 17 to 24 (Table 1).

clearly shows that increasing spacing in time rapidly leads to a decrease both in infection and reconstitution rates (infection: $z = -3.310$, $P < 0.001$; reconstitution: $z = -2.567$, $P = 0.010$). Indeed, with no spacing (0 d), across all segment combinations and replicates, 65 plants were infected and 38 had the full genome reconstituted out of a total of 170 receiver plants, while after 3 d of spacing, only 13 plants were infected and only 1had a reconstituted genome over a total of 174. Most infections and reconstitutions occurred with no or just 1 d spacing, while very little of either occurred with more than 1 d.

None of the order of acquisition comparisons revealed a statistically significant effect except for the N-||U4- condition that resulted in fewer infections ($z = -2.148$, $P = 0.032$).

Incomplete Genome Reconstitution and Loss of Segments. As mentioned earlier, many infections did not result in the reconstitution of the complete FBNSV genome, implying that at least one segment was missing. Here, we summarize the broad patterns, first on the replicates without time spacing and then on the replicates of time spacing under sequential transmission, and relegate the detailed account to *SI Appendix*, Tables S1 and S2.

Among the 385 receiver plants used in all segment combinations and transmission mode (without time spacing), 158 were infected. Only three of those did not carry segment N, indicating that this segment is rarely involved in incomplete or failed reconstitutions. Segment C was absent from 14.5% (23/158) of infected plants, and there was no clear pattern in the distribution of these cases, either between transmission modes or across segment combinations involved. Segment U4 was absent from 33% (52/158) infected plants. It was particularly absent from recipient plants of the parallel transmission mode (63%) but much less so in all sequential transmission mode treatments (27%), with no strong pattern across segment combinations involved.

Time spacing had no effect on segment N that was present in all infected plants except for the plants mentioned above plus two more plants. Segment C was present in ~80% of infected plants for all time-spacing intervals except for 3 d where it was present in only 6/13 of infected plants. Segment U4's presence gradually declined with the time-spacing interval, being present in 75% infected plants at 0-d spacing and down to 38% at 3 d. No clear pattern could be observed across segment combinations involved (*SI Appendix*, Table S2).

Reconstitution May Occur Already in Aphids upon Sequential Acquisition. Using segment-specific fluorescence in situ hybridization (FISH) by specifically labeling the pairs of sequentially acquired segments (segment C then U4 or N then U4), we investigated whether the FBNSV genome could reconstitute in aphid vectors, when subsets of segments were acquired sequentially. Aphids were placed on plants during 7 d for each acquisition step, with 1 d on healthy plants between the two acquisition steps and 1 d of feeding on water through a Parafilm membrane after the second acquisition to eliminate the virus present in the lumen of the aphid midgut.

For segments C and U4, when U4 was acquired first (i.e., C-||U4- treatment), a few intracellular fluorescent foci showed colocalization of the two segments (Fig. 4*A*, white arrows). However, the majority contained only one of the two, rare foci containing only U4 (green), and numerous others containing only C (red). In contrast, when C was acquired first (U4-||C-), a large proportion of the foci contained both segments (Fig. 4*B*). Even more remarkable, for segments N and U4, we observed a near-perfect colocalization in all foci (Fig. 4 *C* and *D*) for both acquisition orders (U4-||N- or N-||U4-). These observations were very consistent across biological replicates; a total of 16, 10, 6, and 14 individual aphids were examined for treatments C-||U4-, U4- ||C-, U4-||N-, and N-||U4-, respectively. At least 10 cells of the anterior midgut (AMG) were zoomed for each individual aphid. For each treatment, all zoomed cells revealed the same pattern, illustrated in Fig. 4. All controls for these experiments, particularly those confirming the segment specificity of the probes, are shown in *SI Appendix*, Fig. S1 (see the without C and without U4 panels).

Altogether, our results demonstrate that full-genome reconstitution can occur at very early stages of the virus cycle within the aphid vectors, soon after entry of the segments into midgut cells.

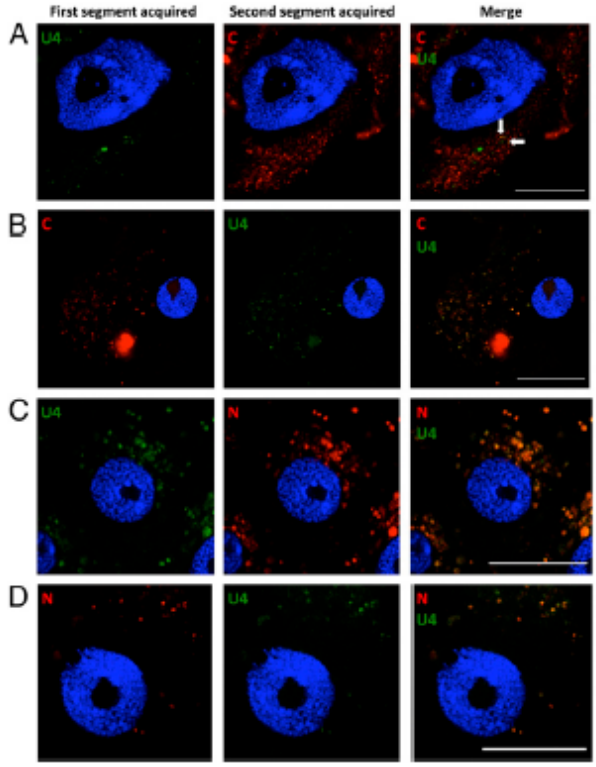


Fig. 4. Localization of sequentially acquired DNA segments of FBNSV in aphid AMG cells. Viral DNA is labeled by FISH in AMG cells of viruliferous aphids and observed by confocal microscopy. The green probe targets segment U4, and the red probes target either C or N in the corresponding panels. The accumulation of FBNSV DNA was similarly revealed in all observed cells (>10 cells per midgut) from 16, 10, 6, and 14 viruliferous aphids for *A*, *B*, *C*, and *D*, respectively. A representative image of each case is shown to illustrate the results. *A* and *B* represent the sequential acquisition of segments C/U4, and *C* and *D* represent that of segments N/U4, in the indicated sequential order. The white arrows show two foci containing both C and U4 segments. All images correspond to single optical sections. Cell nuclei are stained with DAPI (blue). Scale bar, 25 μm.

The Reconstituted Viral Genome Can Be Transmitted as an Integral FBNSV Genome. We finally verified that the FBNSV could be aphid transmitted from plants containing a reconstituted genome in a way similar to that observed from plants initially agro- or aphid inoculated with the complete genome, as detailed in *Materials and Methods*. Over the four transmission assays (one donor plant with a reconstituted genome from each of the sequential transmission treatments C-||U4-, U4-||C-, N-||U4-, and U4-||N-), we obtained 36.36% (28/77) successful transmissions. This transmission rate is close to that we typically observe for transmissions by a single aphid under our experimental conditions, from plants initially infected with the full FBNSV genome. There was variability among treatments, at least partially reflecting the relatively small sample sizes (U4-||C-: 45.45% (10/22); C-||U4-: 31.25% (5/16); U4-||N-: 50% (10/20); N-||U4-: 15.79% (3/19)).

In the 28 infected plants, we observed the absence of 1 or 2 segments in only 3 plants, as follows: segment U4 was absent from 1 plant of the U4-||C- treatment, segments C and N were both absent from 1 plant of the C-||U4- treatment, and segments C and U4 were both absent from another plant of the C-||U4- treatment. Finally, one plant of the U4-||C- treatment tested negative for segments N and U4 and had one qPCR replicate slightly above (cycle threshold [Ct] = 30.85) and one slightly below (Ct = 30.28) the threshold value of segment C (Ct = 30.43). Repeating the qPCRs for this individual plant yielded almost identical values. We decided to consider this plant as not infected, as indicated by most of the evidence.

These results thus confirm that the reconstituted FBNSV can successfully be transmitted by aphids to uninfected plants as the integral viral genome.

Discussion

Our results clearly show that the reconstitution of the complete FBNSV genome through aphid transmission from distinct plants infected with incomplete genomes is possible. This result demonstrates that all FBNSV segments do not need to be acquired and/or inoculated concomitantly and thus has important implications for our understanding of how multipartite viruses may be coping with the so far assumed cost of a multipartite way of life related to the maintenance of their genomic integrity during host-to-host transmission (1, 3–5, 14, 21).

The nonconcomitant transmission of genomic segments could potentially occur in all vector-borne multipartite viruses, even those not following the circulative transmission mode. There is no reason why the parallel mode would be restricted to such viruses. The relevance of the sequential mode would depend on the opportunity of viral particles from different acquisitions to co-occur in the vector. This should not be a problem for semipersistent viruses that may be retained in their vectors from hours to days (e.g., the bipartite criniviruses transmitted by whiteflies (22)) to weeks/months/years (e.g., the bipartite nepoviruses and tobraviruses transmitted by nematodes (23)). Even nonpersistent viruses could potentially use the sequential mode because they may cause alterations of their vectors' behavior that make the vectors visit more plants more rapidly (e.g., the tripartite cucumber mosaic virus (CMV) (24)).

We found that genome reconstitution could occur when incomplete infections are transmitted either by distinct aphids, a mode we termed parallel transmission, or by the same aphid sequentially acquiring viral segment sets from different plants to transmit them together to a receiver plant. These modes represent the two extremes of a continuum; reconstitutions could potentially, and probably most likely, occur in nature through any mixture of these modes.

When transmission is parallel, reconstitution can only occur in the receiver plant. Our results show that the overall infection levels we observed, ~60%, are compatible with the independent transmission of the two incomplete infections by the two aphids that potentially carry them. However, only ~22% of these infected plants ended up with a reconstituted complete genome, with the remaining plants lacking either segment C or, for unknown reasons, more frequently U4. These segments may be missing either because they were not transmitted or because they were lost in the receiver plant after transmission. The former could occur because the aphids that were supposed to transmit them did not acquire or acquired but did not inoculate them to the receiver plant in sufficient amounts. Consistent with this is our previous demonstration that very few viral particles are transmitted from single aphids (5), likely resulting in frequent segment loss sometimes resulting in infection failure. Unfortunately, this previous study quantifying the bottleneck that aphid transmission imposes to FBNSV populations did not specifically monitor the transmission of segments C and U4, and thus, we have no data allowing a direct comparison. If missing segments were lost after transmission, we see no obvious reason explaining why segment U4 should be lost

more often than segment C, as it is more frequent than segment C both in plants (8) and aphids (25). Our results on the transmission from plants with the reconstituted genomes indicate that segment loss is much rarer when donor plants have all eight segments, suggesting that transmission from plants with incomplete infections may be particular in this respect. This matter thus deserves further investigation. Incomplete infections are rarely reported on field isolates of nanoviruses, which could be due to several compatible explanations, as follows: a reporting bias, a lower performance relative to full infections, or complementation of initially incomplete infections through the processes investigated here.

Sequential transmission without time spacing led to infection levels in agreement with previous single-aphid transmission results (5). Reconstitution of the complete genome was much more frequent under this transmission mode, as ~50% of infected plants carried all genome segments, as opposed to only 22% through parallel transmission. It is remarkable that the segments acquired sequentially colocalized within intracellular granules (fluorescent foci) in aphid gut epithelial cells, indicating that under this transmission mode, reconstitution can occur at very early stages of the virus cycle within the body of its vector. This observation suggests similar accumulation sites for the eight FBNSV segments in aphid midgut cells, whether acquired concomitantly (26) or sequentially (this study), which may explain the higher reconstitution rate, when compared to parallel transmission.

We found an unexpected effect due to the order of acquisition; more infections and reconstitutions occurred when segment C was acquired first (U4-||C- treatment) than second (C-||U4 treatment). The fact that we only rarely manage to obtain infections lacking segment C through agroinoculation may suggest that, at least under our conditions, this segment plays an important role that was not revealed in other experimental conditions (20). Alternatively, we cannot exclude that segment C plays a yet unanticipated role in FBNSV transmission in the aphids, resulting into higher transmission if it is acquired early.

A much more remarkable result, however, concerns segment N. As previously stated, it is established that the presence of this segment is mandatory for aphid transmission (20, 26). It is thus considered as the segment producing the helper component required for FBNSV transmission (27). So far, all described helper components in the vector transmission of plant viruses must be acquired together with or before the viral particles (28–31). Our sequential experiment results, however, clearly demonstrate that this is not the case for the N segment of FBNSV; it is possible to have transmission, successful infection, and even full-genome reconstitution even if segment N (and/or its expression products) is acquired after the other segments (i.e., our N-||U4- treatment). This suggests that the segments acquired by the aphid from plants that do not contain the segment N can survive or wait in the aphid's gut until segment N (and/or its expression products) is acquired during the second acquisition step. The single full-genome reconstitution (out of a total of 13 infected plants) we observed with a 3-d spacing shows that this complementation can occur with at least this delay between the 2 acquisition steps. We have previously shown that when the virus is acquired from plants lacking segment N, no viral segment is visible in aphid midgut cells, while when N is present in the infected plant, they all penetrate and accumulate together into cytoplasmic granules (26). Our confocal microscopy results show that the same process occurs during the sequential acquisition, whether N is acquired prior to or after the other segments. The specific mechanisms through which the helper of FBNSV acts need further studies. We can note, however, more to the point of this study, that by being able to potentiate the transmission of viral segments previously or subsequently acquired by aphids, this helper allows for a much larger spectrum of possible reconstitutions of the virus' full genome. Consequently, genome reconstitution through sequential acquisition and transmission may help FBNSV, and potentially other nanoviruses, alleviate the cost of the maintenance of genome integrity upon host-to-host transmission.

Our time-spacing results were very surprising. We observed that infection and reconstitution declined as time spacing increased. FBNSV, as all nanoviruses, is supposed to be transmitted according to the circulative nonpropagative mode (32) and thus travels through the aphid's body, gut, and salivary gland cells without replicating. Depletion of the first acquired infection (drastic reduction of transmission rate) during the short time intervals we used is thus surprising (33). But what is even more surprising is that the second acquired infection should lead to successful infections on its own, even if the first infection was depleted, at a rate that in principle should not depend on the time since the first acquisition. This is clearly not what happened in our experiments. One potential explanation could be that the ability of aphids to successfully transmit viral infections decreases with time. As already noted above, this would be unlikely within the short time intervals we used (33). An alternative explanation could be that the first acquisition somehow triggers a mechanism that inhibits transmission after the second acquisition, and the strength of this inhibitory mechanism would increase with spacing time between the two acquisitions. Because it is beyond the scope of this work, we leave this very interesting phenomenon to the future investigation it deserves We note for the moment that this unanticipated negative effect of time spacing between sequential acquisitions may limit the potential of reconstitutions to help multipartite viruses overcome the cost linked to genomic integrity upon host-to-host transmission.

It is currently impossible to quantitatively evaluate the potential of complete genome reconstitution through nonconcomitant transmission of different segments to alleviate the host-to-host transmission cost of multipartitism. Such evaluation would require the precise knowledge of ecological and behavioral variables, such as the prevalence and density of infected plants, the density of vectors, how often vectors move from plant to plant, and what governs their individual host plant choice, all of which is currently unknown. Moreover, in the introduction, we stated that our study partly addresses this issue. We used this term because we acknowledge that here we demonstrate that reconstitution is possible between incomplete infections that are viable (even though they do not necessarily perform as well as complete genome infections). The potential of genome reconstitution to alleviate the cost of maintenance of genomic integrity would be much larger if it could be shown that such reconstitution occurs also from uptake within aphids or release within plants of segments sets that are not viable on their own. We relegate this to future investigation. This study nevertheless demonstrates that the concomitant transmission of the different genome components is not mandatory. Thus, after showing that multipartite viruses may solve the within-host cost of genomic integrity maintenance through a supracellular, distributed functioning (12), the present study shows that they may solve the between-host cost of genomic integrity maintenance through a supraindividual host, distributed transmission. Furthermore, this functioning suggests that these viruses have an immense potential for reassortment, compared to, e.g., segmented viruses that in order to

reassort need to coinfect the same host cell since parental genotypes could exchange segments without ever having to meet each other, a feature that may greatly influence their evolution.

Materials and Methods

Virus Isolate, Clones, and Plant Agroinoculation. The first FBNSV isolate, discovered on faba beans (*Vicia faba*) in Ethiopia (15), was characterized and cloned by Grigoras and colleagues (34). Each genome segment was inserted as a head-to-tail dimer into the binary plasmid pBin19. Eight plasmids, each containing a dimer of a segment, together constitute the FBNSV infectious clone. The impact on plant infection and plant-to-plant transmission of the absence of any of the segments was investigated (20). The authors showed that the presence of segments R, S, and M was mandatory for infection, but it was possible to infect plants when one of segments C, N, U1, U2, or U4 was missing. They further found that the absence of segments C or U4 did not affect virus accumulation and transmission, while the absence of U1 or U2 decreased symptom severity and virus titer. They also established that the absence of the N segment impedes plant-to-plant aphid transmission, even though it does not seem to affect within-plant viral growth.

In this study, we inoculated 10-d-old faba bean plantlets (*V. faba* cv. Sevilla; Vilmorin) with $FBNSV^{complete}$, $FBNSV^{C-}$, $FBNSV^{N-}$, or $FBNSV^{U4-}$ agrobacteria solutions (strain COR308), as earlier described (34). While $FNSV^{complete}$ corresponds to a mixture at equal ratio of eight agrobacterium cultures, with each containing one of the eight plasmids described above, the indicated culture and thus the indicated segment was omitted in $FBNSV^{C-}$, $FBNSV^{N-}$, or $FBNSV^{U4-}$. Plants were maintained as previously described in ref. 26. Plant infection was controlled by total DNA extraction and qPCR (see below) for 21 d after agroinoculation. Even though we were able to obtain relatively many infected plants through agroinoculation of $FBNSV^{N-}$ and $FBNSV^{U4-}$ at 51.06% (=96/188) and 52.81% (=94/178), respectively, we could only obtain 13.08% (=56/428) infected plants with $FBNSV^{C-}$.

Aphid Transmission. For all experiments, we used aphid colonies of *Acyrthosiphon pisum* (clone 210) maintained in a controlled chamber as previously described in ref. 26. Transmission tests were made under similar conditions using faba beans as host plants.

We tested whether the complete virus genome could be reconstituted if 1) different aphids fed on different plants carrying incomplete infections missing different segments came together on the same recipient plant, a mode we hereafter call parallel transmission, and if the same aphid sequentially fed on different plants carrying incomplete infections missing different segments was then transferred on one recipient plant, a mode we hereafter call sequential transmission. We use these terms for ease of language, although the parallel mode implies separate acquisition as well as separate inoculation of the distinct sets of segments, while the sequential mode implies a sequential acquisition of the distinct sets of segments and may allow their concomitant inoculation. Subsequently 2), we asked whether sequential transmission could successfully reconstitute the complete viral genome if the sequential feedings of the aphid were separated in time (from 0 up to 3 d). Finally, 3) we tested that the reconstituted genomes were successfully transmitted as the genuine parental genotype. Fig. 1 provides a schematic representation of the experimental design.

Parallel vs. sequential transmission. In a first experiment, we tested whether the complete genome could be reconstituted through parallel or sequential transmission and at which rate. For the parallel transmission test (Fig. 1, blue arrows), 2 $FBNSV^{C-}$ and 2 $FBNSV^{U4-}$ plants, 1 of each per replicate, were used as donors and 40 aphids were placed on each plant. Aphids fed on plants during an acquisition access period (AAP) of 3 d to acquire the respective set of segments. Then, we took one aphid from each donor plant and placed them on two different leaflets of the same recipient plant in a dip-cage (cage allowing the maintenance the aphid on one leaflet) during an inoculation access period (IAP) of 3 d.

For sequential transmission, we used one $FBNSV^{C-}$ and two $FBNSV^{U4-}$ donor plants and further tested whether the order of acquisition of the different incomplete infections affected the probability of complete genome reconstitution. Thus, for each condition, 40 aphids were placed on each donor plant (either $FBNSV^{C-}$ or $FBNSV^{U4-}$) for the first acquisition step. Then, we directly transferred aphids on the plants for the second acquisition step with the complementary set of segments. Each AAP step lasted 3 d. Finally, aphids were individually transferred into a dip-cage on a leaflet of a receiver plant (one aphid/receiver) during an IAP of 3 d.

Transmission does not occur in the absence of segment N (20, 26), which is thus considered to encode for a transmission helper component (27, 35). For this reason, we used only incomplete infections lacking either segments C or U4 for the parallel transmission mode. For the sequential transmission mode, we used the following combinations: 1) $FBNSV^{C-}$ then $FBNSV^{U4-}$ (C-||U4-), 2) $FBNSV^{U4-}$ then $FBNSV^{C-}$ (U4-||C-), 3) $FBNSV^{N-}$ then $FBNSV^{U4-}$ (N-||U4-), and 4) $FBNSV^{U4-}$ then $FBNSV^{N-}$ (U4-||N-). The third combination allowed us to test whether the successful transmission of segments can occur even if the N segment, which is necessary for aphid transmission, cannot be acquired during the first acquisition but is acquired during the second acquisition.

The number of receiver plants per transmission mode and replicates are given in Table 1. All replicates for both transmission modes were conducted at the same time.

Sequential transmission spaced in time. We subsequently tested whether the time separating the two sequential acquisitions had an effect on the reconstitution of the complete viral genome.

For this type of experiment (Fig. 1, purple arrows), 200 to 250 aphids were placed on a first donor plant infected with 1 set of segments. Then, 40 aphids were transferred on a second donor plant infected with the complementary set of segments. When the two sequential acquisitions were not spaced in time (0 d), the transfer occurred immediately after the end of the first AAP. The remaining aphids were placed onto a healthy plant. Every 24 h, up to a maximum spacing of 3 d, 40 aphids were taken from the healthy plant and placed on a second donor plant infected with the complementary set of segments. Each AAP lasted 3 d. Finally, at the end of the second AAP, aphids were individually placed in a dip-cage on a leaflet of a receiver plantlet during an IAP of 3 d.

Two replicates were carried out for each combination and spacing time (Table 1). However, because the number of receiver plants involved was very large, we first conducted one replicate with all segment combinations and spacing times, and 2 wk later, conducted the second replicate.

All IAPs were stopped with a spray of insecticide (Pirimor; Certis) except where otherwise mentioned. For all experiments, 21 d after the end of the IAP, the total DNA of symptomatic plants was extracted, and qPCR was performed to test for the presence of segments C, N, and U4 (see qPCR details below).

Transmission of the reconstituted complete FBNSV genome. To verify if the reconstituted complete FBNSV infections can be transmitted to healthy plants, just as the genuine parental genotype, we used the infected faba beans from the sequential transmission experiments with 1 d of time spacing as source plants. In this case, at the end of the IAP of the second replicate, plants were not treated with the insecticide solution. Instead, all aphids were individually removed from the plants and a spray of Marseille soap in water (~1 wt/vol concentration) was applied onto the plants to eliminate potentially remaining aphids. We allowed for 21 d of disease progression, during which plants were isolated to avoid any aphid contamination. After this period, one plant from each of the C-||U4-, U4-||C-, N-||U4-, and U4-||N- sequential transmission experiments with the reconstituted complete FBNSV genome (checked by qPCR) was used as a donor for the transmission of the reconstituted virus. Forty aphids were placed onto one donor plant of each combination. After an AAP of 3 d, 1 aphid was placed onto each of 20 to 24 receiver plantlets for an IAP of 3 d. Finally, 21 d after the end of the IAP, the total DNA of symptomatic receiver plants was extracted and the segments C, N, and U4 were quantified by qPCR.

Viral DNA Extraction and qPCR Detection. Total DNA extraction from plants was performed as described in ref. 36. FBNSV infection causes faba bean leaf rolling and plant stunting in the early stage of infection and chlorosis and necrosis in the late stage. We extracted DNA from all plants showing some signs of leaf rolling and stunting and quantified segments C, N, or U4 in all these plants. On the two upper leaf levels of symptomatic faba beans, three circular areas of a 0.6-cm diameter were squashed onto a Whatman paper disk. Then, we deposited each disk onto the filter of a 200-µL micropipette tip and added 100 µL of modified Edwards buffer (200 mM Tris-HCl [pH 7.5], 25 mM EDTA, 250 mM NaCl, 0.5% SDS, 1% PVP40, and 0.2%

Data Analysis. Based on the Ct threshold values defined above, we determined whether each segment was present or absent from each recipient plant. In all cases but one, both technical replicates of each sample were either above or below the threshold. The only exception concerned segment C of one sample that we conservatively considered as absent.

All asymptomatic plants were considered noninfected. All plants showing symptoms were qPCR processed, and if any of the C, N, or U4 segments was present in the sample, we considered the plant as infected. When all three segments were present in recipient plants, we considered that a complete infection was reconstituted.

To investigate whether transmission modalities (parallel or sequential) affected the probability of reconstitution (P(R)) or infection (P(I)), we ran Generalized Linear Mixed Models considering that P(R) and P(I) were binomially distributed with logit-distributed errors, transmission modality and segments involved as fixed factors, and experimental replicate as a random factor. Spacing time was added as a covariate when relevant. We used the lme4 package in R to run these analyses. Figures were drawn using JMP 13.2.1. The data and R scripts for these analyses can be found at https://entrepot.recherche.data.gouv.fr/citation?persistentId=doi:10.57745/EH9KRW.

Data Availability. Tables with results of qPCR reactions allowing us to infer the presence of viral segments have been deposited in https://entrepot.recherche.data.gouv.fr/dataverse/inrae (https://entrepot.recherche.data.gouv.fr/citation?persistentId=doi:10.57745/EH9KRW).

ACKNOWLEDGMENTS. This work was funded by the Agence nationale de la recherche grants Reassort ANR-20-CE02-0016 and Nanovirus ANR-18-CE92-0028. M.Y. and S.B. acknowledge support from INRAE Département Santé des Plantes et Environnement, Y.M. from CNRS and IRD, and J.L.-Z. from IRD. The authors acknowledge valuable technical help provided by Ms Sophie Le Blaye. Many thanks to Bernard Francq and Thomas Mathew for their insights on tolerance intervals.

Author affiliations: [a]PHIM (Plant Health Institute Montpellier, Université Montpellier, INRAE (Institut national de recherche pour l'agriculture, l'alimentation et l'environnement), CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), IRD (Institut de recherche pour le développement), Institut Agro, Montpellier, F-34398 France; and [b]MIVEGEC (Maladies Infectieuses et Vecteurs: Ecologie, Génétique, Evolution et Contrôle), Université Montpellier, CNRS, IRD (Institut de recherche pour le développement), Montpellier, F-34394 France

1. A. Lucía Sanz, S. Manrubia, Multipartite viruses: Adaptive trick or evolutionary treat? *NPJ Syst. Biol. Appl.* **3**, 34. (2017).
2. S. Ojosnegros *et al.*, Viral genome segmentation can result from a trade-off between genetic content and particle stability. *PLoS Genet.* **7**, e1001344 (2011).
3. A. Sicard, Y. Michalakis, S. Gutiérrez, S. Blanc, The strange lifestyle of multipartite viruses. *PLoS Pathog.* **12**, e1005819 (2016).
4. J. Iranzo, S. C. Manrubia, Evolutionary dynamics of genome segmentation in multipartite viruses. *Proc. Biol. Sci.* **279**, 3812-3819 (2012).
5. R. Gallet *et al.*, Small bottleneck size in a highly multipartite virus during a complete infection cycle. *J. Virol.* **92**, e00139-e18 (2018).
6. B. Gronenborn, Nanoviruses: Genome organisation and protein function. *Vet. Microbiol.* **98**, 103–109 (2004).
7. A. Lal *et al.*, Nanovirus disease complexes: An emerging threat in the modern era. *Front Plant Sci* **11**, 558403 (2020).
8. A. Sicard *et al.*, Gene copy number is differentially regulated in a multipartite virus. *Nat. Commun.* **4**, 2248 (2013).
9. M. Mansourpour *et al.*, Effects of an alphasatellite on life cycle of the nanovirus Faba bean necrotic yellows virus. *J. Virol.* **96**, e0138821 (2022).
10. Z. Hu *et al.*, Genome segments accumulate with different frequencies in Bombyx mori bidensovirus. *J. Basic Microbiol.* **56**, 1338–1343 (2016).
11. B. Wu, M. P. Zwart, J. A. Sánchez-Navarro, S. F. Elena, Within-host evolution of segments ratio for the tripartite genome of Alfalfa mosaic virus. *Sci. Rep.* **7**, 5004 (2017).
12. A. Sicard *et al.*, A multicellular way of life for a multipartite virus. *eLife* **8**, e43599 (2019).
13. M. Betancourt, A. Fereres, A. Fraile, F. García-Arenal, Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. *J. Virol.* **82**, 12416–12421 (2008).
14. Y. Michalakis, S. Blanc, The curious strategy of multipartite viruses. *Annu. Rev. Virol.* **7**, 203–218 (2020).
15. A. Franz, K. M. Makkouk, H. J. Vetten, Host range of faba bean necrotic yellows virus and potential yield loss in infected faba bean. *Phytopathol. Mediterr.* **36**, 94-103 (1997).
16. I. Grigoras *et al.*, Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *J. Gen. Virol.* **95**, 1178–1191 (2014).
17. A. D. Abraham *et al.*, Two distinct nanovirus species infecting faba bean in Morocco. *Arch. Virol.* **155**, 37–46 (2010).
18. M. Lotfipour, S. A. A. Behjatnia, M. Dalil Ara, C. Ratti, The full-length genome characterization and diversity of faba bean necrotic stunt virus in Iran. *Eur. J. Plant Pathol.* **157**, 239-250 (2020).
19. T. Timchenko *et al.*, Infectivity of nanovirus DNAs: Induction of disease by cloned genome components of Faba bean necrotic yellows virus. *J. Gen. Virol.* **87**, 1735–1743 (2006).
20. I. Grigoras *et al.*, Nanovirus DNA-N encodes a protein mandatory for aphid transmission. *Virology* **522**, 281–291 (2018).
21. Y. Michalakis, S. Blanc, Editorial overview: Multicomponent viral systems. *Curr. Opin. Virol.* **33**, vi–ix (2018).
22. A. Y. S. Chen, G. P. Walker, D. Carter, J. C. K. Ng, A virus capsid component mediates virion retention and transmission by its insect vector. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16777–16782 (2011).
23. D. J. F. Brown, W. M. Robertson, D. L. Trudgill, Transmission of viruses by plant nematodes. *Annu. Rev. Phytopathol.* **33**, 223-249 (1995).
24. K. E. Mauck, C. M. De Moraes, M. C. Mescher, Deceptive chemical signals induced by a plant virus attract insect vectors to inferior hosts. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3600–3605 (2010).
25. A. Sicard *et al.*, Circulative nonpropagative aphid transmission of nanoviruses: An oversimplified view. *J. Virol.* **89**, 9719–9726 (2015).
26. J. Di Mattia *et al.*, Route of a multipartite Nanovirus across the body of its aphid vector. *J. Virol.* **94**, e01998-e19 (2020).
27. A. W. E. Franz, F. van der Wilk, M. Verbeek, A. M. Dullemans, J. F. J. M. van den Heuvel, Faba bean necrotic yellows virus (genus Nanovirus) requires a helper factor for its aphid transmission. *Virology* **262**, 210–219 (1999).
28. B. Kassanis, D. A. Govier, New evidence on the mechanism of aphid transmission of potato C and potato aucuba mosaic viruses. *J. Gen. Virol.* **10**, 99–101 (1971).
29. B. Kassanis, D. A. Govier, The role of the helper virus in aphid transmission of potato aucuba mosaic virus and potato virus C. *J. Gen. Virol.* **13**, 221–228 (1971).
30. M. C. Y. Lung, T. P. Pirone, Studies on the reason for differential transmissibility of cauliflower mosaic virus isolates by aphids. *Phytopathology* **63**, 910 (1973).
31. G. Lu *et al.*, Tenuivirus utilizes its glycoprotein as a helper component to overcome insect midgut barriers for its circulative and propagative transmission. *PLoS Pathog.* **15**, e1007655 (2019).
32. Y. Z. A. Gaafar, H. Ziebell, Aphid transmission of nanoviruses. *Arch. Insect Biochem. Physiol.* **104**, e21668 (2020).
33. S. Watanabe, A. Bressan, Tropism, compartmentalization and retention of banana bunchy top virus (Nanoviridae) in the aphid vector Pentalonia nigronervosa. *J. Gen. Virol.* **94**, 209–219 (2013).
34. I. Grigoras *et al.*, Reconstitution of authentic nanovirus from multiple cloned DNAs. *J. Virol.* **83**, 10778–10787 (2009).
35. R. Froissart, Y. Michalakis, S. Blanc, Helper component-transcomplementation in the vector transmission of plant viruses. *Phytopathology* **92**, 576–579 (2002).
36. J. Di Mattia *et al.*, Co-acquired nanovirus and geminivirus exhibit a contrasted localization within their common aphid vector. *Viruses* **12**, 299 (2020).
37. R. Gallet, F. Fabre, Y. Michalakis, S. Blanc, The number of target molecules of the amplification step limits accuracy and sensitivity in ultradeep-sequencing viral population studies. *J. Virol.* **91**, e00561-17 (2017).
38. G. Sharma, T. Mathew, One-sided and two-sided tolerance intervals in general mixed and random effects models using small-sample asymptotics. *J. Am. Stat. Assoc.* **107**, 258–267 (2012).
39. B. G. Francq, D. Lin, W. Hoyer, Confidence, prediction, and tolerance in linear mixed models. *Stat. Med.* **38**, 5603-5622 (2019).
40. M. S. Verneney, E. Pirolles, S. Blanc, A. Sicard, Localizing genome segments and protein products of a multipartite virus in host plant cells. *Bio Protoc.* **9**, e3443 (2019).

PNAS

Supplementary Information for

Non-concomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution.

Jérémy Di Mattia, Babil Torralba, Michel Yvon, Jean-Louis Zeddam, Stéphane Blanc and Yannis Michalakis

Supplementary Information Text

Defining qPCR Ct thresholds beyond which segments are considered absent through tolerance limits.

To reliably claim that full genome reconstitution has occurred, we must be able to ascertain when any specific segment can be considered absent from or present in an infected plant. To do this, for each segment, we performed qPCR using the focal segment's primers on infected plants lacking this segment, that is plants inoculated with all segments but the focal segment. We thus recorded the Ct's obtained for a focal segment when it is itself lacking from an infected plant but when all other segments were present. We then established a Ct distribution for each of the three "missing segments" of interest, C, N and U4 from a number of infected plants with two technical replicates for each sample (sampling and qPCR conditions were performed as described in the main text).
Based on these Ct distributions we calculated for each segment the (0.95, 0.95)-one-sided lower tolerance interval, *i.e.* the Ct value above which would lie 95% of future comparable samples with 95% confidence, given the observed distribution of Ct values. Based on this method, in the genome reconstitution experiments we considered a segment to be absent from a sample if its Ct value lied above the corresponding threshold.

In order to calculate these thresholds, we used the method described by (**Francq et al. 2019**) for the one random factor design (in our case the random factor is the sampled plant, accounting for the fact that the two technical replicates originate from the same sample). We adapted the R script provided by the authors in the appendix of their paper by adjusting the quantiles of the Normal, t and X2 distributions to our desired 95% coverage and 95% confidence proportions, and the sample-size related parameters to our sample sizes for each segment.

The sample sizes to obtain these tolerance limit values were 36 infected plants (2 technical replicates each) for missing-segment N and 56 infected plants (2 technical replicates each) for missing-segment U4. In the case of U4, we did not consider one of the infected plants because its two technical replicates yielded Ct values around 24, and were identified as outliers (below the lower quartile – 1.5 the interquartile range), indicating contamination at some stage (the next Ct values were >26 and these were included in our estimation). Incorporating this sample would have yielded a threshold Ct value of 27.66, instead of 28.29. Including this discarded individual plant in establishing the threshold value changes the status, from containing to non-containing segment U4, of only three plants: one in the U4-||C- parallel transmission, one in the C-||U4- sequential transmission with one day of time spacing, and one in the U4-||N- sequential transmission with one day of time spacing treatments. Hence considering or not this individual does not affect our results and inferences.

As mentioned in the main text, it was very difficult to obtain infected plants without segment C. We only had Ct values from 7 such plants (with two replicates each). Because we did not want to base distributional properties on such a small sample size, we also used Ct values from 15 uninfected plants (with two replicates each). We first ran a mixed model with infection or not as a fixed factor and sampled plant as random factor. This model showed that the infection status did not significantly affect the Ct values of the missing C segment ($p=0.68$), with very similar means (infected/non-infected: 35.13/34.75). Subsequently, we calculated the threshold Ct using the distribution from these 22 plants and performed all our analyses with this threshold value.

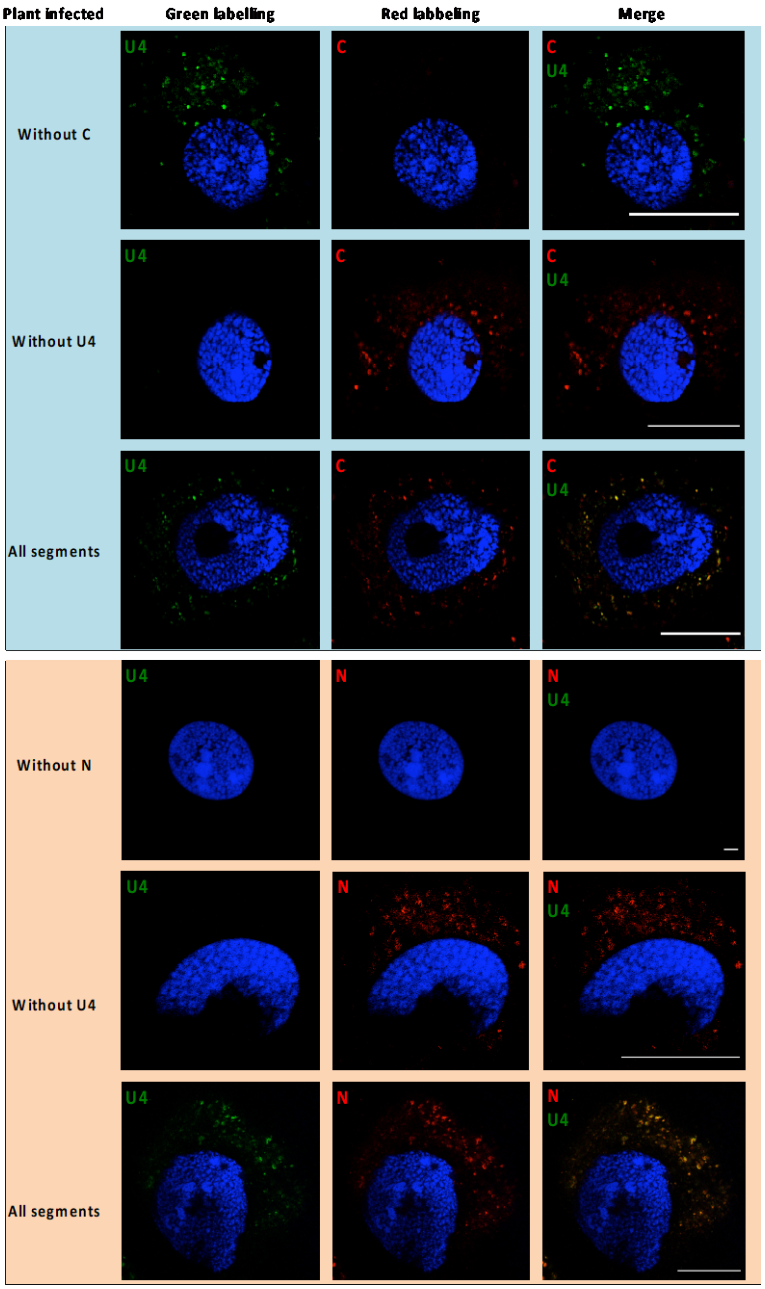


Fig. S1. Localization of FBNSV segments in aphid AMG cells. Viral DNA is labeled by FISH in the AMG of viruliferous aphids and observed by confocal microscopy. The green probe targets segment U4 and the red probes target either C or N in the respective panels. Each panel corresponds to the control experiment of C/U4 (light blue part) and N/U4 (light orange part) of figure 4. The "without C", "without N" or "without U4" panels show midguts of aphids fed on plant lacking the corresponding segments and thus control for probe specificity. The "All segments" panels show the localization of C/U4 or N/U4 in midguts from aphid fed on plant containing all eight segments (concomitant acquisition of all segments). The accumulation of FBNSV DNA was similarly revealed in all observed cells (>10 cells per midgut) from 18, 10, 23, 10, 13 and 26 viruliferous aphids, respectively from top to bottom rows; A representative image of each case is shown to illustrate the results. All images correspond to single optical sections. Cell nuclei are stained with DAPI (4,6-diamidino-2-phenylindole; blue). The scale bar represents 25 µm.

Spacing	N	N_inf	N_C	% C		N_N	% N	N_U4	% U4
0	170	65	51	78		62	95	49	75
1	172	51	37	73		50	98	26	51
2	174	28	24	86		27	96	10	36
3	172	13	6	46		13	100	5	38

Table S1. Missing segments per spacing time. Number of inoculated (N) and infected (N_inf) plants, number of plants containing each segment (N_segment name) and percentage of plants containing each segment among infected plants (% segment name) as a function of spacing time (in days).

spacing	condition	rep	N	N_inf	N_C	% C	N_N	% N	N_U4	% U4
0	U4-||C-	1	21	13	11	85	12	92	11	85
0	U4-||C-	2	20	8	7	88	8	100	4	50
0	C-||U4-	1	22	6	5	83	5	83	4	67
0	C-||U4-	2	21	10	7	70	10	100	7	70
0	U4-||N-	1	21	7	7	100	7	100	7	100
0	U4-||N-	2	23	12	6	50	12	100	8	67
0	N-||U4-	1	21	4	3	75	4	100	4	100
0	N-||U4-	2	21	5	5	100	4	80	4	80
1	U4-||C-	1	22	8	6	75	8	100	6	75
1	U4-||C-	2	20	7	6	86	6	86	3	43
1	C-||U4-	1	20	4	4	100	4	100	3	75
1	C-||U4-	2	24	10	6	60	10	100	7	70
1	U4-||N-	1	21	5	4	80	5	100	3	60
1	U4-||N-	2	24	14	8	57	14	100	3	21
1	N-||U4-	1	21	0						
1	N-||U4-	2	20	3	3	100	3	100	1	33
2	U4-||C-	1	23	6	5	83	6	100	2	33
2	U4-||C-	2	22	4	4	100	4	100	0	0
2	C-||U4-	1	23	1	1	100	1	100	0	0
2	C-||U4-	2	21	4	3	75	4	100	4	100
2	U4-||N-	1	20	1	1	100	0	0	0	0
2	U4-||N-	2	21	5	4	80	5	100	0	0
2	N-||U4-	1	22	6	5	83	6	100	3	50
2	N-||U4-	2	22	1	1	100	1	100	1	100
3	U4-||C-	1	21	0						
3	U4-||C-	2	24	5	2	40	5	100	3	60
3	C-||U4-	1	17	0						
3	C-||U4-	2	20	2	0	0	2	100	2	100
3	U4-||N-	1	23	0						
3	U4-||N-	2	22	3	3	100	3	100	0	0
3	N-||U4-	1	23	2	1	50	2	100	0	0
3	N-||U4-	2	22	1	0	0	1	100	0	0

Table S2. Missing segments per condition, replicate and spacing time. Number of inoculated (N) and infected (N_inf) plants, number of plants containing each segment (N_segment name) and percentage of plants containing each segment among infected plants (% segment name) as a function of spacing time (in days) per condition and replicate (rep).

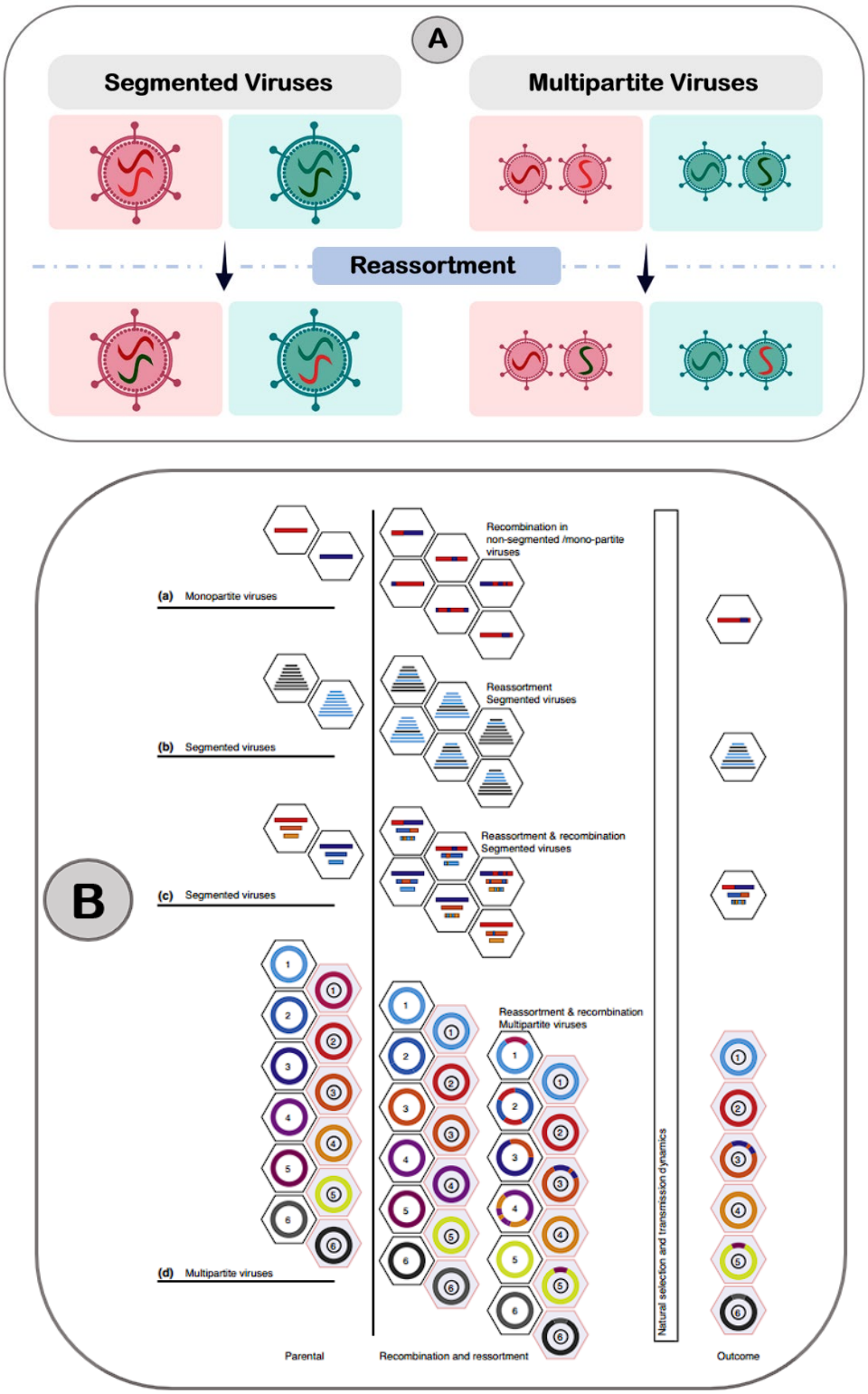


Figure 14 : [A] Illustration of the possible reassortant genotypes that could result following reassortment of both segmented and multipartite green and red genotypes, with top (segmented) and left (multipartite) segment coding for the viral particle. Figure was created using **BioRender.com**. **[B] Illustration of (a) recombination in monopartite viruses, (b) reassortment in segmented viruses, (c) recombination and reassortment in segmented viruses and (d) recombination and reassortment in multipartite viruses.** Adapted from (**Varsani et al. 2018**).

4. Reassortment, a process enhanced by multipartitism?

Establishing the ability for nanoviruses, and potentially other multipartite viruses, to reconstitute their genome via non-concomitant transmission of their segments, both within hosts and in-between hosts, substantially alleviates the previously anticipated enormous cost of genome integrity maintenance. Furthermore, this implies that these viruses possess the capacity to reassort on a much larger spatial scale compared to segmented viruses, which can only undergo reassortment when co-infecting the same cell due to their dependance on genomic segments co-packaging. Indeed, reassortments might occur without parental genotypes co-infecting the same cell (**Sicard et al. 2019**), and even when the parental genotypes do not co-infect the same host (**Di Mattia et al. 2022**). Moreover, if the ecological conditions of geographical sympatry among viruses, hosts and vector species are met, the frequency and viability of nanovirus reassortments might play an additional and significant role in reducing the cost of genomic integrity maintenance. Interestingly, ease for "reassortment of characteristics" of the parental genotypes, such as sex in eukaryotes, have been proposed long ago as a key advantage of the multipartite architecture (**Fulton, 1980**).

At the viral population level, and with regards to genetic diversity, most reassortments probably concern very closely related genotypes that would be confounded with straightforward genome reconstitutions of various degrees of genetic divergence. Despite the challenges associated with detecting reassortments due to the previously mentioned bias in sequence conservation, it is well-established that nanovirus evolution is significantly influenced by recombinations and reassortments (**Grigoras et al. 2014**). Additionally, sympatry has been observed for multiple nanoviruses. This co-occurrence has been documented in various locations, including Morocco, Azerbaijan, Iran, and Ethiopia (**Abraham, Varrelmann and Vetten, 2012; Abraham et al. 2010; Grigoras et al. 2014; Hasanvand et al. 2021; Heydarnejad et al. 2017**). This observation suggests that ecological conditions might not pose significant constraints on inter-specific nanovirus reassortments.

Like recombination, reassortment may disrupt co-adapted gene complexes by generating hybrid genotypes (**Martin et al. 2011a; Martin et al. 2011b; Varsani et al. 2018**) (**Figure 14**) but in return, it can also promote genetic innovation (**Roossinck 1997; Holland and Domingo 1998; Martin et al. 2011a**), potentially playing an important role in evolution and adaptation (**Chao, Tran and Matthews 1992; Martin et al. 2011a**).

While reassortant genotypes may occur frequently (**Matsuzaki et al. 2003**) and can be associated to substantial alterations in viral traits, such as hypervirulence (**Chakraborty et al. 2008**), resistance-breaking (**Tentchev et al. 2011**), or host range expansion (**Idris et al. 2008**), reassortments are also often associated with deleterious effects (**Escriu, Fraile and García-Arenal 2007; Ohshima et al. 2016; Villa and Lässig 2017**). Overall, we presently lack a systematic characterization of the phenotypic effects of reassortments, such that their net effect on viral fitness is unknown for any viral taxon. In addition to raising an intriguing question regarding the evolutionary implications of reassortment for both segmented and multipartite viruses, this characterization would also contribute to our understanding of how these genetic exchanges could impact the way multipartite viruses handle the costs associated genomic integrity maintenance.

THESIS OUTLINE AND OBJECTIVES

This thesis delved into the study of the multipartite genomic architecture, contributing to the previous demonstration that non-concomitant horizontal transmission of viral particles leads to genome reconstitution, and followed by focusing on reassortments and its potential impact on the maintenance of genomic integrity and virus evolution.

The primary objective of the thesis was to investigate reassortments in the *faba bean necrotic stunt virus* (**FBNSV**) by systematically characterizing single-segment reassortants derived from two distinct isolates of the same species. A series of preliminary and intermediate objectives have structured the work and writing of the manuscript.

First, I ought to get familiar with the state of the art. I made use of the first COVID lockdown to write and publish a comprehensive review of the molecular constraints associated with reassortment in single-stranded DNA multipartite viruses. As a complement to this review article, I initiated a genomic study focused on the *Nanovirus* genus, aiming to identify iteron trans-replication determinants and to explore their compatibility based on the detection of reassortant in natural populations. This represents a gap in the current knowledge on nanoviruses, in comparison to babuviruses and geminiviruses where analogous information has been more extensively investigated.

Second, a pre-requisite of the experimental production of reassortant genotypes was the selection of two appropriate isolates, **[AZ;15]** and the hybrid **[AZ;10_12b]**, the construction of corresponding infectious clones and their following thorough characterization, including the existing **[JKI-2000]** isolate, across their associated field hosts and aphid vectors. This extensive phenotypic characterization serves as a foundational study for the subsequent reassortant investigations.

Third, the systematic production and characterization of single-segment reassortants between **[AZ;15]** and **[AZ;10_12b]** across three different host species could be achieved.

Finally, the general discussion section analyzes the results of this research and proposes future prospects based on the published findings and on those that are to be completed in the near future.

Additional annex materials include both unfinished work and preliminary results, some prefiguring the future complementary lines of research that will be pursued by others in the team, all being evoked in the general discussion. A general bibliography and French summary are available at the end of the manuscript.

CHAPTER 1 :

REASSORTMENTS IN ssDNA MULTIPARTITE VIRUSES

1. Context

The earlier evidence that nanoviruses can achieve functional complementation of their genome at a supra-cellular level and reconstitute a complete genome through non-concomitant horizontal transmission has underscored their large potential for reassortment (**Sicard et al. 2019; Di Mattia et al. 2022**).

In addition to their significant influence on virus evolution, reassortment events that result in the reconstitution of complete genomes may also have notable implications for the maintenance of the genomic integrity in these viruses. If the ecological conditions are met and viable reassortments happen frequently, even though reassortments between closely related viral populations might go unnoticed, they would further decrease the chance of losing some segments. Conversely, if reassortments are frequent but predominantly lead to genotypes with reduced fitness, reassortment would impose an additional cost.

With this in mind, I aimed to explore the factors that could influence the viability of reassortants in ssDNA multipartite viruses. In addition to the evident requirement for the reassorted segment to effectively complement its new genomic environment, the reassorted segment also depends on the genomic environment to undergo key processes such as replication, encapsidation, intra-host movement and inter-host transmission. The goal of this bibliographic search was to review existing literature to determine how these crucial processes could influence the viability of reassortment in ssDNA multipartite viruses, and subsequently confront these findings with data related to reassortment in natural populations. Most of the review and conclusions focus on bipartite begomoviruses from the *Geminiviridae* family, on which the most part of the relevant information has been reported, as well as the *Nanoviridae* family, which comprises the *Babuvirus* and *Nanovirus* genera previously introduced in this manuscript.

It is worth noting that ssDNA multipartite viruses have a broad host range, extending beyond typical virus infecting plant host species to include fungi (**Li et al. 2020**) and even animal hosts (**Hu et al. 2013; Pénzes et al. 2023**), which adds an intriguing layer of diversity. Hence, although the review's exploration may seem somewhat limited in scope, it is also important to emphasize that the mechanisms associated with rolling-circle replication (RCR), and as such the molecular constraints associated to replication investigated in this review are applicable to the majority of CRESS viruses that compose most of the recently identified ssDNA multipartite viruses.

2. General results

Our review identified trans-replication of the reassorted segment as the primary constraint on reassortment, with a necessary compatibility in the recognition, binding and recruitment of the host machinery that would enable the segment to replicate. On the contrary, in both bipartite begomoviruses and nanovirids, packaging, viral movement, and transmission appeared to be less stringent, with a reduced emphasis on specific sequence compatibility and a greater reliance on nonspecific binding or interactions. Although RCR is recognized as a significant factor influencing reassortment compatibility and viability, our understanding of iterons, iterative sequences determinant for the binding of the Rep protein, and their interactions with Rep in begomoviruses and nanovirids remains limited, particularly when it comes to nanoviruses, where iterons have yet to be definitively identified and experimentally validated.

The high prevalence of detected reassortment events in natural populations of begomoviruses and nanovirids is particularly noteworthy, especially given the underestimation resulting from sequence conservation and limited data bias in their detection. This aligns with the potential increased capacity to reassort in single-stranded DNA multipartite viruses, at least nanoviruses However, this observation highlights the significant role that reassortments may play in the evolution of these viruses as the detection of reassortants only give us information on the ones that have been effectively maintained throughout the course of evolution and as such they could either happen frequently and/or often be selected.

However, when we compare experimental data predictions for nanovirids to data obtained from natural populations, two significant discrepancies become apparent. Experimental data concerning nanoviruses suggest a relaxed capacity for inter-specific reassortment with shared host and insect vector ranges across the genus, as well as an effective functional complementation of vital viral processes across diverse species. In contrast to this, the vast majority of detected reassortants in natural populations are intra-specific reassortants. The relative scarceness of inter-specific reassortments is even more striking given the relatively large frequency of inter-specific recombination, the latter suggesting that ecological opportunities for inter-specific genetic exchanges are not limiting

Furthermore, the ability of nanoviruses to initiate infections with missing genomic segments in controlled laboratory settings suggests the possibility that genome reconstitution through horizontal transmission might involve the simultaneous complementation of multiple "non-essential" segments. This scenario might facilitate the occurrence of multiple-segment reassortments. Nevertheless, what is predominantly observed are single-segment reassortments with almost no multiple-segments reassortments detected.

These discrepancies indicate that inter-specific and multiple-segment reassortments often lead to significant disruptions in intra-genomic interactions, potentially resulting to reduced fitness in the case of inter-specific and multiple-segment reassortants.

3. Under minor revision review article, referenced as (*Torralba, Blanc and Michalakis 2023*) but now published (*Virus Evolution*, 2024, DOI: 10.1093/ve/veae010)

Reassortments in Single-Stranded DNA Multipartite Viruses: Insights into Molecular Constraints, Predictions and Confrontation to Field Observations

Authors: Babil Torralba1, Stéphane Blanc1*, Yannis Michalakis2*

1 PHIM, Université Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France

2 MIVEGEC, Université Montpellier, CNRS, IRD, Montpellier, France

*These authors contributed equally to this work

Corresponding authors:

Yannis Michalakis (yannis.michalakis@ird.fr), Stéphane Blanc (stephane.blanc@inrae.fr)

Abstract

Single-stranded DNA multipartite viruses partly resolve the cost of genomic integrity maintenance through two remarkable capacities. They are able to systemically infect a host even when their genomic segments are not together in the same host cell, and these segments can be separately transmitted by insect vectors from host to host. These capacities potentially allow such viruses to reassort at a much larger spatial scale, since reassortants could arise from parental genotypes who do not co-infect the same cell or even the same host. In order to assess the limitations affecting reassortment and their implications in genome integrity maintenance, the objective of this review is to identify putative molecular constraints influencing reassorted segments throughout the infection cycle, and to confront predictions based on these constraints with empirical observations. Trans-replication of the reassorted segments emerges as the major constraint, while encapsidation, viral movement, and transmission compatibilities appear more permissive. Confronting the available molecular data and the resulting predictions to field population surveys reveals notable discrepancies, particularly a surprising rarity of inter-specific natural reassortments within the *Nanoviridae* family. These inconsistencies unveil important knowledge gaps in the biology of ssDNA multipartite viruses and call for further investigation on the role of reassortments in their biology.

General Introduction

Depending on their genomic organization and packaging strategy, viruses can be monopartite, segmented, or multipartite. While monopartite viruses encode all their genetic information in a single nucleic acid molecule packaged within a viral particle, the other two categories have fragmented genomes but package their segments either all together in the same (segmented) or in separate particles (multipartite). The physical separation of the genetic information in segmented and multipartite viruses allows the replacement of entire segments by homologous ones from a distinct parental genotype through a genetical exchange mechanism known as reassortment or pseudo-recombination (Roossinck 1997; Holland and Domingo 1998; McDonald et al. 2016). Like recombination, reassortment may disrupt co-adapted gene complexes by generating hybrid genotypes (Martin et al. 2011a; Martin et al. 2011b; Varsani et al. 2018) but in return, it can also promote genetic innovation (Roossinck 1997; Holland and Domingo 1998; Martin et al. 2011a), potentially playing an important role in evolution and adaptation (Chao, Tran and Matthews 1992; Martin et al. 2011a).

While reassortant genotypes may occur frequently (Matsuzaki et al. 2003) and can be associated to substantial alterations in viral traits, such as hypervirulence (Chakraborty et al. 2008), resistance-breaking (Tentchev et al. 2011), or host range expansion (Idris et al. 2008), reassortments are also often associated with deleterious effects (Escriu, Fraile and García-Arenal 2007; Ohshima et al. 2016; Villa and Lässig 2017). Overall, we presently lack a systematic characterization of the phenotypic effects of reassortments, such that their net effect on viral fitness is unknown for any viral taxon.

Although not duly emphasized in the literature (but see Varsani et al. 2018), segmented and multipartite viruses differ in their potential to undergo reassortment. In principle, for segmented viruses reassortment implies co-packaging of heterologous segments, while in multipartite viruses, distinct segments are packaged separately (Figure 1). Recent discoveries suggest this fundamental difference may have far reaching implications. A study showed that the single-stranded DNA (ssDNA) multipartite faba bean necrotic stunt virus (FBNSV, *Nanovirus* genus), exhibits a unique "pluricellular lifestyle" (Sicard et al. 2019). Using sequence specific fluorescence in-situ hybridization (FISH), the authors demonstrated that the different genomic segments rarely co-occur in the same cell, the viral system functioning through complementation across multiple cells (Sicard et al. 2019). A subsequent study showed that it was possible to reconstitute a complete genome from complementary incomplete segment

sets infecting distinct host plants (Di Mattia et al. 2022). To achieve such complementation, recipient plants could be inoculated by distinct aphid vector individuals, each carrying a complementary set of genome segments, or inoculated by a single aphid that sequentially acquired the two complementary segment sets (Di Mattia et al. 2022). The capacity to transmit genomic segments separately between cells and even between hosts implies that, contrary to segmented viruses, at least nanoviruses do not require to co-infect the same cell or even the same host to exchange segments, drastically increasing their potential to reassort (Figure 1).

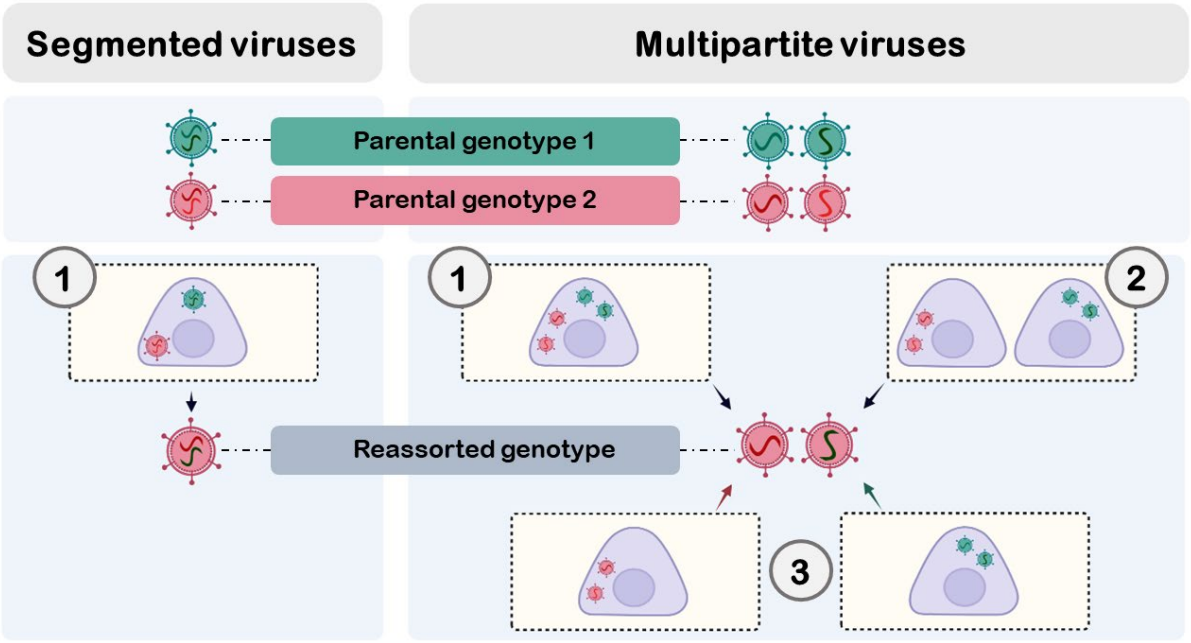


Figure 1: Comparison of the reassortant production processes in multipartite and segmented viruses. **1- Two viral genotypes co-infect the same cell.** A reassortment results from the co-packaging of two segments, each originating from a distinct parental genotype in the case of segmented viruses, and from the production of a population of particles containing complementary segments from two parental genotypes for multipartite viruses. **2- Two viral genotypes co-infect the same host.** Some ssDNA multipartite viruses follow a multicellular infection cycle where functional complementation occurs at a supra-cellular level. When a host is co-infected by two parental genotypes, their distinct segments can infect separate cells and interact through complementation which can result in the formation of a reassorted genotype. **3- Two viral genotypes infect different hosts.** Reconstitution of a reassorted genome can result from the separate transmission of two particles containing complementary genomic segments originating from distinct host and parental genotypes. Light violet triangles represent individual cells; Yellow rectangles represent host individuals. Parental and reassorted genotypes are indicated for both segmented and multipartite viruses. Distinct segments of a given genotype are colored and oriented differently. Figure was created using **BioRender.com**.

Reassortments may thus play a prominent role in the life cycle of these viruses, if the conditions to reassort are met. A favorable ecological context is necessary, which involves parental genotypes sharing host or vector species within the same geographical area. As a result, reassortments tend to occur at geographical hotspots where multiple species or isolates co-exist (Savory, Varma and Ramakrishan 2014; Xavier et al. 2021). In the presence of such diversity, the chances of genome reconstitution would further increase if viable reassortants were frequent. However, reassortments may be subject to multiple molecular constraints. Not only the reassorted segments have to functionally complement their new genomic background but they must also be complemented by it as they undergo replication, packaging, intra- and inter-host movement to generate an infection focus (Figure 2).

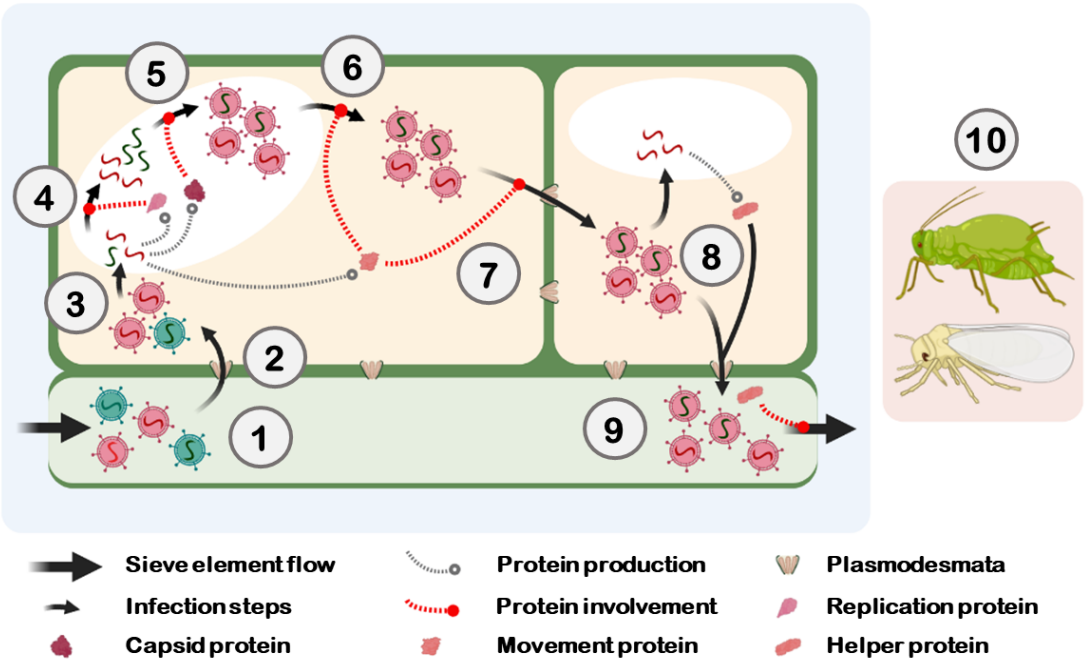


Figure 2: Overview of the infection cycle key steps where molecular constraints may impact segment compatibility upon reassortment of multipartite viruses. **1-** This scenario illustrates a host plant co-infected by two bipartite virus genotypes (green and red) first released into the sieve element. **2-** All types of viral particles can invade companion cells through plasmodesmata. **3-** Decapsidation enables viral DNA to replicate within the nucleus and leads to the transcription and translation of associated proteins. **4-** A reassortant is produced when *e.g.* the replication protein of the green genotype can replicate the complementary segment of the red genotype or *vice versa*. **5-** Reassorted viral DNA is packaged by the capsid protein. There, reassortment might be facilitated by the capacity of the capsid protein to package heterologous ssDNA. **6/7/8-** Viral particles leave the nucleus thanks to the intra- and intercellular movement proteins allowing any ssDNA segment to invade neighboring cells. **9-** Virions travel long-distance in the vasculature. The movement and capsid proteins can be involved for this long-distance progression and thus compatibility between them may be required. **10-** Insect vectors acquire viral particles while feeding on the plant host. The transmission of the reassortant depends on the compatibility between the vector species, the capsid protein and a potential transmission helper protein. Distinct segments of a given genotype are colored and oriented differently. Figure was created using **BioRender.com**.

Given the potential importance of genome integrity maintenance through genome reconstitution demonstrated in nanovirids and the associated significance of reassortant viability, we here review the potential ecological and molecular constraints affecting reassortments in ssDNA multipartite viruses. This literature survey generates predictions on reassortant viability, which are further confronted to the available data on the properties and frequency of reassortants artificially produced or found in field samples. We here assume, awaiting further demonstration, that the property for genomic segments to propagate separately within and between hosts is not unique to nanoviruses, but may in fact be shared by at least other ssDNA multipartite viruses. We thus expand our survey and arguments on other such viruses and primarily on the extensively studied multipartite geminivirus species.

Geminiviridae – Begomovirus

Introduction

Until recently the only members of the family with a non-monopartite genome belonged to the genus *Begomovirus*. The *Olea europea* geminivirus (OEGV) is currently unclassified but is likely to represent its own genus within the *Geminiviridae* family (Materatski et al. 2021). OEGV is bipartite and shares similarities in genome size and DNA-A organization with bipartite begomoviruses, but it differs in its DNA-B organization and overall nucleotide identity. Specifically, its MP is in the virion sense, while a small undetermined ORF is in the complementary sense (Chiumenti et al. 2021). Limited information is available about this virus, except that it appears to be highly prevalent in olive trees, and its isolates show a very close sequence conservation (99% pairwise identity) (Alabi et al. 2021), making it difficult to detect reassortment events. We will thus focus in what follows to begomoviruses.

The largest and most extensively studied group of ssDNA plant viruses is the genus *Begomovirus* comprising both monopartite and bipartite species. Begomoviruses are a significant threat to a wide range of crops, especially in subtropical and tropical regions, and are thus the object of many studies.

Bipartite begomoviruses have a genome consisting of two circular segments named DNA-A and DNA-B, each of approximately 2.7 kb (Rojas et al. 2005) (Figure 3). DNA-A commonly has six open-reading-frames (ORF): in the virion-sense AV1 encoding the only structural protein (CP) and AV2 encoding the pre-coat protein (PCP), involved in movement and gene silencing, that is absent in some species; in the complementary-sense AC1 encoding the replication

associated protein (Rep), AC2 encoding a transcriptional activator protein (TrAP), AC3 encoding a replication enhancer (REn) and AC4 encoding (RepA) also involved in the replication process. DNA-B carries two ORFs: in the viral-sense BV1 encoding a nuclear shuttle protein (NSP), and in the complementary-sense BC1 encoding a movement protein (MP). Recent studies identified new candidate ORFs with putative homologs in bipartite species (Gong et al. 2021; Chiu et al. 2022).

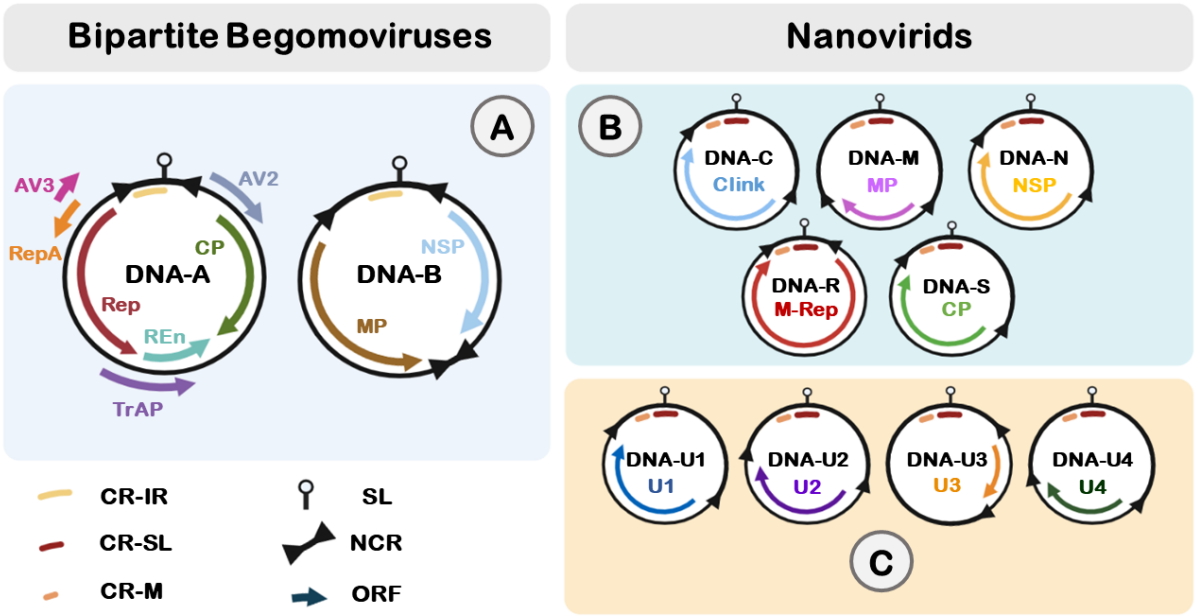


Figure 3: A- Genome organization of a bipartite begomovirus. Both genomic components DNA-A and -B are represented by a single DNA molecule (black circle). The colored arrows represent the identified open reading frames (**ORF**). The corresponding proteins are indicated with the same color as the arrows. Intergenic non-coding regions (**NCR**) are represented overlapping the black circles. At the top of each DNA circle, a highly conserved stem-loop structure is represented. **Rep**: replication-associated protein; **RepA**: replication-associated protein A; **REn**: replication enhancer protein; **TrAP**: transcriptional activator protein; **CP**: capsid protein; **MP**: movement protein; **NSP**: nuclear shuttle protein; **NCR**: intergenic non-coding regions; **CR-IR**: common regions of DNA-A and DNA-B corresponding to the highly conserved 200-nucleotide stretch within the large intergenic region of bipartite begomoviruses. **B and C-** Genome organization of the *Nanoviridae* family. The five genomic segments (DNA-C, -M, -N, -R and -S) shared between nanoviruses and babuviruses are shown in panel **B**. The four genomic components specific to nanoviruses (DNA-U1, -U2 and -U4) and babuviruses (U3) are shown in panel **C**. The name of each component is indicated within the corresponding circle. Colored arrows indicate the approximate size and position of open reading frames (**ORF**) with the corresponding name of the encoded protein accordingly indicated in the same color. At the top of each circle, two conserved regions (**CR-M** and **CR-SL**) including a stem-loop is represented inside the non-coding region (**NCR**). **Rep**: replication-associated protein; **CP**: capsid protein; **MP**: movement protein; **Clink**: cell-cycle link protein; **NSP**: nuclear shuttle protein; **U1**, **U2**, **U3** and **U4**: proteins of unknown functions

Begomoviruses are often associated with alpha-, beta-, or delta- satellites which are facultative subviral agents relying on their helper virus for essential viral functions such as replication, encapsidation and transmission. Delta- and beta-satellites are dependent on their helper virus to replicate (Mubin et al. 2019; Ferro et al. 2021), whereas alpha-satellites are responsible for their own replication and do not replicate the helper's segments (Mansoor et al. 1999; Briddon et al. 2018). Satellites can form associations with new helper species following reassortments (Chen et al. 2009; Mubin et al. 2019). The impact of satellites on their helper viruses is highly heterogeneous and can be beneficial in some instances. For example, some beta-satellites have been described to subtitute for a DNA-B in bipartite species or to increase the ability of their helper virus to infect a different host plant species (Saeed et al. 2007; Singh, Chattopadhyay and Chakraborty 2012; Mubin et al. 2019).

Most begomoviruses are phloem-restricted but a few bipartite species can also infect mesophyll tissues (Morra and Petty 2000). They are all transmitted by whiteflies, most notably of the supervector complex species *Bemisia tabaci* (Gilbertson et al. 2015; Fiallo-Olivé et al. 2020). Since numerous begomoviruses share host and vector species, mixed infections are frequent and the potential to produce intra- and inter-specific reassortants seems high. Consistently, well documented recombinations and reassortments are known to impact begomovirus evolution (Pita et al. 2001; Saunders et al. 2002; Lefeuvre et al. 2009; De Bruyn et al. 2012; Lefeuvre and Moriones 2015; Fiallo-Olivé and Navas-Castillo 2023) even though mutation is reported to predominantly drive begomovirus diversification (Lima et al. 2017). Moreover, due to the availability of multiple infectious clones, experimental reassortants are often produced to study the interaction of various species and isolates.

For brevity we review below molecular constraints that could arise during the fundamental steps of replication, packaging, movement and host-to-host transmission. We acknowledge that constraints could arise during other aspects of the viral cycle. For example, the transcription factor TrAP borne by DNA-A was shown to regulate the expression of CP and MP ORFs borne by DNA-A and -B respectively (Sunter and Bisaro 1992; Hartiz, Sunter and Bisaro 1999) and thus, because of the inter-segment transcription regulation, could represent a significant associated constraint in reassortment.

Replication

DNA-A and -B share one conserved region (CR-IR) in the intergenic non-coding region (NCR) that can form a stem-loop secondary structure with a highly conserved nona-nucleotide loop-sequence (5′ – TAATATTA/C – 3′) acting as the origin of replication (Argüello-Astorga et al. 1994). It also contains a conserved sequence used to prime the synthesis of the complementary strand producing double-stranded DNA (dsDNA) replication intermediates (Saunders, Lucy and Stanley 1992; Bonnamy, Blanc and Michalakis 2023). The origin of replication is cleaved at the slash position in the nona-nucleotide loop by the Rep protein to initiate replication (Laufs et al. 1995; Bonnamy, Blanc and Michalakis 2023). The 3′ end of the cleaved strand serves as primer for the host DNA polymerase (Wu et al. 2021) and the DNA synthesis proceeds while the helicase activity of the Rep protein unwinds the dsDNA intermediate (Choudhury et al. 2006; Clérot and Bernardi 2006; Bonnamy, Blanc and Michalakis 2023). When this polymerizing complex has copied the whole molecule and reaches again the origin of replication, the Rep protein ligates the 5′ and 3′ extremities of the newly formed ssDNA circle and the whole process, designated as Rolling-Circle Replication (RCR), can resume (Laufs et al. 1995; Bonnamy, Blanc and Michalakis 2023).

Before cleaving the loop, the Rep protein plays a key role in initiating replication by binding dsDNA at specific repetitive sequences called "iterons" located in the CR-IR region (Fontes et al. 1994; Rizvi, Choudhury and Tuteja 2014). These sequences act as replication specificity determinants and are present in iterative arrangements around the stem-loop structure (Argüello-Astorga et al. 1994). Over time, the length of iterons has been refined, and now comprises short repeated sequences of 5 to 8 nucleotides (Argüello-Astorga and Ruiz-Medrano 2001). The sequence, number and arrangement of iterons (relative position and spacing between them) can vary between genera and species (Argüello-Astorga et al. 1994; Fontenele et al. 2021).

The Rep protein is structured around several core functional domains. These domains are notably involved in DNA binding (Motif I), helicase (Motif II) and endonuclease activity (Motif III) (Ilyina and Koonin 1992; Orozco and Hanley-Bowdoin 1998). By analyzing more than 100 *Begomovirus* iteron sequences and the corresponding associated Rep amino acid sequences, a hypervariable domain in the N-terminal region of the Rep protein was identified and called the Iteron Recognition Domain (IRD) (Argüello-Astorga and Ruiz-Medrano 2001). Interestingly, this variable Rep domain is conserved in species that share the same iterons, suggesting a role in their recognition (Argüello-Astorga and Ruiz-Medrano 2001). The IRD was defined as a stretch of less than 10 residues located in the 20 first amino acids of Rep that contributes to

the folding of a conserved DNA binding tertiary structure (Argüello-Astorga and Ruiz-Medrano 2001; Ramos et al. 2003; Londoño, Riego-Ruiz and Argüello-Astorga 2010). Acting as a second replication determinant, and also involved in the tertiary structure, two additional residues, approximatively 60 AA downstream of the IRD, have been identified (Londoño, Riego-Ruiz and Argüello-Astorga 2010; Avalos-Calleros et al. 2021). Distinct amino acids at these key positions recognize specific nucleotides, enabling prediction of potential iteron recognition based on the IRD sequence, and as such, of a potential match between the Rep encoded by a given DNA-A and a compatible CR-IR sequence in DNA-B (Argüello-Astorga and Ruiz-Medrano 2001; Gregorio-Jorge et al. 2010; Maliano et al. 2022).

The Iteron-IRD matching predictions should be nuanced, as repeated iterons can be slightly different in the same molecule and also between DNA-A and -B of the same genome (Maliano et al. 2022). Moreover, iterons can differ among closely related species and, conversely, distantly related species can share the same iteron sequences (Argüello-Astorga and Ruiz-Medrano 2001; Gregorio-Jorge et al. 2010). The Rep IRD presumably allows for some variation in the sequences it recognizes, probably affecting its DNA binding (Fontes et al. 1994) and thus the efficiency of the associated DNA-B replication (Chakraborty et al. 2008).

The function of the begomovirus RepA involved in replication has not been clearly established (Rizvi et al. 2014; Sun et al. 2022), whereas the REn protein, although not essential for replication, enhances it through interactions with Rep and host proteins (Sun et al. 2020). Consequently, from what is known so far, these DNA-A borne proteins mostly interact with Rep, also borne by DNA-A, and any interactions these two proteins may have with a reassorted DNA-B remain unknown.

Begomoviruses are reported to use a second replication process called recombination-dependent replication (RDR) (Jeske, Lütgemeier and Preiss 2001; Preiss and Jeske 2003; Bonnamy, Blanc and Michalakis 2023). This replication mechanism involves homologous recombination of a partially replicated ssDNA fragment which inserts between the two strands of a circular dsDNA replication intermediate, anneals with the homologous region of the complementary strand, and primes the elongation process driven by a host polymerase (Bonnamy, Blanc and Michalakis 2023). Depending on how far the elongation proceeds on this circular dsDNA intermediate template, linear dsDNA fragments of variable sizes are produced (Bonnamy, Blanc and Michalakis 2023). These dsDNA linear fragments may then be used to initiate rolling circle replication (RCR) if they encompass two origins of replication (Bonnamy, Blanc and Michalakis 2023). The initial steps of the RDR are not driven by the Rep protein and do not rely on replicative determinants like iterons. The homologous recombination associated

with the RDR process can occur anywhere between two copies of the same segment, but also in the conserved regions between DNA-A and -B. Consistently, recombination hotspots are detected in the CR-IR region (Lefeuvre et al. 2007; Lefeuvre et al. 2009; Martin et al. 2011b). Because iteron or Rep IRD exchanges can enable new associations between DNA-A, -B and beta-satellites (Saunders et al. 2002; Chen et al. 2009; Silva et al. 2014), these hotspots near the origin of replication suggest that recombination and the RDR process can facilitate the production of viable reassortants. In the same way, recombination of DNA-A CR-IR in a newly reassorted DNA-B can result in an ameliorated match between iterons and IRD, facilitating its trans-replication (Hou and Gilbertson 1996; De Bruyn et al. 2012). Interestingly, monopartite begomoviruses, which are homologous to the DNA-A of bipartite species, have been suspected to capture DNA-B after recombination, thereby becoming bipartite (Saunders et al. 2002; Mansoor et al. 2003; Lefeuvre et al. 2007; Briddon et al. 2010; Ouattara et al. 2022).

Overall, a large proportion of the described begomovirus reassortants have close or identical iterons on their two genomic segments (Ramos et al. 2003; Bull et al. 2007; Singh, Chattopadhyay and Chakraborty 2012; Silvia et al. 2014). As a result, the compatibility of iteron sequences and Rep IRD amino acid sequences appears as a good predictor of the production of viable reassortants between different species and isolates (Gregorio-Jorge et al. 2010; Avalos-Calleros et al. 2021; Maliano et al. 2022). However, asymmetric patterns of reassortment where two species with compatible iterons can produce a viable A1-B2 reassortant but not the reciprocal A2-B1 are often observed (Hill et al. 1998; Garrido-Ramirez, Sudarshana and Gilbertson 2000; Idris et al. 2008). Moreover, viable reassortants have been produced despite significant differences in iteron sequences or in iteron spatial arrangement (Garrido-Ramirez, Sudarshana and Gilbertson 2000; Fontenele et al. 2021). For these cases, it should be noted that the structural details of iteron-Rep binding are not fully understood and that distinct combinations of both viral DNA and Rep protein may result in compatible binding and folding allowing trans-replication, likely with variable efficiency. Also intriguing is the fact that some seemingly iteron-independent replication has been reported for a delta-satellite where high-affinity binding between Rep and iterons were severely limited (Lin et al. 2003) implying the putative involvement of additional replication determinants allowing binding (Zhang et al. 2016) and additional use of RDR (Alberter, Ali Rezaian and Jeske 2005).

Finally, there are also examples where viable or competitive reassortants cannot be produced despite similar or strictly identical iterons (Avalos-Calleros et al. 2021). These examples indicate that the iteron-Rep compatibility is necessary but not sufficient. Other factors involved in replication, such as putative additional regulatory sequences, or sequences in the Rep protein

affecting its folding or binding properties, may also impact the stability of Rep-DNA or the recruitment of host polymerases. Unfortunately, unlike iterons, the other factors that could preclude replicative compatibility between two segments are not well characterized.

Packaging

Encapsidation plays a crucial role in the infection cycle of begomoviruses as it protects the viral genome, helps the importation of the ssDNA into the nucleus, and is a prerequisite, at least in most cases (Pooma et al. 1996), for long-distance movement (Jeffrey, Pooma and Petty 1996) and plant-to-plant vector transmission (Czosnek et al. 2017). Therefore, successful packaging and assembly of stable particles are essential for the production of viable reassortants.

Geminiviruses are non-enveloped, and their genome is typically packaged within characteristic quasi-icosahedral geminate particles, resulting from the fusion of two single quasi-icosahedrons together composed of 110 copies of the CP, each of symmetry T=1 (Hesketh et al. 2018). Cryo-EM structure analysis suggested that a 2.7 kb genomic DNA would fill the available internal space of geminate particles (Hesketh et al. 2018). Therefore, two such geminate particles are required to individually package the two segments of bipartite species (Hesketh et al. 2018). A minority of single and double icosahedrons, the latter being unstable and probably rapidly disintegrating in single icosahedrons (Saunders et al. 2020), as well as single and triple quasi-icosahedron particles (Hooker and Salazar 1983; Frischmuth, Ringel and Kocher 2001) have also been reported, all assembled from the same building-block coat protein (Saunders et al. 2020). Single icosahedrons can encapsidate smaller ssDNA circles that are generally defective-interfering or satellite molecules of approximately half the size of a genome component for alpha- and beta- to one quarter of the size for delta-satellites (Frischmuth, Ringel and Kocher 2001; Casado et al. 2004; Jovel et al. 2007; Fiallo-Olivé et al. 2012; Saunders et al. 2020). Geminate particles can also encapsidate satellites (Hesketh et al. 2018; Saunders et al. 2020). What is encapsidated in the larger triple quasi-icosahedrons is unclear although it has been observed that begomoviruses often produce recombining defective ssDNAs of variable size that can exceed that of a full-length genomic segment (Jovel, Preiss and Jeske 2007; Patil et al. 2007). In addition, experimental studies suggest that the begomovirus capsid protein may be more permissive to the encapsidation of DNA molecules of different lengths than that of other geminiviruses, which require more precise genome-size molecules to ensure stable particle production (Saunders et al. 2020). This size-packaging

flexibility in begomoviruses might facilitate the viable association with genomic components or satellites of variable length (Jovel, Preiss and Jeske 2007; Patil et al. 2007; Fiallo-Olivé et al. 2012). Therefore, would the packaging step be a limiting factor for reassortment, it could only be due to a specificity between a given coat protein and the encapsidated sequences, for example through the existence of specific packaging signals.

The N-terminal part of the coat protein of several geminiviruses has been found to contain a DNA binding domain (Liu, Boulton and Davies 1997; Qin Ward and Lazarowitz 1998), which allows the coat protein to bind to both ss- and dsDNA, in a sequence non-specific manner (Ingham, Pascal and Lazarowitz 1995; Liu, Boulton and Davies 1997; Palanichelvam et al. 1998; Hehnle, Wege and Jeske 2004). First identified in the species *African cassava mosaic virus* (ACMV) (Hipp et al. 2017), a ring of basic residues, conserved in begomoviruses (Bennett and Agbandje-McKenna 2020), forms a pocket under the pentameric units interacting with the encapsidated DNA (Hipp et al. 2017). Cryo-EM structural analysis of the ageratum yellow vein virus (AYVV) particle determined that the electron density visible beneath each CP corresponded to a hexa- or hepta-nucleotide sequence (Hesketh et al. 2018). Interestingly, mutational analysis of AA residues within the DNA binding site affected capsid assembly, highlighting the importance of CP-DNA interactions in this process (Hesketh et al. 2018). Encapsidation is believed to begin with no requirement for a specific DNA packaging signal, but uses genomic DNA as a scaffold enabling CP conformational changes necessary for the assembly of viral particles (Hesketh et al. 2018; Xu et al. 2019). The interface between the two quasi-icosahedrons probably acts as an anchoring point for the viral DNA (Xu et al. 2019). The DNA-binding sites within each CP subunit of a geminate particle together represent 110 potential inner contact points for the encapsidated DNA (Hesketh et al. 2018). Being single stranded, segments may adopt variable secondary structures depending on their sequence and local physicochemical conditions. Distinct secondary structures may allow to attach variable proportions of these contact points, and thus some segments might consolidate the viral particle more efficiently. Although this specific aspect has never been investigated, it may affect the particle stability and so eventually impact some of the reassortant properties.

In summary, the CP of begomoviruses can encapsidate a diverse range of ssDNA, both in length and sequence, and specific packaging determinants are lacking suggesting that packaging is not a major constraint for the viability of reassortants.

Viral Movement

Several proteins are involved in the movement of begomoviruses: (NSP)/BV1 and (MP)/BC1 proteins encoded by ORFs borne by DNA-B, and (CP)/AV1 plus (PCP)/AV2 and AV3 (when present) (Ho, Kuchie and Duffy 2014; Moshe et al. 2015; Gong et al. 2022), encoded by ORFs borne by the DNA-A component. Unfortunately, the mode of action of these proteins remains poorly understood.

The functional domains of NSP in begomoviruses have not been extensively studied (Sanderfoot, Ingham and Lazarowitz 1996; Ward and Lazarowitz 1999; Zhou et al. 2007; Patil and Dasgupta 2022). Only two bipartite species have been investigated, *Squash leaf curl virus* (SqLCV) (Sanderfoot, Ingham and Lazarowitz 1996) and *Indian cassava mosaic virus* (ICMV) (Patil and Dasgupta 2022). Two nuclear localization signals (NLS) were identified in the NSP of SqLCV and ICMV, while an additional nuclear export signal (NES) was also found in ICMV NSP (Patil and Dasgupta 2022). Consistently, ICMV NSP was shown to localize both into the nucleus and at the cell periphery (Patil and Dasgupta 2022), whereas SqLCV NSP requires the co-expression of MP to exit the nucleus and move to the cell periphery (Sanderfoot, Ingham and Lazarowitz 1996). The NSP protein is predicted to multimerize and is capable of binding both ssDNA and dsDNA in a size-dependent but sequence-independent manner (Rojas et al. 1998; Hehnle, Wege and Jeske 2004). The role of NSP would be to shuttle the viral ssDNA strand to and from the nucleus. NSP and MP would form a complex either at the cell periphery (ICMV), or in/near the nucleus (SqLCV) and then translocate to the plasmodesmata for further cell-to-cell movement (Pascal et al. 1994; Sanderfoot and Lazarowitz 1995; Sanderfoot et al. 1996; Patil and Dasgupta 2022). Depending on the species, MP is similarly capable to bind ssDNA and dsDNA in a size-dependent but sequence-independent manner (Pascal et al. 1994; Rojas et al. 1998; Hehnle, Wege and Jeske 2004). It has been proposed that MP serves as a membrane anchor at the protoplasmic face of microsomes and plasma membranes, facilitating the movement of the NSP-DNA complex to reach plasmodesmata (Zhang, Ghosh and Jeske 2002). In fact, two models have been proposed to explain the functions of MP in begomoviruses (Rojas et al. 1998; Jeske 2009). For phloem-limited species such as *Abutilon mosaic virus* (AbMV), MP is thought to form a complex with NSP-bound viral DNA and localize along the plasma membrane before transferring NSP-DNA to adjacent cells (Hehnle, Wege and Jeske 2004; Frischmuth et al. 2007; Jeske 2009). For mesophyll-invading begomoviruses like the bipartite species *Bean dwarf mosaic virus* (BDMV), MP is thought to take over the viral DNA

from NSP after being exported from the nucleus to the cytoplasm and deliver it to neighboring cells (Rojas et al. 1998).

Some bipartite begomoviruses also encode an AV2 protein, involved in gene silencing (Roshan et al. 2018; Basu et al. 2021), coat protein production (Bull et al. 2007), viral accumulation (Padidam, Beachy and Fauquet 1996; Bull et al. 2007; Roshan et al. 2018), and movement (Padidam, Beachy and Fauquet 1996; Roshan, Kulshreshtha and Hallan 2017; Roshan et al. 2018), but this latter function is still not fully understood (Rothenstein et al. 2007; Briddon et al. 2010; Roshan, Kulshreshtha and Hallan 2017). Although monopartite begomovirus CP plays a crucial role in viral movement, in bipartite begomoviruses it seems to be mostly involved in long distance spread (Jeffrey, Pooma and Petty 1996; Fondong 2013), probably through the production of virion progenies going through the sieve elements and indirectly due to its role in the protection and accumulation of ssDNA in the nucleus (Qin et al. 1998). The CP of some bipartite species is not necessary for systemic infection in specific hosts, but this remains the exception rather than the rule (Pooma et al. 1996; Sudarshana et al. 1998; Levy and Czosnek 2003).

Among the recently reported DNA-A new ORFs, AV3 might also function as a viral MP in monopartite begomoviruses. AV3 could localize at the plasmodesmata, traffic between cells and was able to partially complement a potyvirus movement deficient mutant (Gong et al. 2022). Interestingly, AV3 was first identified in a monopartite begomovirus, but is conserved in many bipartite species such as ACMV (Gong et al. 2021). Unfortunately, its role in viral trafficking remains largely unknown and appears to act on the host itself, independent of any interaction with other viral proteins. Therefore, its action in the movement of bipartite viruses is unlikely to impact the viability of a newly reassorted segment.

Together, the non-specific-sequence binding of CP, NSP, and MP to ssDNA and dsDNA, and the "autonomous" putative action of AV2 and AV3, suggest that the intra-host movement should not impose significant constraints on the viability of reassortants. However, functional complementation of DNA-B for different species infecting a common host was not always reciprocal suggesting additional functional constraints that could limit reassortment (Frischmuth et al. 1993).

Host-to-Host Vector Transmission

The coat protein has been found to be the sole determinant of vector transmission (Czosnek et al. 2017). Begomoviruses are typically transmitted in a circulative non-propagative manner, where virus particles are assumed to be the viral form cycling through the vector, without replicating (Wang and Blanc 2020). It is then assumed that the reassorted segment only needs to be efficiently packaged to be successfully transmitted.

Recent studies confirmed that the monopartite Tomato yellow leaf curl virus (TYLCV) undergoes replication mainly in whitefly salivary glands (Pakkianathan et al. 2015; He et al. 2020). Would this be true for bipartite species, it might add interaction levels for the production of viable and transmissible reassortants. However, investigations with at least five other geminivirus species failed to reveal replication in their vector (Rosen et al. 2015; Wang and Blanc 2020), suggesting that only a small number of species may be able to replicate in whiteflies, only in specific organs and under certain conditions as this replication is negatively affected by stresses (Wang et al. 2016; He et al. 2020; Wang and Blanc 2020). In addition, while replication seems to contribute to TYLCV viral persistence in the vector (He et al. 2020), it is unclear how the virion progenies produced in insects contribute to transmission. Indeed, as for other species, most of the transmitted virus particles are assumed to be solely crossing the salivary gland cells through transcytosis.

In conclusion, while a DNA-A might not impose additional constraints on a reassorting DNA-B concerning the transmission step, it could nevertheless change some of its traits such as host range and vector specificity due to new CP properties (Höhnle et al. 2001; Idris et al. 2008; Fan et al. 2023) (Table 1).

Natural Population Observations

The role of reassortment in the evolutionary dynamics of bipartite begomoviruses has been much less studied than recombination (Lefeuvre et al. 2009; Martin et al. 2011a; Lefeuvre and Moriones 2015; Crespo-Bellido et al. 2021; Fiallo-Olivé and Navas-Castillo 2023). While numerous studies have investigated the genetic diversity of natural populations (De Bruyn et al. 2016; Crespo-Bellido et al. 2021), few studies undertook reassortment detection (Pita et al., 2001; Saunders et al. 2002; Briddon et al. 2010; De Bruyn et al. 2012; Xavier et al. 2021; Chen et al. 2021). Most of the related research focuses on experimental production of reassortants to evaluate the compatibility of isolates or species of interest in natural populations, and

anticipate potential agricultural threats (Crespo-Bellido et al. 2021; Fontenele et al. 2021; Maliano et al. 2022).

Among the few studies estimating reassortment frequency in natural populations, it has been reported that detecting reassortment events remained difficult and fluctuating depending on the sequence-analysis method (Xavier et al. 2021). This is mostly due to the limited number of sequences in datasets and the high nucleotide identity among exchanged components making reassortments hard to detect. Hence, it seems apparent that reassortments occur preferentially between closely related begomoviruses (De Bruyn et al. 2012). This preference, besides iteron compatibility, would reduce the disruption of intra-genomic interactions. Overall, these results stress the difficulty to identify reassortment events amidst conserved sequences, and most likely leads to an underestimation of their prevalence in natural populations.

Known studies identified 21% to 76% of the described isolates, depending on the species, originating from at least one reassortment event (Briddon et al. 2010; Xavier et al. 2021). Considering the above-mentioned potential underestimation, reassortment appears as a significant phenomenon in the ecological cycle of begomoviruses.

Conclusions

In conclusion, the constraints for reassortment in begomoviruses appear to be primarily related to transreplication compatibility between matching iterons and Rep IRD. Viral movement relies on interactions between two movement proteins, NSP and MP, and their non-sequence-specific binding to viral DNA. Since both NSP and MP are present on the same genomic component (Table 1), their compatibility is not impacted by reassortment. The inconstantly present AV2 and AV3 genes, seem involved in movement but their function is still open to speculation. Encapsidation shows a high degree of permissiveness, as observed associations between different genome sizes and stable capsid assembly do not rely on specific packaging signals. Lastly, vector transmission relies entirely on packaging, so reassortants capable of systemic infection and virion production are unlikely to face significant constraints for transmission, except concerning the inherent interactions between the capsid and vector proteins. Since all begomoviruses share the same whitefly supervector, these interactions are likely complemented by most species. Reassortment surveys in natural populations are scarce but what appears very clear is that the prevalence of reassortants is very high.

1- Begomoviruses

Function	DNA-A	DNA-B
Replication	Rep – RepA - REn	-
Encapsidation	CP	-
Viral movement	AV2 – AV3 - CP	MP - NSP
Transmission	CP	-

2 – Nanovirids

Segment	Protein	Interactions	Function
C	Clink	-	Host manipulation
M	MP	**NSP** – **U4** – (CP)	Viral movement
N	NSP	*CP* – *MP* – **M-Rep**	Transmission
R	M-Rep	**NSP** - **CP**	Viral replication
S	CP	**M-Rep**	Encapsidation
U1	U1	-	Unknown
U2	U2	-	Unknown
U3	U3	-	Unknown
U4	U4	**MP**	Unknown

Table 1: Table of interactions among proteins encoded by ORFs borne on different genomic segments. **1) Begomoviruses**: Table overview of confirmed protein interactions involved in the corresponding viral functions borne either on DNA-A or DNA-B. **2) Nanovirids:** Table overview of confirmed and suspected nanovirid protein interactions. Each column in order corresponds respectively to the concerned genomic segment, associated protein, (suspected) – detected interaction for **nanovirus** or *babuvirus* only, putative associated function.

Nanoviridae* – *Nanovirus* and *Babuvirus

Introduction

The *Nanoviridae* family is composed of two genera. The *Babuvirus* genus comprises three species, namely *Banana bunchy top virus* (BBTV) (Stainton et al. 2015), *Abaca bunchy top virus* (ABTV) (Sharman et al. 2008) and *Cardamom bushy dwarf virus* (CBDV) (Mandal et al. 2013), which infect monocotyledonous hosts from the *Musaceae* (BBTV and ABTV) and *Zingiberaceae* (CBDV) plant families. All known babuviruses are phloem-restricted and exclusively transmitted by aphid vectors, with *Pentalonia nigronervosa* being the most important species (Sharman et al. 2008; Mandal et al. 2013; Safari Murhububa et al. 2021). BBTV is notably responsible for the banana bunchy top disease (BBTD), which is a highly devastating viral disease on banana crops in Asia and currently invading Africa (Dale 1987; Qazi 2016).

The *Nanovirus* genus comprises 12 species that infect dicotyledonous plants, mainly from the *Fabaceae* family (Lal et al. 2020). New species and isolates are reported frequently (Grigoras et al. 2014; Gallet et al. 2018; Vetten et al. 2019; Lotfipour et al. 2020; Hassan-Sheikhi et al, 2020; Sun et al. 2022). Some of them increase the host range of the genus: an isolate of the faba bean necrotic yellows virus (FBNYV) can experimentally infect *Arabidopsis thaliana* from the *Brassicaceae* family (Vega-Arreguín, Gronenborn and Ramírez 2007); several milk vetch dwarf virus (MDV) isolates infect tobacco plants from the *Solanaceae* family (Kamran et al. 2019) and garlic from the *Amaryllidaceae* family (Sun et al. 2022); the new species, *Parsley severe stunt associated virus* (PSSaV), infects parsley from the *Apiaceae* family (Vetten et al. 2019; Hasanvand et al. 2021). This current host range expansion qualifies the genus *Nanovirus* as an emerging threat to the world agriculture (Lal et al. 2020). Just like babuviruses, all known nanoviruses are phloem-restricted and transmitted by several aphid species (Lal et al. 2020).

Nanovirids are the multipartite viruses with the highest number of genomic segments, six and eight for the *Babuvirus* and *Nanovirus* genera, respectively (Varsani et al. 2018) (Figure 3). Each segment is a ssDNA circle of approximately 1 kb. Five segments are conserved in both genera: DNA-C encodes a protein interfering with the host cell-cycle (Clink) (Wanitchakorn et al. 2000; Lageix et al. 2007), DNA-M encodes the movement protein (MP) (Amin et al. 2011; Krenz et al. 2017), DNA-N encodes the helper component mandatory for aphid transmission (NSP) (Wanitchakorn et al. 2000; Grigoras et al. 2018), DNA-R encodes the replication initiator protein (M-Rep) (Timchenko et al. 2000; Horser , Harding and Dale 2001), and DNA-S encodes the capsid protein (CP) (Wanitchakorn, Harding and Dale 1997; Trapani et al. 2023). An additional U3 segment is found in most babuviruses (Savory and Ramakrishnan 2014; Stainton et al. 2015;), while additional U1, U2 and U4 segments are found in almost all known nanoviruses (Grigoras et al. 2014; Knierim et al. 2019; Hasanvand et al. 2021). The function of U1-U4 is presently unknown (Krenz et al. 2017). However, due to their consistent presence in natural isolates and their impact on several viral traits (Timchenko et al. 2006; Grigoras et al. 2014; Grigoras et al. 2018), they are considered as integral parts of the nanoviral genome. Nanoviruses are often found associated with alpha-satellites (Briddon et al. 2018). These are self-replicating segments of approximately 1 kb encoding a protein homologous to M-Rep (Briddon et al. 2018; Kazlauskas, Varsani and Krupovic 2018; Zhao et al. 2019). Nanovirus alpha-satellites are not believed to contribute any function to the other genome components as they do not trans-replicate segments of the helper virus (Timchenko et al. 1999), instead they rely

on it for encapsidation, movement and transmission (Mansourpour et al. 2022; Guyot et al. 2022).

Each nanovirus species investigated has been consistently reported to differentially accumulate its genomic segments within host plants, yielding a pattern of segment frequency distribution designated as the "genome formula" (Sicard et al. 2013; Sicard et al. 2015; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022). This highly reproducible pattern is host-specific (Sicard et al. 2013; Sicard et al. 2015) and has been interpreted as a means to rapidly tune gene expression through gene copy number variation in distinct host plant species (Gutiérrez and Zwart 2018; Zwart and Elena 2020; Gallet et al. 2022).

Infectious clones are available for several legume-infecting species and isolates of the genus *Nanovirus* as one clone per segment which can be inoculated as a mixture of clones via agro-infiltration (Grigoras et al. 2009; Grigoras et al. 2014). Experiments showed that segments R, S and M are mandatory for systemic infection (Timchenko et al. 2006; Grigoras et al. 2018). The remaining segments are dispensable though their absence affects important viral traits: the absence of DNA-C reduces infection rate (Grigoras et al. 2018; Di Mattia et al. 2022), that of DNA-N abolishes aphid transmission (Grigoras et al. 2018; Di Mattia et al. 2020), and that of either U1 or U2 reduces viral accumulation, attenuates symptom severity and may decrease infection rate (Timchenko et al. 2006; Grigoras et al. 2018). Intriguingly, the absence of U4 has no reported effect under laboratory conditions (Timchenko et al. 2006; Grigoras et al. 2018). Thus its function remains a mystery, but as it is always present in field isolates (Grigoras et al. 2014), except in the recently discovered species PSSaV (Vetten et al. 2019; Hasanvand et al. 2021), it is assumed to have a role, likely in field conditions (Grigoras et al. 2018). Any of DNA-C, -N and -U4 can be absent in systemically infected plants with no major phenotypic changes (Timchenko et al. 2006; Grigoras et al. 2018; Di Mattia et al. 2022) and few infected plants lacking these three segments could even be obtained (Timchenko et al. 2006), whereas infected plants with both U1 and U2 omitted at inoculation are extremely rare and show very mild symptoms (Timchenko et al. 2006; Grigoras et al. 2018). Unfortunately, no infectious clones are available for the *Babuvirus* genus and thus analogous investigation of the dispensability of genome segments has not been conducted. Nonetheless, the absence of DNA-N in a field isolate has been reported (Fu et al. 2009), and symptoms were observed in two plants obtained under laboratory conditions, following aphid inoculation, where this segment remained undetected (Guyot et al. 2022). It should be noted that missing segments can be very common in nanovirid detection (Stainton et al. 2015; Knierim et al. 2019), including DNA-R, -S or -M

which were shown to be mandatory in *Nanovirus*. It is very likely that these absences are due to the detection techniques as most of these samples are PCR amplified using degenerated primers which can miss amplification and limit detection (Knierim et al. 2019). Field isolates when deeply sequenced after rolling-circle amplification reveal complete genomes with the exceptions of a few missing DNA-U4 (Knierim et al. 2019; Vetten et al. 2019; Hasavand et al. 2021).

Experimental nanovirus reassortants were produced to study inter-specific complementation. M-Rep proteins of either the FBNSV, MDV or subterranean clover stunt virus (SCSV) were able to trans-replicate DNA-S of any of these three species (Timchenko et al. 2000). A reassorted MDV DNA-S in a FBNYV genomic background was viable (Timchenko et al. 2006). FBNSV and FBNYV reassortants later enabled the identification of NSP as the "helper component" for transmission (Grigoras et al. 2018). The NSP from FBNSV was able to complement the transmission of a divergent (76% amino acid identity between the two NSP) pea necrotic yellow dwarf virus (PNYDV) genomic background, but the reverse did not work. It is currently unknown whether this failure is due to a problem related to the reassorted segment helper capacity or to any other step in the viral life-cycle (Grigoras et al. 2018). Therefore, in laboratory conditions three essential nanovirus functions, replication, packaging and vector-transmission, can be complemented by distinct species, though with variable efficiency, which suggests relatively weak constraints on inter-specific reassortment production.

Replication

All nanovirid genomic segments share two conserved regions in the non-coding region: common-region (CR-SL) that can form a stem-loop with a highly conserved nona-nucleotide sequence (5' "TAGTATTA/C" 3') acting as origin of replication (Hafner et al. 1997; Timchenko et al. 1999), and a major common-region (CR-M) involved in the complementary strand synthesis by the plant machinery to produce dsDNA replication intermediates (Bonnamy, Blanc and Michalakis 2023). No recombination-dependent replication (RDR) (Jeske, Lütgemeier and Preiss 2001) has been reported for nanovirids, but this may reflect a lack of relevant studies. Thus, nanovirids are presently assumed to replicate through the RCR process (Hafner et al. 1997; Timchenko et al. 1999). The Rep proteins of single-stranded DNA viruses share a high structural similarity despite nucleotide and amino acid sequence divergence (Londoño, Riego-Ruiz and Argüello-Astorga 2010; Kazlauskas et al. 2019; Venkataraman and Selvarajan 2019).

Somewhat functionally supporting this structural observation, M-Rep appears to control an RCR mechanism similar to begomoviruses, with the involvement of iteron sequences (Herrera-Valencia et al. 2006; Stainton et al. 2016; Bonnamy, Blanc and Michalakis 2023).

Three putative iteron motifs have been identified in the CR-SL of babuviruses, two located upstream and one downstream of the stem-loop (Herrera-Valencia et al. 2006; Stainton et al. 2016). Mutagenesis studies confirmed that changes in these iteron sequences can decrease or even abolish replication, depending on which iterons are modified (Herrera-Valencia et al. 2006). Alignment of all available babuvirus sequences (Stainton et al. 2016) has revealed some small differences in the iteron sequences not only among species but also among isolates and between segments within an isolate, suggesting that M-Rep may tolerate small sequence variations. Since these sequence differences are small, iterons are nevertheless quite similar across all three babuvirus species, also suggesting the possibility of heterologous trans-replication of most segments and a potentially relaxed constraint on both intra- and inter-specific reassortments (Stainton et al. 2016). Unfortunately, in contrast to babuviruses, iterons have not been experimentally validated for nanoviruses. Nevertheless, repetitive and short palindromic sequences flanking the origin of replication are conserved between segments of the same species and may act as such in combination with an uncharacterized IRD in the M-Rep (Timchenko et al. 2000; Grigoras et al. 2009; Londoño, Riego-Ruiz and Argüello-Astorga 2010). Even though complementation was possible between FBNYV, MDV and SCSV, observed quantitative differences in trans-replication were compatible with candidate iteron sequence divergences (Timchenko et al. 2000) suggesting a quantitative barrier to reassortment. Because proposed nanovirus iterons are not experimentally validated while several are short (3 or 4 nt), it is at this stage impossible to make solid predictions on the constraints of reassortment at this level.

Interactions between M-Rep/NSP and M-Rep/CP have been reported for the PNYDV in a leaf-infiltration system using fusions with reporter proteins (Krenz et al. 2017). Thus far, these putative interactions of M-Rep with other viral proteins are not understood and have not been detected in a natural viral infection context. Would they be validated, they could play a role in the regulation of replication and thereby impose a co-dependence between DNA-R, -N and -S. Finally, it is unclear how the trans-replication of a reassorted segment might affect the genome formula and its consequences on reassortant fitness.

Packaging

With only two structural studies on nanovirids, the ssDNA packaging process is still largely unknown (Venkataraman et al. 2022; Trapani et al. 2023). The structural analysis of FBNSV (Trapani et al. 2023) was performed via cryoelectron-microscopy with a 3D reconstruction model at atomic resolution applying the icosahedral symmetry. Consequently, the DNA densities within particles were averaged and not resolved, apart for the inner contact point with each of the 60 CP subunits which could be localized.

No packaging signal has been reported thus far in the conserved region shared among genomic segments. The genome packaging might thus rely mostly on non-specific interactions between the CP and viral DNA, and impose no severe constraints to the reassorted segments. Comforting this hypothesis, a viable reassorted MDV DNA-S in a FBNYV genomic background was experimentally produced (Timchenko et al. 2006).

Viral Movement

Again, not much is known or empirically confirmed about intra-host movement of nanoviruses. DNA-M has structural similarity with a geminivirus movement protein and thus likely encodes a protein of similar function (Burns, Harding and Dale 1995; Sano et al. 1998). Fusion with fluorescent reporter proteins expressed in agro-infiltrated leaves of the non-host *Nicotiana benthamiana* demonstrated that the MP of a nanovirus (Krenz et al. 2017) as well as that of a babuvirus (Zhuang et al. 2019) localized in cellular membranes, suggesting that it could travel to neighboring cells and to the vasculature via the symplastic route of the plasmodesmata (Zhuang et al. 2019). Very recently, CP amino acid substitutions preventing viral particle assembly, without affecting the CP/DNA interaction, proved to abolish systemic infection of FBNSV (Trapani et al. 2023). The authors concluded that full particle assembly and DNA encapsidation is likely required for long distance movement. The two proteins MP and CP being involved, an interaction between the two may be required and thus the compatibility between DNA-M and -S might impose constraints on the success of reassortments. A better understanding of the process underlying viral movement is thus necessary.

The protein encoded by DNA-N has been named NSP because GFP-fusion experiments showed patterns of re-localization of the products of the DNA-N and DNA-M to the cell periphery (Wanitchakorn et al. 2000), similar to what was observed in begomoviruses (Rojas et al. 1998; Jeske 2009). As previously discussed, the NSP of begomoviruses has been

experimentally shown to act in intracellular movement, in conjunction with the MP (Rojas et al. 1998; Jeske 2009). However, no implication of nanovirid NSP in movement has been experimentally confirmed, although an interaction between NSP and CP has been detected for BBTV relocating the CP to the cytoplasm from the nucleus (Ji et al. 2019). The fact that NSP omission at inoculation does not affect subsequent systemic symptom development and severity (Timchenko et al. 2006; Grigoras et al. 2018) is casting even more doubts on the contribution of NSP in within-host movement. Moreover, another function in vector transmission has been proven for this protein (see next section). In conclusion, the compatibility between MP and NSP or NSP and CP should not impose significant constraints for systemic movement of reassortants within the host plant.

Bimolecular fluorescence complementation (BiFC) analysis of PNYDV-infiltrated *N. benthamiana* leaves detected MP-MP and MP-U4 interactions, inducing in both cases localized fluorescent spots at the nuclear membrane, associated with the endoplasmic reticulum and at the cell periphery close to plasmodesmata (Krenz et al. 2017). MP and U4 are predicted to have a transmembrane domain and their co-localization near plasmodesmata suggests that they may form a complex at these sites, though definitive proof is lacking (Krenz et al. 2017). Due to U4 dispensability in laboratory conditions (Grigoras et al. 2018), whether a compatible interaction between DNA-U4 and -M is required at some step of the infection-cycle of nanoviruses is completely unknown, and whether a putative constraint exists at this level upon reassortment awaits additional functional information.

Host-to-Host Vector Transmission

Nanovirids are transmitted in a circulative non-propagative manner (Hafner, Harding and Dale 1995; Sicard et al. 2015; Wang and Blanc 2020). One peculiarity that differs from begomoviruses, is that the purified nanovirus particles cannot be acquired and transmitted by the insect vectors (Franz et al. 1999). A helper component (HC) produced in infected host plants is necessary to complement the aphid transmission of purified virus particles. This HC has recently been identified as the product of DNA-N, the NSP protein (Wanitchakorn et al. 2000; Grigoras et al. 2018). Aphid vectors fed on BBTV infected plants missing DNA-N were unable to transmit the virus, suggesting a function as HC also for babuviruses (Guyot et al. 2022). The interaction between NSP and CP identified for the BBTV (Ji et al. 2019) remains undetected for any species of the *Nanovirus* genus (Krenz et al. 2017).

All described HCs, in many unrelated viral taxa, connect the virus particles to the receptor in the insect vector, *via* two distinct functional domains, one interacting with the coat protein and the other with the receptor molecule (Di Mattia et al. 2023). Although its putative receptor in aphids is unknown, NSP has been shown to be mandatory for the internalization of the virus particles of FBNSV in midgut cells, where viral DNA, coat protein, and NSP colocalize and accumulate in cytoplasmic membrane-bound inclusions (Di Mattia et al. 2020; Di Mattia et al. 2022). Unlike all known HCs, the nanovirus NSP appears to have a unique property that can impact reassortment (Di Mattia et al. 2022; Di Mattia et al. 2023). Due to their mode of action, first recognizing a receptor in the vector and then bridging the virus particles, all characterized HCs must be acquired prior to (or together with) the virus particles for successful transmission (Di Mattia et al. 2023). Intriguingly, both acquisition orders proved efficient for the transmission of FBNSV enabling complementation of the transmission of genomic components acquired either a few days before or after NSP (Di Mattia et al. 2022). This phenomenon opens a "window" of several days where the acquisition of a DNA-N renders the aphid competent for acquiring additional segment sets on other host plants, even if these are not transmissible on their own (Di Mattia et al. 2022; Di Mattia et al. 2023).

Natural Population Observations

Investigations in babuvirus natural populations have detected multiple isolates with shared reassortment events (Stainton et al. 2012; Savory and Ramakrishan 2014; Stainton et al. 2015) (Table 2). Studies of BBTV and CBDV isolates established that 40% had undergone at least one reassortment (Savory and Ramakrishan 2014; Stainton et al. 2015). Despite frequent reassortment events for CBDV and BBTV, there is a noticeable absence of inter-specific reassortants (Stainton et al. 2012; Savory and Ramakrishan 2014; Stainton et al. 2015). Interestingly, most of the detected intra-segment recombination events involve different species, indicating that the rarity of inter-specific reassortments cannot be attributed to limited encounter of parental genotypes.

Only two comprehensive sequencing studies were conducted on natural populations of nanoviruses. One compared multiple isolates of eight species and identified 12 reassortment events, 11 involving only one segment and one involving two segments (Grigoras et al. 2014) (Table 2). Only two reassortment events involved different species, indicating that, as already noted for babuviruses, successful reassortment events are more likely to occur within a species.

This study also detected 23 recombination events, 18 occurring between different species, again suggesting that encounters between parental genotypes of distinct species cannot explain the relative paucity of interspecific reassortants. The second study compared 16 complete isolates, mostly of FBNYV (Kraberger et al. 2017) and detected 10 reassortment events, mainly intra-specific, and only one involving two segments. For both genera, the survey size was limited. In all cases, however, multiple species co-occurred in the same geographical area and shared host species (Stainton et al. 2012; Savory and Ramakrishan 2014; Grigoras et al. 2014; Stainton et al. 2015; Kraberger et al. 2018). In this regard, the relative scarceness of inter-specific reassortments is remarkable and should be investigated further.

In babuviruses, although all segments were involved in at least one reassortment event, some segments were more often involved than others (Table 2). Over several studies (Stainton et al. 2012; Savory and Ramakrishan 2014; Stainton et al. 2015) DNA-M, -N and -U3 proved more prone to reassort, than DNA-R and-S. DNA-C was the segment which reassorted less in all cases (Table 2). In nanoviruses, all segments except DNA-C were detected in reassortants (Table 2). Specifically, DNA-M reassorted most, followed by DNA-N, -R and -U2, while DNA-U1 and U4 were sometimes not involved in reassortments at all. Since this is based on only two studies with a small number of genomes (Grigoras et al. 2014; Kraberger et al. 2017) the conclusions are limited and do not account for the recent expansion of available nanovirus sequences. However, in both genera, DNA-M and -C are highly and poorly reassorting, respectively.

Genus	Complete Genomes	Events	C	M	N	R	S	U1	U2	U3	U4	Ref
Babuvirus	**17** BBTV	**8**	**0** (0 %)	**3** (38 %)	**2** (25 %)	**2** (25 %)	**1** (13 %)	-	-	**1** (13 %)	-	Stainton et al., 2012
Babuvirus	**121** BBTV – **2** ABTV	**40**	**5** (13 %)	**8** (20 %)	**7** (18 %)	**2** (5 %)	**7** (18 %)	-	-	**11** (28 %)	-	Stainton et al., 2015
Babuvirus	**163** CBDV	**23**	**2** (9 %)	**7** (30 %)	**7** (30 %)	**2** (9 %)	**2** (9 %)	-	-	**2** (9 %)	-	Savory and Ramakrishnan, 2014
Nanovirus	**3** BMLRV - **6** FBNSV – **13** FBNYV – **1** FBYLV – **1** MDV – **1** PYSV – **2** PNYDV – **2** SCSV	**12**	**0** (0 %)	**4** (33 %)	**2** (17 %)	**2** (17 %)	**0** (0 %)	**0** (0 %)	**4** (33 %)	-	**1** (8 %)	Grigoras et al., 2014
Nanovirus	**16** FBNYV	**10**	**0** (0 %)	**3** (30 %)	**2** (20 %)	**1** (10 %)	**1** (10 %)	**3** (30 %)	**1** (10 %)	-	**0** (0 %)	Kraberger et al., 2017

Table 2: Table summary of nanovirid natural population surveys and reassortment detections. Columns from left to right correspond to the genus, the number of complete genomes per surveyed species, the number of individual reassortment events detected in the study, how many were associated with the indicated genomic segment and, finally, the corresponding publication.

Genomic segment sequence conservation across species and genera might affect the reassortment frequencies as most conserved segments probably reassort more easily. Nanovirus studies showed three groups of segment conservation: DNA-R, -S, -N are highly conserved at the species (~90% pairwise identity) and genera levels, followed by DNA-C, -U1 and DNA-M, -U2, -U4. Thus, at the intra-specific level sequence divergence does not seem to explain reassortment prevalence as DNA-C is more conserved than -M but has not been involved in any reassortment even described so far contrary to DNA-M which is the most involved segment e. DNA-R stands as the most conserved segment in the genus with around 76% pairwise identity between the two most divergent species. This conservation is less contrasted for the other genomic components, which share approximatively 60% pairwise identity between the most divergent species (Grigoras et al. 2014; Kraberger et al. 2017). From a large BBTV study, in average DNA-R (>88 %) and -C (>85 %) appear to be the most conserved inside the species followed by DNA-N (>83 %), -S (>82 %), -M (>82 %) and DNA-U3 (>74 %) (Stainton et al. 2015). Again, the degree of sequence conservation does not match sequence implication in reassortments (Table 2).

Conclusion

Although our understanding of the molecular processes in nanovirids is currently limited, reassortment constraints likely depend heavily on the trans-replication of the reassorted segment through compatibility between M-Rep and iterons. However, despite the limited diversity of babuviruses, the replication complementation between divergent nanovirus species and the small range of identified babuvirus iteron sequences may allow for more relaxed constraints than in begomoviruses. We presently do not know how replication of distinct segments with small differences in iteron sequences might affect intra-genomic interactions. Packaging constraints in nanovirids remain largely unknown, as no studies have been conducted on the interaction between the coat protein and DNA or packaging itself. However, frequent association with alpha- satellites and experimental observations showing the viability of an inter-specific reassortment of DNA-S suggest a permissive process. Concerning viral movement, interactions involving DNA-M, -S, -N, and -U4 may be necessary and may significantly limit the viability of reassortants (Table 1). However, the frequent loss of DNA-N and -U4 in experimental conditions, along with the absence of detected interactions between DNA-M and -S, suggest that these constraints may not be as important as initially thought. Regarding transmission, nanovirids heavily rely on both DNA-S for stable virion production and DNA-N for the helper component NSP required to penetrate the aphid midgut

barrier. Reassortment of DNA-N might therefore impose a severe constraint on reassortant transmission. However, once again, the experimental functional complementation between divergent species and the shared aphid vectors indicate degrees of permissiveness. Furthermore, the window for complementation of vector transmission opened by the peculiar mode of action of NSP likely enhances virus capacity for reassortment. Although natural observations confirm a relatively high frequency of reassortment in nanovirids, a striking discrepancy is present between predictions of relaxed constrains on inter-specific reassortments and their rarity in the field. Similarly, the scarcity of reassortments involving more than one segment is intriguing, and further commented in the Discussion section.

Other multipartite ssDNA viruses

The other ssDNA viruses that are known or strongly suspected to have a multipartite genome organization can be distinguished in two categories. The first one regroups viral genomes probably replicated by a protein of the Rep family according to the RCR mechanism, similar to the begomo- and nano-viruses detailed above. These include the plant-infecting *Coconut foliar decay virus* (CFDV), sole member of the *Cofodevirus* genus within the *Metaxyviridae* family (Gronenborn et al. 2018), and the first ssDNA multipartite virus described in fungi, *Fusarium graminearum gemytripvirus 1* (FgGMTV1), sole member of the *Gemytripvirus* genus within the *Genomoviridae* family (Li et al. 2020). Both of these viruses have three circular ssDNA genomic segments with a conserved stem-loop region analogous to the origin of replication of gemini- and nano-viruses, but very little information is available on the molecular biology of their infection cycle. Thus, apart from the necessary match between the Rep protein and the conserved iteron-containing region encompassing the origin of replication (Gronenborn et al. 2018; Li et al. 2020), how the other putative limitations to reassortment detailed above apply to these viruses is impossible to predict. Additional sequences become available every day and it is most likely that countless similar multipartite circular Rep-encoding ssDNA (CRESS) viral species will be characterized in the future (Male et al. 2016; Kraberger et al. 2019), perhaps infecting a wider range of hosts where the role or consequences of reassortment will be most interesting to address.

The second category of multipartite ssDNA virus is represented by two animal infecting species: *Bombyx mori bidensovirus* (BnBDV), only representative species of the *Bidensovirus* genus in the *Bidnaviridae* family (Hu et al. 2013), and *Acheta domesticus segmented densovirus* (AdSV) representative of the *Brevihamaparvovirus* genus in the *Parvoviridae* family (Pénzes et

al. 2023). The genomes of these two viruses consist of two linear ssDNA genomic segments (Wang et al. 2006; Pénzes et al. 2023). Molecular processes of the infection cycle of bidensoviruses and brevihamaparvoviruses, encompassing replication (Tijssen and Bergoin 1995; Hu et al. 2016), packaging (Pan et al. 2014; Lü et al. 2017), movement and transmission, are poorly documented and could be very different from those of the plant-infecting CRESS viruses previously described. Unfortunately, despite evidence of co-occurrence of distinct BnBDV isolates (Gani et al. 2021), the existence and properties of putative reassortants have never been investigated, neither in laboratory conditions nor in natural populations.

Discussion

Among the reviewed molecular processes, replication compatibility appears to be the primary factor determining the production of viable reassortants. This is because trans-replication of heterologous segments mostly depends on specific recognition of iterons by the Rep protein for CRESS viruses, *i.e.* all ssDNA multipartite viruses up to date except for bidensoviruses and brevihamaparvoviruses. Updating the precise correspondence between iterons and the Rep IRD domain in known reassortants would greatly enhance our capacity to forecast potential associations. This is particularly significant for begomoviruses, as no IRD has yet been confirmed on the M-Rep protein of nanovirids, although it probably exists (Londoño, Riego-Ruiz and Argüello-Astorga 2010). Additionally, the characterization of nanovirus iterons remains incomplete, restricting our understanding of their trans-replication capacity. While babuvirus iterons have been more extensively characterized, with only three known species conclusions on their diversity are limited.

In contrast to replication, packaging and intra-host movement do not seem to impose major constraints on reassortment since they involve mostly non-specific binding to viral DNA. However, despite the permissiveness of their CP to the encapsidation of variable size DNA components, begomoviruses incur limitations to their genome size imposed by NSP, MP and plasmodesmata (Gilbertson et al. 2003) as, both NSP and MP trafficking is optimized around the size of the genome molecules (Rojas et al. 1998; Gilbertson et al. 2003). Limitations related to intra-cellular movement could arise if interactions among viral proteins carried by separate segments were involved. Current knowledge suggests that for begomoviruses the focus is mainly on interactions between NSP and MP, both of which are present on DNA-B, and potentially CP, carried by DNA-A (Table 1). Therefore, except when the CP is also involved, this step of the viral life cycle should not represent a significant impediment to reassortment.

However, the functional complementation of DNA-A movement by a DNA-B from a different species demonstrated certain limitations as reciprocal complementation for multiple species was not possible (Frischmuth et al. 1993). In the case of nanovirids, compatibility between MP, U4 and CP may be required but the processes involved in movement necessitate more investigations.

Host-to-host transmission is another life-cycle step that should not represent a major hurdle to reassortment. For begomoviruses, transmission predominantly relies on viral particles, making it largely reliant on the assembly and packaging of virions. Therefore, if a reassortant begomovirus can successfully complete the packaging step, transmission should not pose additional problems. In contrast, nanovirids require a helper component, the NSP protein, but inter-specific complementation of the NSP function is possible in laboratory conditions (Grigoras et al. 2018). A relevant and so far unique aspect of the helper function of nanovirids is that NSP can complement the transmission of viral particles that have been acquired by the vector at least a few days earlier (Di Mattia et al. 2022), greatly enhancing the temporal range of complementation and thus successful reassortment.

Confronting begomovirus predictions with field data is a limited exercise because these viruses have a bipartite organization. Identified functional molecular constraints in reassortments are mitigated by the fact that many functions are encoded on the same segment (Table 1). This situation seems different from the highly multipartite nanoviruses with mostly one function-one segment organization. Nevertheless, numerous reassortment events in the few available surveys reveal high reassortment prevalence which may suggest a relevant role in genome reconstitution, that is the possibility that DNA-A and DNA-B may be often transmitted separately from cell-to-cell or even host-to-host.

The nanovirus experimental functional complementations, the limited diversity of known babuvirus iterons, and the existing ecological opportunities would predict relatively high prevalence of inter-specific reassortments in nanovirids. However, this is not consistent with field data where inter-specific reassortants are puzzlingly rare relative to the abundance of inter-specific recombinants. Because recombination requires co-infections at the individual cell level, while reassortment in nanovirids may occur even in the absence of co-infection of the same individual host, the relative prevalence of inter-specific recombinants strongly suggests that the rarity of reassortants is not due to a lack of opportunity of encounter between viral genotypes but rather to significant constraints on inter-specific reassortant viability or competitivity. This striking imbalance between inter-specific reassortants vs recombinants is

intriguing and begs for an explanation. The availability of infectious clones for several nanovirus species, and the possibility to produce new ones for additional isolates of the same species, opens the opportunity to further investigate and compare the properties of inter- and intra-specific reassortment in the lab. Beyond viability, a prominent question would then be the comparison of their fitness relative to that of parental genotypes.

As incomplete sets of segments can be independently transmitted between hosts, missing segments, with *nanoviruses* having at least three non-essential segments in laboratory conditions, should allow for the complementation of multiple segments at once. Yet, another intriguing aspect of *nanovirid* reassortments is the almost exclusive representation of single-segment reassortment events in most studies, which seems to indicate significant genome disruption when more than one segment is involved. This would limit the scope of complementation of incomplete genomes, and could be experimentally investigated by the evaluation of multiple-segment reassortant fitness.

As mentioned earlier in this review, there is a discrepancy between the frequently observed loss of non-essential segments under laboratory conditions and the prevalence of complete genomes in deeply sequenced field samples (Knierim et al. 2019). One possible explanation could be that some important conditions in maintaining all segments in the field are relaxed in the lab. For example, nanoviruses may switch host very often *in natura* with some segments being mandatory depending on the host. Another possible explanation could be that field ecological dynamics involving hosts, vectors and viruses lead to frequent reconstitutions, rendering the existence of incomplete infections extremely transient and masking a recurrent but reversible loss of segments. In this regard segment reassortment frequency might shed some light on the complementation dynamics in natural populations and should in principle reveal more frequent reassortments for dispensable segments. From the limited number of relevant studies, all segments except DNA-C have been involved in reassortment events (Table 2). In particular there is no bias toward non-essential segment reassortments. This may also suggest that delayed complementation, enhanced by the action of nanovirus NSP, could occur and potentially rescue incomplete "latent" infections missing essential segments, if these viral particles can "wait" in their host and vector for a sufficient amount of time. This possibility of a latent phase for incomplete sets of segments lacking an essential function warrants further investigation as it could further enhance the potential to reassort and further reduce the cost of maintaining genomic integrity.

An additional point where nanovirid predictions and field observations do not match concerns how molecular constraints might affect the reassortment frequency of associated segments. Replication stands out as a key step limiting reassortment and we would thus expect DNA-R to be the least reassorting segment. Surprisingly, DNA-R does not stand out as a poorly reassorting segment, while DNA-M, another 'essential' segment, for which an interaction with CP is suspected is the most frequently reassorting in both genera (Table 2). The third essential segment, DNA-S, which is expected to require compatible interactions with NSP and MP is roughly reassorting as frequently as DNA-R (Table 2). The case of the non-essential DNA-C is even more intriguing. As it is supposed to modulate the host cell-cycle it should not impose serious limitations for reassortments. Yet, DNA-C has so far never been found reassorting in nanoviruses and to a limited extent in babuviruses (Table 2). Highly conserved segments could reassort more easily than expected from their implication in molecular constraints. This could be the case for the very conserved DNA-R which is found relatively frequently in reassortants (Grigoras et al. 2014; Stainton et al. 2015; Kraberger et al. 2018) despite its implication in trans-replication, the largest molecular constraint. However, sequence conservation cannot explain why DNA-M and -C differ so much in their reassortment prevalence as DNA-C is more conserved but reassorts less.

In conclusion, reassortment appears to be a significant phenomenon in ssDNA multipartite viruses. However, experimental predictions, especially for nanovirids, do not align with natural population observations, indicating that there are likely major gaps in our understanding of the molecular processes, the ecological dynamics, and the phenotypic effects. Conducting a large and systematic study on the fitness of reassortants of ssDNA multipartite viruses and their parental genotypes would help bridge this gap.

ACKNOWLEDGMENTS

This work was funded by the French ANRs Nanovirus ANR-18-CE92-0028, Reassort ANR-20-CE02-0016-01, as well as by the University of Montpellier MUSE, project MULTIVIR. SB acknowledges support from INRAE and Department SPE. YM acknowledges support from CNRS and IRD. We thank Philippe Roumagnac for useful comments and suggestions.

Bibliography

Alberter, B., Ali Rezaian, M., & Jeske, H. (2005). 'Replicative intermediates of Tomato leaf curl virus and its satellite DNAs', *Virology*, 331/2: 441–8. DOI: 10.1016/j.virol.2004.10.043

Amin, I., Ilyas, M., Qazi, J., Bashir, R., Yadav, J. S., Mansoor, S., Fauquet, C. M., et al. (2011). 'Identification of a major pathogenicity determinant and suppressors of RNA silencing encoded by a South Pacific isolate of Banana bunchy top virus originating from Pakistan', *Virus Genes*, 42/2: 272–81. DOI: 10.1007/s11262-010-0559-3

Argüello-Astorga, G. R., Guevara-González, R. G., Herrera-Estrella, L. R., & Rivera-Bustamante, R. F. (1994). 'Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication', *Virology*, 203/1: 90–100. DOI: 10.1006/viro.1994.1458

Argüello-Astorga, G. R., & Ruiz-Medrano, R. (2001). 'An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach', *Archives of Virology*, 146/8: 1465–85. DOI: 10.1007/s007050170072

Avalos-Calleros, J. A., Pastor-Palacios, G., Bolaños-Martínez, O. C., Mauricio-Castillo, A., Gregorio-Jorge, J., Martínez-Marrero, N., Bañuelos-Hernández, B., et al. (2021). 'Two strains of a novel begomovirus encoding Rep proteins with identical β1 strands but different β5 strands are not compatible in replication', *Archives of Virology*, 166/6: 1691–709. DOI: 10.1007/s00705-021-05066-2

Basu, S., Singh, A. K., Singh, D., Sahu, S. K., & Chakraborty, S. (2021). 'Role of viral suppressors governing asymmetric synergism between tomato-infecting begomoviruses', *Applied Microbiology and Biotechnology*, 105/3: 1107–21. DOI: 10.1007/s00253-020-11070-4

Bennett, A., & Agbandje-McKenna, M. (2020). 'Geminivirus structure and assembly', *Advances in Virus Research*, 108: 1–32. DOI: 10.1016/bs.aivir.2020.09.005

Briddon, R. W., Martin, D. P., Roumagnac, P., Navas-Castillo, J., Fiallo-Olivé, E., Moriones, E., Lett, J.-M., et al. (2018). 'Alphasatellitidae: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites', *Archives of Virology*, 163/9: 2587–600. DOI: 10.1007/s00705-018-3854-2

Briddon, R. W., Patil, B. L., Bagewadi, B., Nawaz-ul-Rehman, M. S., & Fauquet, C. M. (2010). 'Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses', *BMC evolutionary biology*, 10: 97. DOI: 10.1186/1471-2148-10-97

Bull, S. E., Briddon, R. W., Sserubombwe, W. S., Ngugi, K., Markham, P. G., & Stanley, J. (2007). 'Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses', *The Journal of General Virology*, 88/Pt 5: 1624–33. DOI: 10.1099/vir.0.82662-0

Burns, T. M., Harding, R. M., & Dale, J. L. (1995). 'The genome organization of banana bunchy top virus: analysis of six ssDNA components', *The Journal of General Virology*, 76 (Pt 6): 1471–82. DOI: 10.1099/0022-1317-76-6-1471

Casado, C. G., Javier Ortiz, G., Padron, E., Bean, S. J., McKenna, R., Agbandje-McKenna, M., & Boulton, M. I. (2004). 'Isolation and characterization of subgenomic DNAs encapsidated in

"single" T = 1 isometric particles of Maize streak virus', *Virology*, 323/1: 164–71. DOI: 10.1016/j.virol.2004.02.014

Chakraborty, S., Vanitharani, R., Chattopadhyay, B., & Fauquet, C. M. (2008). 'Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India', *The Journal of General Virology*, 89/Pt 3: 818–28. DOI: 10.1099/vir.0.82873-0

Chao, L., Tran, T., & Matthews, C. (1992). 'MULLER'S RATCHET AND THE ADVANTAGE OF SEX IN THE RNA VIRUS ϕ6', *Evolution; International Journal of Organic Evolution*, 46/2: 289–99. DOI: 10.1111/j.1558-5646.1992.tb02038.x

Chen, L.-F., Rojas, M., Kon, T., Gamby, K., Xoconostle-Cazares, B., & Gilbertson, R. L. (2009). 'A severe symptom phenotype in tomato in Mali is caused by a reassortant between a novel recombinant begomovirus (Tomato yellow leaf curl Mali virus) and a betasatellite', *Molecular Plant Pathology*, 10/3: 415–30. DOI: 10.1111/j.1364-3703.2009.00541.x

Chen, Y.-J., Lai, H.-C., Lin, C.-C., Neoh, Z. Y., & Tsai, W.-S. (2021). 'Genetic Diversity, Pathogenicity and Pseudorecombination of Cucurbit-Infecting Begomoviruses in Malaysia', *Plants (Basel, Switzerland)*, 10/11: 2396. DOI: 10.3390/plants10112396

Chiu, C.-W., Li, Y.-R., Lin, C.-Y., Yeh, H.-H., & Liu, M.-J. (2022). 'Translation initiation landscape profiling reveals hidden open-reading frames required for the pathogenesis of tomato yellow leaf curl Thailand virus', *The Plant Cell*, 34/5: 1804–21. DOI: 10.1093/plcell/koac019

Choudhury, N. R., Malik, P. S., Singh, D. K., Islam, M. N., Kaliappan, K., & Mukherjee, S. K. (2006). 'The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase', *Nucleic Acids Research*, 34/21: 6362–77. DOI: 10.1093/nar/gkl903

Clérot, D., & Bernardi, F. (2006). 'DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus', *Journal of Virology*, 80/22: 11322–30. DOI: 10.1128/JVI.00924-06

Crespo-Bellido, A., Hoyer, J. S., Dubey, D., Jeannot, R. B., & Duffy, S. (2021). 'Interspecies Recombination Has Driven the Macroevolution of Cassava Mosaic Begomoviruses', *Journal of Virology*, 95/17: e0054121. DOI: 10.1128/JVI.00541-21

Czosnek, H., Hariton-Shalev, A., Sobol, I., Gorovits, R., & Ghanim, M. (2017). 'The Incredible Journey of Begomoviruses in Their Whitefly Vector', *Viruses*, 9/10: 273. DOI: 10.3390/v9100273

Dale, J. L. (1987). 'Banana Bunchy Top: An Economically Important Tropical Plant Virus Disease'. Maramorosch K., Murphy F. A., & Shatkin A. J. (eds) *Advances in Virus Research*, Vol. 33, pp. 301–25. Academic Press. DOI: 10.1016/S0065-3527(08)60321-8

De Bruyn, A., Harimalala, M., Zinga, I., Mabvakure, B. M., Hoareau, M., Ravigné, V., Walters, M., et al. (2016). 'Divergent evolutionary and epidemiological dynamics of cassava mosaic geminiviruses in Madagascar', *BMC evolutionary biology*, 16/1: 182. DOI: 10.1186/s12862-016-0749-2

De Bruyn, A., Villemot, J., Lefeuvre, P., Villar, E., Hoareau, M., Harimalala, M., Abdoul-Karime, A. L., et al. (2012). 'East African cassava mosaic-like viruses from Africa to Indian ocean islands:

molecular diversity, evolutionary history and geographical dissemination of a bipartite begomovirus', *BMC evolutionary biology*, 12: 228. DOI: 10.1186/1471-2148-12-228

Di Mattia, J., Torralba, B., Yvon, M., Zeddam, J.-L., Blanc, S., & Michalakis, Y. (2022). 'Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution', *Proceedings of the National Academy of Sciences of the United States of America*, 119/32: e2201453119. DOI: 10.1073/pnas.2201453119

Di Mattia, J., Vernerey, M.-S., Yvon, M., Pirolles, E., Villegas, M., Gaafar, Y., Ziebell, H., et al. (2020). 'Route of a Multipartite Nanovirus across the Body of Its Aphid Vector', *Journal of Virology*, 94/9: e01998-19. DOI: 10.1128/JVI.01998-19

Di Mattia, J., Zeddam, J.-L., Uzest, M., & Blanc, S. (2023). 'The helper strategy in vector-transmission of plant viruses', *Peer Community Journal*, 3. DOI: 10.24072/pcjournal.258

Escriu, F., Fraile, A., & García-Arenal, F. (2007). 'Constraints to genetic exchange support gene coadaptation in a tripartite RNA virus', *PLoS pathogens*, 3/1: e8. DOI: 10.1371/journal.ppat.0030008

Fan, Y.-Y., Zhong, Y.-W., Pan, L.-L., Wang, X.-W., Ding, M., & Liu, S.-S. (2023). 'A shift of vector specificity acquired by a begomovirus through natural homologous recombination', *Molecular Plant Pathology*. DOI: 10.1111/mpp.13351

Ferro, C. G., Zerbini, F. M., Navas-Castillo, J., & Fiallo-Olivé, E. (2021). 'Revealing the Complexity of Sweepovirus-Deltasatellite-Plant Host Interactions: Expanded Natural and Experimental Helper Virus Range and Effect Dependence on Virus-Host Combination', *Microorganisms*, 9/5: 1018. DOI: 10.3390/microorganisms9051018

Fiallo-Olivé, E., Martínez-Zubiaur, Y., Moriones, E., & Navas-Castillo, J. (2012). 'A novel class of DNA satellites associated with New World begomoviruses', *Virology*, 426/1: 1–6. DOI: 10.1016/j.virol.2012.01.024

Fiallo-Olivé, E., & Navas-Castillo, J. (2023). 'The Role of Extensive Recombination in the Evolution of Geminiviruses', *Current Topics in Microbiology and Immunology*, 439: 139–66. DOI: 10.1007/978-3-031-15640-3_4

Fiallo-Olivé, E., Pan, L.-L., Liu, S.-S., & Navas-Castillo, J. (2020). 'Transmission of Begomoviruses and Other Whitefly-Borne Viruses: Dependence on the Vector Species', *Phytopathology*, 110/1: 10–7. DOI: 10.1094/PHYTO-07-19-0273-FI

Fondong, V. N. (2013). 'Geminivirus protein structure and function', *Molecular Plant Pathology*, 14/6: 635–49. DOI: 10.1111/mpp.12032

Fontenele, R. S., Bhaskara, A., Cobb, I. N., Majure, L. C., Salywon, A. M., Avalos-Calleros, J. A., Argüello-Astorga, G. R., et al. (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. DOI: 10.3390/v13050810

Fontes, E. P., Eagle, P. A., Sipe, P. S., Luckow, V. A., & Hanley-Bowdoin, L. (1994). 'Interaction between a geminivirus replication protein and origin DNA is essential for viral replication', *The Journal of Biological Chemistry*, 269/11: 8459–65.

Franz, A. W., van der Wilk, F., Verbeek, M., Dullemans, A. M., & van den Heuvel, J. F. (1999). 'Faba bean necrotic yellows virus (genus Nanovirus) requires a helper factor for its aphid transmission', *Virology*, 262/1: 210–9. DOI: 10.1006/viro.1999.9904

Frischmuth, S., Wege, C., Hülser, D., & Jeske, H. (2007). 'The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast', *Protoplasma*, 230/1–2: 117–23. DOI: 10.1007/s00709-006-0223-x

Frischmuth, T., Roberts, S., von Arnim, A., & Stanley, J. (1993). 'Specificity of bipartite geminivirus movement proteins', *Virology*, 196/2: 666–73. DOI: 10.1006/viro.1993.1523

Frischmuth, Thomas, Ringel, M., & Kocher, C. (2001). 'The size of encapsidated single-stranded DNA determines the multiplicity of African cassava mosaic virus particles', *The Journal of General Virology*, 82/Pt 3: 673–6. DOI: 10.1099/0022-1317-82-3-673

Fu, H.-C., Hu, J.-M., Hung, T.-H., Su, H.-J., & Yeh, H.-H. (2009). 'Unusual events involved in Banana bunchy top virus strain evolution', *Phytopathology*, 99/7: 812–22. DOI: 10.1094/PHYTO-99-7-0812

Gallet, R., Di Mattia, J., Ravel, S., Zeddam, J.-L., Vitalis, R., Michalakis, Y., & Blanc, S. (2022). 'Gene copy number variations at the within-host population level modulate gene expression in a multipartite virus', *Virus Evolution*, 8/2: veac058. DOI: 10.1093/ve/veac058

Gallet, R., Kraberger, S., Filloux, D., Galzi, S., Fontes, H., Martin, D. P., Varsani, A., et al. (2018). 'Nanovirus-alphasatellite complex identified in Vicia cracca in the Rhône delta region of France', *Archives of Virology*, 163/3: 695–700. DOI: 10.1007/s00705-017-3634-4

Gani, M., Senger, S., Lokanath, S., Saini, P., Bali, K., Gupta, R., Sivaprasad, V., et al. (2021). 'Patterns in Genotype Composition of Indian Isolates of the Bombyx mori Nucleopolyhedrovirus and Bombyx mori Bidensovirus', *Viruses*, 13/5: 901. DOI: 10.3390/v13050901

Garrido-Ramirez, E. R., Sudarshana, M. R., & Gilbertson, R. L. (2000). 'Bean golden yellow mosaic virus from Chiapas, Mexico: Characterization, Pseudorecombination with Other Bean-Infecting Geminiviruses and Germ Plasm Screening', *Phytopathology*, 90/11: 1224–32. DOI: 10.1094/PHYTO.2000.90.11.1224

Gilbertson, R. L., Batuman, O., Webster, C. G., & Adkins, S. (2015). 'Role of the Insect Supervectors Bemisia tabaci and Frankliniella occidentalis in the Emergence and Global Spread of Plant Viruses', *Annual Review of Virology*, 2/1: 67–93. DOI: 10.1146/annurev-virology-031413-085410

Gilbertson, R. L., Sudarshana, M., Jiang, H., Rojas, M. R., & Lucas, W. J. (2003). 'Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: insights into DNA trafficking', *The Plant Cell*, 15/11: 2578–91. DOI: 10.1105/tpc.015057

Gong, P., Tan, H., Zhao, S., Li, H., Liu, H., Ma, Y., Zhang, X., et al. (2021). 'Geminiviruses encode additional small proteins with specific subcellular localizations and virulence function', *Nature Communications*, 12/1: 4278. DOI: 10.1038/s41467-021-24617-4

Gong, P., Zhao, S., Liu, H., Chang, Z., Li, F., & Zhou, X. (2022). 'Tomato yellow leaf curl virus V3 protein traffics along microfilaments to plasmodesmata to promote virus cell-to-cell movement', *Science China. Life Sciences*, 65/5: 1046–9. DOI: 10.1007/s11427-021-2063-4

Gregorio-Jorge, J., Bernal-Alcocer, A., Bañuelos-Hernández, B., Alpuche-Solís, A. G., Hernández-Zepeda, C., Moreno-Valenzuela, O., Frías-Treviño, G., et al. (2010). 'Analysis of a new strain of Euphorbia mosaic virus with distinct replication specificity unveils a lineage of begomoviruses with short Rep sequences in the DNA-B intergenic region', *Virology Journal*, 7: 275. DOI: 10.1186/1743-422X-7-275

Grigoras, I., Ginzo, A. I. D. C., Martin, D. P., Varsani, A., Romero, J., Mammadov, A. C., Huseynova, I. M., et al. (2014). 'Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe', *The Journal of General Virology*, 95/Pt 5: 1178–91. DOI: 10.1099/vir.0.063115-0

Grigoras, I., Timchenko, T., Katul, L., Grande-Pérez, A., Vetten, H.-J., & Gronenborn, B. (2009). 'Reconstitution of authentic nanovirus from multiple cloned DNAs', *Journal of Virology*, 83/20: 10778–87. DOI: 10.1128/JVI.01212-09

Grigoras, I., Vetten, H.-J., Commandeur, U., Ziebell, H., Gronenborn, B., & Timchenko, T. (2018). 'Nanovirus DNA-N encodes a protein mandatory for aphid transmission', *Virology*, 522: 281–91. DOI: 10.1016/j.virol.2018.07.001

Gronenborn, B., Randles, J. W., Knierim, D., Barrière, Q., Vetten, H. J., Warthmann, N., Cornu, D., et al. (2018). 'Analysis of DNAs associated with coconut foliar decay disease implicates a unique single-stranded DNA virus representing a new taxon', *Scientific Reports*, 8/1: 5698. DOI: 10.1038/s41598-018-23739-y

Gutiérrez, S., & Zwart, M. P. (2018). 'Population bottlenecks in multicomponent viruses: first forays into the uncharted territory of genome-formula drift', *Current Opinion in Virology*, 33: 184–90. DOI: 10.1016/j.coviro.2018.09.001

Guyot, V., Rajeswaran, R., Chu, H. C., Karthikeyan, C., Laboureau, N., Galzi, S., Mukwa, L. F. T., et al. (2022). 'A newly emerging alphasatellite affects banana bunchy top virus replication, transcription, siRNA production and transmission by aphids', *PLoS pathogens*, 18/4: e1010448. DOI: 10.1371/journal.ppat.1010448

Hafner, G. J., Harding, R. M., & Dale, J. L. (1995). 'Movement and transmission of banana bunchy top virus DNA component one in bananas', *The Journal of General Virology*, 76 (Pt 9): 2279–85. DOI: 10.1099/0022-1317-76-9-2279

Hafner, G. J., Stafford, M. R., Wolter, L. C., Harding, R. M., & Dale, J. L. (1997). 'Nicking and joining activity of banana bunchy top virus replication protein in vitro', *The Journal of General Virology*, 78 (Pt 7): 1795–9. DOI: 10.1099/0022-1317-78-7-1795

Hartitz, M. D., Sunter, G., & Bisaro, D. M. (1999). 'The tomato golden mosaic virus transactivator (TrAP) is a single-stranded DNA and zinc-binding phosphoprotein with an acidic activation domain', *Virology*, 263/1: 1–14. DOI: 10.1006/viro.1999.9925

Hasanvand, V., Heydanejad, J., Massumi, H., Kleinow, T., Jeske, H., Fontenele, R. S., Kraberger, S., et al. (2021). 'Genome characterization of parsley severe stunt-associated virus in Iran', *Virus Genes*, 57/3: 293–301. DOI: 10.1007/s11262-021-01835-y

Hassan-Sheikhi, P., Heydarnejad, J., Massumi, H., Kraberger, S., & Varsani, A. (2020). 'Novel nanovirus and associated alphasatellites identified in milk vetch plants with chlorotic dwarf disease in Iran', *Virus Research*, 276: 197830. DOI: 10.1016/j.virusres.2019.197830

He, Y.-Z., Wang, Y.-M., Yin, T.-Y., Fiallo-Olivé, E., Liu, Y.-Q., Hanley-Bowdoin, L., & Wang, X.-W. (2020). 'A plant DNA virus replicates in the salivary glands of its insect vector via recruitment of host DNA synthesis machinery', *Proceedings of the National Academy of Sciences of the United States of America*, 117/29: 16928–37. DOI: 10.1073/pnas.1820132117

Hehnle, S., Wege, C., & Jeske, H. (2004). 'Interaction of DNA with the movement proteins of geminiviruses revisited', *Journal of Virology*, 78/14: 7698–706. DOI: 10.1128/JVI.78.14.7698-7706.2004

Herrera-Valencia, V. A., Dugdale, B., Harding, R. M., & Dale, J. L. (2006). 'An iterated sequence in the genome of Banana bunchy top virus is essential for efficient replication', *The Journal of General Virology*, 87/Pt 11: 3409–12. DOI: 10.1099/vir.0.82166-0

Hesketh, E. L., Saunders, K., Fisher, C., Potze, J., Stanley, J., Lomonossoff, G. P., & Ranson, N. A. (2018). 'The 3.3 Å structure of a plant geminivirus using cryo-EM', *Nature Communications*, 9/1: 2369. DOI: 10.1038/s41467-018-04793-6

Hill, J. E., Strandberg, J. O., Hiebert, E., & Lazarowitz, S. G. (1998). 'Asymmetric infectivity of pseudorecombinants of cabbage leaf curl virus and squash leaf curl virus: implications for bipartite geminivirus evolution and movement', *Virology*, 250/2: 283–92. DOI: 10.1006/viro.1998.9366

Hipp, K., Grimm, C., Jeske, H., & Böttcher, B. (2017). 'Near-Atomic Resolution Structure of a Plant Geminivirus Determined by Electron Cryomicroscopy', *Structure (London, England: 1993)*, 25/8: 1303-1309.e3. DOI: 10.1016/j.str.2017.06.013

Ho, E. S., Kuchie, J., & Duffy, S. (2014). 'Bioinformatic analysis reveals genome size reduction and the emergence of tyrosine phosphorylation site in the movement protein of New World bipartite begomoviruses', *PloS One*, 9/11: e111957. DOI: 10.1371/journal.pone.0111957

Höhnle, M., Höfer, P., Bedford, I. D., Briddon, R. W., Markham, P. G., & Frischmuth, T. (2001). 'Exchange of three amino acids in the coat protein results in efficient whitefly transmission of a nontransmissible Abutilon mosaic virus isolate', *Virology*, 290/1: 164–71. DOI: 10.1006/viro.2001.1140

Holland, J., & Domingo, E. (1998). 'Origin and Evolution of Viruses', *Virus Genes*, 16/1: 13–21. DOI: 10.1023/A:1007989407305

Hooker, W. J., & Salazar, L. F. (1983). 'A new plant virus from the high jungle of the Eastern Andes; Solanum apical leaf curling virus (SALCV)', *Annals of Applied Biology*, 103/3: 449–54. DOI: 10.1111/j.1744-7348.1983.tb02782.x

Horser, C. L., Harding, R. M., & Dale, J. L. (2001). 'Banana bunchy top nanovirus DNA-1 encodes the "master" replication initiation protein', *The Journal of General Virology*, 82/Pt 2: 459–64. DOI: 10.1099/0022-1317-82-2-459

Hou, Y. M., & Gilbertson, R. L. (1996). 'Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination.', *Journal of Virology*, 70/8: 5430–6.

Hu, ZhaoYang, Li, G., Li, G., Yao, Q., & Chen, K. (2013). 'Bombyx mori bidensovirus: The type species of the new genus Bidensovirus in the new family Bidnaviridae', *Chinese Science Bulletin*, 58/36: 4528–32. DOI: 10.1007/s11434-013-5876-1

Hu, Zhaoyang, Zhang, X., Liu, W., Zhou, Q., Zhang, Q., Li, G., & Yao, Q. (2016). 'Genome segments accumulate with different frequencies in Bombyx mori bidensovirus', *Journal of Basic Microbiology*, 56/12: 1338–43. DOI: 10.1002/jobm.201600120

Idris, A. M., Mills-Lujan, K., Martin, K., & Brown, J. K. (2008). 'Melon chlorotic leaf curl virus: characterization and differential reassortment with closest relatives reveal adaptive virulence in the squash leaf curl virus clade and host shifting by the host-restricted bean calico mosaic virus', *Journal of Virology*, 82/4: 1959–67. DOI: 10.1128/JVI.01992-07

Ilyina, T. V., & Koonin, E. V. (1992). 'Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria', *Nucleic Acids Research*, 20/13: 3279–85. DOI: 10.1093/nar/20.13.3279

Ingham, D. J., Pascal, E., & Lazarowitz, S. G. (1995). 'Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity', *Virology*, 207/1: 191–204. DOI: 10.1006/viro.1995.1066

Jeffrey, J. L., Pooma, W., & Petty, I. T. (1996). 'Genetic requirements for local and systemic movement of tomato golden mosaic virus in infected plants', *Virology*, 223/1: 208–18. DOI: 10.1006/viro.1996.0469

Jeske, H. (2009). 'Geminiviruses', *Current Topics in Microbiology and Immunology*, 331: 185–226. DOI: 10.1007/978-3-540-70972-5_11

Jeske, Holger, Lütgemeier, M., & Preiß, W. (2001). 'DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus', *The EMBO Journal*, 20/21: 6158–67. DOI: 10.1093/emboj/20.21.6158

Ji, X.-L., Yu, N.-T., Qu, L., Li, B.-B., & Liu, Z.-X. (2019). 'Banana bunchy top virus (BBTV) nuclear shuttle protein interacts and re-distributes BBTV coat protein in Nicotiana benthamiana', *3 Biotech*, 9/4: 121. DOI: 10.1007/s13205-019-1656-1

Jovel, J., Preiss, W., & Jeske, H. (2007). 'Characterization of DNA intermediates of an arising geminivirus', *Virus Research*, 130/1–2: 63–70. DOI: 10.1016/j.virusres.2007.05.018

Kamran, A., Hou, H., Xie, Y., Zhao, C., Wei, X., Zhang, C., Yu, X., et al. (2019). 'Full genome sequence analysis and putative host-shifting of Milk vetch dwarf virus infecting tobacco (Nicotiana tabacum) in China', *Virology Journal*, 16/1: 38. DOI: 10.1186/s12985-019-1129-5

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. DOI: 10.1038/s41467-019-11433-0

Kazlauskas, D., Varsani, A., & Krupovic, M. (2018). 'Pervasive Chimerism in the Replication-Associated Proteins of Uncultured Single-Stranded DNA Viruses', *Viruses*, 10/4: 187. DOI: 10.3390/v10040187

Kraberger, S., Kumari, S. G., Najar, A., Stainton, D., Martin, D. P., & Varsani, A. (2018). 'Molecular characterization of faba bean necrotic yellows viruses in Tunisia', *Archives of Virology*, 163/3: 687–94. DOI: 10.1007/s00705-017-3651-3

Kraberger, S., Schmidlin, K., Fontenele, R. S., Walters, M., & Varsani, A. (2019). 'Unravelling the Single-Stranded DNA Virome of the New Zealand Blackfly', *Viruses*, 11/6: 532. DOI: 10.3390/v11060532

Krenz, B., Schießl, I., Greiner, E., & Krapp, S. (2017). 'Analyses of pea necrotic yellow dwarf virus-encoded proteins', *Virus Genes*, 53/3: 454–63. DOI: 10.1007/s11262-017-1439-x

Lageix, S., Catrice, O., Deragon, J.-M., Gronenborn, B., Pélissier, T., & Ramírez, B. C. (2007). 'The nanovirus-encoded Clink protein affects plant cell cycle regulation through interaction with the retinoblastoma-related protein', *Journal of Virology*, 81/8: 4177–85. DOI: 10.1128/JVI.02103-06

Lal, A., Vo, T. T. B., Sanjaya, I. G. N. P. W., Ho, P. T., Kim, J.-K., Kil, E.-J., & Lee, S. (2020). 'Nanovirus Disease Complexes: An Emerging Threat in the Modern Era', *Frontiers in Plant Science*, 11: 558403. DOI: 10.3389/fpls.2020.558403

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S. G., Schell, J., & Gronenborn, B. (1995). 'In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus.', *Proceedings of the National Academy of Sciences of the United States of America*, 92/9: 3879–83.

Lefeuvre, P., Lett, J.-M., Varsani, A., & Martin, D. P. (2009). 'Widely conserved recombination patterns among single-stranded DNA viruses', *Journal of Virology*, 83/6: 2697–707. DOI: 10.1128/JVI.02152-08

Lefeuvre, P., Martin, D. P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., et al. (2007). 'Begomovirus "melting pot" in the south-west Indian Ocean islands: molecular diversity and evolution through recombination', *The Journal of General Virology*, 88/Pt 12: 3458–68. DOI: 10.1099/vir.0.83252-0

Lefeuvre, Pierre, & Moriones, E. (2015). 'Recombination as a motor of host switches and virus emergence: geminiviruses as case studies', *Current Opinion in Virology*, 10: 14–9. DOI: 10.1016/j.coviro.2014.12.005

Levy, A., & Czosnek, H. (2003). 'The DNA-B of the non-phloem-limited bean dwarf mosaic virus (BDMV) is able to move the phloem-limited Abutilon mosaic virus (AbMV) out of the phloem, but DNA-B of AbMV is unable to confine BDMV to the phloem', *Plant Molecular Biology*, 53/6: 789–803. DOI: 10.1023/B:PLAN.0000023662.25756.43

Li, P., Wang, S., Zhang, L., Qiu, D., Zhou, X., & Guo, L. (2020). 'A tripartite ssDNA mycovirus from a plant pathogenic fungus is infectious as cloned DNA and purified virions', *Science Advances*, 6/14: eaay9634. DOI: 10.1126/sciadv.aay9634

Lima, A. T. M., Silva, J. C. F., Silva, F. N., Castillo-Urquiza, G. P., Silva, F. F., Seah, Y. M., Mizubuti, E. S. G., et al. (2017). 'The diversification of begomovirus populations is predominantly driven by mutational dynamics', *Virus Evolution*, 3/1: vex005. DOI: 10.1093/ve/vex005

Lin, B., Akbar Behjatnia, S. A., Dry, I. B., Randles, J. W., & Rezaian, M. A. (2003). 'High-affinity Rep-binding is not required for the replication of a geminivirus DNA and its satellite', *Virology*, 305/2: 353–63. DOI: 10.1006/viro.2002.1671

Liu, H., Boulton, M. I., & Davies, J. W. (1997). 'Maize streak virus coat protein binds single- and double-stranded DNA in vitro', *The Journal of General Virology*, 78 (Pt 6): 1265–70. DOI: 10.1099/0022-1317-78-6-1265

Londoño, A., Riego-Ruiz, L., & Argüello-Astorga, G. R. (2010). 'DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs', *Archives of Virology*, 155/7: 1033–46. DOI: 10.1007/s00705-010-0674-4

Lotfipour, M., Behjatnia, S. A. A., Dall'Ara, M., & Ratti, C. (2020). 'The full-length genome characterization and diversity of faba bean necrotic stunt virus in Iran', *European Journal of Plant Pathology*, 157/2: 239–50. DOI: 10.1007/s10658-020-01973-3

Lü, P., Xing, Y., Hu, Z., Yang, Y., Pan, Y., Chen, K., Zhu, F., et al. (2017). 'A characterization of structural proteins expressed by Bombyx mori bidensovirus', *Journal of Invertebrate Pathology*, 144: 18–23. DOI: 10.1016/j.jip.2016.12.008

Male, M. F., Kraberger, S., Stainton, D., Kami, V., & Varsani, A. (2016). 'Cycloviruses, gemycircularviruses and other novel replication-associated protein encoding circular viruses in Pacific flying fox (Pteropus tonganus) faeces', *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 39: 279–92. DOI: 10.1016/j.meegid.2016.02.009

Maliano, M. R., Rojas, M. R., Macedo, M. A., Barboza, N., & Gilbertson, R. L. (2022). 'The invasion biology of tomato begomoviruses in Costa Rica reveals neutral synergism that may lead to increased disease pressure and economic loss', *Virus Research*, 317: 198793. DOI: 10.1016/j.virusres.2022.198793

Mandal, B., Shilpi, S., Barman, A. R., Mandal, S., & Varma, A. (2013). 'Nine novel DNA components associated with the foorkey disease of large cardamom: evidence of a distinct babuvirus species in Nanoviridae', *Virus Research*, 178/2: 297–305. DOI: 10.1016/j.virusres.2013.09.027

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R., et al. (1999). 'Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan', *Virology*, 259/1: 190–9. DOI: 10.1006/viro.1999.9766

Mansoor, Shahid, Briddon, R. W., Zafar, Y., & Stanley, J. (2003). 'Geminivirus disease complexes: an emerging threat', *Trends in Plant Science*, 8/3: 128–34. DOI: 10.1016/S1360-1385(03)00007-4

Mansourpour, M., Gallet, R., Abbasi, A., Blanc, S., Dizadji, A., & Zeddam, J.-L. (2022). 'Effects of an Alphasatellite on the Life Cycle of the Nanovirus Faba Bean Necrotic Yellows Virus', *Journal of Virology*, 96/3: e0138821. DOI: 10.1128/JVI.01388-21

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., & Varsani, A. (2011). 'Recombination in Eukaryotic Single Stranded DNA Viruses', *Viruses*, 3/9: 1699–738. DOI: 10.3390/v3091699

Martin, D. P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.-Y., Dijoux, B., Vincent, C., et al. (2011). 'Complex Recombination Patterns Arising during Geminivirus Coinfections Preserve and Demarcate Biologically Important Intra-Genome Interaction Networks', *PLoS Pathogens*, 7/9: e1002203. DOI: 10.1371/journal.ppat.1002203

Matsuzaki, Y., Mizuta, K., Sugawara, K., Tsuchiya, E., Muraki, Y., Hongo, S., Suzuki, H., et al. (2003). 'Frequent Reassortment among Influenza C Viruses', *Journal of Virology*, 77/2: 871–81. DOI: 10.1128/JVI.77.2.871-881.2003

McDonald, S. M., Nelson, M. I., Turner, P. E., & Patton, J. T. (2016). 'Reassortment in segmented RNA viruses: mechanisms and outcomes', *Nature Reviews. Microbiology*, 14/7: 448–60. DOI: 10.1038/nrmicro.2016.46

Morra, M. R., & Petty, I. T. (2000). 'Tissue specificity of geminivirus infection is genetically determined', *The Plant Cell*, 12/11: 2259–70. DOI: 10.1105/tpc.12.11.2259

Moshe, A., Belausov, E., Niehl, A., Heinlein, M., Czosnek, H., & Gorovits, R. (2015). 'The Tomato yellow leaf curl virus V2 protein forms aggregates depending on the cytoskeleton integrity and binds viral genomic DNA', *Scientific Reports*, 5: 9967. DOI: 10.1038/srep09967

Mubin, M., Ijaz, S., Nahid, N., Hassan, M., Younus, A., Qazi, J., & Nawaz-Ul-Rehman, M. S. (2020). 'Journey of begomovirus betasatellite molecules: from satellites to indispensable partners', *Virus Genes*, 56/1: 16–26. DOI: 10.1007/s11262-019-01716-5

Ohshima, K., Matsumoto, K., Yasaka, R., Nishiyama, M., Soejima, K., Korkmaz, S., Ho, S. Y. W., et al. (2016). 'Temporal analysis of reassortment and molecular evolution of Cucumber mosaic virus: Extra clues from its segmented genome', *Virology*, 487: 188–97. DOI: 10.1016/j.virol.2015.09.024

Orozco, B. M., & Hanley-Bowdoin, L. (1998). 'Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein', *The Journal of Biological Chemistry*, 273/38: 24448–56. DOI: 10.1074/jbc.273.38.24448

Ouattara, A., Tiendrébéogo, F., Becker, N., Urbino, C., Thébaud, G., Hoareau, M., Allibert, A., et al. (2022). 'Synergy between an emerging monopartite begomovirus and a DNA-B component', *Scientific Reports*, 12/1: 695. DOI: 10.1038/s41598-021-03957-7

Padidam, M., Beachy, R. N., & Fauquet, C. M. (1996). 'The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus', *Virology*, 224/2: 390–404. DOI: 10.1006/viro.1996.0546

Pakkianathan, B. C., Kontsedalov, S., Lebedev, G., Mahadav, A., Zeidan, M., Czosnek, H., & Ghanim, M. (2015). 'Replication of Tomato Yellow Leaf Curl Virus in Its Whitefly Vector, Bemisia tabaci', *Journal of Virology*, 89/19: 9791–803. DOI: 10.1128/JVI.00779-15

Palanichelvam, K., Kunik, T., Citovsky, V., & Gafni, Y. (1998). 'The capsid protein of tomato yellow leaf curl virus binds cooperatively to single-stranded DNA', *The Journal of General Virology*, 79 (Pt 11): 2829–33. DOI: 10.1099/0022-1317-79-11-2829

Pan, X., Lü, P., Zhang, M., Hu, Z., Li, G., Ma, S., Feng, F., et al. (2014). 'Expression analysis of Bombyx mori bidensovirus structural proteins and assembly of virus-like particles in insect cells', *Current Microbiology*, 69/4: 567–73. DOI: 10.1007/s00284-014-0613-9

Pascal, E., Sanderfoot, A. A., Ward, B. M., Medville, R., Turgeon, R., & Lazarowitz, S. G. (1994). 'The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus', *The Plant Cell*, 6/7: 995–1006. DOI: 10.1105/tpc.6.7.995

Patil, B. L., & Dasgupta, I. (2022). 'Characterization of the functional domains of nuclear shuttle protein (NSP) of Indian cassava mosaic virus using green fluorescent protein as reporter', *Virus Genes*, 58/4: 308–18. DOI: 10.1007/s11262-022-01909-5

Patil, B. L., Dutt, N., Briddon, R. W., Bull, S. E., Rothenstein, D., Borah, B. K., Dasgupta, I., et al. (2007). 'Deletion and recombination events between the DNA-A and DNA-B components of Indian cassava-infecting geminiviruses generate defective molecules in Nicotiana benthamiana', *Virus Research*, 124/1–2: 59–67. DOI: 10.1016/j.virusres.2006.10.003

Pénzes, J. J., Pham, H. T., Chipman, P., Smith, E. W., McKenna, R., & Tijssen, P. (2023). 'Bipartite genome and structural organization of the parvovirus Acheta domesticus segmented densovirus', *Nature Communications*, 14/1: 3515. DOI: 10.1038/s41467-023-38875-x

Pita, J. S., Fondong, V. N., Sangaré, 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', *The Journal of General Virology*, 82/Pt 3: 655–65. DOI: 10.1099/0022-1317-82-3-655

Pooma, W., Gillette, W. K., Jeffrey, J. L., & Petty, I. T. (1996). 'Host and viral factors determine the dispensability of coat protein for bipartite geminivirus systemic movement', *Virology*, 218/1: 264–8. DOI: 10.1006/viro.1996.0189

Preiss, W., & Jeske, H. (2003). 'Multitasking in Replication Is Common among Geminiviruses', *Journal of Virology*, 77/5: 2972–80. DOI: 10.1128/JVI.77.5.2972-2980.2003

Qazi, J. (2016). 'Banana bunchy top virus and the bunchy top disease', *Journal of General Plant Pathology*, 82/1: 2–11. DOI: 10.1007/s10327-015-0642-7

Qin, S., Ward, B. M., & Lazarowitz, S. G. (1998). 'The bipartite geminivirus coat protein aids BR1 function in viral movement by affecting the accumulation of viral single-stranded DNA', *Journal of Virology*, 72/11: 9247–56. DOI: 10.1128/JVI.72.11.9247-9256.1998

Ramos, P. L., Guevara-González, R. G., Peral, R., Ascencio-Ibañez, J. T., Polston, J. E., Argüello-Astorga, G. R., Vega-Arreguín, J. C., et al. (2003). 'Tomato mottle Taino virus pseudorecombines with PYMV but not with ToMoV: Implications for the delimitation of cis- and trans-acting replication specificity determinants', *Archives of Virology*, 148/9: 1697–712. DOI: 10.1007/s00705-003-0136-3

Rizvi, I., Choudhury, N. R., & Tuteja, N. (2015). 'Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication', *Archives of Virology*, 160/2: 375–87. DOI: 10.1007/s00705-014-2297-7

Rojas, M. R., Noueiry, A. O., Lucas, W. J., & Gilbertson, R. L. (1998). 'Bean Dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner', *Cell*, 95/1: 105–13. DOI: 10.1016/s0092-8674(00)81786-9

Rojas, Maria 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: 361–94. DOI: 10.1146/annurev.phyto.43.040204.135939

Roossinck, M. J. (1997). 'Mechanisms of plant virus evolution', *Annual Review of Phytopathology*, 35: 191–209. DOI: 10.1146/annurev.phyto.35.1.191

Rosen, R., Kanakala, S., Kliot, A., Cathrin Pakkianathan, B., Farich, B. A., Santana-Magal, N., Elimelech, M., et al. (2015). 'Persistent, circulative transmission of begomoviruses by whitefly vectors', *Current Opinion in Virology*, 15: 1–8. DOI: 10.1016/j.coviro.2015.06.008

Roshan, P., Kulshreshtha, A., & Hallan, V. (2017). 'Identification of host cellular targets of AC4 and AV2 proteins of tomato leaf curl palampur virus and their sub-cellular localization studies', *Virusdisease*, 28/4: 390–400. DOI: 10.1007/s13337-017-0405-5

Roshan, P., Kulshreshtha, A., Kumar, S., Purohit, R., & Hallan, V. (2018). 'AV2 protein of tomato leaf curl Palampur virus promotes systemic necrosis in Nicotiana benthamiana and interacts with host Catalase2', *Scientific Reports*, 8/1: 1273. DOI: 10.1038/s41598-018-19292-3

Rothenstein, D., Krenz, B., Selchow, O., & Jeske, H. (2007). 'Tissue and cell tropism of Indian cassava mosaic virus (ICMV) and its AV2 (precoat) gene product', *Virology*, 359/1: 137–45. DOI: 10.1016/j.virol.2006.09.014

Saeed, M., Zafar, Y., Randles, J. W., & Rezaian, M. A. (2007). 'A monopartite begomovirus-associated DNA beta satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection', *The Journal of General Virology*, 88/Pt 10: 2881–9. DOI: 10.1099/vir.0.83049-0

Safari Murhububa, I., Tougeron, K., Bragard, C., Fauconnier, M.-L., Bisimwa Basengere, E., Walangululu Masamba, J., & Hance, T. (2021). 'Banana Tree Infected with Banana Bunchy Top Virus Attracts Pentalonia nigronervosa Aphids Through Increased Volatile Organic Compounds Emission', *Journal of Chemical Ecology*, 47/8–9: 755–67. DOI: 10.1007/s10886-021-01298-3

Sanderfoot, A. A., Ingham, D. J., & Lazarowitz, S. G. (1996). 'A viral movement protein as a nuclear shuttle. The geminivirus BR1 movement protein contains domains essential for interaction with BL1 and nuclear localization', *Plant Physiology*, 110/1: 23–33. DOI: 10.1104/pp.110.1.23

Sanderfoot, A. A., & Lazarowitz, S. G. (1995). 'Cooperation in Viral Movement: The Geminivirus BL1 Movement Protein Interacts with BR1 and Redirects It from the Nucleus to the Cell Periphery', *The Plant Cell*, 7/8: 1185–94. DOI: 10.1105/tpc.7.8.1185

Sano, Y., Wada, M., Hashimoto, Y., Matsumoto, T., & Kojima, M. (1998). 'Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome', *The Journal of General Virology*, 79 (Pt 12): 3111–8. DOI: 10.1099/0022-1317-79-12-3111

Saunders, K., Lucy, A., & Stanley, J. (1992). 'RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus', *Nucleic Acids Research*, 20/23: 6311–5. DOI: 10.1093/nar/20.23.6311

Saunders, Keith, Richardson, J., Lawson, D. M., & Lomonossoff, 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. DOI: 10.3390/v12111235

Saunders, Keith, Salim, N., Mali, V. R., Malathi, V. G., Briddon, R., Markham, P. G., & Stanley, J. (2002). 'Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus', *Virology*, 293/1: 63–74. DOI: 10.1006/viro.2001.1251

Savory, F. R., & Ramakrishnan, U. (2014). 'Asymmetric patterns of reassortment and concerted evolution in Cardamom bushy dwarf virus', *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 24: 15–24. DOI: 10.1016/j.meegid.2014.02.012

Savory, Fiona R., Varma, V., & Ramakrishnan, U. (2014). 'Identifying geographic hot spots of reassortment in a multipartite plant virus', *Evolutionary Applications*, 7/5: 569–79. DOI: 10.1111/eva.12156

Sharman, M., Thomas, J. E., Skabo, S., & Holton, T. A. (2008). 'Abacá bunchy top virus, a new member of the genus Babuvirus (family Nanoviridae)', *Archives of Virology*, 153/1: 135–47. DOI: 10.1007/s00705-007-1077-z

Sicard, A., Pirolles, E., Gallet, R., Vernerey, M.-S., Yvon, M., Urbino, C., Peterschmitt, M., et al. (2019). 'A multicellular way of life for a multipartite virus', *eLife*, 8: e43599. DOI: 10.7554/eLife.43599

Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakis, Y., Gutierrez, S., & Blanc, S. (2013). 'Gene copy number is differentially regulated in a multipartite virus', *Nature Communications*, 4: 2248. DOI: 10.1038/ncomms3248

Sicard, A., Zeddam, J.-L., Yvon, M., Michalakis, Y., Gutiérrez, S., & Blanc, S. (2015). 'Circulative Nonpropagative Aphid Transmission of Nanoviruses: an Oversimplified View', *Journal of Virology*, 89/19: 9719–26. DOI: 10.1128/JVI.00780-15

Silva, F. N., Lima, A. T. M., Rocha, C. S., Castillo-Urquiza, G. P., Alves-Júnior, M., & Zerbini, F. M. (2014). 'Recombination and pseudorecombination driving the evolution of the begomoviruses Tomato severe rugose virus (ToSRV) and Tomato rugose mosaic virus (ToRMV): two recombinant DNA-A components sharing the same DNA-B', *Virology Journal*, 11: 66. DOI: 10.1186/1743-422X-11-66

Singh, A. K., Chattopadhyay, B., & Chakraborty, S. (2012). 'Biology and interactions of two distinct monopartite begomoviruses and betasatellites associated with radish leaf curl disease in India', *Virology Journal*, 9: 43. DOI: 10.1186/1743-422X-9-43

Stainton, D., Kraberger, S., Walters, M., Wiltshire, E. J., Rosario, K., Halafihi, M., Lolohea, S., et al. (2012). 'Evidence of inter-component recombination, intra-component recombination and reassortment in banana bunchy top virus', *The Journal of General Virology*, 93/Pt 5: 1103–19. DOI: 10.1099/vir.0.040337-0

Stainton, D., Martin, D. P., Collings, D. A., & Varsani, A. (2017). 'Comparative analysis of common regions found in babuviruses and alphasatellite molecules', *Archives of Virology*, 162/3: 849–55. DOI: 10.1007/s00705-016-3168-1

Stainton, D., Martin, D. P., Muhire, B. M., Lolohea, S., Halafihi, M., Lepoint, P., Blomme, G., et al. (2015). 'The global distribution of Banana bunchy top virus reveals little evidence for frequent recent, human-mediated long distance dispersal events', *Virus Evolution*, 1/1: vev009. DOI: 10.1093/ve/vev009

Sudarshana, M. R., Wang, H. L., Lucas, W. J., & Gilbertson, R. L. (1998). 'Dynamics of Bean Dwarf Mosaic Geminivirus Cell-to-Cell and Long-Distance Movement in Phaseolus vulgaris Revealed, Using the Green Fluorescent Protein', *Molecular Plant-Microbe Interactions®*, 11/4: 277–91. Scientific Societies. DOI: 10.1094/MPMI.1998.11.4.277

Sun, M., Jiang, K., Li, C., Du, J., Li, M., Ghanem, H., Wu, G., et al. (2020). 'Tobacco curly shoot virus C3 protein enhances viral replication and gene expression in Nicotiana benthamiana plants', *Virus Research*, 281: 197939. DOI: 10.1016/j.virusres.2020.197939

Sun, S., Ren, Y., Wang, D., Farooq, T., He, Z., Zhang, C., Li, S., et al. (2022). 'A group I WRKY transcription factor regulates mulberry mosaic dwarf-associated virus-triggered cell death in Nicotiana benthamiana', *Molecular Plant Pathology*, 23/2: 237–53. DOI: 10.1111/mpp.13156

Sun, X., Xiao, L., Cheng, C., Shi, Z., Sun, Z., Ma, Y., & Zhu, X. (2022). 'Genome-wide analysis of milk vetch dwarf virus and related alphasatellites from garlic and aphids in Shandong Province, China', *Archives of Virology*, 167/10: 2071–7. DOI: 10.1007/s00705-022-05490-y

Sunter, G., & Bisaro, D. M. (1992). 'Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription', *The Plant Cell*, 4/10: 1321–31. DOI: 10.1105/tpc.4.10.1321

Tentchev, D., Verdin, E., Marchal, C., Jacquet, M., Aguilar, J. M., & Moury, B. (2011). 'Evolution and structure of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes', *The Journal of General Virology*, 92/Pt 4: 961–73. DOI: 10.1099/vir.0.029082-0

Tijssen, P., & Bergoin*, M. (1995). 'Densonucleosis viruses constitute an increasingly diversified subfamily among the parvoviruses', *Seminars in Virology*, 6/5: 347–55. DOI: 10.1006/smvy.1995.0041

Timchenko, T., Katul, L., Aronson, M., Vega-Arreguín, J. C., Ramirez, B. C., Vetten, H. J., & Gronenborn, B. (2006). 'Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus', *The Journal of General Virology*, 87/Pt 6: 1735–43. DOI: 10.1099/vir.0.81753-0

Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H. J., & Gronenborn, B. (2000). 'The master rep concept in nanovirus replication: identification of missing genome components and potential for natural genetic reassortment', *Virology*, 274/1: 189–95. DOI: 10.1006/viro.2000.0439

Timchenko, T., de Kouchkovsky, F., Katul, L., David, C., Vetten, H. J., & Gronenborn, B. (1999). 'A single rep protein initiates replication of multiple genome components of faba bean necrotic yellows virus, a single-stranded DNA virus of plants', *Journal of Virology*, 73/12: 10173–82. DOI: 10.1128/JVI.73.12.10173-10182.1999

Trapani, S., Bhat, E. A., Yvon, M., Lai-Kee-Him, J., Hoh, F., Vernerey, M.-S., Pirolles, E., et al. (2023). 'Structure-guided mutagenesis of the capsid protein indicates that a nanovirus requires

assembled viral particles for systemic infection', *PLOS Pathogens*, 19/1: e1011086. DOI: 10.1371/journal.ppat.1011086

Varsani, A., Lefeuvre, P., Roumagnac, P., & Martin, D. (2018). 'Notes on recombination and reassortment in multipartite/segmented viruses', *Current Opinion in Virology*, 33: 156–66. DOI: 10.1016/j.coviro.2018.08.013

Vega-Arreguín, J. C., Gronenborn, B., & Ramírez, B. C. (2007). 'Arabidopsis thaliana is a host of the legume nanovirus Faba bean necrotic yellows virus', *Virus Research*, 128/1–2: 81–7. DOI: 10.1016/j.virusres.2007.04.011

Venkataraman, S., & Selvarajan, R. (2019). 'Recent advances in understanding the replication initiator protein of the ssDNA plant viruses of the family Nanoviridae', *Virusdisease*, 30/1: 22–31. DOI: 10.1007/s13337-019-00514-9

Venkataraman, S., Selvarajan, R., Subramanian, S. S., & Handanahalli, S. S. (2022). 'Insights into the capsid structure of banana bunchy top virus', *3 Biotech*, 12/7: 144. DOI: 10.1007/s13205-022-03204-4

Vetten, H. J., Knierim, D., Rakoski, M. S., Menzel, W., Maiss, E., Gronenborn, B., Winter, S., et al. (2019). 'Identification of a novel nanovirus in parsley', *Archives of Virology*, 164/7: 1883–7. DOI: 10.1007/s00705-019-04280-3

Villa, M., & Lässig, M. (2017). 'Fitness cost of reassortment in human influenza', *PLoS pathogens*, 13/11: e1006685. DOI: 10.1371/journal.ppat.1006685

Wang, L.-L., Wang, X.-R., Wei, X.-M., Huang, H., Wu, J.-X., Chen, X.-X., Liu, S.-S., et al. (2016). 'The autophagy pathway participates in resistance to tomato yellow leaf curl virus infection in whiteflies', *Autophagy*, 12/9: 1560–74. DOI: 10.1080/15548627.2016.1192749

Wang, X.-W., & Blanc, S. (2021). 'Insect Transmission of Plant Single-Stranded DNA Viruses', *Annual Review of Entomology*, 66: 389–405. DOI: 10.1146/annurev-ento-060920-094531

Wang, Y.-J., Yao, Q., Chen, K.-P., & Han, X. (2006). '[Organization and transcription strategy of genome of Bombyx mori bidensovirus (China isolate) VD1]', *Sheng Wu Gong Cheng Xue Bao = Chinese Journal of Biotechnology*, 22/5: 707–12. DOI: 10.1016/s1872-2075(06)60052-6

Wanitchakorn, R., Hafner, G. J., Harding, R. M., & Dale, J. L. (2000). 'Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6', *The Journal of General Virology*, 81/Pt 1: 299–306. DOI: 10.1099/0022-1317-81-1-299

Wanitchakorn, R., Harding, R. M., & Dale, J. L. (1997). 'Banana bunchy top virus DNA-3 encodes the viral coat protein', *Archives of Virology*, 142/8: 1673–80. DOI: 10.1007/s007050050188

Ward, B. M., & Lazarowitz, S. G. (1999). 'Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay', *The Plant Cell*, 11/7: 1267–76. DOI: 10.1105/tpc.11.7.1267

Wu, M., Wei, H., Tan, H., Pan, S., Liu, Q., Bejarano, E. R., & Lozano-Durán, R. (2021). 'Plant DNA polymerases α and δ mediate replication of geminiviruses', *Nature Communications*, 12/1: 2780. DOI: 10.1038/s41467-021-23013-2

Xavier, C. A. D., Godinho, M. T., Mar, T. B., Ferro, C. G., Sande, O. F. L., Silva, J. C., Ramos-Sobrinho, R., et al. (2021). 'Evolutionary dynamics of bipartite begomoviruses revealed by complete genome analysis', *Molecular Ecology*, 30/15: 3747–67. DOI: 10.1111/mec.15997

Xu, X., Zhang, Q., Hong, J., Li, Z., Zhang, X., & Zhou, X. (2019). 'Cryo-EM Structure of a Begomovirus Geminate Particle', *International Journal of Molecular Sciences*, 20/7: 1738. DOI: 10.3390/ijms20071738

Yu, N.-T., Xie, H.-M., Zhang, Y.-L., Wang, J.-H., Xiong, Z., & Liu, Z.-X. (2019). 'Independent modulation of individual genomic component transcription and a cis-acting element related to high transcriptional activity in a multipartite DNA virus', *BMC genomics*, 20/1: 573. DOI: 10.1186/s12864-019-5901-0

Zhang, S. C., Ghosh, R., & Jeske, H. (2002). 'Subcellular targeting domains of Abutilon mosaic geminivirus movement protein BC1', *Archives of Virology*, 147/12: 2349–63. DOI: 10.1007/s00705-002-0880-9

Zhang, T., Xu, X., Huang, C., Qian, Y., Li, Z., & Zhou, X. (2016). 'A Novel DNA Motif Contributes to Selective Replication of a Geminivirus-Associated Betasatellite by a Helper Virus-Encoded Replication-Related Protein', *Journal of Virology*, 90/4: 2077–89. DOI: 10.1128/JVI.02290-15

Zhao, L., Rosario, K., Breitbart, M., & Duffy, S. (2019). 'Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range', *Advances in Virus Research*, 103: 71–133. DOI: 10.1016/bs.aivir.2018.10.001

Zhou, Y.-C., Garrido-Ramirez, E. R., Sudarshana, M. R., Yendluri, S., & Gilbertson, R. L. (2007). 'The N-terminus of the Begomovirus nuclear shuttle protein (BV1) determines virulence or avirulence in Phaseolus vulgaris', *Molecular plant-microbe interactions: MPMI*, 20/12: 1523–34. DOI: 10.1094/MPMI-20-12-1523

Zhuang, J., Lin, W., Coates, C. J., Shang, P., Wei, T., Wu, Z., & Xie, L. (2019). 'Cleavage of the Babuvirus Movement Protein B4 into Functional Peptides Capable of Host Factor Conjugation is Required for Virulence', *Virologica Sinica*, 34/3: 295–305. DOI: 10.1007/s12250-019-00094-4

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*, 2/1: 161–79. DOI: 10.1146/annurev-virology-100114-055135

SPECIES	ISOLATES	GRIGORAS 2014	KRABERGER 2017
FBNSV	12	6	1
FBNYV	20	13	16
MDV	9	1	0
FBYLV	1	1	0
MVCDV	2	0	0
BMLRV	3	3	0
CvLV	1	0	0
PYSV	1	1	0
PNYDV	11	2	0
SYSaV	2	0	0
PSSaV	3	0	0
SCSV	2	2*	0

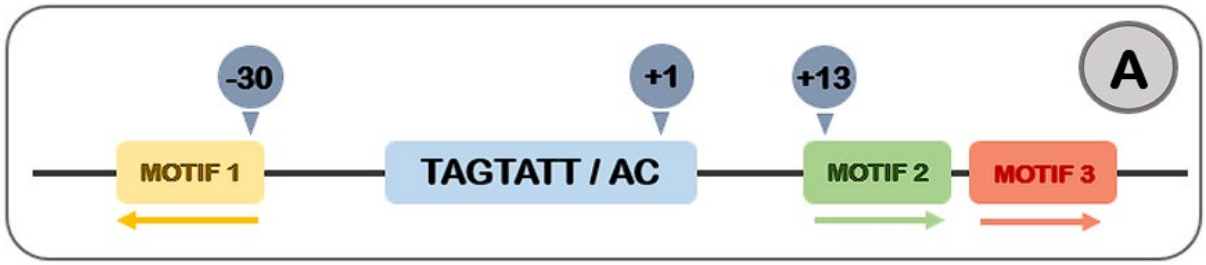
Table 1 : Summary table comparing the number of available complete nanovirus isolates (with all segments) in 2021 (Isolates column) with previous nanovirus genomic studies. Colored numbers in black (full), orange (partial), and red (none), indicate how they compare to the currently available data. The asterisk (*) associated with the SCSV genomes compared in Grigoras et al. 2014, indicates that this study involved two SCSV isolates with only 6 segments, which are not integrated into the complete genome of the previous "Isolates" column.

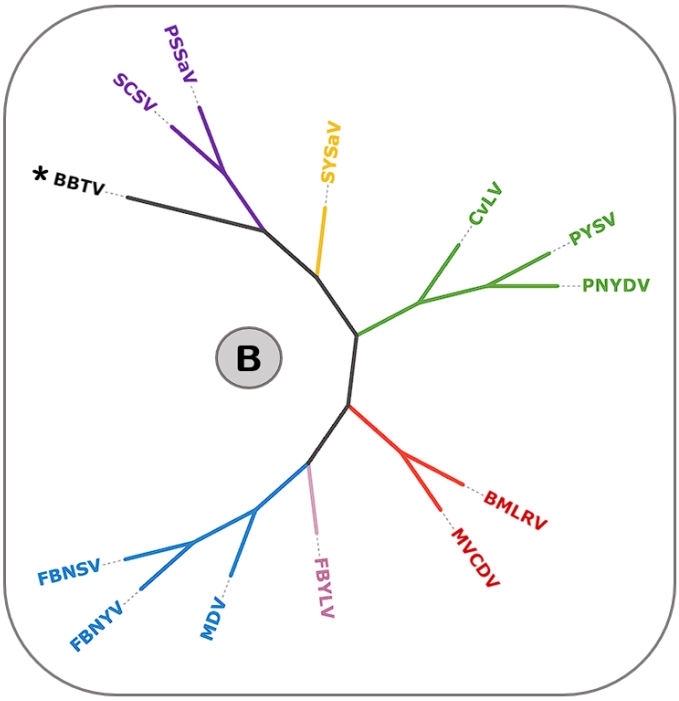
4. Genomics

As mentioned in the previous review article (**Torralba, Blanc and Michalakis 2023**), the definition and characterization of nanovirus trans-replication specificity determinants (SPDs), primarily composed of iterons, lag behind to the state of the art in babuviruses. In this regard, I have initiated a genomic study of the *Nanovirus* genus with the aim of identifying potential candidate sequences for both iterons and the Iteron Recognition Domain (IRD) within the M-Rep protein, similar to the approach undertaken for begomoviruses (**Argüello-Astorga and Ruiz-Medrano 2001; Londoño, Riego-Ruiz and Argüello-Astorga 2010; Avalos-Calleros et al. 2021**). Additionally, given the continuous discovery of new nanovirus species and isolates, it has become pertinent to update prior genomic studies of the whole genus (**Grigoras et al. 2014; Kraberger et al. 2017**) (**Table 1**).

This endeavor not only allows us to detect new reassortment events but may also provide valuable insights on the compatibility of M-Rep and iterons among different species and isolates. We compiled all available genetic information for the *Nanovirus* genus from GenBank and conducted a comprehensive alignment comparative analysis. This comprehensive dataset offered us the opportunity to explore their genetic diversity (**Figure 15**) and to identify short, repetitive sequences, typically 5 to 6 nucleotides in length, suggesting iterons.

Alignment of the non-coding regions of each segment from the nanovirus complete genome available has allowed for the identification of several highly conserved positions, particularly in the vicinity of the nonanucleotide conserved region, acting as the origin of replication (**TAGTATT/AC**). Notably, positions **+13** and **-30** from the cleavage site exhibit a very high degree of conservation among all nanovirus sequences, featuring a G at the first position of a variable hexanucleotide motif that could be shared among both closely and distantly related species (**Figure 15**). The arrangement of these motifs closely resembles what has been observed in babuvirus iterons (**Herrera-Valencia et al. 2006; Stainton et al. 2016**).





SPECIES	MOTIF 1	MOTIF 2	MOTIF 3
FBNSV	GGATCA	GGATCA GGTTCA	GGGTCA GAGTCA
FBNYV	GAGTCA GGATCA	GGTTCA	GAGTCA
MDV	GGATCA GAGTCA	GGATCA GGTTCA	GAGTCA
FBYLV	GGATCA	GGTTCA GGCTCA	GGATCA
MVCDV	GGATCA	GGATCA	GAGTCA
BMLRV	GGATCA GAGTCA	GGATCG GGATCA	GAGTCA
CvLV	GGATCA	GGATCA GGATCG	GAGTCA
PYSV	GGATCA	GGATCA	GAGCGC GAGTCA
PNYDV	GGATCA	GGATCA	GGATCA GAGTCA
SYSaV	GATACA GGAACA	GGACGG GGTACC	GATACA GGTACA GGAACA
PSSaV	GCTACA GGATCA	GGCGCT GGTTCA	GCTACA GGATCA
SCSV	AGACAT	GGATCA GGGTCA	AGACAT

Figure 15: [A] Illustration of the genomic organization of the identified iteron-like motifs around the origin of replication sequence, TAGTATT/AC. The line represents the DNA sequence. Following the M-Rep cleavage site, C is considered the first nucleotide position +1. The numbers +13 and -30 indicate the position relative to the +1 of the motifs 2 and 1 respectively. Arrows indicate the reading sense of the motif sequences. The space between motifs 1 and 2 illustrates the optional presence and varying length of a spacer sequence. **[B] Nanovirus phylogenetic tree based on updated genomic data (2021) and concatenated genomes (C-M-N-R-S-U1-U2-U4 / U3)**. **Nanovirus phylogenetic tree based on updated genomic data (2021) and concatenated genomes (C-M-N-R-S-U1-U2-U4 / U3)**. (*) represent the BBTV Babuvirus [Malawi73] isolate used as outgroup. The following nanovirus isolates were used as reference for BMLRV [AZ;47], CvLV [Sambuc_2010], FBNSV [Et; Hol-1997], FBNYV [Es; Mu29D]; FBYLV [Eth;231]; MDV [VF]; MVCDV [G53]; PNDYV [DE;15]; PSSaV [DE; Pa21]; PYSV [AT;15]; SCSV [MyallVale 2534B] and SYSaV [Har_H13_Soph17] respectively. Branch lengths are not proportionate to real genetic distances and no bootstrap was done for this illustrative phylogenetic tree. **Table displaying candidate motif sequences for the entire *Nanovirus* genus**. The table illustrates the variability of motifs across all isolates from a given species, without displaying specific associations or the presence and nature of spacer sequences. Matching motifs are indicated by the same color.

We have initiated the compilation of motifs identified for each segment and isolate. These insights could be instrumental in our efforts to predict or understand the viability and severity of certain observed intra- or inter-specific reassortant phenotypes, particularly if there are discernible patterns of compatibility associated with these motifs. It is worth noting different isolates from the same species can have divergent iterons and that genomic segments from a given isolate may not necessarily share identical iterons or spacer sequences from one another as it was indicated in (**Torralba, Blanc and Michalakis 2023**).

Additionally, I employed the MEME Suite software (**Bailey et al. 2015**) to identify motifs that were previously suspected, although these results are in a very preliminary stage. The preliminary findings confirm the detection of the three candidate motifs, and they also reveal the presence of two additional motifs that warrant further investigation.

Furthermore, applying the RDP5 software (**Martin et al. 2020**) to our recently updated genome database will allow us to detect new reassortment and recombination events. Subsequently, we will assess whether the previously recognized discrepancies related to inter-specific and multi-segment reassortments predicted from lab experiment versus those observed in the field persist. Additionally, this analysis will provide us with a clearer understanding regarding which specific genomic segments experience a higher degree of reassortment.

Ultimately, having access to comprehensive information about reassortment events and the SPDs of the corresponding parental genotypes could offer valuable insights into the compatibility between SPDs and M-Rep IR and may enhance our understanding of phenotypic characteristics displayed by experimental or emerging natural reassortants such as trans-replication efficiency.

CHAPTER 2 :

PHENOTYPIC CHARACTERIZATION OF THREE FBNSV ISOLATES

1. Context

To investigate intra-specific single-segment reassortments between two FBNSV isolates, we made the deliberate choice to create two new infectious clones based on specific criteria. Our objective was to choose isolates that were co-circulating in the same geographical area, meaning that the resulting reassortants might plausibly emerge under natural conditions. Additionally, we opted for isolates from distinct phylogenetic clades so that we could discriminate them, showcasing indirectly a degree of genetic diversity. Finally, we selected isolates identified from different field hosts, enabling us to explore putative host range adaptation, as reassortments have been documented to be associated with host shifts and host range expansion.

The genomic investigation in **Chapter 1** provided us with an updated phylogenetic structure of the FBNSV species. As previously observed (**Grigoras et al. 2014**), FBNSV is organized into three primary phylogenetic clades that we arbitrarily designated as "Red," "Blue," and "Green." An FBNSV infectious clone, **[JKI-2000]**, which was developed in 2009, was already available at the start of my PhD (**Grigoras et al. 2009**). This clone was instrumental in identifying and defining the genome formula (**Sicard et al. 2013; Sicard et al. 2015; Gallet et al. 2022**) and demonstrating supra-cellular functional complementation (**Sicard et al. 2019**), as well as the virus ability to reconstitute its genome through non-concomitant transmission (**Di Mattia et al. 2022**).

The **[JKI-2000]** infectious clone originated from an Ethiopian isolate, **[ET;Holetta-1997]**, which belonged to the "Blue" clade. This isolate was collected from faba beans in 1997 and subsequently maintained on faba beans through a series of aphid inoculations using the aphid vector *Acyrthosiphon pisum*. This isolate was lost in 2004, and as a result, the **[JKI-2000]** infectious clone was generated using frozen faba bean samples preserved since 2000. While the original isolate was initially collected from a faba bean field, it may have further evolved/specialized for three years in the lab with successive transmissions on faba bean cv. Sirocco and Condor and through *A. pisum* (Harris) transmission (**Grigoras et al. 2009**).

In this regard and because no other FBNSV isolate was reported from Ethiopia, we chose to focus on several Azeri isolates that had been previously identified in a nanovirus genomic study (**Grigoras et al. 2014**). These isolates were categorized into either the "Red" or "Green" clade. Notably, these isolates exhibited frequent reassortments between the two clades (**Grigoras et al. 2014**), a topic that will be elaborated further when justifying our choice of isolates in the subsequent article. Therefore, we specifically chose **[AZ;15]** ("Red") isolated from *Vicia sativa*

(vetch), and a derivative (further explained in the corresponding article) of **[AZ;10]** ("Green") isolated from *Lens culinaris* (lentil) for the production of the **[AZ;15]** and **[AZ;10_12b]** infectious clones.

The **[JKI-2000]** infectious clone was created through the insertion of head-to-tail tandem repeats of the corresponding genomic segments into pBin19 binary plasmids (**Grigoras et al. 2009**). This cloning strategy differs from our "bitmer" approach (**Urbino et al. 2008; Yepes et al. 2018**), where each segment is cloned as a partial tandem repeat, encompassing the whole segment so that the conserved region CR-SL associated with the origin of replication is repeated twice, once at each end of the inserted sequence. This allows the release of the segment by the M-Rep protein and the initiation of the rolling-circle replication (**Bonnamy, Blanc and Michalakis 2023**). The bitmers were synthesized, cloned into the pCAMBIA2300 binary plasmid, and subsequently confirmed through Sanger sequencing by Agate Bioservices company (Mialet, France). The FBNSV **[AZ;15]** and **[AZ;10_12b]** infectious clones consist of two sets of eight plasmids, each containing a genomic segment bitmer.

Upon receiving the infectious clones, I assessed their viability by introducing them into three distinct strains of *Agrobacterium tumefaciens*, namely COR308, C58C1, and LBA4404. This was done in an effort to enhance the variable and somewhat unreliable infection rates we had previously encountered using our agro-inoculation method involving stem-piercing. The outcomes of these experiments, when applied to faba beans, confirmed the viability of the infectious clones, which resulted in the development of typical FBNSV symptoms (leaf curling, dwarfing, chlorosis). These experiments also led to the selection of STR COR308 for subsequent use, based on an apparent more efficient agro-inoculation success, and the modification of the agro-inoculation method, from stem-piercing to infiltration of very young immature leaves, thanks to the efforts of a team colleague (results detailed in the annex).

While we did observe symptomatic plants, there was uncertainty regarding whether all genomic segments were indeed effectively replicated and maintained during the infection. We addressed this concern using the created pairs of primers specific to each segment of each isolate used to discriminate them. These primers were tested in cross-amplification experiments, using both pCAMBIA-segment constructs of **[AZ;15]** and **[AZ;10_12b]**, as well as infected plant material from **[JKI-2000]**, **[AZ;15]**, and **[AZ;10_12b]** and were able to specifically discriminate segments of **[AZ;15]** and **[AZ;10_12b]** only. The resulting primers and some amplification profiles are provided in the **Annex section**.

Before starting the systematic phenotypic characterization of single-segment reassortants between **[AZ;15]** and **[AZ;10_12b]** on their respective field hosts (**Chapter 3**), *Vicia sativa* (vetch) and *Lens culinaris* (lentil), we first laid out a detailed characterization plan for these two forthcoming parental genotypes, along with the existing **[JKI-2000]** isolate, in this **Chapter 2**.

2. General results

One of our primary findings underscored the ability of all genotypes to efficiently infect all the assessed hosts, including faba beans, lentils, and vetches, and to be transmitted both to and from these three plant host species. This observation has significant implications for the ecological and evolutionary dynamics of nanoviruses that seem to share broad host and aphid vector ranges. It indicates that these viruses possess a considerable potential for frequent mixed infections and the resulting genetic exchanges, including both recombinations and, more prominently, reassortments, which can occur on an even larger spatial scale as previously discussed (**Di Mattia et al. 2022; Torralba, Blanc and Michalakis 2023 – Chapter 1**).

Interestingly, we witnessed a range of phenotypes associated with the three genotypes depending on the host plant species, providing us with a first documentation of the phenotypic diversity of this nanovirus species. In particular, we observed substantial differences in both plant and aphid genome formulas among all genotypes establishing for the first time that the genome formula exhibits intraspecific variation. This underscores that the genome formula reacts differentially depending on the isolate, the plant host, and the aphid vector.

By using plant viral load and transmission rates as proxies for viral fitness, the **[AZ;10_12b]** isolate emerged as the better fit genotype in our experimental setting, exhibiting the highest transmission rate across all hosts. In contrast, the **[JKI-2000]** isolate was found to be the less fit genotype, displaying notably low viral load and a reduced transmission rate in comparison to the other two. We did not observe a correlation between the severity of symptoms and our viral fitness assessment. For instance, **[JKI-2000]** exhibited the most severe symptoms on faba beans despite having a very low viral load and a lower transmission rate, while **[AZ;15]** was less severe on vetches and less fit than **[AZ;10_12b]**. Consequently, we could not identify adaptive properties associated with the field hosts and their respective isolates in terms of fitness nor symptom severity, as the hierarchy of fitnesses among the three isolates remained consistent across all plant host species.

Furthermore, we observed sporadic segment loss in all isolates across multiple host plant species, with a higher frequency of such loss occurring primarily in faba beans, following agro-

inoculation or aphid inoculation. While we acknowledge that our CT thresholds could potentially yield false positives or negatives, this observation, especially following aphid transmission, though very infrequent, suggests that the loss of genomic segments in nanoviruses might be a plausible scenario in natural settings, although probably selected against at multiple scales, rather than solely an experimental artifact (**Knierim et al. 2019**).

3. Article status

This experiment was lengthy and was immediately succeeded by the even more extensive single-segment reassortment phenotypic characterization (**Chapter 3**). Therefore, as the article is still in the preparation phase, the results are not yet fully analyzed and interpreted. Consequently, in the subsequent articles, the data are only partially exploited and extensive analysis and conclusions will be comprehensively explored in the near future.

Notably, delving deeper into the statistical analysis of the correlations between biological traits and genome formula may help answer the intriguing question of whether the genome formula represents an optimized adaptive state. In this scenario, the signature of the genomic segment relative frequencies would be associated with an optimized adaptive state of the viral genotype. Deviating from this established formula might potentially incur costs to the viral infection that is probably detectable with our experimental setup.

4. Article in preparation, referenced as (*Torralba et al. 2024a*)

Phenotypic characterization of three isolates of the faba bean necrotic stunt virus (FBNSV) in multiple host species and common vector

Authors

Babil Torralba[1,2], Andy Brousse[1,2], Prune Lacôte-Popovic[1,2], Stéphane Blanc[2] and Yannis Michalakis[1]

Affilations

[1] *MIVEGEC, CNRS, IRD, Univ Montpellier, Montpellier, France*

[2] *PHIM, Univ Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France*

Authors contribution to data collection and analyzes

Andy Brousse was fully involved in all aspects of the laboratory experiments, while Prune Lacôte-Popovic had a more limited participation. I led the majority of the experimental work, conducted the data statistical analysis and comparative genomic study.

ABSTRACT

As emerging worldwide phytopathogens, the phenotypic diversity of ssDNA multipartite nanoviruses has not been well established. Because only few infectious clones are available for experimental investigations, the data regarding nanovirus biology are restricted. In this study, we present the production of two newly engineered infectious clones of the *faba bean necrotic stunt virus* (FBNSV), belonging to the *Nanovirus* genus, specifically the **[AZ;15]** and **[AZ;10_12b]** sympatric Azeri isolates, which were synthesized from sequences accessible in open-source existing database. We detail their phenotypic characteristics, in addition to the already existing FBNSV **[JKI-2000]** isolate. The three isolates were subjected to phenotypic characterization on their respective field hosts, namely *Vicia sativa* (vetches), *Lens culinaris* (lentils), and *Vicia faba* (faba beans), and in a common aphid vector, *Aphis craccivora*. We quantified various biological traits, including height, weight, number of branches, foliar levels, symptom severity, symptom progression over time, viral accumulation in both the host plants and vector, identification of missing genomic segments, determination of genome formulas, and assessment of aphid transmission efficiency. This emphasizes that each isolate possesses its own distinctive genome formula that varies specifically when interacting with both hosts and vectors. The availability of these two novel infectious clones of nanoviruses will greatly broaden the range of tools for further research on the biology, ecology and evolution of these intriguing pathogens.

INTRODUCTION

In recent years, nanovirids have become an emerging worldwide threat as member species of this family were recognized to infect a large variety of crops of prime importance such as legumes, banana and cardamom (**Lal et al. 2020**). The International Committee on Taxonomy of Viruses (ICTV) divides the family *Nanoviridae* in two genera, *Babuvirus* and *Nanovirus*. The diversity appears quite low in the *Babuvirus* genus, with only three reported species: *Cardamom bushy dwarf virus* (CBDV) (**Mandal et al. 2013**), *Abaca bunchy top virus* (ABTV) (**Sharman et al. 2008**) and *Banana bunchy top virus* (BBTV) (**Stainton et al. 2015**) respectively infecting *Zingiberaceae* and *Musaceae* plant families (**Lal et al. 2020**) mostly in Asia. The *Nanovirus* genus shows a higher diversity and currently comprises 12 species (**Heydarnejad et al. 2017; Gallet et al. 2018; Vetten et al. 2019; Hassan-Sheikhi et al. 2020; Lal et al. 2020**). Most infect plants of the *Fabaceae* family with an increasing number of viral species being described in legumes (**Grigoras et al. 2014; Gallet et al. 2018**), while others are being

described from new botanical families such *Brassicaceae*, *Solanaceae*, *Amaryllidaceae*, and *Apiaceae* (**Vega-Arreguín, Gronenborn and Ramírez 2007; Kamran et al. 2019; Vetten et al. 2019; Sun et al. 2022**). Nanoviruses outbreaks have been reported in Africa (**Tadesse et al. 1999**), the Middle-East (**Kumari et al. 2004**), Asia (**Zhang et al. 2017**), Oceania (**Chu et al. 1993**), as well as in Europe (**Babin et al. 2000**), where numerous species and isolates have recently been discovered in Spain, Austria, Sweden, Azerbaijan (**Grigoras et al. 2014**), Germany (**Grigoras, Gronenborn and Vetten 2010; Grigoras et al. 2014**), and France (**Gallet et al. 2018**). Intriguingly, no nanovirus has been detected thus far in the American continent (**Lal et al. 2020**).

Nanoviruses are multipartite viruses with a genome consisting of eight circular closed ssDNA segments of approximately 1kb, each encoding a single gene and each encapsidated individually (**Gronenborn 2004; Lal et al. 2020**). Segment C encodes the cell cycle-linked protein (Clink) (**Lageix et al. 2007**), M encodes the movement protein (MP) (**Krenz et al. 2017**), N encodes the nuclear shuttle protein (NSP) mandatory for aphid transmission (**Grigoras et al. 2018**), S encodes the coat protein (CP) (**Trapani et al. 2023**), R encodes the master replication initiator protein (M-Rep) (**Timchenko et al. 2000**), and U1, U2, and U4 encode proteins of unknown functions (**Krenz et al. 2017**).

An intriguing aspect of the biology of nanoviruses (**Sicard et al. 2013; Sicard et al. 2015; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022**), and more generally of multipartite viruses (**Sánchez-Navarro, Zwart and Elena 2013; Hu et al. 2016**), is the "genome formula": the genomic segments accumulate very reproducibly at a specific frequency distribution within viral populations (**Sicard et al. 2013; Sánchez-Navarro, Zwart and Elena, 2013; Sicard et al. 2015; Hu et al. 2016; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022; Gallet et al. 2022**) which is host dependent (**Sicard et al. 2013; Wu et al. 2017; Sicard et al. 2015; Gallet et al. 2022**) and whose variations tune gene expression in changing environments (**Gutiérrez and Zwart 2018; Zwart and Elena 2020; Gallet et al. 2022**).

The genome formula of distinct nanovirus species is different (**Sicard et al. 2013; Mansourpour et al. 2022**), but possible variations within species, in between isolates, have not been evaluated. The first isolate of the species *Faba bean necrotic stunt virus* (FBNSV) was discovered on *Vicia faba* (faba beans) in Ethiopia (**Franz et al. 1997**) and cloned as the isolate **[JKI-2000]** (**Grigoras et al. 2009**). Until now, despite an increasing number of sequences of FBNSV isolates in the databases (**Abraham et al. 2010; Grigoras et al. 2014; Loftipour et al. 2020**), **[JKI-2000]** remains the only FBSNV isolate for which an infectious clone has been produced. Consequently, while there is available data on the genetic diversity of FBNSV, there

is a lack of formal characterization of its corresponding phenotypic diversity. Different isolates may exhibit distinct host and/or insect vector ranges, genome formula and pathogenicity. Further study in this direction is required to improve our understanding of factors that contribute to viral evolution, adaptation and putative emergence of these viruses.

Here we report the production of two synthetic clones based on the complete sequences of FBNSV genotypes available in databases. These genotypes named **[AZ;15]** and **[AZ;10_12b]** were sampled in the same geographical area but belong to distinct clades (**Grigoras et al. 2014**). Details on these isolates, including the rationale for choosing them, are provided in the Results section. We quantified several biological traits: symptom severity, symptom kinetics, viral accumulation, genome formula and aphid transmission, first in faba beans to allow direct comparison with the original **[JKI-2000]** isolate. We further evaluated the phenotypic traits of the two novel isolates in the legume field hosts from which they were isolated: *Vicia sativa* (vetches) and *Lens culinaris* (lentils). To be able to characterize completely and compare the reaction norm of the three available FBNSV isolates, we also evaluated the original **[JKI-2000]** on lentils and vetches. These results, by establishing the across hosts reaction norm of important viral traits, inform on the properties of distinct isolates circulating in sympatry and potentially sharing hosts and vectors. In particular, they provide the first demonstration of intraspecific variations of the genome formula of a multipartite virus.

METHODS

1. Overview of the experiment

The characterization and comparison of the FBNSV infectious clones was designed as follows (**Figure 1**). First, faba beans (FB) were agro-inoculated with either one of the infectious clones of the three isolates and several phenotypic traits were estimated: height, number of branches, number of leaf levels, fresh weight, dry weight, symptom severity and kinetics, viral load, and genome formula. The properties of the three isolates were compared in this single host species, faba bean (original field host of **[JKI-2000]**). Then, the infected faba beans were used as source plants for acquisition of the virus by aphid vectors followed by transmission to recipient faba beans, lentils or vetches. The transmission from faba bean to faba bean was used to evaluate the transmission rate FB>FB. In contrast, the transmission of **[JKI-2000]**, **[AZ;10_12b]** and **[AZ;15]** to lentils and vetches was used to generate infected plants for completing the characterization of the three isolate phenotypes in the other two field hosts due to our inability to efficiently agroinoculate any of the infectious clones directly into lentils (L; original field host of **[AZ;10_12b]**) and vetches (V; original field host of **[AZ;15]**). Finally, infected lentils and vetches were used as source plants for further acquisition and transmission by aphids to lentils and vetches respectively to establish the L>L and V>V transmission rate for each of the three isolates. In sum, this experiment characterizes and compares the phenotypic properties of each of the three isolates in each of their respective field host species.

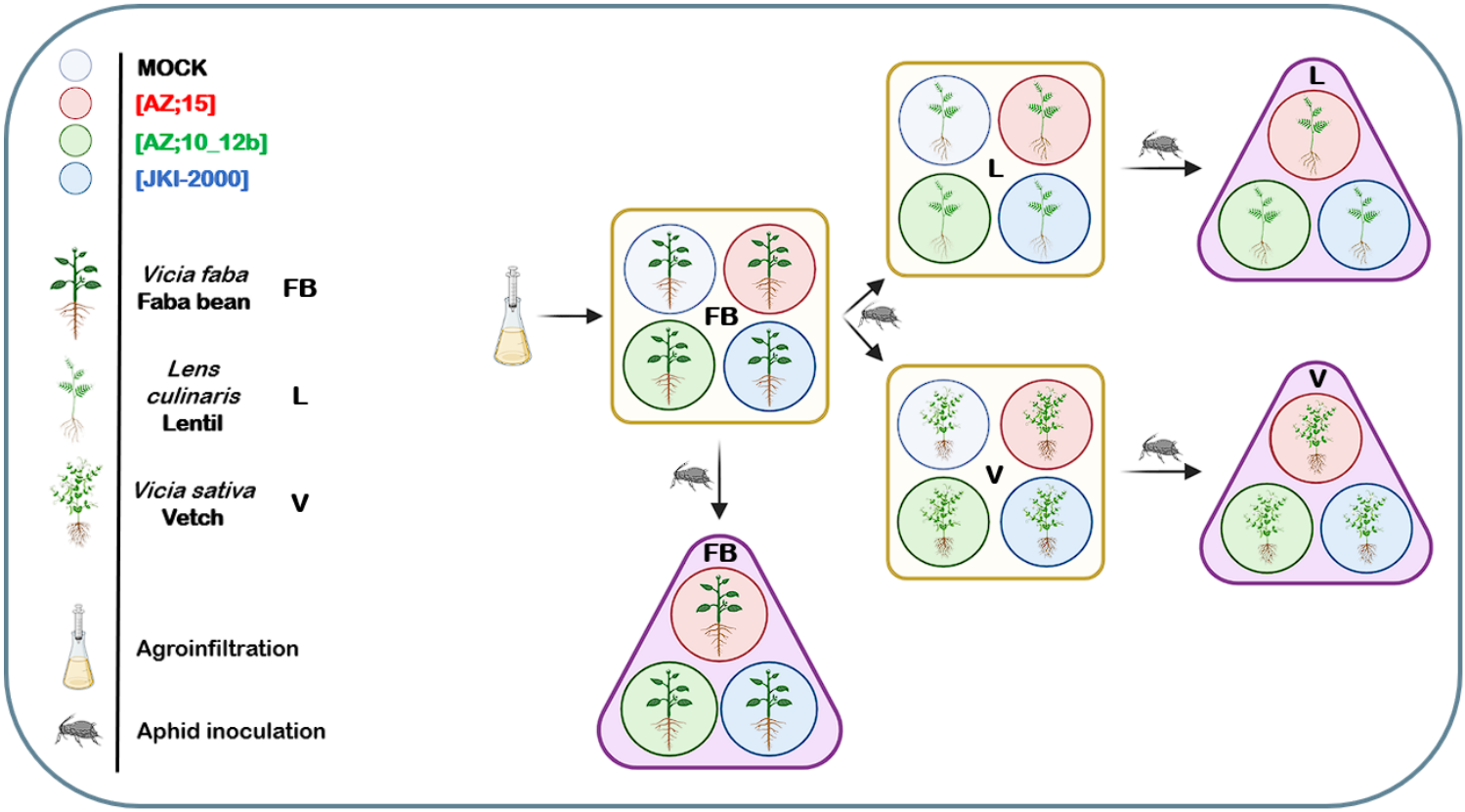


Figure 1 : Overview of the phenotypic characterization of the three FBNSV isolates. Sequential steps are indicated by the black arrows. Yellow square represents the phenotypical characterization step, whereas purple triangle represents the transmission assay where infected plants were determined based on symptoms alone and qPCR for ambiguous ones.

2. FBNSV isolate phylogenetic analysis

All the genomic sequences of FBNSV isolates used in this study were collected from GenBank NCBI and are listed in (**Supplementary Table 1**). All protein sequences were determined using Geneious 8.0 (**Kearse et al. 2012**). Complete genomes were assembled by concatenating all eight genomic segments as C-M-N-R-S-U1-U2-U4. Alignments were performed using ClustalW. Maximum-likelihood phylogenetic trees were inferred using PhyML 3.0 (**Guindon et al. 2010; Lefort, Longueville, Gascuel 2017**). Branches with <60 % bootstrap support were collapsed with iToL (**Letunic and Bork 2021**). Pairwise identity matrices of amino-acid sequences were determined using SDT v1.2 (**Muhire, Varsani and Martin 2014**).

3. Infectious clone design

Each genomic segment of the earlier reported **[JKI-2000]** infectious clone is inserted as a head-to-tail dimer into the binary plasmid pBin19. Eight plasmids, each containing a dimer of a segment, together constitute the FBNSV **[JKI-2000]** infectious clone (**Grigoras et al. 2009**).

[AZ;15] and **[AZ;10_12b]** infectious clones were designed as synthetic bitmers (**Urbino et al. 2008; Yepes et al. 2018**): each segment is cloned as a partial tandem repeat, encompassing the whole segment so that the conserved region CR-SL associated with the origin of replication is repeated twice, one at each end of the inserted sequence. This allows the release of the segment by the M-Rep protein and the initiation of rolling circle replication (**Bonnamy, Blanc and Michalakis 2023**). The bitmers were synthesized, cloned into the binary plasmid pCAMBIA2300, and verified by Sangers sequencing on both strands by Agate Bioservices company (Mialet, France). Two batches of eight plasmids, each containing a genomic segment bitmer, respectively constitute the FBNSV **[AZ;15]** and **[AZ;10/12b]** infectious clones. All plasmids were individually transformed by electroporation in *Agrobacterium tumefaciens* STR COR308 for agro-inoculation as later described.

4. Plants and growth conditions

The three legume host plants used in this study are *Vicia faba* (faba bean; var. Sevilla, Vilmorin, France), *Vicia sativa* (vetch; var. « José », Baumaux, France) and *Lens culinaris* (lentil; var. « Coralie », Baumaux, France). Faba beans were seeded directly in N2 soil medium (NEUHAUS, Humin-Substrat N2) in a greenhouse under a 13.5/10.5 day/night photoperiod at 25/18°C temperature and 70% hygrometry whereas lentils and vetches were seeded in the same medium in a growth chamber under a 13/11 day/night photoperiod at a temperature of 25/18°C day/night and 70% hygrometry for better germination. Once germinated, plantlets of faba beans, lentils and vetches were transplanted in individual pots (7x7x7 cm) and all moved to the greenhouse. After inoculation (see below) and for the rest of the growing period, lentils and vetches were bounded in Aracones (Arasystem), whereas faba beans were tutored when 14 days old.

5. Agroinoculation and aphid-inoculation

Agroinoculation concerns only faba beans as we never succeeded agroinoculating lentils or vetches, neither in leaflets nor in stems. Each *A. tumefaciens* colony containing one cloned segment was maintained at -80°C as preculture aliquots of 40 µL, DO=2.0. One aliquot was used to initiate an overnight culture in a 500mL deflected Erlenmeyer flask containing 50 mL of supplemented NZY medium (0.1% NZ amine, 0.5% yeast extract, 0.5% NaCl, 12.5 mM MgCl2, 12.5 mM MgSO4 and 0.4% glucose at pH 7.5) supplemented with kanamycine 50 $\mu g.mL^{-1}$, tetracycline 5 $\mu g.mL^{-1}$, gentamycin 25 $\mu g.mL^{-1}$, 200 µM acetosyringone and 10 mM MES pH 5.5

(final concentrations). Each bacterial culture was grown at 28°C/150 RPM-agitation, and stopped and adjusted at an OD_{600} = 2.0. All cultures were then centrifugated at 18°C/1000*g for 30 minutes, the supernatant was discarded and the bacterial pellet was recovered in 5 mL of a supplemented Murashige-Skoog (MS) medium solution (for 50 mL: 0.1 g MS (M5524, Sigma-Aldrich.), 800 µM acetosyringone, MES 10 mM). For each FBNSV isolate, a mixture of 8 x 5mL of each bacterial clone (one per segment) was adjusted to 50 mL and incubated at room temperature for 1.5 hour before infiltration in young expanding leaves of 9 days-old seedlings. The young leaves were punctured prior to infiltration and 0.5 mL of the bacteria mixture was used for each plant.

For all transmission experiments we used individuals of the aphid species *Aphis craccivora* from a clonal population established and maintained in a controlled environment as previously described (**Ryckebusch et al. 2021; Di Mattia et al. 2020**). Infected faba beans used for the FB>L and FB>V transmission were used as plant source at, lentils and vetches were used as source plants at 21 dpi whereas faba beans used for the FB>FB transmission were used as plant source at 26 dpi for aphid inoculation. Infected lentils and vetches were used as plant source at 20 dpi for aphid inoculation. Aphids were allowed an acquisition access period (AAP) of to 4 (F>L/V, L>L and V>V) to 5 (F>F) days on infected plants and then mixed and transferred randomly onto recipient plants for an inoculation access period (IAP) of 3 days. Two aphids were used to inoculate F>F and F>L/V whereas only one aphid was used to inoculate L>L and V>V per recipient plants. The recipient plants were 7-, 6-, and 5-days old plantlets for faba bean, lentil and vetch, respectively. In all cases, leaflets were harvested from the source plant, extracted and checked by qPCR for the presence of all eight segments prior to the AAP period since only plants containing all segments were used as donors for aphid transmission. To assess the viral load and genome formula within the insect vector, aphids were collected after the AAP (FB) or IAP (L and V) for qPCR analysis. Otherwise, they were killed by 4 successive sprays of ERADICOAT (Certis) (FB>L or V) or by one spray of 0.2% PIRIMOR G (Syngenta) (FB>FB, L>L and V>V) solution after the aphid IAP.

6. Total DNA extraction from plants and insects

Total DNA extraction from faba beans (*Vicia faba*) was performed exactly as earlier described (**Di Mattia et al. 2020**). All plants were sampled at 30 and/or 31 dpi due to logistics on their upper expanded leaf level. Total DNA extraction from vetches and lentils was performed following Edwards' method (**Edwards, Johnstone and Thompson 1991**). The leaflets are too

small on these plants to collect leaf discs as for faba beans. We thus collected a piece of leaf including several leaflets from the upper expanded leaf-level. Leaf samples were frozen at -80°C for at least 1 hour before being extracted. Four hundred µL of extraction buffer (200 mM Tris-HCl ; 250 nM NaCl ; 25 mM EDTA ; 0.5% SDS ; 1% polyvinylpyrrolidone 40 (PVP40)) were added to each sample which was then vortexed for 20 seconds before centrifugation for 2 minutes at 5000*g at room temperature. Then, 300 µL of supernatant were recovered and mixed with 300 µL of isopropanol before incubation for 30 minutes on ice. Samples were then centrifugated for 5 minutes at 5000*g at room temperature. The supernatant was discarded and the pellet washed with 300 µL of EtOH 70% before drying at 60°C for 1 hour. The pellet was finally resuspended in 100 µL of distilled water. Total DNA was also extracted from individual aphids also according to Edwards' method (**Edwards, Johnstone and Thompson 1991**). In all cases, the total amount of DNA extracted was estimated by NanoDrop2000 (ThermoScientific) and used for normalizing the qPCR quantification in each sample.

7. Phenotypic trait estimates

For faba beans, plant height, number of branches and leaf levels were noted every 6 days post-inoculation. Symptom score was determined every three days starting at 6 dpi, because all plants are asymptomatic during the first week after inoculation. Symptom scoring was visually estimated according to the criteria described in (**Supplementary Table 2**). Height and fresh weight of green tissues were measured 35 days after infiltration, at the end of the experiment. The same plant material was then dried in a Poupinel sterilizing oven at 50°C for 10 days to estimate the dry weight.

For vetches, seven days after the beginning of the aphid IAP, the monitoring of plant height within the Aracones and symptom scoring was initiated and repeated every 3 to 4 days. Symptom scoring was visually estimated according to the criteria described in (**Supplementary Table 2**). At the end of the experiment, at 31 dpi, Aracones were removed and the last measures of final plant height, as well as the number of branches were taken directly on the plants. Fresh and dry weight were then estimated exactly as for faba bean.

For lentils, symptom score was monitored at 7, 14, 21, and 28 days after the beginning of the aphid IAP. Symptom scoring was visually estimated according to the criteria described in (**Supplementary Table 2**). Plant height was measured only once, at the end of the experiment, at 30 dpi, and the total green tissues were subsequently dried in a Poupinel sterilizing oven at 50°C for 10 days just as for faba beans and vetches.

8. qPCR, genome formula and viral load

For the qPCR quantification of all segments, each extracted sample was diluted 10 times in distilled water and 2 µL were used as template. The qPCR analysis was carried out using the LightCycler 480 thermocycler (Roche, Indianapolis, Ind, USA) with the LightCycler FastStart DNA Master Plus SYBR Green I kit (Roche) according to the manufacturer's instruction. Primers (**Supplementary Table 3**) were used at a final concentration of 0.3 µM to 1.5 µM depending on the associated isolate/segment (**Supplementary Table 3**). Primers were designed to function using the same cycle parameters (40 cycles at 95°C for 10 s, 60°C for 10 s and 72°C for 10 s) for all three isolates and to be highly specific of their respective segment/isolate. For all samples we analyzed two qPCR technical replicates. Post-PCR data were analyzed as described in (**Gallet et al. 2022**) with the LinRegPCR program (**Rujiter et al. 2009**). Genome formulas were then calculated as described in (**Sicard et al. 2013**).

To confirm the presence of a segment we defined the value of threshold cycle (CT), beyond which the corresponding segment was considered absent. To establish these thresholds, we performed qPCRs on samples extracted from mock-inoculated plants, *i.e.* using agroinfiltration mix without bacteria for faba beans and aviruliferous aphids for lentils and vetches. As usual with the SYBR green qPCR approach, we obtained a weak aspecific amplification level associated with each primer couple. We then calculated for each segment the lower tolerance threshold so that 95% of potential future samples not containing the segment would have higher values with a 95% probability (**Sharma and Mathew 2012; Francq et al. 2019; Di-Mattia et al. 2022**). Ct thresholds are detailed in (**Supplementary Table 4**).

Due to the differential development of sampled leaves, the number of cells vary and probably affects the viral load. We decided to normalize the total viral DNA by the total DNA from the extracted samples. Nucleic acids were dosed using a NanoDrop2000 (ThermoScientific). The sum of the concentration of all segments below the CT threshold quantified by qPCR was then divided by the nanodrop-estimated total nucleic acids in each extracted sample, thus obtaining a normalized ratio of the total viral DNA/total DNA concentration used to express the viral load.

9. Statistical analysis

All analyses were performed using R version 4.3.0. To assess the biological traits associated with our viral isolates on each respective host (faba bean, lentil, vetches, aphids and transmissions), we used distinct datasets, each of which was separately analyzed.

9-1. Defining qPCR Ct thresholds for genomic segment absence

We chose to infer the presence/absence of genomic segments using the CT parameters of qPCR for their high sensitivity. To confirm the presence of the genomic segments and analyze fully infected plants, we needed to establish a dependable CT threshold that would determine when segments are considered absent using tolerance limits. To achieve this, we conducted qPCR for each segment on corresponding MOCK plants that had only been exposed to the medium used for agro-inoculation without any bacteria (for faba beans) or by non-viruliferous aphids (for lentils and vetches). We thus recorded the CTs obtained for a focal segment, when there are no viral segments in the plant, as two technical replicates for each sample.

After filtering the samples that did not amplified anything (no CT value) or had an aberrant one (CT>50 considered as a technical artifact), we established a CT distribution for all segments. Because the distribution of most segments strongly deviated from normality, we used the method described by (**Young and Matthew 2014**) to estimate non-parametric tolerance intervals based on interpolated and extrapolated order statistics. We used the nptol.int function of the "tolerance" R library (**Young 2010**) with the "YM" Young-Matthew method to estimate 1-side tolerance intervals based on linear interpolation/extrapolation of order statistics (OS-Based). This method does not take into account structural hierarchy of the data (technical replicates). As such we decided to keep only one technical replicate, selecting the lower CT from each sample. Based on these CT distributions we determined the (0.95, 0.95)-one-sided lower tolerance interval, *i.e.* the CT value above which would lie 95% of future comparable samples with 95% confidence based on the given distributions. In summary, a segment is considered absent from a sample if its CT value lies above the previously determined threshold. The threshold values and sample sizes to obtain them are detailed in (**Supplementary Table 4**).

For all biological trait evaluations, we first filtered out plants that did not contain all eight genomic segments according to the thresholds previously determined.

9-1. Plant height, fresh weight, dry weight

To account for the direct impact on the host plant species, we calculated the standardized ratio of the plant heigh, fresh and dry weight biological traits according to the following formula:

$$r = \frac{\text{mean asymptomatic trait} \; - \; \text{individual symptomatic trait}}{\text{mean asymptomatic trait}}$$

If infection has no effects on the trait of symptomatic plants, the ratio will be zero, if the effect on height important, the ratio will approach one. After calculating the ratio, we identified outliers for each genotype - host combination (below the lower quartile – 1.5 times the interquartile range or above the upper quartile + 1.5 times the interquartile range) and excluded them from the analyses.

We modeled the standardized ratio as a function of genotype as fixed effect in a linear model (lm function from lme4 package). To test the significance of the predictors included in the models, type-III Sum of Squares were used (Anova function from car package) (**Supplementary Table 5-A**). Next, we performed a post-hoc multiple pairwise comparisons test by comparing the predicted standardized height ratio of a reassortant with the predicted ratio of all isolates (contrast function from emmeans package specifying "mvt" – multivariate test – as a method for adjustment for multiple comparisons) (**Supplementary Table 5-B**).

9-2. Symptom severity, branches and foliar levels

To analyze biological kinetics in infected plants, we used standard procedures in plant pathology on the analysis of qualitative symptoms, using nonparametric analysis of ordinal data with repeated measures following the procedures outlined by (**Shah and Madden 2004**). Specifically, to assess the statistical significance of the effect of genotype we used nparLD: an R-language library for non-parametric analysis of longitudinal data with factorial design (**Noguchi et al. 2012**). Analyzing only symptomatic plants that were evaluated with the disease index, we created models using f1.ld.f1 function that refers to the design with one between-subject factors (genotype) and one within-subject factor (time). For each combination of between and within factor levels, f1.ld.f1 models produce relative treatment effects (RTE): the probability that the value of a randomly selected individual measurement with this combination of factor levels is larger (*e.g.* has a larger symptomatic score or measurements in our case) than the value of a randomly selected individual measurement across all possible factor levels. The modified ANOVA-type statistic with Box's approximation

(ANOVA.test.mod.Box) was computed for the between-subject factor (genotype) (**Supplementary Figures 9 – 10**). Since there were no significant P.values for the between-subject factor, no additional post-hoc comparisons were conducted.

9-3. Viral load

Regarding viral load, we excluded outliers for each genotype-host combination if the standardized viral load value fell below the lower quartile minus 1.5 times the interquartile range or above the upper quartile plus 1.5 times the interquartile range. We modelized the standardized viral load as linear model with genotype as fixed effect (lm function from lme4 package). Concerning the aphid dataset, we modelized the standardized viral load as a linear model with genotype and host as fixed effect and the interactions between them. We observed heteroscedasticity and deviation from normality of residuals in all individual datasets (faba beans, vetches, lentils and aphids). We applied Box-Cox transformation (λ-FB=0.303, λ-L=0.424, λ-V=0.465, λ-Aph=0.061) to address this issue. Following Box-Cox transformation, homoscedasticity and normality of residuals were observed for all models except the one associated to the lentil dataset.

To assess the significance of the predictors included in the models, we employed type-III Sum of Squares (using the Anova function from the car package) (**Supplementary Tables 6-A** and **6-C**). We also conducted a post-hoc multiple pairwise comparisons test, comparing the predicted mean relative viral load of each isolate. For this analysis, we used the contrast function from the emmeans package, specifying "mvt" (multivariate test) as the method for adjustment for multiple comparisons. The data was reverse-transformed before comparison using the regrid function from the emmeans package (**Supplementary Tables 6-B** and **6-D**).

9-4. Genome formula

Regarding the genome formulas, outliers of all genomic segments were excluded if they fell below the lower quartile minus 1.5 times the interquartile range or above the upper quartile plus 1.5 times the interquartile range. We compared genome formulas of the three genotypes. To investigate whether different segments accumulated at varying frequencies according to the infecting genotype, we modeled the logit of segment frequency. This model included the identity of the segment, the infecting genotype and all possible interactions between them as fixed effects. The lm function from the lme4 package was used.

The logit transformation was necessary to meet the assumptions of the linear model. Significance of fixed effects and their interactions was evaluated using type-III Sum of Squares, with the Anova function from the car package (**Supplementary Figure 1**). We also conducted a post-hoc multiple pairwise comparisons test, comparing the predicted mean relative of genomic segment frequency. For this analysis, we used the contrast function from the emmeans package, specifying "sidak" as the method for adjustment of multiple comparisons and to calculate confidence intervals for estimated differences in frequencies.

9-5. Transmission rates

To assess the transmission rates of the three isolates, distinct datasets associated with each host were individually analyzed. We modelized the infectivity status (1 or 0) of recipient plants in a generalized linear model as a function of genotype as fixed effect (glm function from lme4 package, binomial distribution). This model uses a reference genotype for comparisons that was always **[AZ;15]** (**Supplementary Table 6-E**).

RESULTS

1. Selection of the isolates

Following the first comprehensive nanovirus genomic study (**Grigoras et al. 2014**), subsequent research has identified numerous new nanovirus species and isolates. For an update with data available up to early 2020, we conducted an additional phylogenetic investigation of FBNSV. By examining the concatenated genomes of all accessible FBNSV isolates and conducting pairwise sequence comparisons, our analysis confirmed the division of FBNSV into three distinct clades, each diverging by approximatively 20% in nucleotide identity (**Figure 2**). We arbitrary named these clades « Red », « Blue » and « Green ». Phylogenetic analysis of individual segments revealed extensive genetical exchanges among isolates of the "Green" clade (**Supplementary Figures 2-8**). Among the three Azeri isolates primarily congruent with the "Green" clade, **[AZ;10]** has only DNA-U2 that belongs to the "Red" clade, while **[AZ;1]** has both DNA-N and DNA-U2 and **[AZ;12b]** has both DNA-M and DNA-N (**Supplementary Figures 2-8**).

To conduct our phenotypic comparison, besides the already available **[JKI-2000]** infectious clone belonging to the "Blue" clade and originally sampled on faba bean (*Vicia faba*) in

Ethiopia, we chose to synthesize two new clones with the following criteria: the two should belong to distinct clades and not belong to the "Blue" clade, originate from the same geographical area, and be sampled on different host species. We thus focused on the Azeri isolates **[AZ;]** (**Grigoras et al. 2014**). We first chose **[AZ;15]** with all segments belonging to the "Red" clade and sampled on vetch (*Vicia sativa*). For the second isolate we chose **[AZ;10]** sampled on *Lens culinaris* (from a plant located just a few meters away from the sampled plant where **[AZ;15]** was isolated) as all segments of this isolate except DNA-U2 belong to the "Green" clade, while DNA-U2 is clearly a reassorting segment from a "Red" clade parental genotype (**Figure 3**). To have a complete genome belonging to the "Green" clade we thus decided to use all segments of **[AZ;10]** except DNA-U2 for which we chose the DNA-U2 of **[AZ;12b]** isolate also sampled on lentil (*Lens culinaris*) in Azerbaijan, but whose segment U2 belongs to the "Green" clade. **[AZ;12b]** had even less segments belonging to the "Green" clade than **[AZ;10]** overall. They were both isolated at different locations and years (2010 – 2011). We named this hybrid genotype **[AZ;10_12b]** isolate.

When comparing protein sequences from these three isolates, three groups can be distinguished based on their amino-acid (AA) identity (**Figure 4**). NSP (DNA-N), M-Rep (DNA-R) and CP (DNA-S) are the most conserved proteins, their AA identity ranging between 94 and 99 %. Clink (DNA-C) and MP (DNA-M) are the second most conserved group, ranging between 86 and 89% of AA identity. Finally, U1 (DNA-U1), U2 (DNA-U2) and U4 (DNA-U4) are the least conserved group, ranging between 70 and 80% AA identity, except for U1 proteins of **[AZ;10_12b]** and **[JKI-2000]** which are more conserved with 88 % AA identity.

The eight segments can alternatively be divided in two groups based on the conservation disparity between coding and non-coding sequences (**Figure 4**). MP (DNA-M), NSP (DNA-N), M-Rep (DNA-R) and CP (DNA-S) have an AA coding sequence more conserved than the whole segment, whereas it is the opposite situation for Clink (DNA-C), U1 (DNA-U1), U2 (DNA-U2) and U4 (DNA-U4) (**Figure 4**). The overall genetic distance among the three isolates appears to be equivalent, with the closest isolates varying depending on the genomic segment (**Figure 4**).

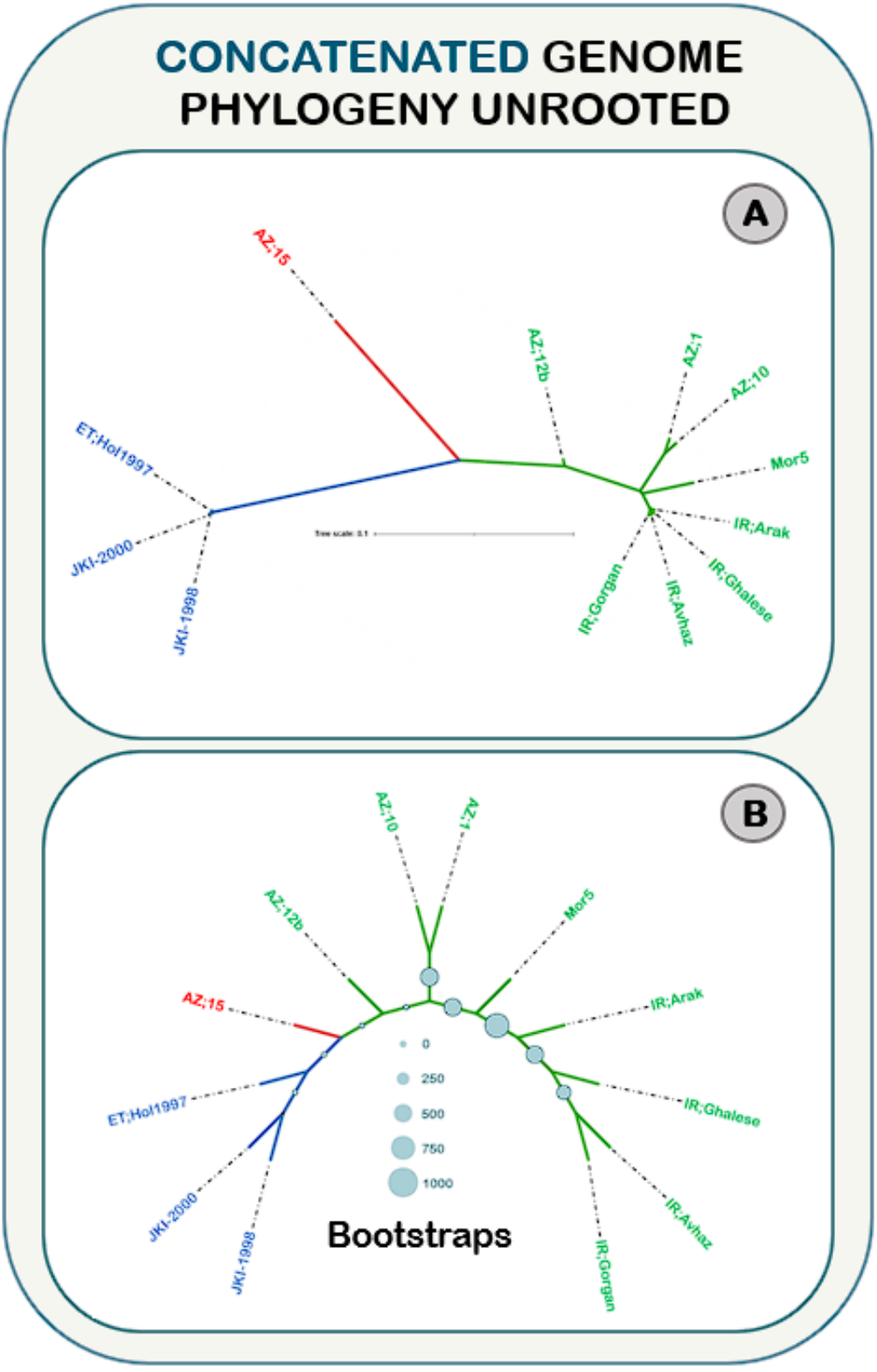


Figure 2 : [A] Unrooted phylogenetic tree of FBNSV concatenated genomic segments (C-M-N-R-S-U1-U2-U4) from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to. **[B] Unrooted cladogram of FBNSV concatenated genomic segments (C-M-N-R-S-U1-U2-U4) from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.

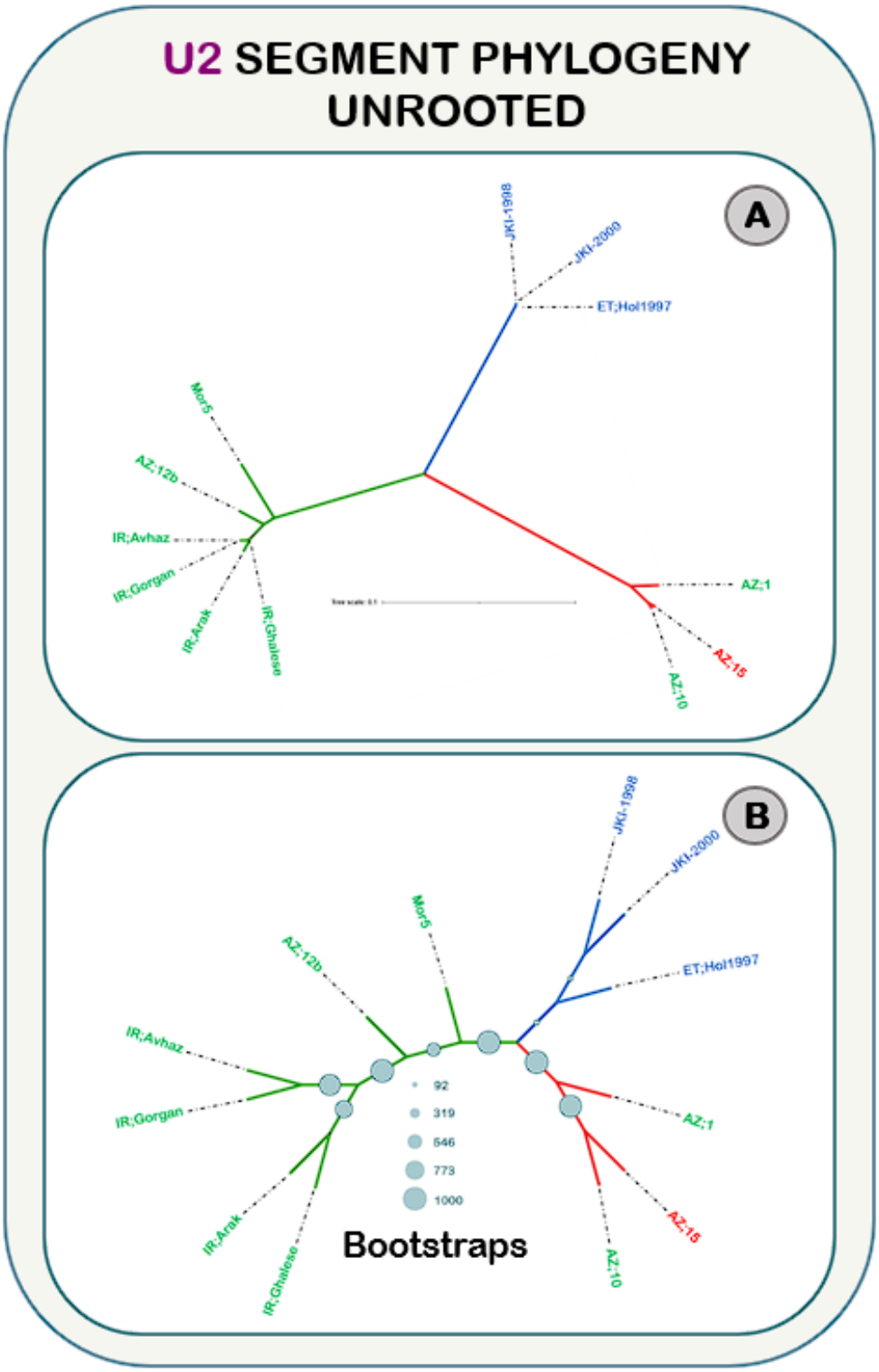


Figure 3 : [A] Unrooted phylogenetic tree of FBNSV U2 genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV U2 genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.

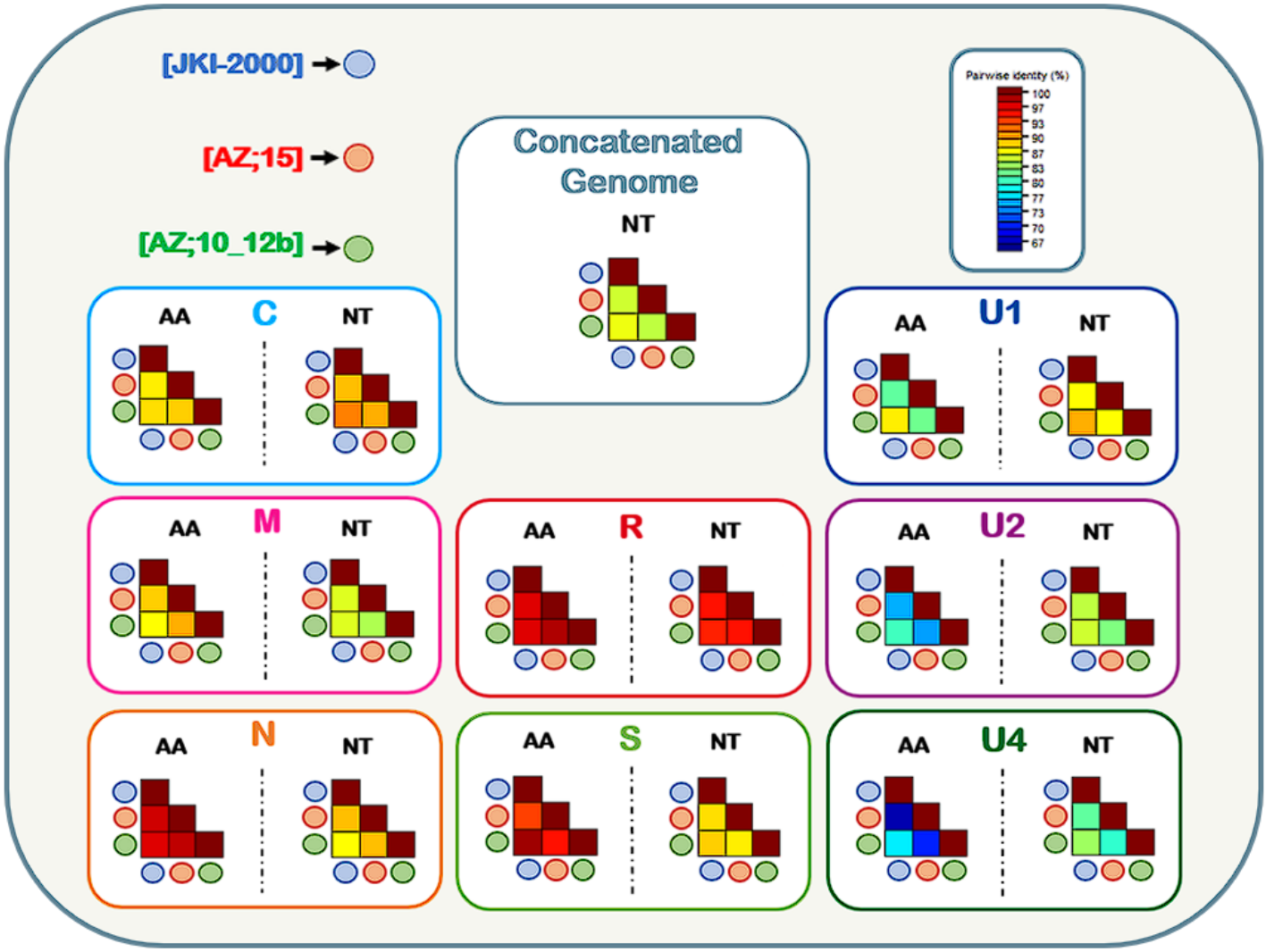


Figure 4 : Graphical representations of nucleotide and amino-acid pairwise identity percentage of [JKI-2000], [AZ;15] and [AZ;10_12b] genomic segments and concatenated genome (C-M-N-R-S-U1-U2-U4) as a color-coded SDT matrix. Colors of the matrix are associated with the corresponding-colored scale from 65 % (navy blue) to 100 % (burgundy) pairwise identities.

2. Viability and transmissibility of the clones

The phenotype characterization and comparison of the three clones rely primarily on the viability of the two new infectious clones in faba bean. Because these have been synthesized from GenBank database sequences, we do not know if all segments encode a functional protein. Additionally, they have been isolated from lentils and vetch, while to date, nanovirus infectious clones have been systematically agro-inoculated with success only on faba bean and pea plant hosts. (**Grigoras et al. 2014**). We thus agro-infiltrated in parallel the three infectious

clones to faba bean plants. Both new isolates **[AZ;10_12b]** and **[AZ;15]** developed typical symptoms (dwarfing, apical leaf-curling and chlorosis) of FBNSV, visually undistinguishable from those induced by **[JKI-2000]**, and all eight genome segments could be detected by qPCR from the infected samples for the three isolates.

We then assessed vector transmission to young faba bean plantlets, as described in the Methods section, but using five distinct aphid species and five aphids per recipient plant. Both new isolates were successfully transmitted by *Aphis craccivora*, *Acyrthosiphum pisum*, and *Myzus persicae*, with *A. craccivora* appearing as the most efficient vector, though this experiment was not designed to formally quantify the transmission rates (**Supplementary Table 7**). We thus decided to use *A. craccivora* to transmit each of **[AZ;10_12b]** and **[AZ;15]** infectious clones to both lentils and vetches using five aphids per recipient plantlets. We aphid-inoculated multiple varieties of lentils and vetches to choose the cultivars for further in-depth characterization. As a result, we selected *Lens culinaris* var. "Coralie" and *Vicia sativa* var. "José" that were infected with good efficiency and displayed discernable symptoms by both **[AZ;10_12b]** and **[AZ;15]** isolates (**Supplementary Table 7**).

3. Quantification of symptoms induced by the three isolates in the faba bean, lentil and vetch.

3-1. Symptom characterization

All three isolates developed overall similar symptoms on both lentils and vetches. We observed leaf curling of the last foliar levels and dwarfing that developed into chlorosis and bushing (**Figure 5**). At later stages of infections, we observed plant death or extreme cases of dwarfing and bushing (**Supplementary Table 2**).

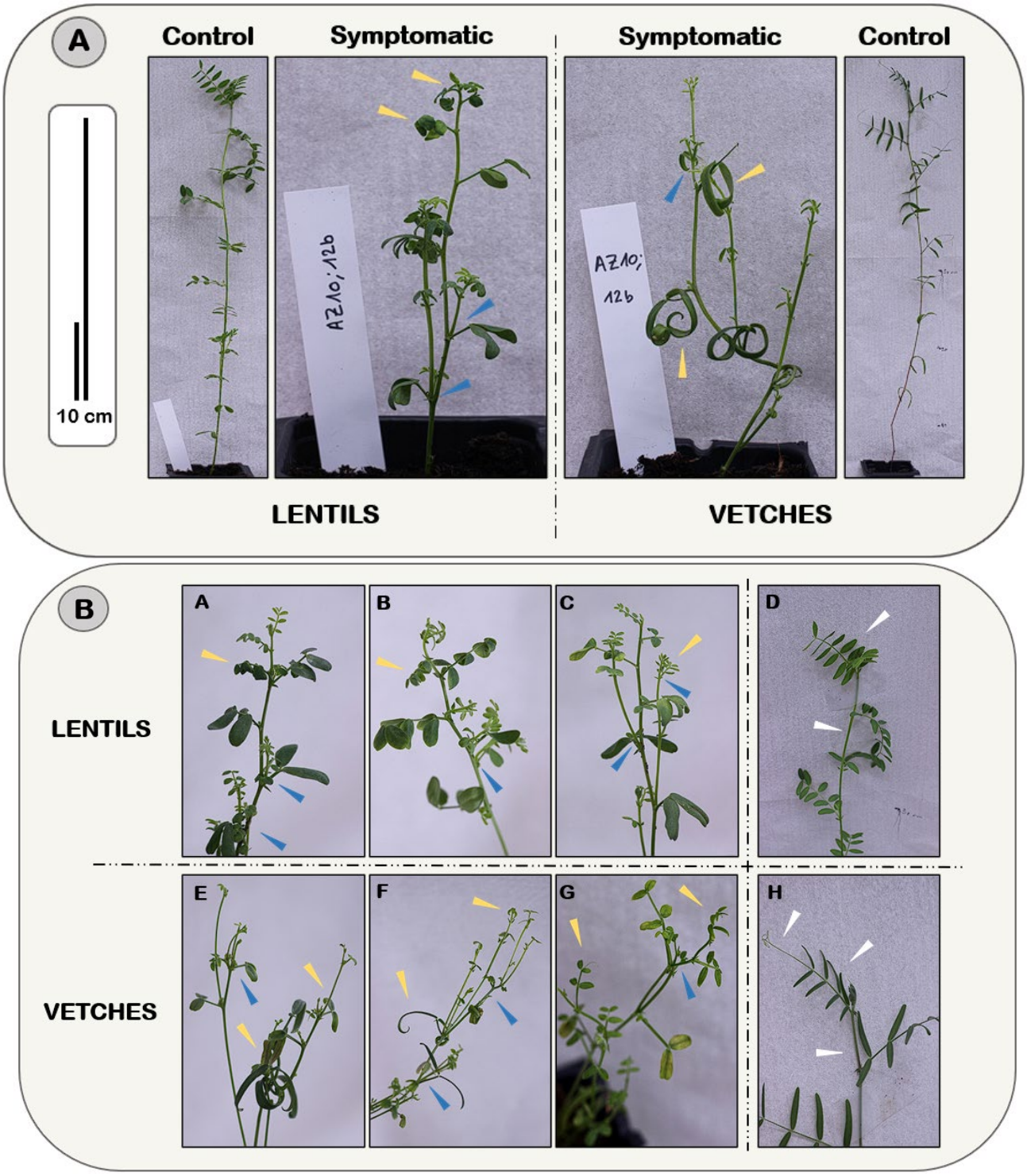


Figure 5 : [A] Symptoms and variations in plant height resulting from [AZ;10_12b] infection in both lentils and vetches (28 dpi). On the far left and right, non-infected healthy lentil and vetch are photographed as controls. Visible associated symptoms of the infection are leaf curling (yellow darts), dwarfing and bushing (blue darts) of the plants. **[B] Detailed symptoms of infected lentils and vetches (28 dpi).** Yellow darts point out leaf deformations, while blue darts indicate increased branching and bushing in symptomatic lentils (A, B, C) and vetches (E, F, G). Healthy lentil and vetch of the same age typically featuring one or two main stems with leaves having numerous leaflets and tendrils, as shown by the white darts on the control images (D and H).

3-2. Formal estimates of biological traits kinetics

A

Traits	Host	Source	Statistic	Df1	Df2	P.Value
Symptoms	FB	Genotype	1.567	1.951	104.153	0.214
Symptoms	V	Genotype	0.336	1.821	24.187	0.698
Symptoms	L	Genotype	1.018	1.971	83.335	0.365
Height	FB	Genotype	2.029	1.992	113.880	0.136
Height	V	Genotype	2.657	1.895	28.009	0.090
Foliar Levels	FB	Genotype	1.747	1.964	228.134	0.177
Branches	FB	Genotype	0.617	1.996	120.017	0.541

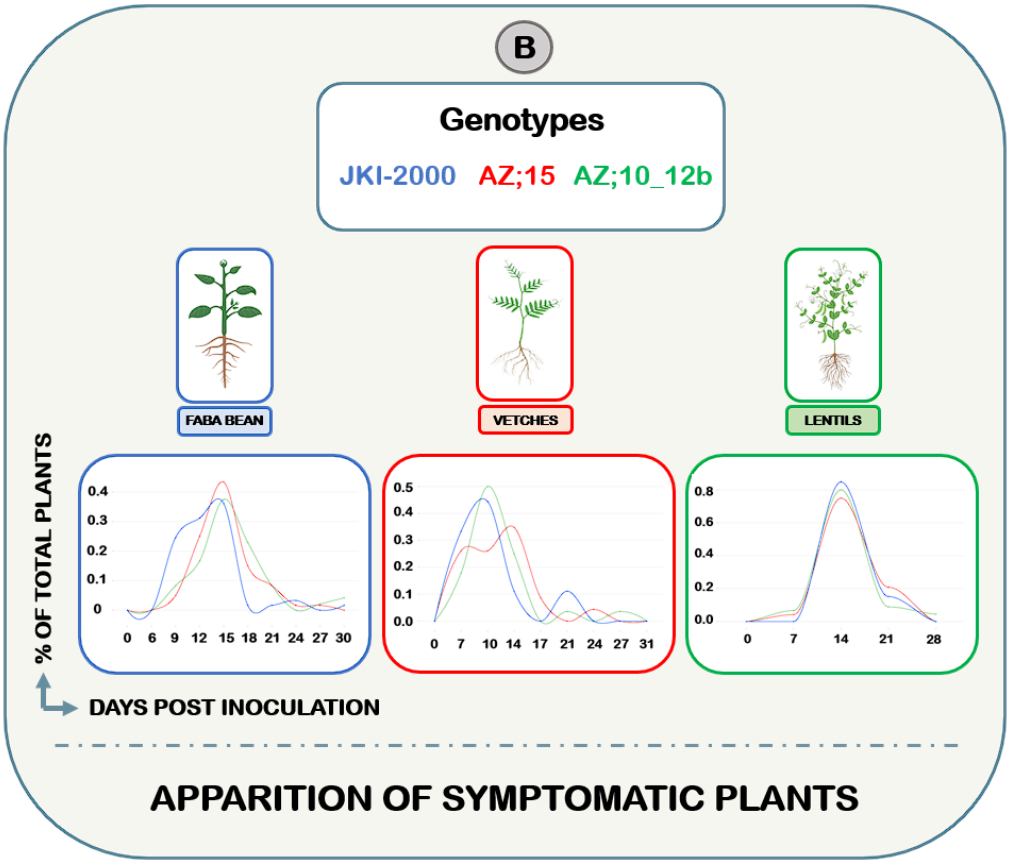


Figure 6 : [A] Coefficients of the statistical F1.LD.F1 models (Supp Figure 9-10) per evaluated biological traits using genotype as between-subject factor and time as within-subject factor. Significative P.values are colored in green (<0.05). **[B] Graphical plots of the symptom severity (top panels) and infection developments (bottom panels) kinetics.** X-axis indicates the days post inoculation. Y-axis indicates the percentage of total plants with symptoms. Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored lines described in the genotypes panel.

Symptom severity and progression were comparable for all hosts on the three isolates (**Figure 6-A**). When we looked at the apparition of the first symptomatic plants, we observed an earlier onset of infection for **[JKI-2000]** in faba beans and a delayed one observed with **[AZ;15]** in vetches (**Figure 6-B**) although these differences were not evaluated statistically yet. In lentils,

the limited number of monitoring time points may have obscured potential differences in this regard.

Kinetics of phenotypic traits were mainly assessed on faba beans (plant height, number of branches, number of foliar levels) and vetches (plant height). The three isolates did not statistically significantly differ from any of the biological traits estimated (**Figure 6-A**), although a general trend of increased number of foliar levels in faba bean and a reduced impact on vetch height could be observed for **[AZ;15]** (**Supplementary Figure 9**).

3-3. Evaluation of plant heights, fresh and dry weights

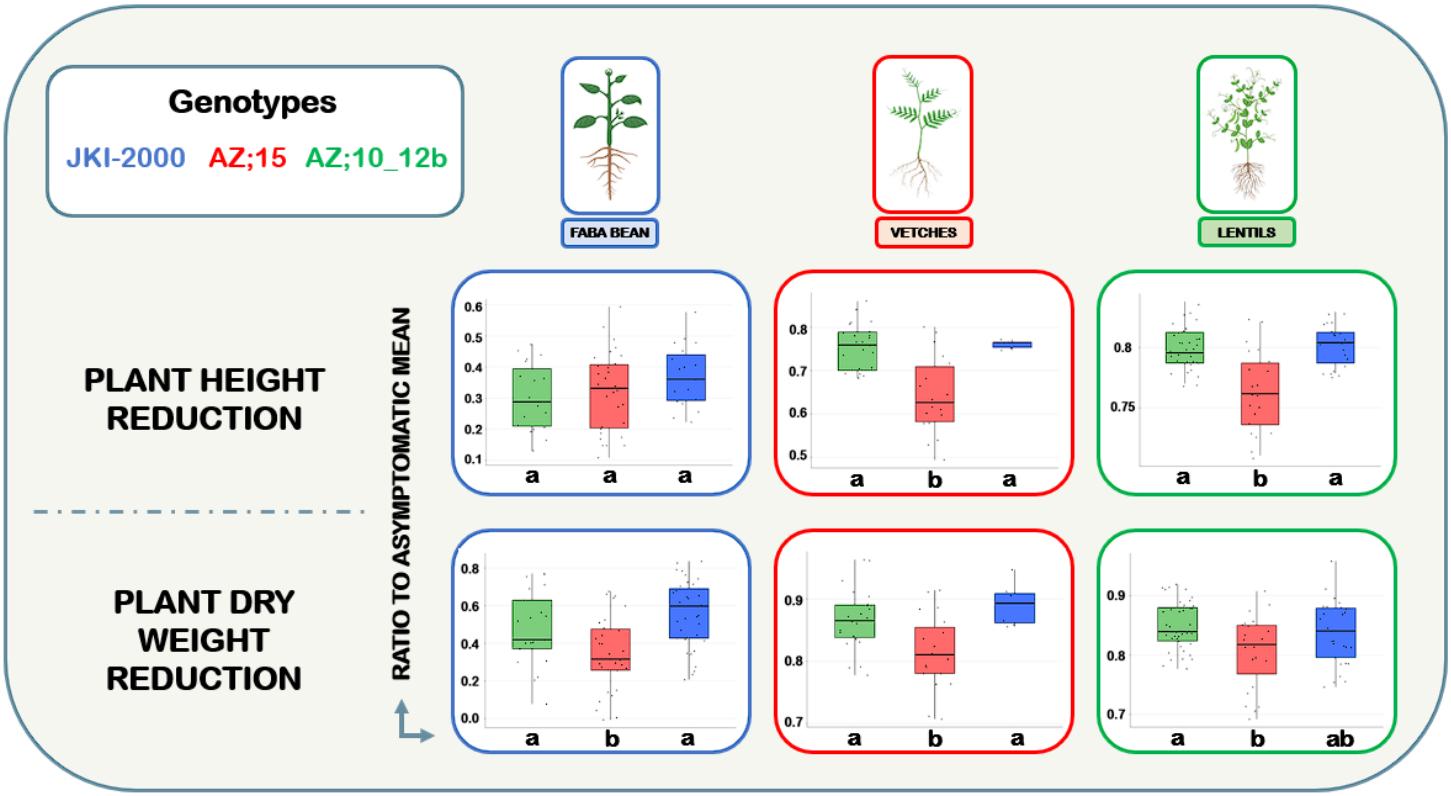


Figure 7 : Graphical plots of the standardized plant height (top panels) and dry weight (bottom panels) effects. Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored boxplots described in the genotype panel. Statistical groups of significance are displayed at the bottom of each panel.

Height, fresh weight and dry weight were measured for all isolates on the three hosts at the end of the experiment. Across all host plants, **[AZ;15]** exhibited a statistically significant decrease in its influence fresh weight, and dry weight compared to the other isolates (**Figure 7** – bottom panels). It also demonstrated a statistically significant reduction in its impact on plant height in both lentils and vetches, confirming the previous non-statistically significant

general trend observed on vetch height kinetic (**Figure 7** – top panels). However, in the case of faba beans, the impact on plant height was comparable among all isolates. **[AZ;10_12b]** and **[JKI-2000]** appeared to have similar effects on all hosts for height and weight parameters (**Figure 7**). It's important to mention that only six vetches infected with **[JKI-2000]** were included in the final dataset. Strong correlations were observed between these three biological traits for each host and isolate (**Supplementary Figure 11**).

4. Viral loads and transmission rates

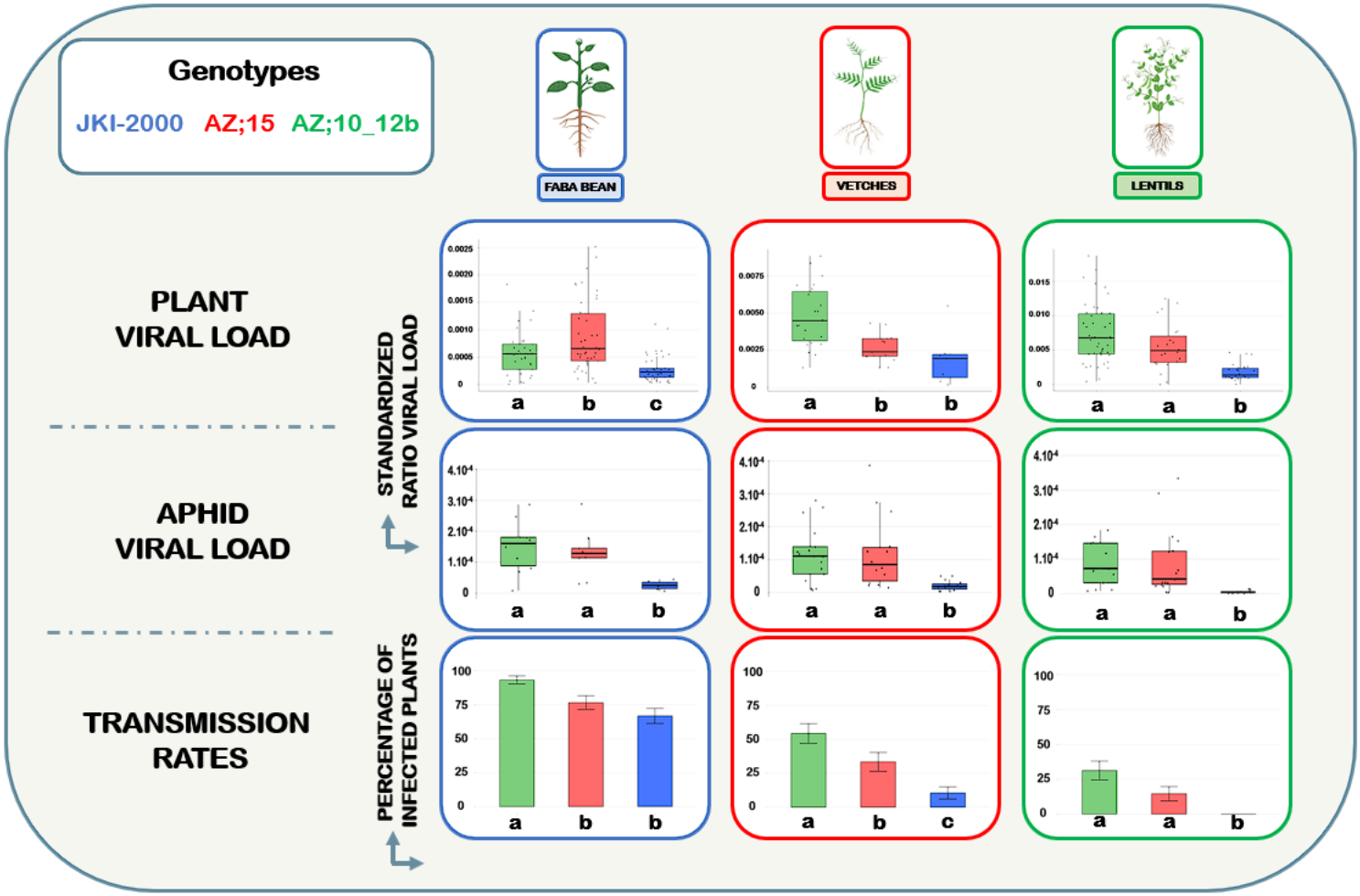


Figure 8 : Graphical plots of the normalized plant viral load (top panels), aphid viral load (mid panels) and transmission rates (bottom panels) effects. Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored boxplots described in the genotypes panel. Statistical groups of significance are displayed at the bottom of each panel.

The estimation of the viral loads (**Figure 8** – top and middle panels), standardized by the total extracted DNA as indicated in the Methods section showed a general trend of lower accumulation for **[JKI-2000]**, and this general trend was statistically significant in the three host plants as well as corresponding aphids that fed on these plants. **[AZ;10_12b]** accumulated

statistically significantly less than **[AZ;15]** in faba beans, more in vetches and equally in lentils. However, their viral load differences in corresponding aphids were not statistically significant.

[AZ;10_12b] was better transmitted than the other isolates, except in lentils where the difference with **[AZ;15]** was not statistically significant (**Figure 8** - lower panels). **[JKI-2000]** was the least transmitted isolate in all hosts, apparently consistent with a lower accumulation in the three host species and in aphids fed on these plants. Although no plants were infected during the **[JKI-2000]** L>L transmission, the isolate proved transmissible on lentils in previous experiments.

5. Genome formulas

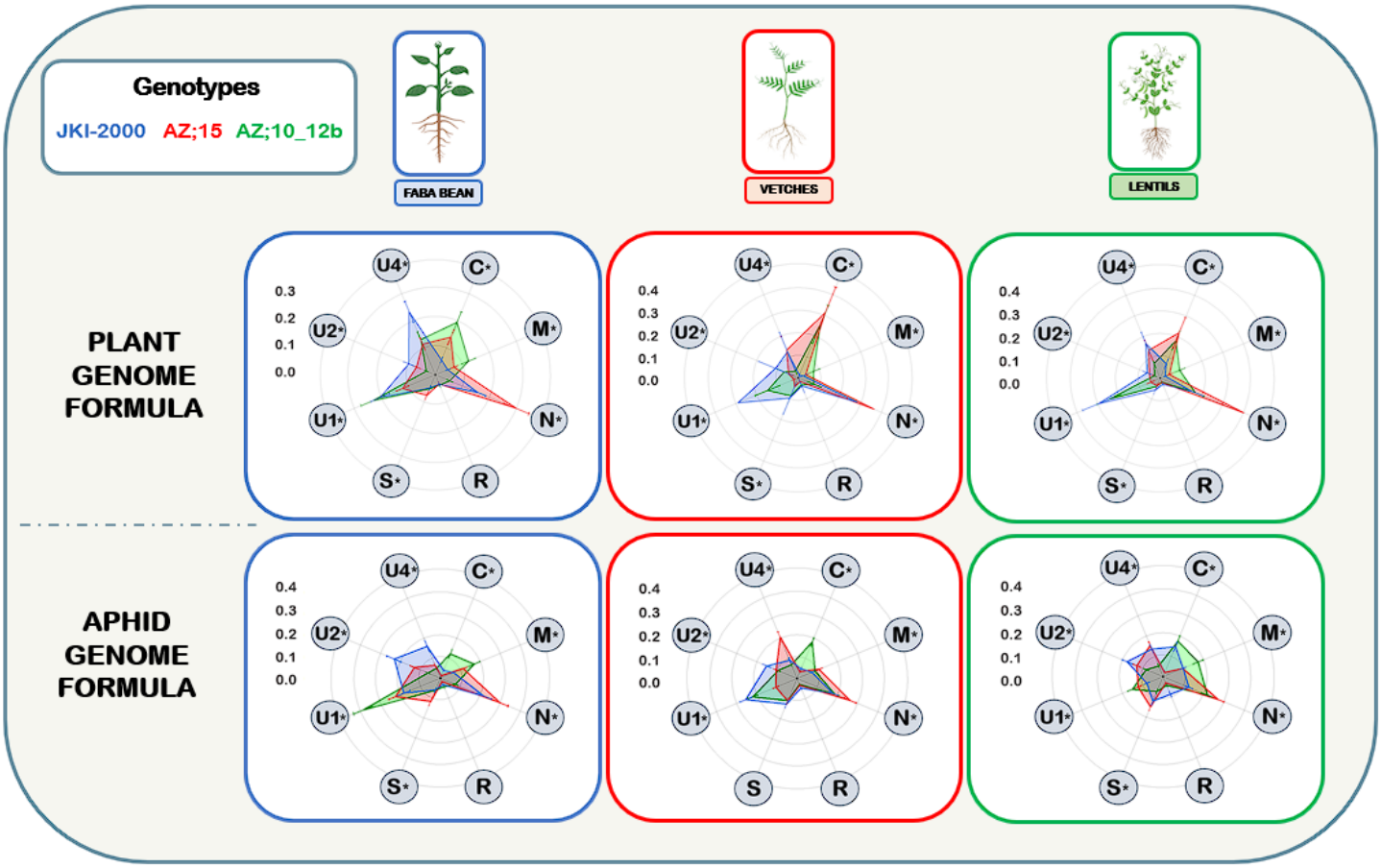


Figure 9 : Graphical representation of the plant (top panels) and aphid (bottom panels) genome formulas. The overall pattern of the genome formula of the three isolates is depicted as radar plots for each of the three plant host species, whereas the statistically significant differences among isolates for a corresponding segment is illustrated by a *. Each circle on the radar plot corresponds to the relative frequency indicated on the left scale. Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored area described in the genotypes panel.

While numerous studies earlier reported that the genome formula of multipartite viruses varies across host plant species (**Sicard et al. 2013; Wu et al. 2017; Sicard et al. 2015; Gallet et al.**

2022), the effect of viral sequence variations within a viral species has so far never been concretely documented (**Yu et al. 2019**). The three FBNSV isolates showed distinct genome formulas on any given host species (**Figure 9** – top panels). The genome formula of each of the three isolates varied across hosts (**Figure 9** – top panels) confirming the already observed across-host species variation.

The genome formula of the three isolates was also greatly different in the aphid vectors, and this is true whether we compare the genome formula in aphids which fed on the same host species infected by different viral isolates or when we compare aphids fed on different host plants infected by the same isolate (**Figure 9** – bottom panels). This was expected as we earlier demonstrated that the genome formula in aphids was only partially determined by that within the source plant (**Sicard et al. 2015**). Consistently, across all host plant species, aphids fed on **[AZ;10_12b]** infected plants had the highest frequencies of DNA-C, -M and -U1, and the lowest frequencies of -U2 and -U4, and low frequencies of -N. Aphids fed on **[AZ;15]** infected plants displayed the highest frequency of DNA-N and -S, the lowest frequency of -C. Finally, aphids fed on **[JKI-2000]** infected plants had the highest frequencies of DNA-R and -U2 but low frequencies of DNA-M.

6. Missing segments

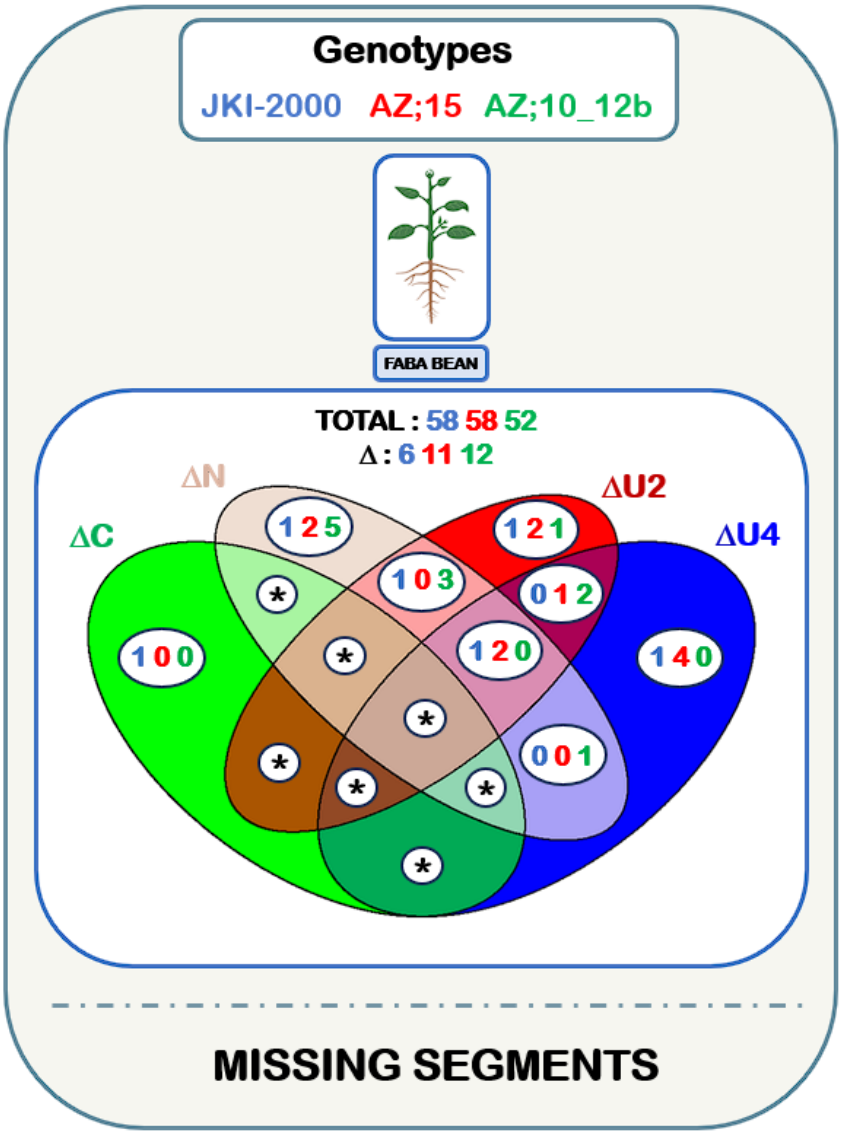


Figure 10 : Venn diagram of the missing segments detected in infected faba beans. The total number of infected plants is displayed, along with the number (Δ) of plant missing at least one segment per genotype. Corresponding genotypes are associated with the same color as indicated in the genotypes panel for both Venn diagram and plant numbers. Asterisks (*) indicate that no plants were detected missing the corresponding combination of genomic segments for any isolate.

With our CT thresholds in place, we occasionally observed the absence of segments across all isolates, in all host plants, and aphid vectors. However, these absences were much more frequently detected in faba beans which might be explained by the use of agro-inoculation method instead of the more reliable aphid inoculations. We quantified this phenomenon in detail in this specific host (**Figure 10**). Among the infected characterized faba beans, those lacking at least one segment accounted for 10% in the case of **[JKI-2000]**, 19% for **[AZ;15]**, and 23% for **[AZ;10_12b]**.

Across all hosts, out of the five segments categorized as "non-essential" for the establishment of FBNSV systemic infections (**Grigoras et al. 2018**), DNA-U1 was never absent, and DNA-C was absent on only one occasion associated with **[JKI-2000]** in faba bean. Further, out of the 29/168 faba bean plants missing at least one segment of any of the three isolates, 21 missed a single segment, five missed two segments, three missed simultaneously DNA-N, -U2 and -U4, and none missed more than three segments. The low per viral isolate number of plants missing at least one segment (min = 6, max = 12) did not allow for a detailed statistical investigation of across isolate distributions of the segment loss.

DISCUSSION

Reaffirming the bitmer design (**Urbino et al. 2008; Yepes et al. 2018**) in generating infectious clones of CRESS DNA viruses, which uses partial tandem repeats enabling rolling-circle replication contingent upon the presence of flanking origin of replication within the conserved regions CR (**Bonnamy, Blanc and Michalakis 2023; Torralba, Blanc and Michalakis 2023** – **Chapter 1**), we produced two novel infectious clones of FBNSV, designated as **[AZ;15]** and **[AZ;10_12b]**, which are now available for further research questions.

These, combined with the already available **[JKI-2000]** clone, comprehensively represent the three main phylogenetic clades that structure the known diversity of FBNSV (**Grigoras et al. 2014**). While each of these isolates were sampled from different legume hosts, the selection of **[AZ;15]** and **[AZ;10_12b]** was also influenced by their occurrence within the same geographical area, suggesting that both isolates could have access to common host and vector species. The extensive characterization of **[AZ;15]**, **[AZ;10_12b]**, and **[JKI-2000]** confirmed the capacity of all corresponding isolates to infect identical hosts and to be transmitted by the same aphid vectors. Among these vectors, *A. craccivora* stood out due to its efficiency in transmitting all three isolates. Notably, *A. craccivora* is a prevalent aphid species of grassland legumes in Ethiopia (**Bekele et al. 2005**) and Iran (**Mokhtari et al. 2012**), and it is conceivable that its populations are also prevalent in nearby Azerbaijan in regard to its overall geographical distribution. The fact that all three isolates share both hosts and vectors, although **[JKI-2000]** was isolated far from the other genotypes, may have substantial implications for nanovirus mixed infections in geographical hotspot of sympatry, with a great potential for genetic recombinations and reassortments.

In the current literature, a significant discrepancy emerges between inter-specific recombinations and reassortments, with the latter being notably rare compared to the former,

despite experimental indications suggesting a higher expected prevalence of reassortments (**Torralba, Blanc, and Michalakis 2023 – Chapter 1**). The availability of infectious clones for these new diverse nanovirus isolates (FBNSV **[AZ;15]**, and **[AZ;10_12b]** – this study) and species (PNYDV, BMLRV, FBNYV) (**Grigoras et al. 2014**), most of which share common hosts and vectors (**Grigoras, Gronenborn, and Vetten 2010; Gaafar et al. 2016; Grigoras et al. 2018; Mansourpour et al. 2022**), will provide good opportunities for further studies concerning intra- and inter-specific reassortments, eventually providing explanation for this discrepancy.

The genome formula stands as general characteristic of multipartite viruses (**Sicard et al. 2013; Sánchez-Navarro, Zwart and Elena 2013; Sicard et al. 2015; Hu et al. 2016; Yu et al. 2019; Mansourpour et al. 2022; Guyot et al. 2022; Gallet et al. 2022**), likely enabling swift adjustments in gene expression by the modification of gene copy numbers in distinct environments (**Gutiérrez and Zwart 2018; Zwart and Elena 2020; Gallet et al. 2022**). While alterations in the genome formula are recognized to occur between different viral species and plant hosts (**Sicard et al. 2013; Wu et al. 2017; Sicard et al. 2015**), the assessment of these changes among viral isolates has remained relatively unexplored (**Yu et al. 2019**). Our analysis of the three FBNSV isolates unequivocally demonstrates that the genome formula is distinct for each genotype, likely due to nucleotide modifications in the genomic sequences. Furthermore, alterations in the genome formula are already recognized to take place between hosts and vectors. Once again, our findings unequivocally highlight the modification of the genome formula within aphids, contingent upon the isolate and the host plant from which they acquired it, unveiling a reaction norm of the genome formula which extends to both host plants and vectors and shows genetic variation across viral genotypes.

Overall, we observed qualitatively comparable symptoms for the three isolates in the three hosts. As already reported for faba beans, lentils and vetches exhibited severe stunting, leaf deformations, chlorosis, branching and bushing of the apical portions and sometimes premature death. These manifestations could have a severe detrimental impact on the yield of these leguminous crops. Consequently, close monitoring of nanovirus outbreaks in legume might be beneficial in the future, as they clearly have the potential to emerge as a new agricultural threat.

Although qualitative symptoms were comparable, each of the three isolates displayed distinct phenotypical dynamics, depending on the host, underscoring the phenotypic diversity and important reaction norm variations within the FBNSV species. **[JKI-2000]** exhibited earlier onset of disease symptoms in faba beans. Counterintuitively, it exhibited relatively low viral

loads in both plants and aphids and was associated with the lowest transmission rate across all hosts. Conversely, **[AZ;15]** was the least symptom causing isolate in all hosts. It also seemed to have a weaker impact on the number of foliar levels of its faba bean host and a slower disease progression in vetches. As already indicated, the milder **[AZ;15]** isolate presented substantially higher viral loads than **[JKI-2000]** in all plant species and aphids which suggest that higher viral loads did not induce more severe symptoms.

Interestingly, although **[JKI-2000]** and **[AZ;10_12b]** exhibited comparable symptom severity in all host, they were the least and better transmitted isolates respectively. Conversely, the milder **[AZ;15]** isolate still outperformed **[JKI-2000]** for transmission across all hosts. This observation indicates that there was no general connection between symptom severity and transmission rates between the isolates, although a formal statistical analysis has not been conducted yet to confirm this and to evaluate it at the intra-host level.

While the association between individual aphids used to evaluate the transmission rate and recipient plants was not recorded, it is important to highlight that the analysis of the virus content in aphids revealed that all genomic segments were retrieved in most aphids, strongly suggesting that the absence of infection in the recipient plants is likely attributed to the loss of segments during transmission bottlenecks, rather than the failure to acquire a complete set of segments. As earlier reported (**Sicard et al 2015**), we here note a modification of the genome formula within aphids where some segments may become very rare and may be inoculated in very few copies, increasing their possibility of loss.

While DNA-C, -N, -U1, -U2 and -U4 can be individually omitted in experimental inoculation and the virus still produce systemic infections, the omission of any of these segments except U4 significantly impacts at least one important viral trait (**Grigoras et al. 2018**). The loss of one of these "dispensable segments" is only exceptionally observed in field population (**Grigoras et al. 2014; Knierim et al. 2019; Vetten et al. 2019; Hasanvand et al. 2021**) indicating that specific processes (perhaps acting at higher scale) might naturally select against these segment losses in natural contexts. While it is acknowledged that our CT thresholds can lead to false positives and negatives, the detection of missing segments across all three hosts and with all three isolates (**Supplementary Table 8**) suggests that this phenomenon is not solely an experimental artifact but probably reflects a biological reality. Interestingly, we observed fewer missing segments in lentils and vetches compared to faba beans, which can likely be attributed to differences in inoculation methods and could indicate a bias when agro-inoculating faba beans (versus systematic aphid-inoculation in vetches and lentils). Despite earlier experimental production of plants missing DNA-U1 segments (**Grigoras et al. 2018**),

we did not identify any plants missing this segment. Similarly, only one plant was confirmed to be missing DNA-C, which was infected with **[JKI-2000]** following agroinoculation, as the isolate with the lowest frequency of DNA-C in faba beans. This observation strongly reinforces the importance of DNA-U1 and -C in the establishment of the infection (**Grigoras et al. 2018**).

While we could intuitively anticipate that isolates would exhibit superior fitness on their corresponding field hosts, the results revealed a different trend. Based on our viral fitness proxies, **[AZ;10_12b]** appears as the overall better fit isolate across all three hosts whether we look at plant viral load or transmission rates, followed by **[AZ;15]**. This observation raises the possibility that these viral genotypes might be better adapted to different plant host species as nanovirus host range appears quite extensive among legumes. However, this is mostly relevant for **[AZ;15]** and **[AZ;10_12b]** isolates, as **[JKI-2000]** has undergone three years of successive passages on faba bean, albeit on different cultivars (**Grigoras et al. 2009**), which likely contributed to its host adaptation. However, due to the absence of phenotypic data of the progenitor **[ET; Hol-1997]** isolate, we cannot assess the phenotypic consequences of this possible within lab host-adaptation. Finally, it is important to highlight that the **[AZ;10_12b]** infectious clone itself is a single-segment reassortant between **[AZ;10]** and **[AZ;12b]**, a genetic exchange that could lead to altered fitness compared to its major parental field **[AZ;10]** isolate.

Our study showcased important reaction norms contingent upon viral genotype in all phenotypical traits evaluated in our study. Particularly with respect to the genome formula, which is suspected to enable swift adaptation to changing environments, it would be worthwhile to further explore the connections between genome formula variations and other biological traits. Delving into the relationships between genome formula variations and overall protein expressions may provide valuable insight in nanovirus biology and help us understand the full scope of the reaction norms we have observed.

In conclusion, the characterization of distinct FBNSV genotypes has unveiled notable variations in both the diversity of phenotypes and reaction norms across different host plant species. This is of paramount importance in increasing our understanding of the ecological dynamics of these viruses that could have profound significance as emerging worldwide pathogens.

Bibliography

Abraham, A. D., Bencharki, B., Torok, V., Katul, L., Varrelmann, M., & Josef Vetten, H. (2010). 'Two distinct nanovirus species infecting faba bean in Morocco', *Archives of Virology*, 155/1: 37–46. DOI: 10.1007/s00705-009-0548-9

Babin, M., Ortíz, V., Castro, S., & Romero, J. (2000). 'First Detection of Faba bean necrotic yellow virus in Spain', *Plant Disease*, 84/6: 707. DOI: 10.1094/PDIS.2000.84.6.707B

Bekele, B., Kumari, S. G., Ali, K., Yusuf, A., Makkouk, K. M., Aslake, M., Ayalew, M., et al. (2005). 'Survey of viruses affecting legume crops in the Amhara and Oromia Regions of Ethiopia', *Phytopathologia Mediterranea*, 44/3: 235–46. [Firenze University Press, Mediterranean Phytopathological Union].

Bonnamy, M., Blanc, S., & Michalakis, Y. (2023). 'Replication mechanisms of circular ssDNA plant viruses and their potential implication in viral gene expression regulation', *mBio*, e0169223. DOI: 10.1128/mbio.01692-23

Chu, P. W., Keese, P., Qiu, B. S., Waterhouse, P. M., & Gerlach, W. L. (1993). 'Putative full-length clones of the genomic DNA segments of subterranean clover stunt virus and identification of the segment coding for the viral coat protein', *Virus Research*, 27/2: 161–71. DOI: 10.1016/0168-1702(93)90079-3

Di Mattia, J., Torralba, B., Yvon, M., Zeddam, J.-L., Blanc, S., & Michalakis, Y. (2022). 'Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution', *Proceedings of the National Academy of Sciences of the United States of America*, 119/32: e2201453119. DOI: 10.1073/pnas.2201453119

Di Mattia, J., Vernerey, M.-S., Yvon, M., Pirolles, E., Villegas, M., Gaafar, Y., Ziebell, H., et al. (2020). 'Route of a Multipartite Nanovirus across the Body of Its Aphid Vector', *Journal of Virology*, 94/9: e01998-19. DOI: 10.1128/JVI.01998-19

Francq, B. G., Lin, D., & Hoyer, W. (2019). 'Confidence, prediction, and tolerance in linear mixed models', *Statistics in Medicine*, 38/30: 5603–22. DOI: 10.1002/sim.8386

Franz, A. W., van der Wilk, F., Verbeek, M., Dullemans, A. M., & van den Heuvel, J. F. (1999). 'Faba bean necrotic yellows virus (genus Nanovirus) requires a helper factor for its aphid transmission', *Virology*, 262/1: 210–9. DOI: 10.1006/viro.1999.9904

Gaafar, Y., Grausgruber-Gröger, S., & Ziebell, H. (2016). 'Vicia faba, V. sativa and Lens culinaris as new hosts for Pea necrotic yellow dwarf virus in Germany and Austria', *New Disease Reports*, 34/1: 28–28. DOI: 10.5197/j.2044-0588.2016.034.028

Gallet, R., Di Mattia, J., Ravel, S., Zeddam, J.-L., Vitalis, R., Michalakis, Y., & Blanc, S. (2022). 'Gene copy number variations at the within-host population level modulate gene expression in a multipartite virus', *Virus Evolution*, 8/2: veac058. DOI: 10.1093/ve/veac058

Gallet, R., Kraberger, S., Filloux, D., Galzi, S., Fontes, H., Martin, D. P., Varsani, A., et al. (2018). 'Nanovirus-alphasatellite complex identified in Vicia cracca in the Rhône delta region of France', *Archives of Virology*, 163/3: 695–700. DOI: 10.1007/s00705-017-3634-4

Grigoras, I., Gronenborn, B., & Vetten, H. J. (2010). 'First Report of a Nanovirus Disease of Pea in Germany', *Plant Disease*, 94/5: 642. DOI: 10.1094/PDIS-94-5-0642C

Grigoras, Ioana, Ginzo, A. I. D. C., Martin, D. P., Varsani, A., Romero, J., Mammadov, A. C., Huseynova, I. M., et al. (2014). 'Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe', *The Journal of General Virology*, 95/Pt 5: 1178–91. DOI: 10.1099/vir.0.063115-0

Grigoras, Ioana, Timchenko, T., Katul, L., Grande-Pérez, A., Vetten, H.-J., & Gronenborn, B. (2009). 'Reconstitution of authentic nanovirus from multiple cloned DNAs', *Journal of Virology*, 83/20: 10778–87. DOI: 10.1128/JVI.01212-09

Grigoras, Ioana, Vetten, H.-J., Commandeur, U., Ziebell, H., Gronenborn, B., & Timchenko, T. (2018). 'Nanovirus DNA-N encodes a protein mandatory for aphid transmission', *Virology*, 522: 281–91. DOI: 10.1016/j.virol.2018.07.001

Gronenborn, B. (2004). 'Nanoviruses: genome organisation and protein function', *Veterinary Microbiology*, 98/2: 103–9. DOI: 10.1016/j.vetmic.2003.10.015

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: 307–21. DOI: 10.1093/sysbio/syq010

Gutiérrez, S., & Zwart, M. P. (2018). 'Population bottlenecks in multicomponent viruses: first forays into the uncharted territory of genome-formula drift', *Current Opinion in Virology*, 33: 184–90. DOI: 10.1016/j.coviro.2018.09.001

Guyot, V., Rajeswaran, R., Chu, H. C., Karthikeyan, C., Laboureau, N., Galzi, S., Mukwa, L. F. T., et al. (2022). 'A newly emerging alphasatellite affects banana bunchy top virus replication, transcription, siRNA production and transmission by aphids', *PLoS pathogens*, 18/4: e1010448. DOI: 10.1371/journal.ppat.1010448

Hasanvand, V., Heydanejad, J., Massumi, H., Kleinow, T., Jeske, H., Fontenele, R. S., Kraberger, S., et al. (2021). 'Genome characterization of parsley severe stunt-associated virus in Iran', *Virus Genes*, 57/3: 293–301. DOI: 10.1007/s11262-021-01835-y

Hassan-Sheikhi, P., Heydarnejad, J., Massumi, H., Kraberger, S., & Varsani, A. (2020). 'Novel nanovirus and associated alphasatellites identified in milk vetch plants with chlorotic dwarf disease in Iran', *Virus Research*, 276: 197830. DOI: 10.1016/j.virusres.2019.197830

Heydarnejad, J., Kamali, M., Massumi, H., Kvarnheden, A., Male, M. F., Kraberger, S., Stainton, D., et al. (2017). 'Identification of a Nanovirus-Alphasatellite Complex in Sophora alopecuroides', *Virus Research*, 235: 24–32. DOI: 10.1016/j.virusres.2017.03.023

Hu, Z., Zhang, X., Liu, W., Zhou, Q., Zhang, Q., Li, G., & Yao, Q. (2016). 'Genome segments accumulate with different frequencies in Bombyx mori bidensovirus', *Journal of Basic Microbiology*, 56/12: 1338–43. DOI: 10.1002/jobm.201600120

Kamran, A., Hou, H., Xie, Y., Zhao, C., Wei, X., Zhang, C., Yu, X., et al. (2019). 'Full genome sequence analysis and putative host-shifting of Milk vetch dwarf virus infecting tobacco (Nicotiana tabacum) in China', *Virology Journal*, 16/1: 38. DOI: 10.1186/s12985-019-1129-5

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., et al. (2012). 'Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data', *Bioinformatics (Oxford, England)*, 28/12: 1647–9. DOI: 10.1093/bioinformatics/bts199

Knierim, D., Barrière, Q., Grigoras, I., Winter, S., Vetten, H.-J., Schwinghamer, M., Thomas, J., et al. (2019). 'Subterranean Clover Stunt Virus Revisited: Detection of Two Missing Genome Components', *Viruses*, 11/2: 138. DOI: 10.3390/v11020138

Krenz, B., Schießl, I., Greiner, E., & Krapp, S. (2017). 'Analyses of pea necrotic yellow dwarf virus-encoded proteins', *Virus Genes*, 53/3: 454–63. DOI: 10.1007/s11262-017-1439-x

Kumari, S. G., Makkouk, K. M., Attar, N., Ghulam, W., & Lesemann, D.-E. (2004). 'First Report of Chickpea Chlorotic Dwarf Virus Infecting Spring Chickpea in Syria', *Plant Disease*, 88/4: 424. DOI: 10.1094/PDIS.2004.88.4.424C

Lageix, S., Catrice, O., Deragon, J.-M., Gronenborn, B., Pélissier, T., & Ramírez, B. C. (2007). 'The nanovirus-encoded Clink protein affects plant cell cycle regulation through interaction with the retinoblastoma-related protein', *Journal of Virology*, 81/8: 4177–85. DOI: 10.1128/JVI.02103-06

Lal, A., Vo, T. T. B., Sanjaya, I. G. N. P. W., Ho, P. T., Kim, J.-K., Kil, E.-J., & Lee, S. (2020). 'Nanovirus Disease Complexes: An Emerging Threat in the Modern Era', *Frontiers in Plant Science*, 11: 558403. DOI: 10.3389/fpls.2020.558403

Lefort, V., Longueville, J.-E., & Gascuel, O. (2017). 'SMS: Smart Model Selection in PhyML', *Molecular Biology and Evolution*, 34/9: 2422–4. DOI: 10.1093/molbev/msx149

Letunic, I., & Bork, P. (2021). 'Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation', *Nucleic Acids Research*, 49/W1: W293–6. DOI: 10.1093/nar/gkab301

Lotfipour, M., Behjatnia, S. A. A., Dall'Ara, M., & Ratti, C. (2020). 'The full-length genome characterization and diversity of faba bean necrotic stunt virus in Iran', *European Journal of Plant Pathology*, 157/2: 239–50. DOI: 10.1007/s10658-020-01973-3

Mandal, B., Shilpi, S., Barman, A. R., Mandal, S., & Varma, A. (2013). 'Nine novel DNA components associated with the foorkey disease of large cardamom: evidence of a distinct babuvirus species in Nanoviridae', *Virus Research*, 178/2: 297–305. DOI: 10.1016/j.virusres.2013.09.027

Mansourpour, M., Gallet, R., Abbasi, A., Blanc, S., Dizadji, A., & Zeddam, J.-L. (2022). 'Effects of an Alphasatellite on the Life Cycle of the Nanovirus Faba Bean Necrotic Yellows Virus', *Journal of Virology*, 96/3: e0138821. DOI: 10.1128/JVI.01388-21

Mokhtari, A., Nozari, J., Rezwani, A., Rasolian, G., Petrović-Obradović, O., & Rakhshani, E. (n.d.). 'APHIDS (HEMIPTERA: APHIDIDAE) ASSOCIATED WITH GRASSLANDS OF CENTRAL ALBORZ, IRAN',.

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den Hoff, M. J. B., & Moorman, A. F. M. (2009). 'Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data', *Nucleic Acids Research*, 37/6: e45. DOI: 10.1093/nar/gkp045

Ryckebusch, F., Peterschmitt, M., Granier, M., & Sauvion, N. (2021). 'Alfalfa leaf curl virus is efficiently acquired by its aphid vector Aphis craccivora but inefficiently transmitted', *The Journal of General Virology*, 102/2: 001516. DOI: 10.1099/jgv.0.001516

Sánchez-Navarro, J. A., Zwart, M. P., & Elena, S. F. (2013). 'Effects of the number of genome segments on primary and systemic infections with a multipartite plant RNA virus', *Journal of Virology*, 87/19: 10805–15. DOI: 10.1128/JVI.01402-13

Sharma, G., & Mathew, T. (2012). 'One-Sided and Two-Sided Tolerance Intervals in General Mixed and Random Effects Models Using Small-Sample Asymptotics', *Journal of the American Statistical Association*, 107/497: 258–67. Taylor & Francis. DOI: 10.1080/01621459.2011.640592

Sharman, M., Thomas, J. E., Skabo, S., & Holton, T. A. (2008). 'Abacá bunchy top virus, a new member of the genus Babuvirus (family Nanoviridae)', *Archives of Virology*, 153/1: 135–47. DOI: 10.1007/s00705-007-1077-z

Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakis, Y., Gutierrez, S., & Blanc, S. (2013). 'Gene copy number is differentially regulated in a multipartite virus', *Nature Communications*, 4: 2248. DOI: 10.1038/ncomms3248

Sicard, A., Zeddam, J.-L., Yvon, M., Michalakis, Y., Gutiérrez, S., & Blanc, S. (2015). 'Circulative Nonpropagative Aphid Transmission of Nanoviruses: an Oversimplified View', *Journal of Virology*, 89/19: 9719–26. DOI: 10.1128/JVI.00780-15

Stainton, D., Martin, D. P., Muhire, B. M., Lolohea, S., Halafihi, M., Lepoint, P., Blomme, G., et al. (2015). 'The global distribution of Banana bunchy top virus reveals little evidence for frequent recent, human-mediated long distance dispersal events', *Virus Evolution*, 1/1: vev009. DOI: 10.1093/ve/vev009

Sun, X., Xiao, L., Cheng, C., Shi, Z., Sun, Z., Ma, Y., & Zhu, X. (2022). 'Genome-wide analysis of milk vetch dwarf virus and related alphasatellites from garlic and aphids in Shandong Province, China', *Archives of Virology*, 167/10: 2071–7. DOI: 10.1007/s00705-022-05490-y

Tadesse, N., Ali, K., Gorfu, D., Abraham, A., Lencho, A., Ayalew, M., Yusuf, A., et al. (1999). 'First Report of Soybean Dwarf Virus Infecting Lentil and Beet Western Yellows Virus Infecting Lentil and Chickpea Crops in Ethiopia', *Plant Disease*, 83/6: 589. DOI: 10.1094/PDIS.1999.83.6.589B

Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H. J., & Gronenborn, B. (2000). 'The master rep concept in nanovirus replication: identification of missing genome components and potential for natural genetic reassortment', *Virology*, 274/1: 189–95. DOI: 10.1006/viro.2000.0439

Trapani, S., Bhat, E. A., Yvon, M., Lai-Kee-Him, J., Hoh, F., Vernerey, M.-S., Pirolles, E., et al. (2023). 'Structure-guided mutagenesis of the capsid protein indicates that a nanovirus requires assembled viral particles for systemic infection', *PLOS Pathogens*, 19/1: e1011086. DOI: 10.1371/journal.ppat.1011086

Urbino, C., Thébaud, G., Granier, M., Blanc, S., & Peterschmitt, M. (2008). 'A novel cloning strategy for isolating, genotyping and phenotyping genetic variants of geminiviruses', *Virology Journal*, 5: 135. DOI: 10.1186/1743-422X-5-135

Vega-Arreguín, J. C., Gronenborn, B., & Ramírez, B. C. (2007). 'Arabidopsis thaliana is a host of the legume nanovirus Faba bean necrotic yellows virus', *Virus Research*, 128/1–2: 81–7. DOI: 10.1016/j.virusres.2007.04.011

Vetten, H. J., Knierim, D., Rakoski, M. S., Menzel, W., Maiss, E., Gronenborn, B., Winter, S., et al. (2019). 'Identification of a novel nanovirus in parsley', *Archives of Virology*, 164/7: 1883–7. DOI: 10.1007/s00705-019-04280-3

Wu, B., Zwart, M. P., Sánchez-Navarro, J. A., & Elena, S. F. (2017). 'Within-host Evolution of Segments Ratio for the Tripartite Genome of Alfalfa Mosaic Virus', *Scientific Reports*, 7/1: 5004. DOI: 10.1038/s41598-017-05335-8

Yepes, L. M., Cieniewicz, E., Krenz, B., McLane, H., Thompson, J. R., Perry, K. L., & Fuchs, M. (2018). 'Causative Role of Grapevine Red Blotch Virus in Red Blotch Disease', *Phytopathology®*, 108/7: 902–9. Scientific Societies. DOI: 10.1094/PHYTO-12-17-0419-R

Yu, N.-T., Xie, H.-M., Zhang, Y.-L., Wang, J.-H., Xiong, Z., & Liu, Z.-X. (2019). 'Independent modulation of individual genomic component transcription and a cis-acting element related to high transcriptional activity in a multipartite DNA virus', *BMC genomics*, 20/1: 573. DOI: 10.1186/s12864-019-5901-0

Zhang, C., Zheng, H., Yan, D., Han, K., Song, X., Liu, Y., Zhang, D., et al. (2017). 'Complete genomic characterization of milk vetch dwarf virus isolates from cowpea and broad bean in Anhui province, China', *Archives of Virology*, 162/8: 2437–40. DOI: 10.1007/s00705-017-3348-7

Zwart, M. P., & Elena, S. F. (2020). 'Modeling multipartite virus evolution: the genome formula facilitates rapid adaptation to heterogeneous environments†', *Virus Evolution*, 6/1: veaa022. DOI: 10.1093/ve/veaa022

Supplementary Information for

Phenotypic characterization of single-segment reassortant between two isolates of the faba bean necrotic stunt virus (FBNSV) in multiple host species

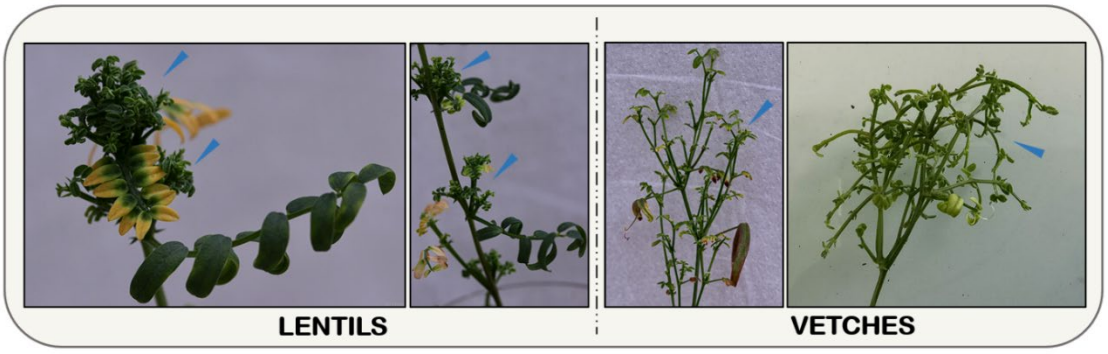
Babil Torralba, Andy Brousse, Prune Lacôte-Popovic, Stéphane Blanc and Yannis Michalakis

Species	Isolate	Segment	Number
FBNSV	Arak	C	MH626075.1
FBNSV	Arak	M	MH626079.1
FBNSV	Arak	N	MH626083.1
FBNSV	Arak	R	MH626087.1
FBNSV	Arak	S	MH626091.1
FBNSV	Arak	U1	MH626095.1
FBNSV	Arak	U2	MH626098.1
FBNSV	Arak	U4	MH626103.1
FBNSV	Avhaz	C	MH626072.1
FBNSV	Avhaz	M	MH626076.1
FBNSV	Avhaz	N	MH626080.1
FBNSV	Avhaz	R	MH626084.1
FBNSV	Avhaz	S	MH626088.1
FBNSV	Avhaz	U1	MH626092.1
FBNSV	Avhaz	U2	MH626096.1
FBNSV	Avhaz	U4	MH626100.1
FBNSV	AZ;1	C	KC978968.1
FBNSV	AZ;1	M	KC978969.1
FBNSV	AZ;1	N	KC978970.1
FBNSV	AZ;1	R	KC978966.1
FBNSV	AZ;1	S	KC978967.1
FBNSV	AZ;1	U1	KC978971.1
FBNSV	AZ;1	U2	KC978972.1
FBNSV	AZ;1	U4	KC978973.1
FBNSV	AZ;10	C	KC978976.1
FBNSV	AZ;10	M	KC978977.1
FBNSV	AZ;10	N	KC978978.1
FBNSV	AZ;10	R	KC978974.1
FBNSV	AZ;10	S	KC978975.1
FBNSV	AZ;10	U1	KC978979.1
FBNSV	AZ;10	U2	KC978980.1
FBNSV	AZ;10	U4	KC978981.1
FBNSV	AZ;12b	C	KC978984.1
FBNSV	AZ;12b	M	KC978985.1
FBNSV	AZ;12b	N	KC978986.1
FBNSV	AZ;12b	R	KC978982.1
FBNSV	AZ;12b	S	KC978983.1
FBNSV	AZ;12b	U1	KC978987.1
FBNSV	AZ;12b	U2	KC978988.1
FBNSV	AZ;12b	U4	KC978989.1
FBNSV	AZ;15	C	KC978994.1
FBNSV	AZ;15	M	KC978995.1
FBNSV	AZ;15	N	KC978996.1
FBNSV	AZ;15	R	KC978992.1
FBNSV	AZ;15	S	KC978993.1
FBNSV	AZ;15	U1	KC978997.1
FBNSV	AZ;15	U2	KC978998.1
FBNSV	AZ;15	U4	KC978999.1

Species	Isolate	Segment	Number
FBNSV	Ghaleseyed	C	MH626074.1
FBNSV	Ghaleseyed	M	MH626078.1
FBNSV	Ghaleseyed	N	MH626082.1
FBNSV	Ghaleseyed	R	MH626086.1
FBNSV	Ghaleseyed	S	MH626090.1
FBNSV	Ghaleseyed	U1	MH626094.1
FBNSV	Ghaleseyed	U2	MH626097.1
FBNSV	Ghaleseyed	U4	MH626102.1
FBNSV	Gorgan	C	MH626073.1
FBNSV	Gorgan	M	MH626077.1
FBNSV	Gorgan	N	MH626081.1
FBNSV	Gorgan	R	MH626085.1
FBNSV	Gorgan	S	MH626089.1
FBNSV	Gorgan	U1	MH626093.1
FBNSV	Gorgan	U2	MH626099.1
FBNSV	Gorgan	U4	MH626101.1
FBNSV	Hol-1997	C	GU983868.1
FBNSV	Hol-1997	M	GU983869.1
FBNSV	Hol-1997	N	GU983870.1
FBNSV	Hol-1997	R	GU983866.1
FBNSV	Hol-1997	S	GU983867.1
FBNSV	Hol-1997	U1	GU983871.1
FBNSV	Hol-1997	U2	GU983872.1
FBNSV	Hol-1997	U4	GU983873.1
FBNSV	JKI-1998/99	C	AJ749900.2
FBNSV	JKI-1998/99	M	AJ749896.2
FBNSV	JKI-1998/99	N	AJ749899.2
FBNSV	JKI-1998/99	R	AJ749894.2
FBNSV	JKI-1998/99	S	AJ749897.2
FBNSV	JKI-1998/99	U1	AJ749895.2
FBNSV	JKI-1998/99	U2	AJ749898.2
FBNSV	JKI-1998/99	U4	AJ749901.2
FBNSV	JKI-2000	C	NC_013096.1
FBNSV	JKI-2000	M	NC_013097.1
FBNSV	JKI-2000	N	NC_013098.1
FBNSV	JKI-2000	R	NC_013094.1
FBNSV	JKI-2000	S	NC_013095.1
FBNSV	JKI-2000	U1	NC_013099.1
FBNSV	JKI-2000	U2	NC_013100.1
FBNSV	JKI-2000	U4	NC_013101.1
FBNSV	Mor5	C	GQ274031.1
FBNSV	Mor5	M	GQ274035.1
FBNSV	Mor5	N	GQ274038.1
FBNSV	Mor5	R	GQ274033.1
FBNSV	Mor5	S	GQ274036.1
FBNSV	Mor5	U1	GQ274034.1
FBNSV	Mor5	U2	GQ274037.1
FBNSV	Mor5	U4	GQ274032.1

Supplementary Table 1 : Summary table containing the GenBank accession numbers for all the sequences of genomic segments used in this study.

Host plant	Score	Associated symptoms
Vicia faba	0	Physiological leaf (small curvature from tip to stem)
Vicia faba	0.5	Physiological leaf but small anomalies and curvatures, suspicions of infection
Vicia faba	1	Leaf curvature along the length of the leaf
Vicia faba	1.5	Pronounced leaf curvature, high suspicion of infection
Vicia faba	2	Obvious leaf deformation on last foliar level, confirmed infection
Vicia faba	2.5	Leaves are highly deformed, expanding to second foliar level
Vicia faba	3	Very highly deformed leaves, concerning last and second foliar level
Vicia faba	3.5	Very highly deformed leaves, expanding to previous foliar levels
Vicia faba	4	Very highly deformed leaves, multiple foliar levels are symptomatic
Vicia faba	4.5	Chlorosis of foliar level in addition to deformations
Vicia faba	5	Multiple foliar levels are symptomatic, deformed leaves, chlorosis installed, sometimes necrosis
Vicia sativa	0	No apparent symptoms, leaves can be slightly curved
Vicia sativa	0.5	Suspected curvatures of the leaves or small leaves development
Vicia sativa	1	Apparent symptoms, apex leaflets are curving on themselves
Vicia sativa	1.5	Multiple foliar levels are curving and start to twist
Vicia sativa	2	Pronounced leaves deformations. New branches start to develop at the nodes
Vicia sativa	2.5	New branches are developing and are symptomatic (small and/or twisted leaves)
Vicia sativa	3	Multiple branches are well developed and clearly symptomatic, bushy appearance starting
Vicia sativa	3.5	Bushing development
Vicia sativa	4	Extreme case of bushing
Lens culinaris	0	No apparent symptoms, leaves can be slightly curved
Lens culinaris	0.5	Suspected curvatures of the leaves or small leaves development
Lens culinaris	1	Apparent symptoms, apex leaflets are curving on themselves
Lens culinaris	1.5	Multiple foliar levels are curving and start to twist
Lens culinaris	2	New branches start to develop at the nodes
Lens culinaris	2.5	New branches are developing and are symptomatic (small and/or curled leaves)
Lens culinaris	3	Multiple branches are well developed and clearly symptomatic, bushy appearance starting
Lens culinaris	3.5	Bushing development
Lens culinaris	4	Extreme case of bushing



<u>Supplementary Table 2</u> : Symptom severity evaluation grid for the assessment of the progression of disease index scoring on *Vicia faba* (faba bean), *Vicia sativa* (vetch) and *Lens culinaris* (lentil). Bottom panel display extreme cases of bushing for both lentils and vetches with plants usually >30 dpi. Blue darts indicates the developments of multiple branches in stark contrast to the usual 2 to 3 stems of healthy plants.

ISOLATE	SEGMENT	SEQUENCE	SENSE	LENGTH	[ ] (µM)
JKI-2000	C - For	GACCCATCAACTGAAGACCTGTTGTCTC	F	28	0,3
JKI-2000	C - Rev	AGCAACTCTCTACATGCAAGGTCTTCCA	R	28	
JKI-2000	M - For	ACGCTGCGTATCAAGACGACGG	F	22	0,3
JKI-2000	M - Rev	TCCCGGAAGATAGCATGCAAGCA	R	23	
JKI-2000	N - For	GCATGAAAGACAAGCTCAACG	F	21	0,5
JKI-2000	N - Rev	GTTCCTGTTTCCACCATAGAAACTAC	R	26	
JKI-2000	R - For	GGAGAACTGAAGAAGAAGAAAGCTAT	F	26	0,5
JKI-2000	R - Rev	AAGCAAAGCATCAACCTCG	R	19	
JKI-2000	S - For	AGAAAGATGACGTGTCTGGTAG	F	22	0,3
JKI-2000	S - Rev	TCAGATAATTCACCAACTCTCCCG	R	24	
JKI-2000	U1 - For	AGCTATTGATTTGGAAGACAGAGT	F	24	0,5
JKI-2000	U1 - Rev	CGACAGCCAAGAAGCGT	R	17	
JKI-2000	U2 - For	ATTAGAGTAGATGAATGCAAAACTTGTAT	F	29	0,5
JKI-2000	U2 - Rev	TAAACCTTCTAGACTGCCATCGT	R	23	
JKI-2000	U4 - For	CCAGATTAAAGAAGTTACTGTCAGGTAATA	F	30	0,5
JKI-2000	U4 - Rev	ATTCAGCATCCTCGAACGG	R	19	
AZ;10_12b	C - For	ACCCATCAACTGAAGATTTG	F	20	1
AZ;10_12b	C - Rev	ACCATACTTAACTTGGAAGGG	R	21	
AZ;10_12b	M - For	AGGAACGATGTCGGATT	F	17	0,8
AZ;10_12b	M - Rev	ACCTATCATGTACAGAGCCT	R	20	
AZ;10_12b	N - For	AGTGTTAGCTTCAACGGGA	F	19	0,5
AZ;10_12b	N - Rev	ACGAACGTTTCTGTTTCCTC	R	20	
AZ;10_12b	R - For	GGAGTATGGAGAGTTCATTGTT	F	22	0,8
AZ;10_12b	R - Rev	GAACTCTCTTAATGTGCCTGAA	R	22	
AZ;10_12b	S - For	GTTTACAGCACCGGCTT	F	17	0,8
AZ;10_12b	S - Rev	ATACAGCCTTAACAGCCTCA	R	20	
AZ;10_12b	U1 - For	AGTTAGATGCCGTTAGTCG	F	19	0,8
AZ;10_12b	U1 - Rev	CAGACAACAGATCTGCGT	R	18	
AZ;10_12b	U2 - For	CTGGTTTTAATGGCTACCG	F	19	0,5
AZ;10_12b	U2 - Rev	GGTGCGTAACTGGGTAAC	R	18	
AZ;10_12b	U4 - For	TTCGAGGATGTTGATGCAG	F	19	0,5
AZ;10_12b	U4 - Rev	TCGATTCCCATAGGCGTT	R	18	
AZ;15	C - For	CAAGATGCTTAGTCAAGTGGAA	F	22	1,5
AZ;15	C - Rev	CAGTGAATAACTTATCTAATTCACTGC	R	27	
AZ;15	M - For	ATCAAGACGACGGAGATG	F	18	1
AZ;15	M - Rev	CTTCTTGACAAATCCAGGTAG	R	21	
AZ;15	N - For	GTTAAATGATTCAAGGAAGGTGTTAC	F	26	1,5
AZ;15	N - Rev	GACCACGAACATTTCTATTTCCG	R	23	
AZ;15	R - For	ATGGTGAGTTCGTTACGTC	F	19	1
AZ;15	R - Rev	CCTCTAAATTCTCTCAACGTACTA	R	24	
AZ;15	S - For	ACGACGATATTCAAGGCCA	F	19	1
AZ;15	S - Rev	CCTGCAACACCATTAGATTCC	R	21	
AZ;15	U1 - For	CTCTGGTTAGTCGAGCTTC	F	19	1
AZ;15	U1 - Rev	CTGAGAATCATCTTCACAACATGG	R	21	
AZ;15	U2 - For	CTTCATTTCATCTTGTTCTATTCGC	F	25	1
AZ;15	U2 - Rev	CCAAACGAGTAGGAGCATATG	R	21	
AZ;15	U4 - For	GTTGTTAAATCCTTCTTTGGTGG	F	23	1
AZ;15	U4 - Rev	CGTCTTCTTTCTTCTTACCCATG	R	23	

Supplementary Table 3 : Summary table presenting the pairs of primers employed for the targeted amplification of the genomic segments from [JKI-2000], [AZ;15] and [AZ;10_12b] isolates. The "SENSE" column specifies whether the corresponding primer is in the forward (F) or reverse (R) orientation. The final column [ ] indicates the recommended primer concentration (µm) for use in the qPCR mix.

Isolate	Segment	Host	CT	n
JKI-2000	C	Faba bean	31.672	28
JKI-2000	M	Faba bean	30.264	23
JKI-2000	N	Faba bean	32.185	31
JKI-2000	R	Faba bean	32.481	29
JKI-2000	S	Faba bean	32.073	21
JKI-2000	U1	Faba bean	30.707	40
JKI-2000	U2	Faba bean	28.813	34
JKI-2000	U4	Faba bean	32.172	42
JKI-2000	C	Lentil	30.863	21
JKI-2000	M	Lentil	31.337	16
JKI-2000	N	Lentil	32.646	21
JKI-2000	R	Lentil	29.442	22
JKI-2000	S	Lentil	35.436	17
JKI-2000	U1	Lentil	31.615	22
JKI-2000	U2	Lentil	35.265	21
JKI-2000	U4	Lentil	33.853	20
JKI-2000	C	Vetch	25.712	21
JKI-2000	M	Vetch	31.774	22
JKI-2000	N	Vetch	25.697	22
JKI-2000	R	Vetch	30.610	22
JKI-2000	S	Vetch	32.604	21
JKI-2000	U1	Vetch	26.463	22
JKI-2000	U2	Vetch	32.819	22
JKI-2000	U4	Vetch	26.689	22

Isolate	Segment	Host	CT	n
AZ;10_12b	C	Faba bean	31.281	18
AZ;10_12b	M	Faba bean	31.744	9
AZ;10_12b	N	Faba bean	31.733	37
AZ;10_12b	R	Faba bean	33.685	33
AZ;10_12b	S	Faba bean	34.000	41
AZ;10_12b	U1	Faba bean	33.032	27
AZ;10_12b	U2	Faba bean	33.851	34
AZ;10_12b	U4	Faba bean	32.410	42
AZ;10_12b	C	Lentil	32.041	19
AZ;10_12b	M	Lentil	32.800	18
AZ;10_12b	N	Lentil	33.520	22
AZ;10_12b	R	Lentil	30.544	19
AZ;10_12b	S	Lentil	31.271	22
AZ;10_12b	U1	Lentil	31.470	20
AZ;10_12b	U2	Lentil	32.047	17
AZ;10_12b	U4	Lentil	33.090	22
AZ;10_12b	C	Vetch	29.572	21
AZ;10_12b	M	Vetch	30.494	20
AZ;10_12b	N	Vetch	31.040	22
AZ;10_12b	R	Vetch	30.673	22
AZ;10_12b	S	Vetch	27.917	22
AZ;10_12b	U1	Vetch	29.873	22
AZ;10_12b	U2	Vetch	30.571	16
AZ;10_12b	U4	Vetch	31.877	22
AZ;15	C	Faba bean	28.273	24
AZ;15	M	Faba bean	34.018	9
AZ;15	N	Faba bean	31.090	23
AZ;15	R	Faba bean	31.128	7
AZ;15	S	Faba bean	26.337	9
AZ;15	U1	Faba bean	33.791	41
AZ;15	U2	Faba bean	29.027	42
AZ;15	U4	Faba bean	30.543	34
AZ;15	C	Lentil	32.815	15
AZ;15	M	Lentil	30.380	12
AZ;15	N	Lentil	29.272	19
AZ;15	R	Lentil	31.128	7
AZ;15	S	Lentil	30.745	12
AZ;15	U1	Lentil	32.447	21
AZ;15	U2	Lentil	30.499	22
AZ;15	U4	Lentil	33.492	21
AZ;15	C	Vetch	27.054	22
AZ;15	M	Vetch	27.874	22
AZ;15	N	Vetch	23.530	22
AZ;15	R	Vetch	29.020	21
AZ;15	S	Vetch	26.395	22
AZ;15	U1	Vetch	28.228	22
AZ;15	U2	Vetch	27.913	22
AZ;15	U4	Vetch	26.637	22

Supplementary Table 4 : Threshold CT values that were statistically determined for each genomic segment of the respective [JKI-2000], [AZ;10_12b] and [AZ;15] isolates across all assessed plant host species. The "Host" column specifies the host plant species (FB = faba bean, L = lentil, V = vetch) for which the thresholds (CT) were statistically determined. The column labeled 'n' indicates the number of plants used to establish these thresholds.

A

Traits	Host	Source	SumSq	Df	F.Value	Pr(>F)
Height	FB	(Intercept)	7,189	1,000	556,104	<0,001
		Genotype	0,048	2,000	1,862	0,164
		Residuals	0,853	66,000	NA	NA
Height	V	(Intercept)	15,730	1,000	3536,464	<0,001
		Genotype	0,139	2,000	15,579	<0,001
		Residuals	0,200	45,000	NA	NA
Height	L	(Intercept)	46,575	1,000	92562,520	<0,001
		Genotype	0,021	2,000	21,020	<0,001
		Residuals	0,039	77,000	NA	NA
Fresh Weight	FB	(Intercept)	4,803	1,000	134,577	<0,001
		Genotype	0,030	2,000	0,424	0,656
		Residuals	2,213	62,000	NA	NA
Fresh Weight	V	(Intercept)	28,954	1,000	7894,478	<0,001
		Genotype	0,056	2,000	7,670	0,001
		Residuals	0,161	44,000	NA	NA
Dry Weight	FB	(Intercept)	17,659	1,000	500,129	<0,001
		Genotype	0,759	2,000	10,741	<0,001
		Residuals	3,037	86,000	NA	NA
Dry Weight	V	(Intercept)	27,674	1,000	10409,355	<0,001
		Genotype	0,041	2,000	7,795	0,001
		Residuals	0,122	46,000	NA	NA
Dry Weight	L	(Intercept)	60,727	1,000	24183,585	<0,001
		Genotype	0,029	2,000	5,718	0,005
		Residuals	0,226	90,000	NA	NA

B

Traits	Host	Contrast	Diff	Lower	Upper	P.Value	SE	Df
Height	V	AZ;10_12b – AZ;15	0,110	0,060	0,159	<0,001	0,021	45,000
		AZ;15 – JKI-2000	-0,117	-0,198	-0,036	0,004	0,034	45,000
		AZ;10_12b – JKI-2000	-0,007	-0,085	0,072	0,976	0,033	45,000
Height	L	AZ;10_12b – AZ;15	0,035	0,021	0,050	<0,001	0,006	77,000
		AZ;15 – JKI-2000	-0,039	-0,055	-0,023	<0,001	0,007	77,000
		AZ;10_12b – JKI-2000	-0,004	-0,018	0,011	0,829	0,006	77,000
Fresh Weight	V	AZ;10_12b – AZ;15	0,059	0,012	0,105	0,010	0,019	44,000
		AZ;15 – JKI-2000	-0,091	-0,156	-0,027	0,004	0,027	44,000
		AZ;10_12b – JKI-2000	-0,033	-0,097	0,031	0,433	0,026	44,000
Dry Weight	FB	AZ;10_12b – AZ;15	0,133	0,006	0,259	0,038	0,053	86,000
		AZ;15 – JKI-2000	-0,211	-0,320	-0,102	<0,001	0,046	86,000
		AZ;10_12b – JKI-2000	-0,079	-0,201	0,043	0,279	0,051	86,000
Dry Weight	V	AZ;10_12b – AZ;15	0,052	0,013	0,090	0,006	0,016	46,000
		AZ;15 – JKI-2000	-0,076	-0,131	-0,021	0,005	0,023	46,000
		AZ;10_12b – JKI-2000	-0,024	-0,078	0,029	0,519	0,022	46,000
Dry Weight	L	AZ;10_12b – AZ;15	0,044	0,013	0,075	0,003	0,013	90,000
		AZ;15 – JKI-2000	-0,032	-0,065	0,002	0,065	0,014	90,000
		AZ;10_12b – JKI-2000	0,012	-0,017	0,041	0,577	0,012	90,000

Supplementary Table 5 : [A] Type III test of fixed effects for standardized traits ratio model with genotype as fixed effect. Significative p.values are colored in green (<0.05). **[B] Pairwise comparisons of standardized traits ratio between [AZ;10_12b], [AZ;15] and [JKI-2000] using Emmeans.** Negative "Diff" values mean a more pronounced effect on plant height (more reduction) of the reassortant, compared to the major parent (control) in the same host. Significative p.values are colored in green (<0.05).

PLANT VIRAL LOAD

A

Host	Source	Df	F	Pr(>F)
	(Intercept)	1,000	150469,882	0,000
FB	Genotype	2,000	22,070	<0,001
	Residuals	121,000	NA	NA
	(Intercept)	1,000	70369,464	0,000
V	Genotype	2,000	10,308	<0,001
	Residuals	42,000	NA	NA
	(Intercept)	1,000	77370,493	0,000
L	Genotype	2,000	41,833	<0,001
	Residuals	94,000	NA	NA

B

Host	Contrast	Diff	Lower	Upper	P.Value	Df
	AZ;15 - AZ;10_12b	0,000	0,000	0,001	0,019	121,000
FB	AZ;15 - JKI-2000	0,000	0,000	0,001	<0,001	121,000
	AZ;10_12b - JKI-2000	0,000	0,000	0,000	0,012	121,000
	AZ;15 - AZ;10_12b	-0,002	-0,003	-0,001	0,003	42,000
V	AZ;15 - JKI-2000	0,001	0,000	0,002	0,152	42,000
	AZ;10_12b - JKI-2000	0,003	0,002	0,004	<0,001	42,000
	AZ;15 - AZ;10_12b	-0,002	-0,004	0,000	0,075	94,000
L	AZ;15 - JKI-2000	0,003	0,001	0,005	<0,001	94,000
	AZ;10_12b - JKI-2000	0,005	0,003	0,007	<0,001	94,000

APHID VIRAL LOAD

C

Host	Df	F	Pr(>F)
(Intercept)	1,000	21805,700	<0,001
Genotype	2,000	51,764	<0,001
Host	2,000	9,274	<0,001
Genotype:Host	4,000	0,917	0,457
Residuals	106,000	NA	NA

D

Host	Contrast	Estimate	SE	Df	Lower	Upper	T.Ratio	P.Value
	AZ;15 - AZ;10_12b	-6,59E-06	4,84E-05	1,06E+02	-1,21E-04	1,07E-04	-1,36E-01	9,89E-01
F	AZ;15 - JKI-2000	8,62E-05	3,54E-05	1,06E+02	2,74E-06	1,70E-04	2,43E+00	4,13E-02
	AZ;10_12b - JKI-2000	9,28E-05	3,55E-05	1,06E+02	9,10E-06	1,77E-04	2,61E+00	2,62E-02
	AZ;15 - AZ;10_12b	-6,96E-06	1,97E-05	1,06E+02	-5,33E-05	3,94E-05	-3,52E-01	9,30E-01
L	AZ;15 - JKI-2000	4,62E-05	1,20E-05	1,06E+02	1,81E-05	7,43E-05	3,86E+00	6,65E-04
	AZ;10_12b - JKI-2000	5,32E-05	1,58E-05	1,06E+02	1,60E-05	9,04E-05	3,36E+00	2,72E-03
	AZ;15 - AZ;10_12b	9,02E-07	2,76E-05	1,06E+02	-6,41E-05	6,59E-05	3,27E-02	9,99E-01
V	AZ;15 - JKI-2000	6,45E-05	2,10E-05	1,06E+02	1,52E-05	1,14E-04	3,08E+00	7,05E-03
	AZ;10_12b - JKI-2000	6,36E-05	1,89E-05	1,06E+02	1,91E-05	1,08E-04	3,36E+00	2,86E-03

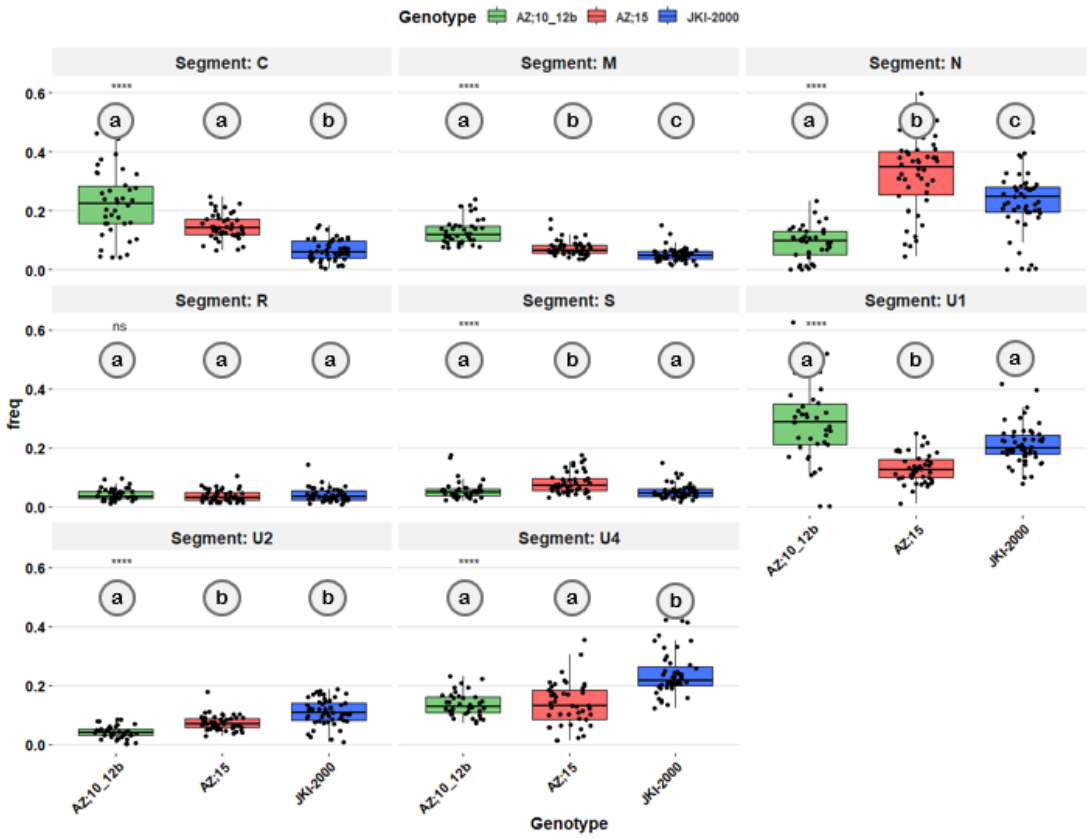
TRANSMISSION RATES

E

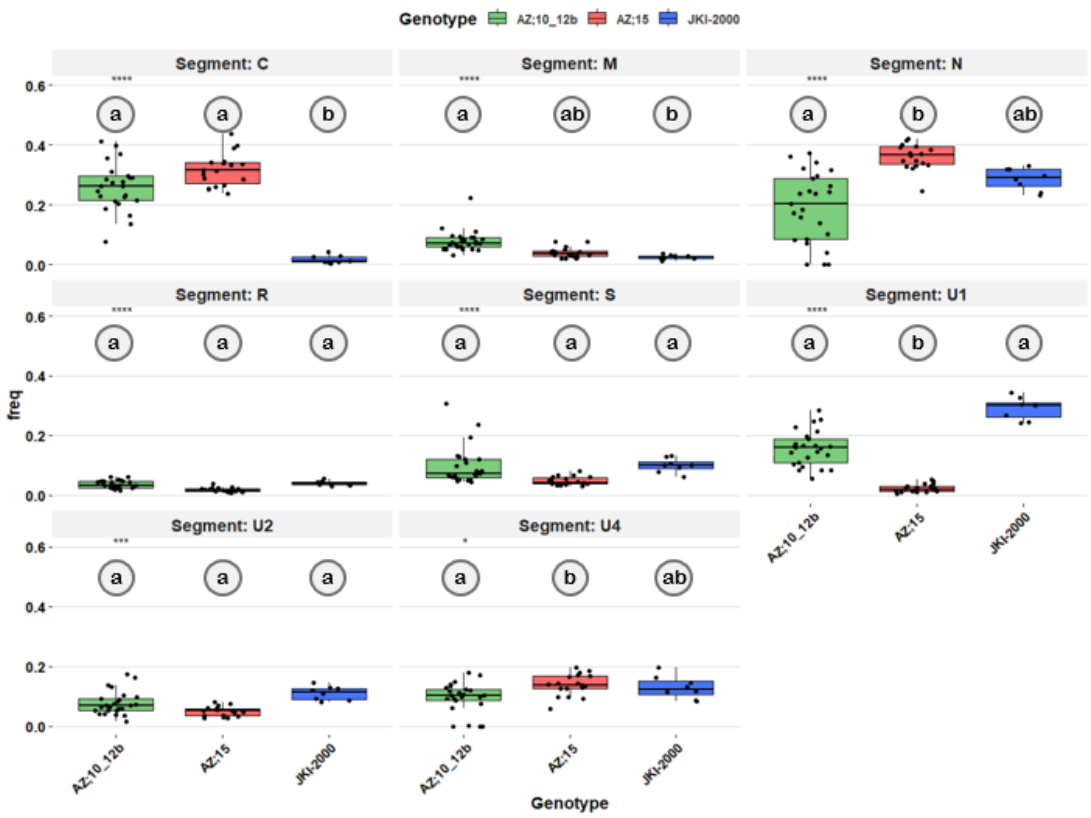
Traits	Host	Estimate	SD	Z.Value	Pr(>|Z|)
	(Intercept) AZ;15	1,174	0,277	4,231	<0,001
FB	AZ;10_12b	1,421	0,540	2,630	0,009
	JKI-2000	-0,481	0,374	-1,288	0,198
	(Intercept) AZ;15	-0,693	0,306	-2,264	0,024
V	AZ;10_12b	0,860	0,422	2,041	0,041
	JKI-2000	-1,459	0,563	-2,591	0,010
	(Intercept) AZ;15	-1,768	0,409	-4,322	0,000
L	AZ;10_12b	0,979	0,514	1,905	0,057
	JKI-2000	-2,082	1,090	-1,910	5,610

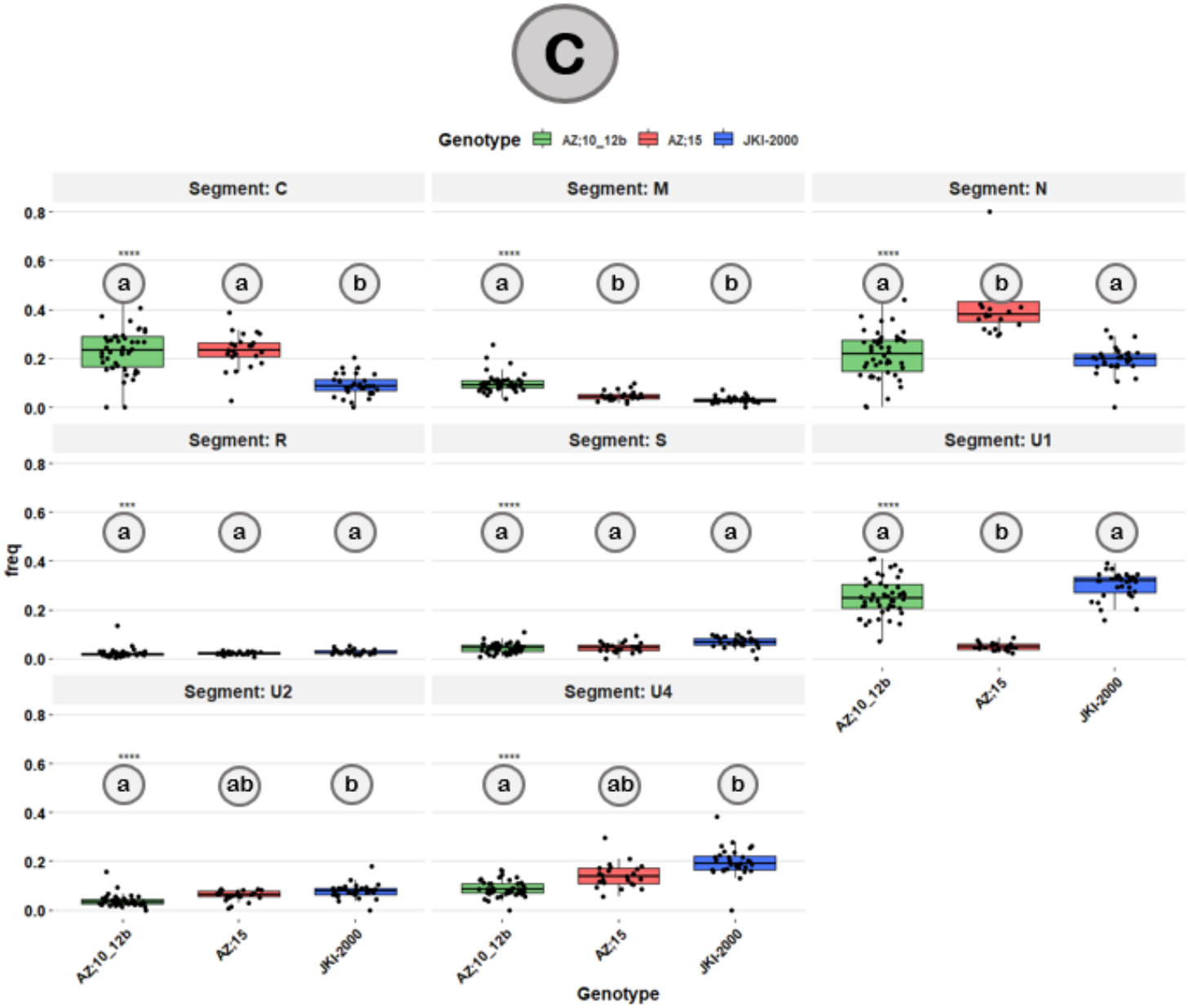
<u>Supplementary Table 6 :</u> [A, C] Type III test of fixed effects for standardized viral load model with genotype as fixed effect only for plant viral load, genotype and host as fixed effect for aphid viral load. Significative p.values are colored in green (<0.05). **[B, D] Pairwise comparisons of standardized traits ratio between [AZ;10_12b], [AZ;15] and [JKI-2000] using Emmeans.** Negative "Diff" values mean a more pronounced effect on plant height (more reduction) of the reassortant, compared to the major parent (control) in the same host. Significative p.values are colored in green (<0.05).

A

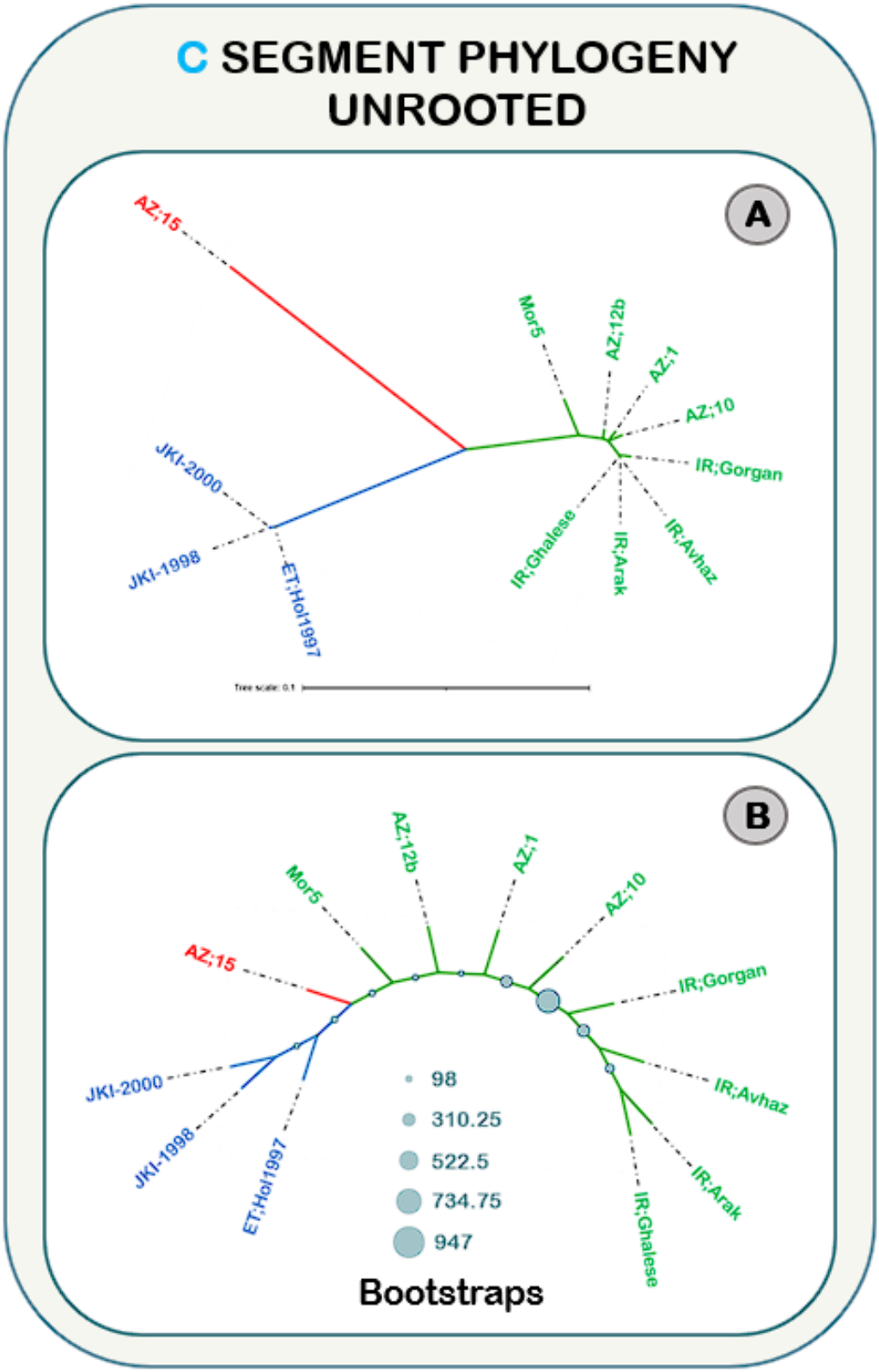

Genotype
AZ;10_12b
AZ;15
JKI-2000
Segment: C
Segment: M
Segment: N
Segment: R
Segment: S
Segment: U1
Segment: U2
Segment: U4
freq
Genotype

B

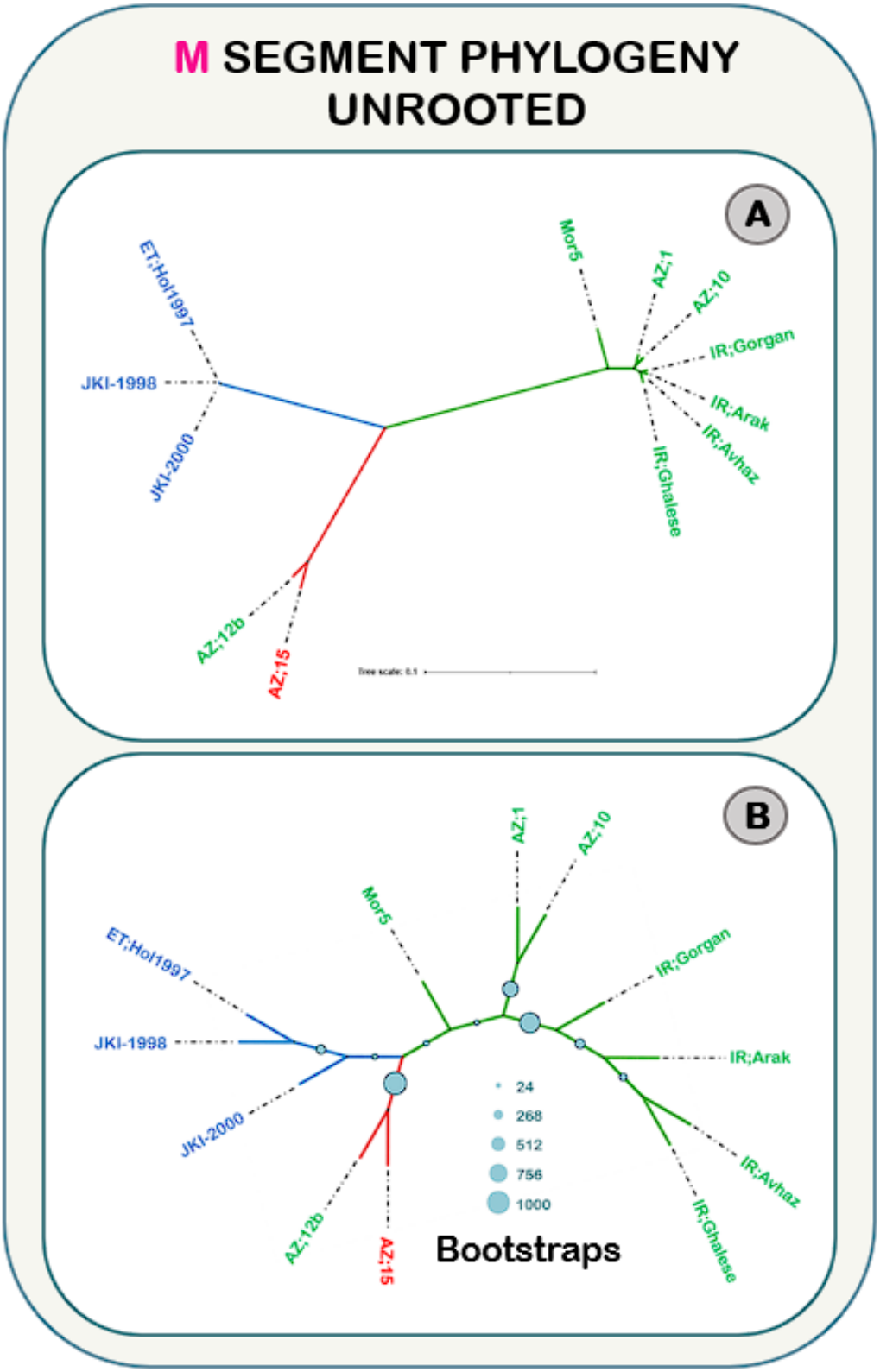

Genotype
AZ;10_12b
AZ;15
JKI-2000
Segment: C
Segment: M
Segment: N
Segment: R
Segment: S
Segment: U1
Segment: U2
Segment: U4
freq
Genotype



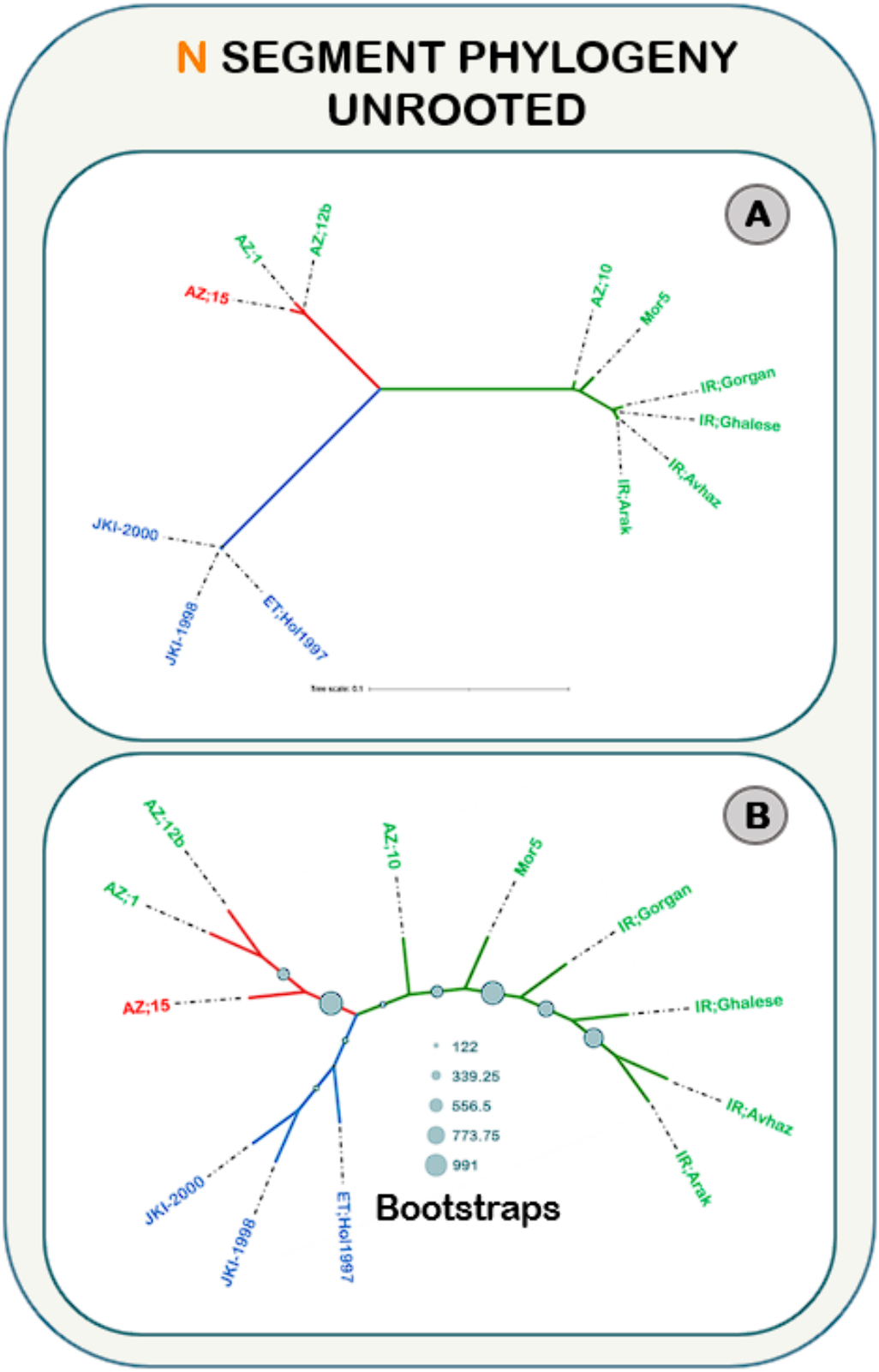
<u>Supplementary Figure 1 :</u> **[A] Genome formula comparisons in faba bean as boxplots of each genomic segment relative frequency between all three isolates.** The outcomes of the statistical analysis are represented by statistical groups labeled as (a,b). **[B] Genome formula comparisons in vetch as boxplots of each genomic segment relative frequency between all three isolates.** The outcomes of the statistical analysis are represented by statistical groups labeled as (a,b). **[C] Genome formula comparisons in lentils as boxplots of each genomic segment relative frequency between all three isolates.** The outcomes of the statistical analysis are represented by statistical groups labeled as (a,b).



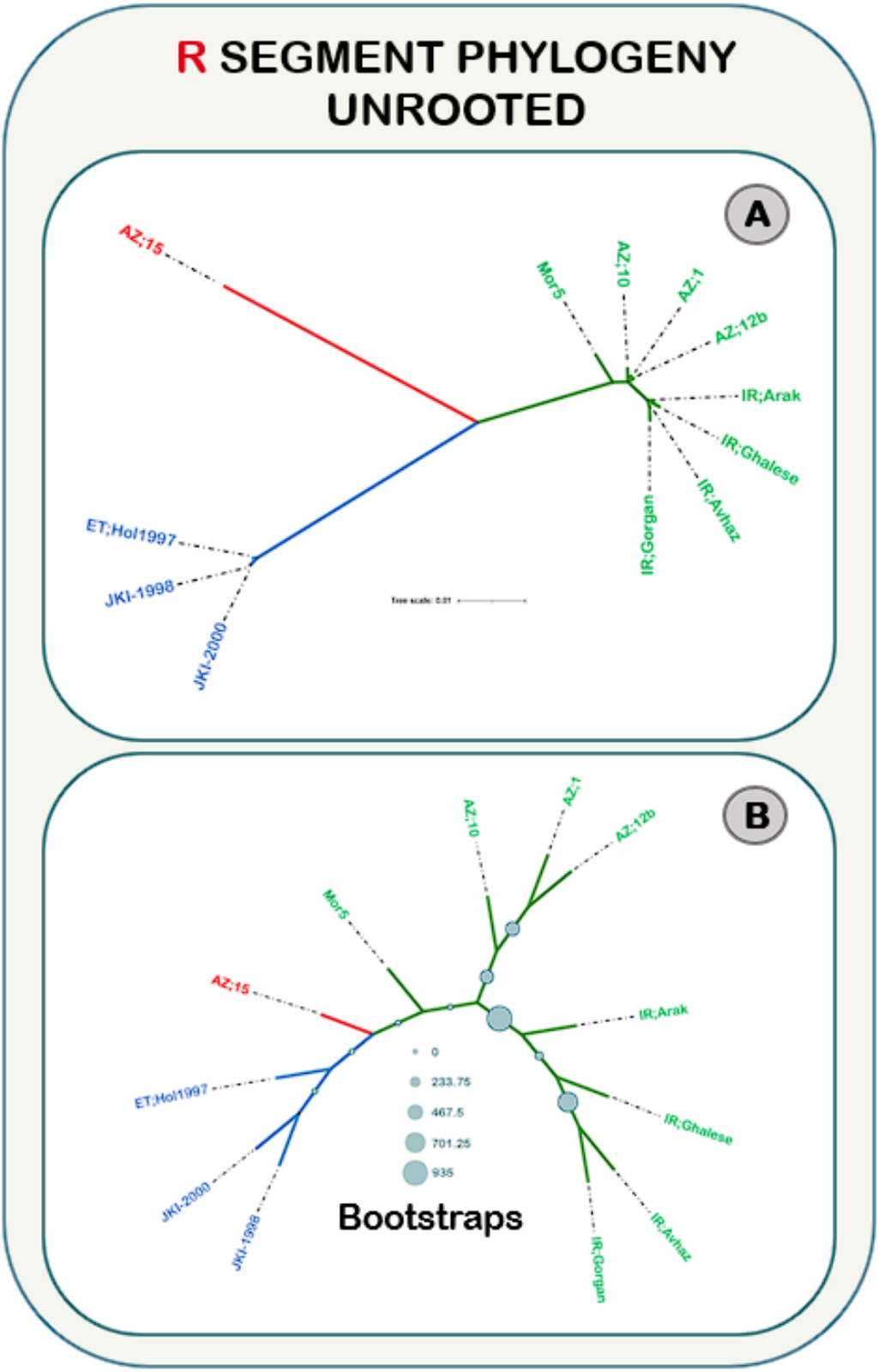
Supplementary Figure 2 : [A] Unrooted phylogenetic tree of FBNSV C genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV C genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



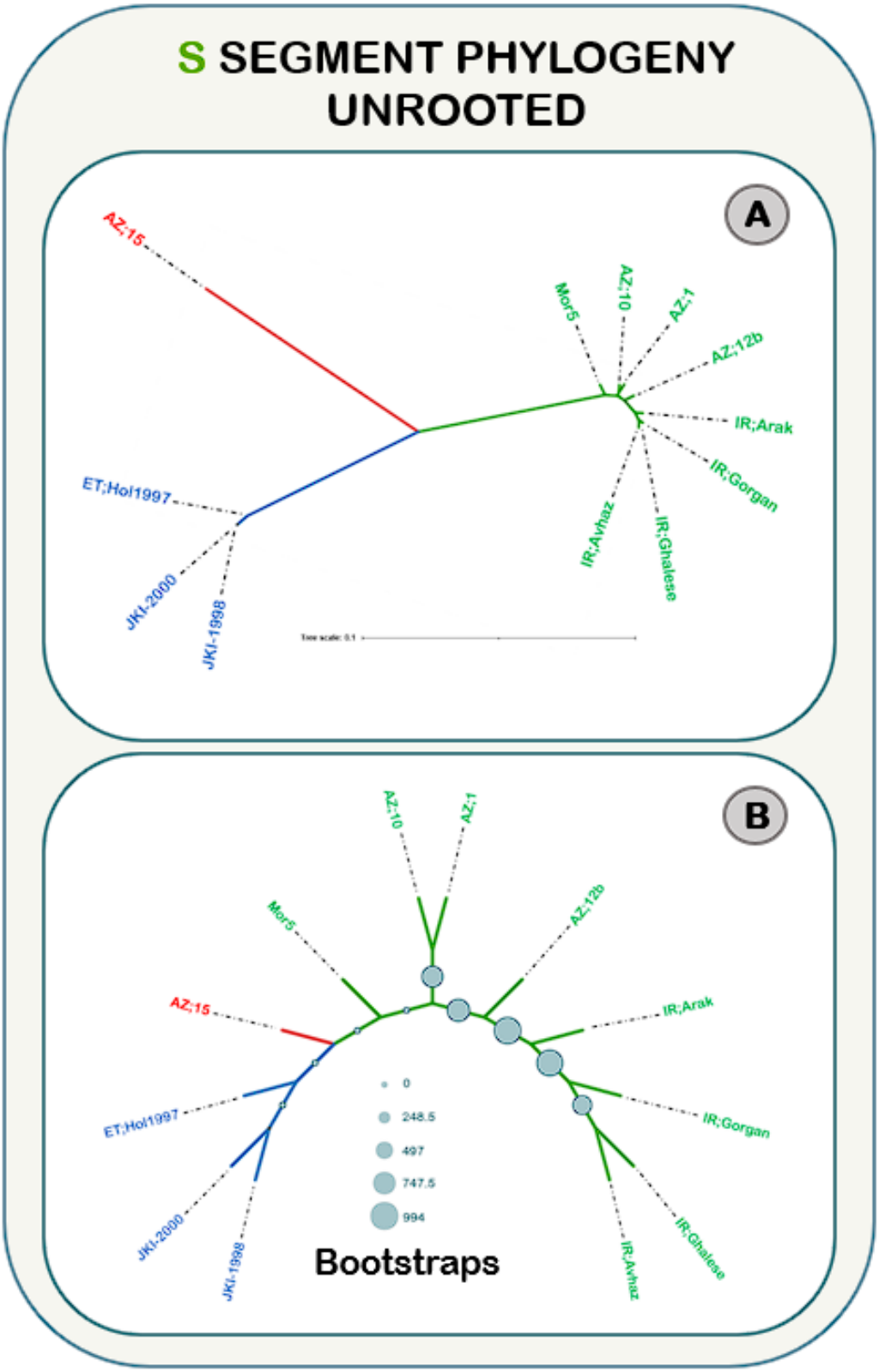
Supplementary Figure 3 : [A] Unrooted phylogenetic tree of FBNSV M genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV M genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



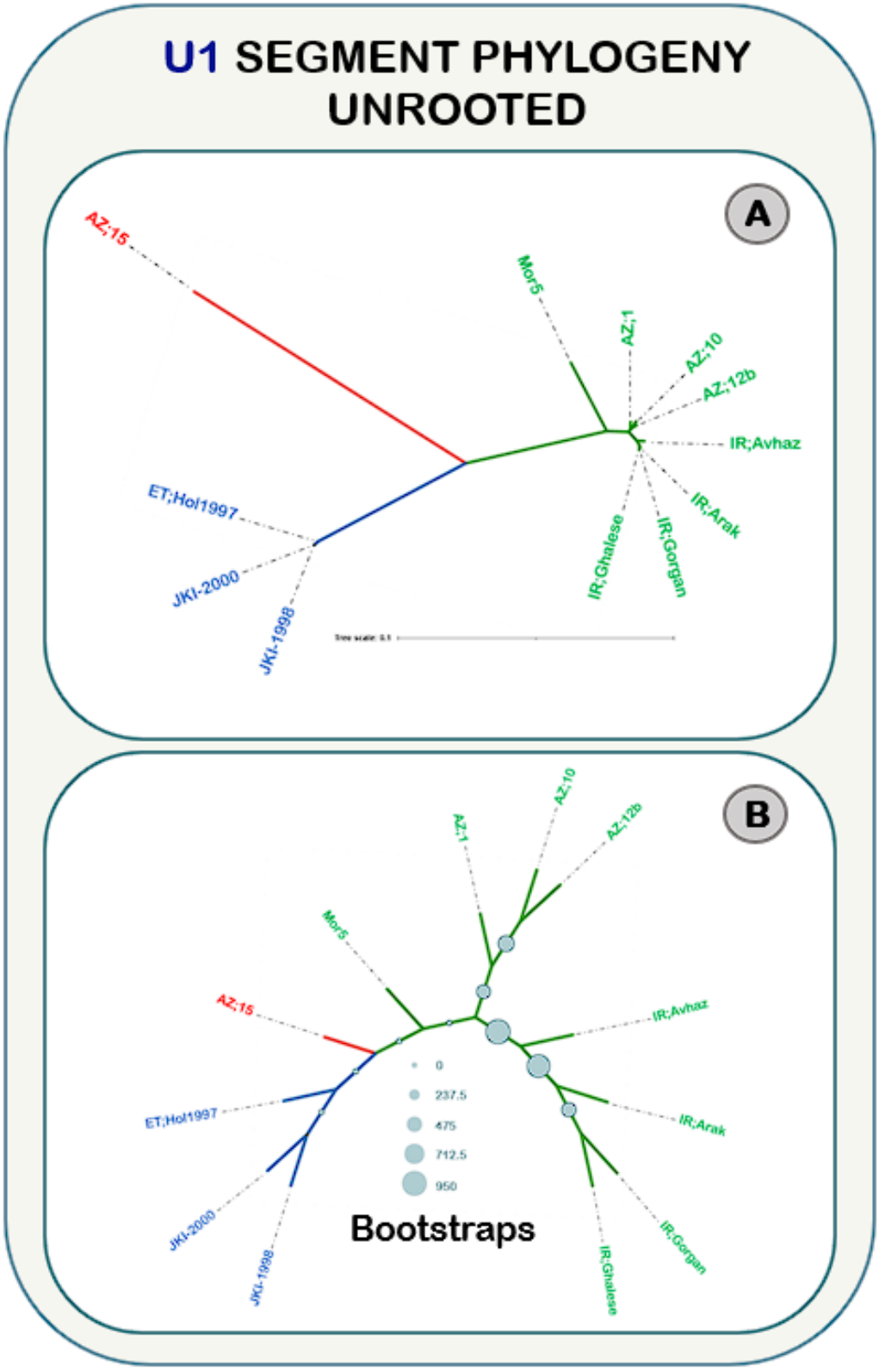
Supplementary Figure 4 : [A] Unrooted phylogenetic tree of FBNSV N genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV N genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



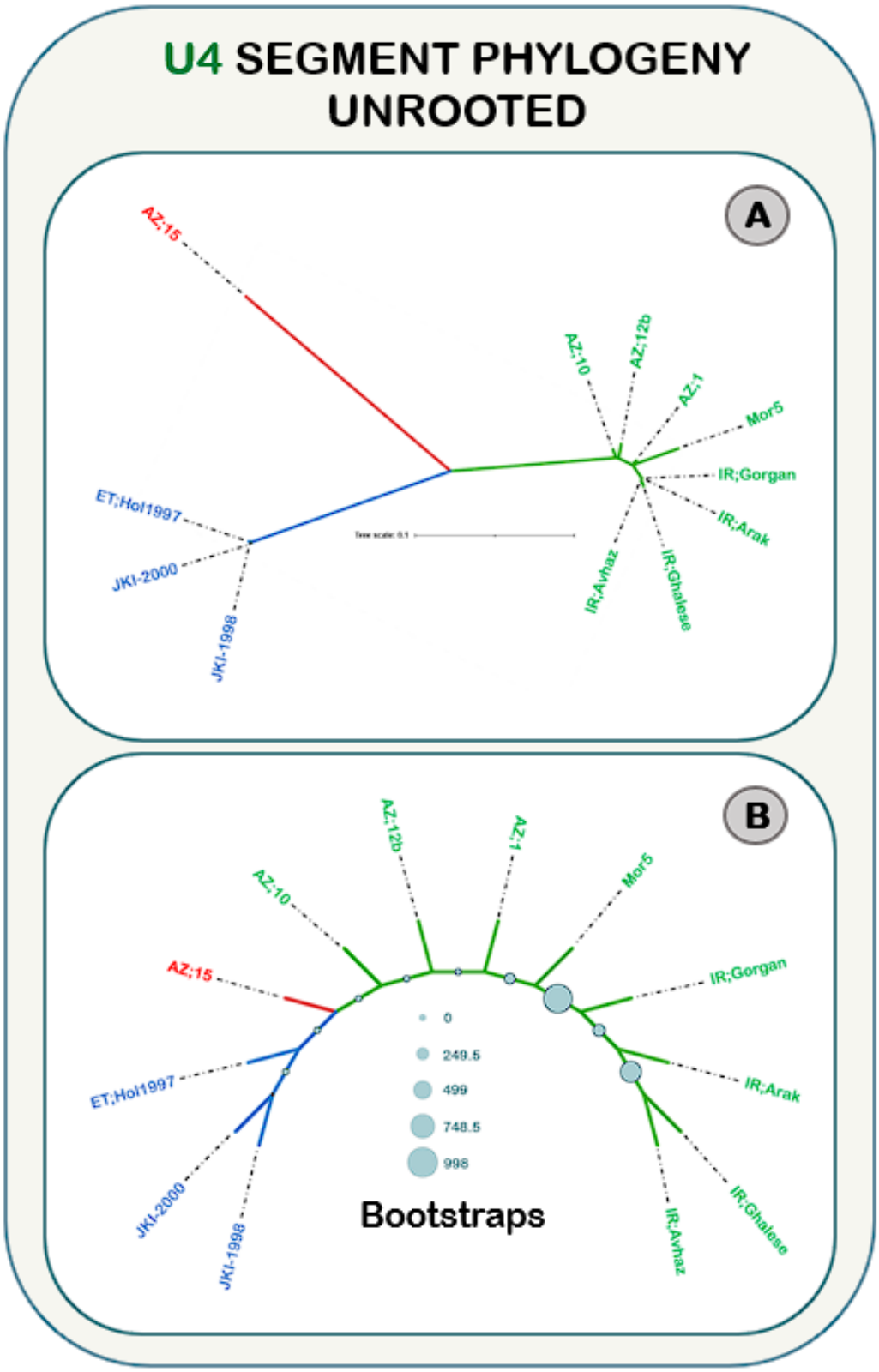
Supplementary Figure 5 : [A] Unrooted phylogenetic tree of FBNSV R genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV R genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



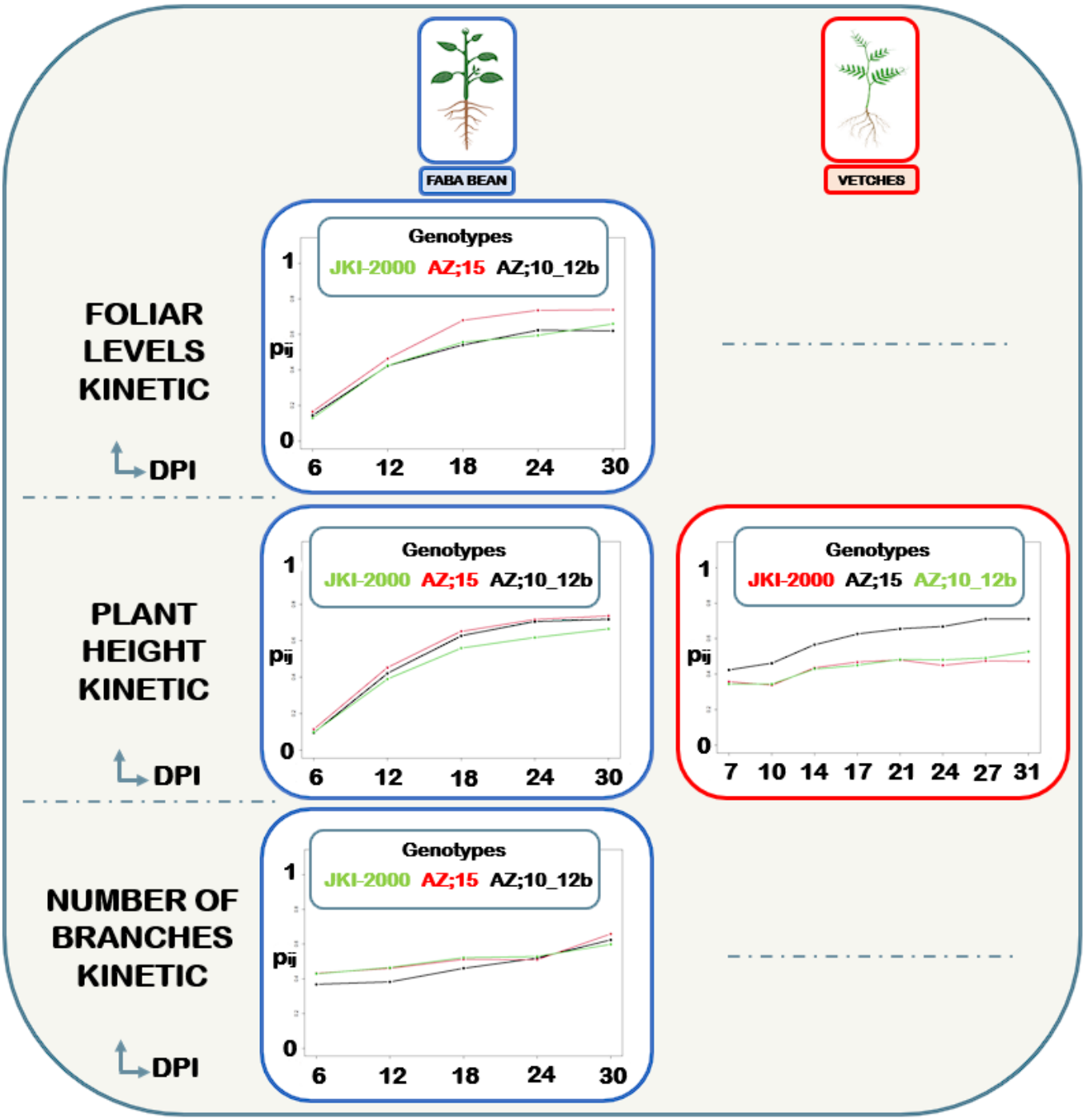
Supplementary Figure 6 : [A] Unrooted phylogenetic tree of FBNSV S genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV S genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



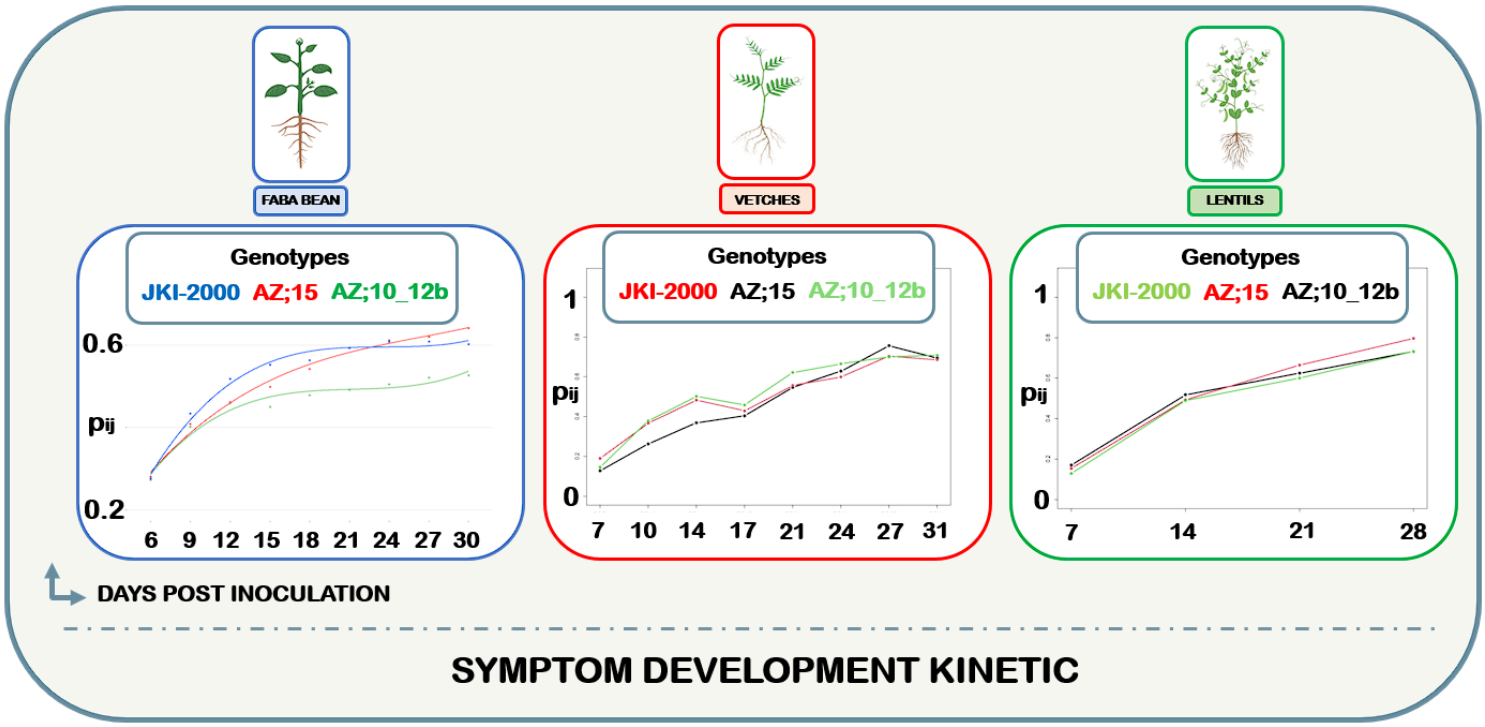
Supplementary Figure 7 : [A] Unrooted phylogenetic tree of FBNSV U1 genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV U1 genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



Supplementary Figure 8 : [A] Unrooted phylogenetic tree of FBNSV U4 genomic segments from available complete isolates. Main clades are distinctively represented in green, red and blue line colors based on their grouping when taking into account genetic distance. The labels of isolates are color-coded according to the clade they belong to when their genomes are concatenated. **[B] Unrooted cladogram of FBNSV U4 genomic segments from available complete isolates.** Bootstrap analysis (n=1000) is represented on branches as circles proportional to their respective values.



<u>Supplementary Figure 9</u> : Graphical plots of the number of foliar levels, branches and plant height kinetics. X-axis indicates the days post inoculation (DPI). Y-axis indicates the produce relative treatment effects (see Methods) . Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored lines described in the genotype panels.

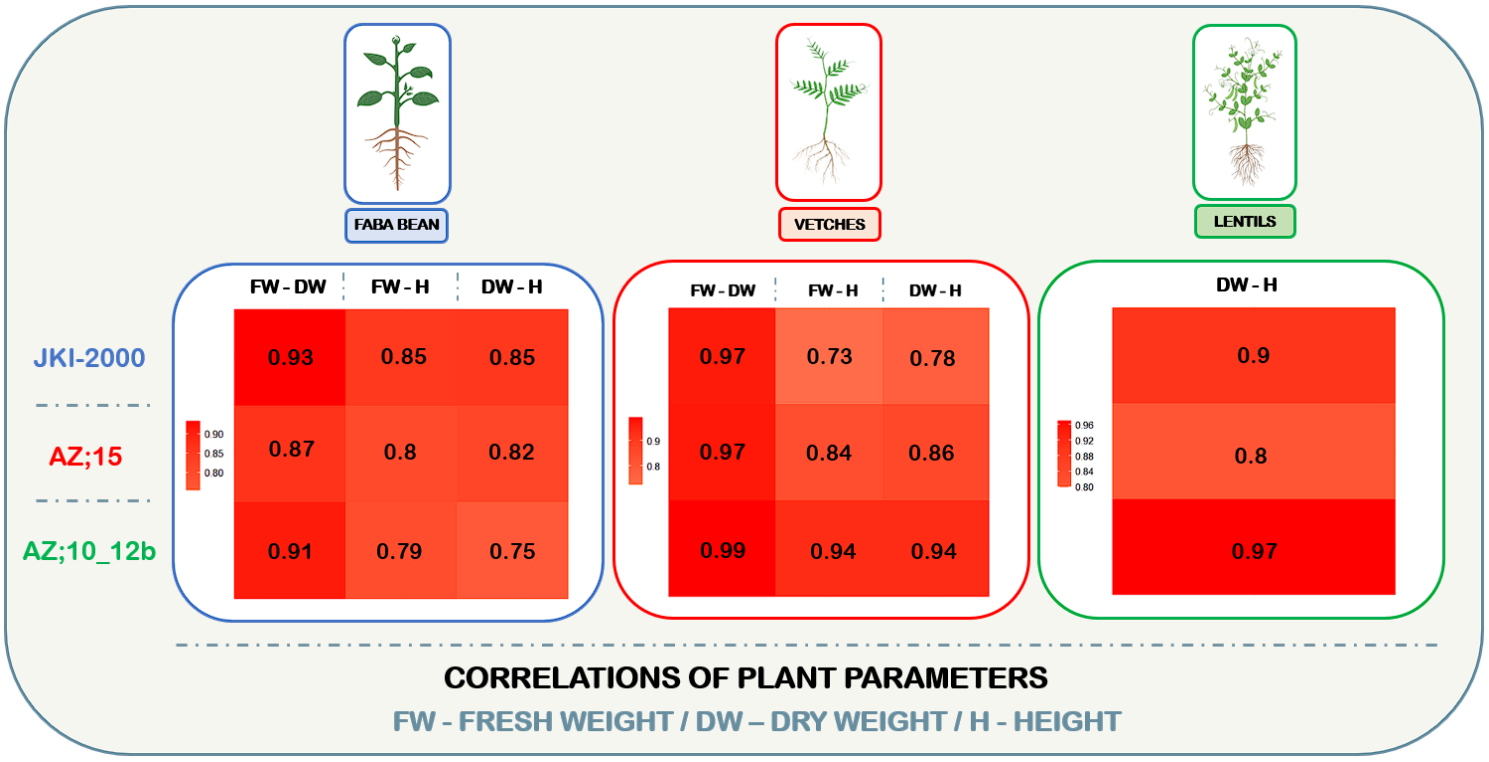


<u>Supplementary Figure 10</u> : Graphical plots of symptom development kinetics based on the evaluation grid from (Supp Table 2). X-axis indicates the days post inoculation. Y-indicates the produce relative treatment effects (see Methods). Panel colors indicate the host in which the data was recovered. Each panel compares all three isolates with the corresponding-colored lines described in the genotype panels.

ISOLATE	HOST	VAR	INOC	PLANTS	SUS	SUS IR	OBS SYMPT
AZ;15	*Vicia sativa*	JOSE	*Aphis craccivora*	11	6	55,00 %	Dwarfing - Curling
AZ;10_12b	*Vicia sativa*	JOSE	*Aphis craccivora*	11	11	100,00 %	Dwarfing – Curling
AZ;15	*Lens culinaris*	ROSANA	*Aphis craccivora*	12	0	0,00 %	Dwarfing - Yellowing
AZ;10_12b	*Lens culinaris*	ROSANA	*Aphis craccivora*	10	9	90,00 %	Dwarfing – Curling - Death
AZ;15	*Lens culinaris*	CORALIE	*Aphis craccivora*	10	6	60,00 %	Dwarfing - Curling
AZ;10_12b	*Lens culinaris*	CORALIE	*Aphis craccivora*	8	8	100,00 %	Dwarfing – Curling - Bush

ISOLATE	HOST	VAR	INOC	PLANTS	SUS	SUS IR	OBS SYMPT
AZ;15	*Vicia faba*	SEVILLE	*Agrobacterium tumefaciens*	37	11	30,00 %	Dwarfing – Curling - Yellowing
AZ;10_12b	*Vicia faba*	SEVILLE	*Agrobacterium tumefaciens*	43	22	50,00 %	Dwarfing – Curling - Yellowing
AZ;15	*Vicia sativa*	JOSE	*Aphis craccivora*	11	6	55,00 %	Dwarfing - Curling
AZ;10_12b	*Vicia sativa*	JOSE	*Aphis craccivora*	11	11	100,00 %	Dwarfing – Curling
AZ;15	*Vicia sativa*	VIGILE	*Aphis craccivora*	12	5	42,00 %	Dwarfing - Curling
AZ;15	*Vicia sativa*	SPIDO	*Aphis craccivora*	12	9	75,00 %	Dwarfing - Curling
AZ;15	*Vicia sativa*	RUBIA	*Aphis craccivora*	12	6	50,00 %	Dwarfing - Curling
AZ;15	*Vicia sativa*	NACRE	*Aphis craccivora*	10	4	40,00 %	Dwarfing - Curling
AZ;15	*Vicia sativa*	ARDENTE	*Aphis craccivora*	11	6	55,00 %	Dwarfing - Curling
AZ;15	*Vicia sativa*	CARBURE	*Aphis craccivora*	12	5	42,00 %	Dwarfing - Curling
AZ;15	*Lens culinaris*	ROSANA	*Aphis craccivora*	12	0	0,00 %	Dwarfing - Yellowing
AZ;10_12b	*Lens culinaris*	ROSANA	*Aphis craccivora*	10	9	90,00 %	Dwarfing – Curling - Death
AZ;15	*Lens culinaris*	CORALIE	*Aphis craccivora*	10	6	60,00 %	Dwarfing - Curling
AZ;10_12b	*Lens culinaris*	CORALIE	*Aphis craccivora*	8	8	100,00 %	Dwarfing – Curling - Bush
AZ;10_12b	*Lens culinaris*	FLORA	*Aphis craccivora*	11	4	36,00 %	Dwarfing – Curling - Bush
AZ;10_12b	*Lens culinaris*	ANICIA	*Aphis craccivora*	10	3	30,00 %	Dwarfing – Curling - Bush

Supplementary Table 7 : **Agroinoculations and aphid inoculations tests for the selection of lentil and vetch varieties**. "Sus" column indicates suspicions based on symptoms. "Sus IR" indicates the associated infection rate.



<u>Supplementary Figure 11 :</u> Coefficients of Pearson correlations between dry weight, fresh weigh and plant height for the three isolates (AZ,10_12b, AZ;15 and JKI-2000) in all hosts (faba bean, lentil and vetch). DW stands for dry weight, HW stands for fresh weight and H or Height stands for plant height. All correlations were statistically significant although not displayed here.

FB	C	N	U1	U2	U4	Total Plants
AZ;10_12b	0	9	0	6	3	52
AZ;15	0	4	0	5	7	58
JKI-2000	1	3	0	3	2	58

L	C	N	U1	U2	U4	Total Plants
AZ;10_12b	0	0	0	0	0	46
AZ;15	0	0	0	1	0	24
JKI-2000	2	0	0	1	0	33

V	C	N	U1	U2	U4	Total Plants
AZ;10_12b	0	0	0	0	3	25
AZ;15	0	0	0	1	1	20
JKI-2000	0	0	0	0	0	8

Supplementary Table 8 : Overview of missing segments observed for agro-inoculated faba beans (FB), aphid inoculated lentils (L) and aphid inoculated vetches (V). Total number of infected plants is indicated in the last column.

CHAPTER 3 :

SYSTEMATIC PHENOTYPIC CHARACTERIZATION OF SINGLE-SEGMENT REASSORTANTS BETWEEN TWO FBNSV ISOLATES

1. Context

After completing the phenotypic characterization of both the **[AZ;15]** and **[AZ;10_12b]** genotypes, we had all the necessary components in place to start the systematic phenotypic characterization of single-segment reassortants between the two isolates on their respective field hosts, vetches and lentils. This experiment had the primary objective to evaluate the costs and benefits associated to single-segment reassortments.

We examined various biological traits, such as plant height, which served as a proxy for assessing the genotype influence on overall plant growth. This choice was based on previous findings that established a correlation between plant height and other infection related traits such as plant weight (**Torralba et al. 2024a – Chapter 2**), symptom progression, plant viral load, transmission rate, and plant genome formula.

Due to logistical constraints in this extensive experiment, it was not feasible to process all reassortants simultaneously with the parental genotypes as controls. Therefore, the experiment was divided into multiple batches, for either the phenotypical characterization or transmission assays, with both parental genotypes included as controls in each batch. Additionally, direct agro-inoculation of vetches and lentils being technically impossible, we had to initially produce both reassortants and parental genotypes through agro-inoculation in faba beans and from them inoculated the viruses onto the field hosts (vetches and lentils) through aphid transmission using *Aphis craccivora*.

2. General results

One of the primary findings derived from this experiment was the viability of all single-segment reassortants, infecting all experimental hosts. As previously mentioned, the additional steps we discussed confirmed that all reassortants remained viable, not only on their respective field hosts but also on faba beans. Furthermore, they proved to be transmissible from and to all host species tested. This has noteworthy implications concerning the ease with which reassortments can occur and the potential advantages this may offer in terms of reducing the cost associated with maintaining genomic integrity through genome reconstitutions, particularly in sympatric favorable contexts.

Reassortants exhibited a wide range of effects on all the assessed phenotypic traits, with several phenotypes remaining significantly unaffected, while others were positively or negatively affected. Furthermore, we observed distinct patterns in these effects, depending on

the major parental genotype and whether the infected host on which the reassortant was evaluated was a "new" host or the field host of the major parent.

Notably, we observed substantial alterations in the genome formula for multiple reassortants, which may suggest significant interactions between some segments, like DNA-C and DNA-U1. There were significant opposed variations of their relative frequencies, in the associated reassortants, and this trend was in opposite directions in the two major parental genotypes. This suggests an unforeseen interaction between these two segments or a common function mediated by their respective proteins. Additionally, the drastic decrease of the relative frequency to almost 0% of certain segments following reassortment raises questions on whether it signifies a severe replication inefficiency, an imbalance in gene copy numbers to compensate for different protein and or RNA interactions, a substantial degradation of these segments, or other yet unknown phenomenon.

Digging deeper into the variations observed in the genome formula could potentially provide insights into whether these changes were a result of disruptions in intra-genomic interactions, leading to a rebalancing of gene copy numbers, or if they were driven by more mechanistic effects such as replication or degradation.

Finally, using plant viral load and transmission rates as proxies for viral fitness, we were able to assess the overall impact on fitness resulting from single-segment reassortment between these two FBNSV isolates and consider its potential implications in a natural context.

3. Article status

As previously stated in (**Torralba et al. 2024a – Chapter 2**), this experiment was a lengthy undertaking, and given its comprehensive nature, it was challenging to fully analyze, employ, and interpret the data in a timely manner. Consequently, the article remains in the preparatory phase, with only a portion of the data being exploited. In the near future, we will develop the thorough and comprehensive analysis to enrich the existing assessment and conclusions.

4. Article in preparation, referenced as (*Torralba et al. 2024b*)

Systematic phenotypic characterization of single-segment reassortments between two isolates of the faba bean necrotic stunt virus (FBNSV) in multiple host species

Authors

Babil Torralba[1,2], Aleksandra Greshnova, Andy Brousse[1,2], Prune Lacôte-Popovic[1,2], Stéphane Blanc[2] and Yannis Michalakis[1]

Affilations

[1] *MIVEGEC, CNRS, IRD, Univ Montpellier, Montpellier, France*

[2] *PHIM, Univ Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France*

Authors contribution to data collection and analyzes

Aleksandra Greshnova (master's degree) contributed partially to the laboratory experiments but primarily conducted the statistical analysis (excluding the evaluation of transmission rates and the assessment of qPCR CT thresholds). Andy Brousse and Prune Lacôte-Popovic were actively involved in all aspects of the laboratory experiments. I participated and drove most of the experimental work, set up the qPCR CT thresholds, analyzed the transmission rates, and updated Aleksandra statistical analyzes with the final data.

ABSTRACT

Genome segmentation enables genetic exchanges through reassortment, where entire genomic segments are substituted by homologous ones from distinct genotypes. Recent discoveries revealed the capacity of single-stranded DNA (ssDNA) nanoviruses to produce reassortants from parental genotypes that do not infect the same cell or even the same host. This phenomenon has implications for the nanovirus ecological dynamics, the evolution, and the maintenance of genomic integrity, contingent upon the frequency and fitness of the reassortants. While reassortment-associated adaptive changes are well-documented, deleterious effects have also been observed. Overall, no systematic study has been conducted on viral fitness in response to reassortments, in any viral taxa. We here present a comprehensive phenotypic analysis of all possible combinations of single-segment reassortants between two isolates of the octopartite nanovirus *faba bean necrotic stunt virus* (FBNSV), **[AZ;15]** and **[AZ;10_12b]**. These two isolates, originating from the same geographical area but different field hosts, exhibit an overall nucleotide-level divergence of approximately 20%. Despite this genetic distinction, they can infect and be transmitted by both shared hosts and vectors. All 16 single-segment reassortants were characterized on the respective field hosts of the parental genotypes. One major finding is that all reassortants are viable on the two hosts and readily aphid-transmissible. Additionally, the monitoring of a number of phenotypic traits showed that most are rather robust against reassortments while the relative frequency of the distinct genome segments, the genome formula, is highly variable. Finally, when looking at traits that might be considered as proxies of the fitness (viral load and transmission rate), the reassortments between these two FBNSV isolates appear as often deleterious than beneficial.

INTRODUCTION

Viruses can be associated to three distinct genomic architectures and packaging strategies : the monopartite architecture, where the entire genetic information is encoded by a single nucleic acid molecule; the segmented architecture, where the genetic information is divided into multiple molecules, called genomic segments, co-encapsidated in one viral particle; and the multipartite architecture, where the genetic information is also partitioned in genomic segments separately encapsidated. This physical fragmentation of the genetic information in segmented and multipartite viruses enables the substitution of entire segments by

homologous segments from a different genotype, a genetic exchange named reassortment or pseudo-recombination (**Roossinck 1997; Holland and Domingo 1998; McDonald et al. 2016**).

Similar to recombination, reassortment has the potential to disrupt co-adapted gene complexes (**Martin et al. 2011a; Martin et al. 2011b; Varsani et al. 2018**) and lead to genetic innovations (**Roossinck 1997; Holland and Domingo 1998; Martin et al. 2011a**) which may play a significant role in evolution and adaptation (**Chao, Tran, and Matthews 1992; Martin et al. 2011a**). In contrast, reassortment does not affect intra-segment interactions. The hybrid genotypes are produced by a major parental genotype, which provides the majority of segments, along with reassorted segments from a minor parental genotype.

Recent discoveries have sparked a renewed interest in studying reassortment in multipartite viruses (**Sicard et al. 2019; Di Mattia et al. 2022; Torralba, Blanc, and Michalakis 2023**). In the octopartite single-stranded DNA (ssDNA) faba bean necrotic stunt virus (FBNSV, genus *Nanovirus*, family *Nanoviridae*), it has been demonstrated that genomic segments do not accumulate together in a single cell but can exchange gene products, mRNA or proteins, across cells, thereby functionally complementing the viral system at a supra-cellular level (**Sicard et al. 2019**). Even more intriguing, we have recently shown that the distinct FBNSV segments can be acquired and/or transmitted separately by insect vectors and reconstitute a complete genome in a recipient host (**Di Mattia et al. 2022**). These phenomena increase the potential to produce reassortants by expanding the spatial scale at which these viruses can undergo reassortment compared to segmented viruses (**Di Mattia et al. 2022; Torralba, Blanc, and Michalakis 2023**). While these latter viruses require a-priori co-infection of the same cell by two parental genotypes for reassortment to occur due to co-packaging of all genomic segments, the ssDNA multipartite nanoviruses can reassort from parental genotypes that do not co-infect the same cell or even the same host (**Sicard et al. 2019; Di Mattia et al. 2022; Torralba, Blanc, and Michalakis 2023**).

In addition to an increased potential for genetic exchanges, depending on the frequency and fitness of the reassortants, viruses that can reconstitute a complete viable genome from incomplete segment sets may relax the maintenance cost of their genomic integrity. Such reconstitutions will often generate reassortants subject to multiple constraints (**Torralba, Blanc, and Michalakis 2023**) that would influence their fitness. Frequent reassortment events leading to hybrid genotypes less adapted than their parental counterparts would result in the decline of viral populations when ecological conditions favor reassortment. Conversely, if

viable reassortants are frequently produced the costs due to potential loss of genomic integrity could be alleviated through genome reconstitutions between different hosts.

It has been observed that reassortant genotypes can emerge either frequently (**Matsuzaki et al. 2003**) or rarely (**Fraile et al. 1997**), depending on the viral taxon. While some may provide adaptive changes, such as increased virulence (**Chakraborty et al. 2008**), the ability to overcome host resistance (**Tentchev et al. 2011**), or an expansion of the viral host range (**Idris et al. 2008**), others can be negatively affected with a drastically reduced fitness (**Escriu, Fraile, and García-Arenal 2007; Ohshima et al. 2016; Villa and Lässig 2017**). Despite the myriads of available data on the properties of individual reassortants, a systematic characterization of the phenotypic effects of reassortments is currently lacking, whatever the viral taxon considered.

We have recently reported the detailed phenotypic characterization of two novel infectious clones of FBNSV, corresponding to isolates **[AZ;15]** and **[AZ;10_12b]**, collected in the same geographical area but from distinct host plant species (**Torralba et al. 2024a**). Considering their overall sequence divergence of approximatively 20%, and their demonstrated ability to share hosts and vectors (**Torralba et al. 2024a**), they appear as perfect candidates to fill the gap alluded above by implementing a comprehensive study of the phenotypic effects of reassortment. Because investigating all possible combinations of 8 segments from two genotypes required too much logistics, we set to characterize all possible 16 combinations of single-segment reassortants. The phenotypic characterization has been carried out on the respective hosts of both parental genotypes, vetches (*Vicia sativa*) and lentils (*Lens culinaris*) respectively, thereby assessing potential adaptive effects between the "field" hosts and the "new" hosts. Based on previous evaluation of the properties of **[AZ;15]** and **[AZ;10_12b]** isolates (**Torralba et al. 2024a**), we monitored plant height as an indicator of overall plant growth, symptom severity and progression, viral load in plants, genomic segment frequency distribution in plants, hereafter called plant genome formula, and aphid transmission rate.

These results, which establish the reaction norm of the reassortants across different host plant species, provide insights into the characteristics of all possible single-segment reassortants and their impact on viral fitness. Remarkably, all reassortants proved viable in all hosts and transmissions and none was strongly deleterious. This observation has important consequences for the biology, life cycle and spatial dynamics of nanoviruses and the dissemination of their genomic segments in their ecosystems.

METHODS

1. Overview of the experiment

A graphical summary of the whole experiment is shown in (**Figure 1**). Similar to the prior phenotypic characterization of the parental genotypes **[AZ;15]** and **[AZ;10_12b]** (**Torralba et al. 2024a**), our experiment started with the agro-infiltration of faba beans with either one of the parental genotypes or one of single-segment reassortants. Infected faba beans were subsequently used as the primary source plants for aphid-inoculation of both vetches and lentils, respectively representing the field hosts of the parental genotypes. These plants were then subjected to a comprehensive phenotypic assessment, which involved measuring their height and tracking symptom development. At later stages, they were harvested to determine viral load and genome formula. To accommodate logistical constraints, the experiment was organized into four distinct characterization batches. Each batch consisted of four reassortants, consisting of two segments each reciprocally introduced in each genomic background and the two parental genotypes as controls. Due to logistic limitations the transmission rates were evaluated separately in new transmission batches. In these batches the phenotypic characterization was not repeated but the plants were used at an earlier stage of the infection for estimation of the transmission rates, following the verification of the presence of the 8 segments through qPCR.

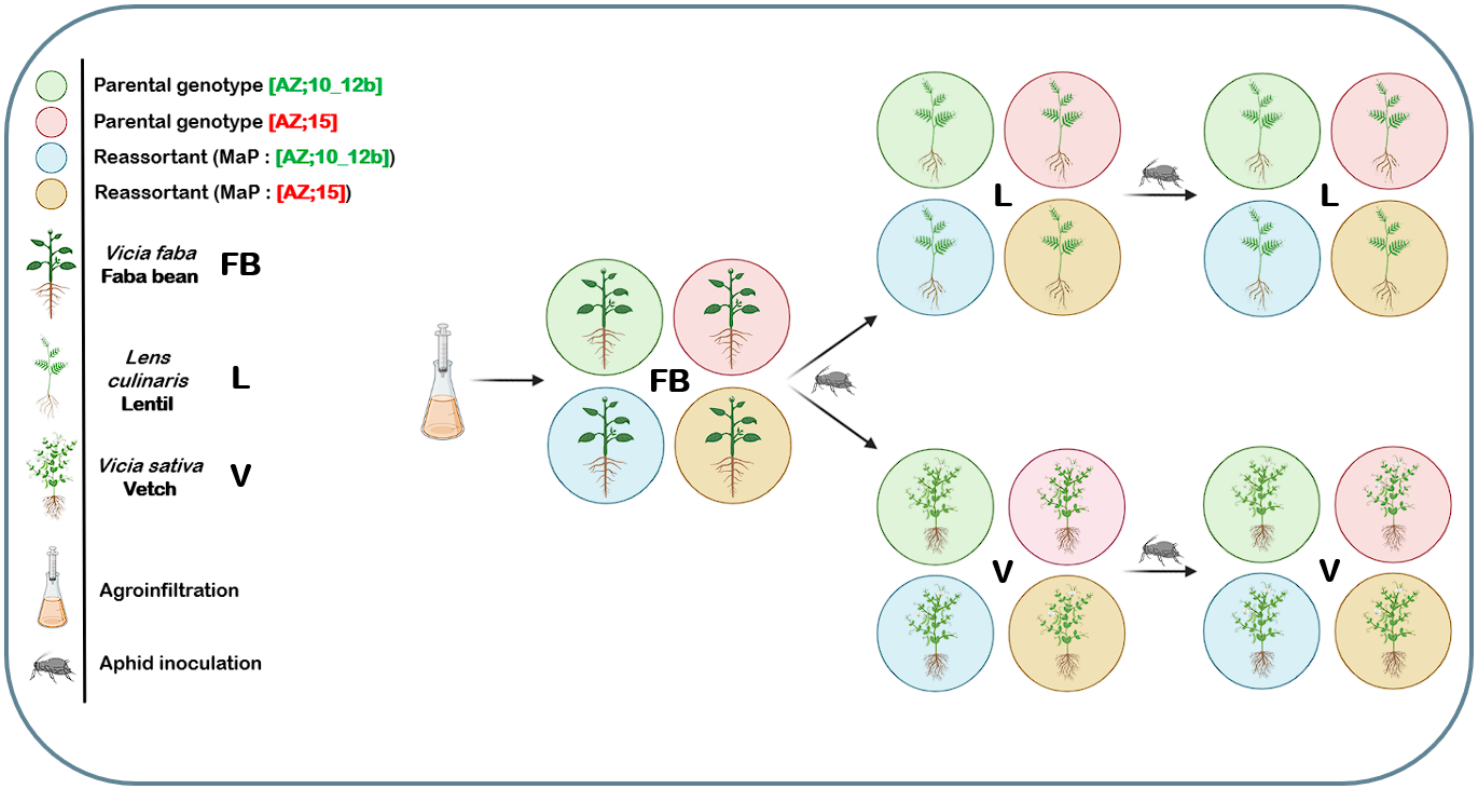


Figure 1 : **Overview of the systematic phenotypic characterization of single-segment reassortants**. Sequential steps are indicated by the black arrows and numbers. 1- Bacteria mix of corresponding agrobacterium clones constituting the various genotypes are infiltrated in faba beans. 2- After four weeks, infected faba beans are extracted and verified by qPCR for the presence of all eight segments. 3- Aphid inoculation, using infected faba beans as source plants, to young lentils and vetches. 4- Phenotypic characterization of viral traits in lentils and vetches. 5- After verification of the presence of the 8 segments, aphid inoculation from lentils to lentils and from vetches to vetches. 6- Evaluation of the transmission rates. MaP: major parental genotype.

2. Infectious clones, isolates and reassortant nomenclature

The FBNSV genome consists of eight circular single-stranded DNA segments, each approximately 1 kb in length, and comprising only one open reading frame (**Gronenborn 2004; Lal et al. 2020**). Among these segments, C encodes the cell cycle-linked protein (Clink), M encodes the movement protein (MP), N encodes the nuclear shuttle protein (NSP) mandatory for aphid transmission, S encodes the coat protein (CP), R encodes the master replication-associated protein (M-Rep) enabling replication, while U1, U2, and U4 encode proteins with unknown functions (**Gronenborn 2004**). For this study, we used **[AZ;15]** and **[AZ;10_12b]** infectious clones based on isolates respectively isolated from vetches (*Vicia sativa*) and lentils (*Lens culinaris*) as previously described (**Grigoras et al. 2014; Torralba et**

al. 2024a). Vetches are then considered "field" host for **[AZ;15]** and "new" host for **[AZ;10_12b]** and reciprocally for lentils. The genomic sequences of all FBNSV isolates used in this study are accessible on GenBank NCBI (**Supplementary Table 1**). Single-segment reassortants are named based on their minor parent *i.e.* C-15 designates the reassortant with its DNA-C originating from the minor parent **[AZ;15]** and the rest of its genome from **[AZ;10_12b]**. Conversely, C-10 designates the reassortant with its DNA-C sourced from the minor parent **[AZ;10_12b]** while the remainder of its genome is from **[AZ;15]**.

4. Plants and growth conditions

The three legume host plants used in this study are *Vicia faba* (faba bean; var. Sevilla, Vilmorin, France), *Vicia sativa* (vetch; var. « José », Baumaux, France) and *Lens culinaris* (lentil; var. « Coralie », Baumaux, France). Faba beans were seeded directly in N2 soil medium (NEUHAUS, Humin-Substrat N2) in a greenhouse under a 13.5/10.5 day/night photoperiod at 25/18°C temperature and 70% hygrometry. Lentils and vetches were seeded in the same medium in a growth chamber under a 13/11 day/night photoperiod at a temperature of 25/18°C day/night and 70% hygrometry for better germination. Once germinated, plantlets of faba beans, lentils and vetches were transplanted in individual pots (7x7x7 cm) and all were moved to the greenhouse. After inoculation (see below) and for the rest of the experiment, lentils and vetches were bounded in Aracones (Arasystem), whereas faba beans were tutored at 14 days old and on.

5. Agroinoculation and aphid-inoculation

Bacteria culture and agroinfiltration in faba beans were executed as previously described (**Torralba et al. 2024a**). A mix containing 8 x 5mL of each segment and adjusted to 50 mL for parental genotypes and a mix containing 8 x 2.5 mL of each corresponding reassortant segment and adjusted to 25 mL for reassortants was incubated at ambient temperature for 1 hour 30 before infiltration in available leaves of 9 days old seedlings. Leaves were pierced prior infiltration and around 0.5 mL were used by seedling. Two plates of 24 seedlings were infiltrated for each genotype.

For all transmission experiments we used individuals of the aphid species *Aphis craccivora* from a clonal population established and maintained in a controlled environment as previously described (**Ryckebusch et al. 2021; Di Mattia et al. 2020**). Infected faba beans used for the FB>L and FB>V transmission were used as plant source at 28 dpi, lentils and vetches were used as source plants at 21 dpi whereas faba beans used for the FB>FB transmission were used as plant source at 26 dpi for aphid inoculation. Infected lentils and vetches were used as plant

source at 20 dpi for aphid inoculation. Source plants were all verified for the presence of the 8 segments by qPCR. Five faba beans were selected for each genotype to be used as source plants for aphid inoculation to vetches and lentils. Aphids acquired the virus for 3 days on faba beans and were then mixed together. Two aphids were put for 3 days on 5 days old *Vicia sativa* var. "José" (vetches) seedlings and 6 days old *Lens culinaris* var. "Coralie" (lentils) seedlings as previously stated in (**Torralba et al. 2024a**). To determine L>L and V>V transmission rates fifty aphids were put on each plant to acquire the virus for four days, then collected and mixed together. For each genotype only one aphid per receiver plant was used to inoculate 48 lentils and 48 vetches. Successful infections were determined based on visual symptoms at 30 dpi and plants with unclear symptoms were additionally verified by qPCR. Aphids were killed by 4 successive sprays of ERADICOAT (Certis) for FB>L and FB>V transmissions or by one spray of 0.2% PIRIMOR G (Syngenta) solution for L>L and V>V transmissions after the aphid inoculation access period (IAP).

6. Total DNA extraction from plants

Total DNA extractions from faba beans, vetches and lentils were performed as described in (**Torralba et al. 2024a**). Faba beans were extracted at 28 days dpi. Lentils and vetches used in the characterization batches were extracted at 30 days dpi. Lentils and vetches used as source plants in the transmission batches were extracted at 15 days dpi.

7. Phenotypic trait estimates

For lentils and vetches, the symptom score was visually monitored every seven days until 28 days after the beginning of the aphid IAP. Symptom scoring was visually estimated according to the criteria described in (**Supplementary Table 2**). Aracones were removed at 30 dpi and the final plant height was measured.

8. qPCR, genome formula and viral load

For the qPCR quantification of all segments, each extraction sample was diluted 10 times in distilled water. All qPCR analyses were carried out using the LightCycler 480 thermocycler (Roche, Indianapolis, Ind, USA) with the LightCycler FastStart DNA Master Plus SYBR Green I kit (Roche) according to the manufacturer's instructions. Primers were used at a final concentration of 0.5 µM to 1.5 µM depending on the associated isolate/segment and are described in (**Torralba et al. 2024a**) (**Supplementary Table 3**). All primers were designed to

function using the same cycle parameters (40 cycles at 95°C for 10 s, 60°C for 10 s and 72°C for 10 s) for both parental genotypes and to be highly specific of their respective segment. For all samples we analyzed two qPCR technical replicates. Post-PCR data were treated as described in (**Gallet et al. 2022**) with the LinRegPCR program (**Rujiter et al. 2009**). Genome formulas were established as previously described in (**A. Sicard et al. 2013**).

To confirm the presence of a segment we defined the value of threshold cycle (CT) (**Torralba et al. 2024a**), beyond which the corresponding segment was considered absent. To define the thresholds for each segment in this study, we performed qPCRs using a primer pair targeting one segment of one isolate on samples extracted from plants infected with a reassortant containing the homologous segment from the other isolate and the rest of the genomic background from the major parent. The methodology and the resulting threshold values are detailed in (**Supplementary Table 4**).

Due to the inevitable variable development of sampled leaves or leaflets, the number of cells harvested likely varied and thereby could affect our estimation of the viral load. To account for this potential problem, we normalized the viral load, with the total DNA extracted from a given sample, as previously described (**Torralba et al. 2024a**). Total DNA concentration within each extract was evaluated using a NanoDrop2000 (ThermoScientific).

9. Statistical analysis

All analyses were performed using R version 4.3.0.

9-1. Defining qPCR CT thresholds for genomic segment absence

To confirm the presence of the reassorted segments and the absence of their homologous ones and to take into account possible loss of segments as was observed in (**Torralba et al. 2024a**), we needed to establish a dependable threshold that would determine when segments are considered absent using tolerance limits. We used the same method described in (**Torralba et al. 2024a**) but we used samples from plants infected by the reciprocal reassortant of the focal segment for which we wanted to establish the threshold *i.e.* to establish the threshold for **[AZ;10_12b]** DNA-C, we used samples from the C-15 reassortant, which inherently lacks the **[AZ;10_12b]** DNA-C, having been substituted with the **[AZ;15]** DNA-C. This approach allowed us to include all potential segments along with their homologous counterparts from the second genotype, ensuring the establishment of the most conservative thresholds possible.

For all biological trait evaluations, we first filtered out plants that did not contain all eight genomic segments according to the thresholds previously determined.

9-2. Plant height

Among other symptoms caused by FBNSV, the stunted growth of infected plants is very characteristic. To account for the direct impact on plant height of host plant species, we calculated the standardized ratio according to the following formula:

$$r = \frac{\text{mean asymptomatic height} - \text{individual symptomatic height}}{\text{mean asymptomatic height}}$$

If infection has no effects on the height of symptomatic plants, the ratio will be zero, if the effect on height important, the ratio will approach one. After calculating the ratio, we identified outliers for each genotype - host combination (below the lower quartile – 1.5 times the interquartile range or above the upper quartile + 1.5 times the interquartile range) and excluded them from the analyses (across all groups, 70 outliers were identified and excluded). We first compared the standardized height ratio of parental genotypes in each host within each batch using a t-test. To study the effect of reassortment on plant height, we modeled the standardized ratio as a function of genotype, host species, and the interaction between them as fixed factors and batch as a random effect (lmer function from lme4 package, sum-to-zero contrasts). To test the significance of the predictors included in the models, type-III Sum of Squares were used (Anova function from car package) (**Supplementary Table 5-A**). Next, we performed a post-hoc multiple pairwise comparisons test by comparing the predicted standardized height ratio of a reassortant with the predicted ratio of the major parent per host (contrast function from emmeans package specifying "mvt" – multivariate test – as a method for adjustment for multiple comparisons) (**Supplementary Table 5-B**). The average sample per group for all the above-mentioned analysis was ~26.

9-3. Symptom severity

To analyze symptom severity development in infected plants, we used standard procedures in plant pathology on the analysis of qualitative symptoms, using nonparametric analysis of ordinal data with repeated measures following the procedures outlined by (**Shah and Madden 2004**). Specifically, to assess the statistical significance of the effect of genotype, host plant, and their interaction on the symptom development dynamics, we used nparLD: an R-language library for non-parametric analysis of longitudinal data with factorial design (**Noguchi et al. 2012**). We first tested whether batch effects were indeed observed in the measurement of the symptom dynamics for parental genotypes from different batches. For this, we created a model for each parental genotype separately using f2.ld.f1 function that refers to the design with two between-subject factors (batch and host) and one within-subject factor (time). For each

combination of between and within factor levels, f2.ld.f1 models produce relative treatment effects (RTE): the probability that the value of a randomly selected individual measurement with this combination of factor levels is larger (*e.g.* has a larger symptomatic score in our case) than the value of a randomly selected individual measurement across all possible factor levels. The modified ANOVA-type statistic with Box approximation (ANOVA.test.mod.Box) was reported for the between-subject factors (batch and host) and their interaction (**Supplementary Table 6-A**). We observed differences between measurements of the plants from different batches infected with the same parental genotype.

To account for between-batch differences, we created a model per batch using the f2.ld.f1 function with genotype and host as between-subject factors and time as within-subject factors (the maximum number of between-subjects factors that is possible to analyze with the nparLD library is two, and thus we could not include batch as one of the factors) (**Supplementary Table 6-B**). To investigate the pairwise differences between reassortants and parental genotypes, we applied multiple comparisons as described in (**Noguchi et al. 2012**) by using the f1.ld.f1 function, where the between-subject factor was genotype with two levels – the reassortant and the corresponding parental genotype – and the within-subject factor was time, followed by Bonferroni correction for multiple comparisons adjustment (**Supplementary Table 6-C**). The average sample per group for all the above-mentioned analysis was ~26.

To assess the extent of differentiation between each reassortant and the parental genotype, we computed the difference between their respective RTE values generated by the f1.ld.f1 model at each time point and summed these differences. A negative difference signifies that the reassortant induced overall milder symptoms compared to the control (major parent), while a positive difference indicates that the reassortant caused overall more severe symptoms than the control (**Supplementary Table 6-D**).

9-4. Viral load

Regarding viral load, we excluded outliers for each genotype-host combination if they fell below the lower quartile minus 1.5 times the interquartile range or above the upper quartile plus 1.5 times the interquartile range. Across all groups, 62 outliers were removed. Since we observed significant differences between batches, we incorporated the batch effect into the model.

We could not employ a mixed model due to observed heteroscedasticity of residuals. To address the observed heteroscedasticity (all batches) and deviations from normality (batches 3 and 4) of residuals, we separately modeled each batch using weighted least-squares, with relative viral load as the response variable and genotype, host, and genotype-host interaction as predictors, and applied a Box-Cox transformation (λ1=0.303, λ2=0.101, λ3=0.303, λ4=0.303) to the data. Weights were determined using the following formula:

$$\text{weights} = 1 / \text{resid}(\text{model.batch}_i)\text{^}2$$

Where residuals (model.batch.i) are the residuals of the linear model for batch i, with genotype, host, and genotype-host interaction as fixed effects and relative viral load as the response variable (Box-Cox transformed). To assess the significance of the predictors included in the models, we employed type-III Sum of Squares (using the Anova function from the car package) (**Supplementary Table 7-A**). We also conducted a post-hoc multiple pairwise comparisons test, comparing the predicted mean relative viral load of a reassortant with that of the major parent per host. For this analysis, we used the contrast function from the emmeans package, specifying "mvt" (multivariate test) as the method for adjustment for multiple comparisons. The data was reverse-transformed before comparison using the regrid function from the emmeans package (**Supplementary Table 7-B**). Althoug greatly reduced, persistent residual heteroscedasticity (all batches) and deviation from normality (batches 3 and 4) will be addressed in the near future.

9-5. Genome formula

Regarding the genome formulas, outliers of all genomic segments were excluded if they fell below the lower quartile minus 1.5 times the interquartile range or above the upper quartile plus 1.5 times the interquartile range. A total of 434 outliers were excluded across all groups, with the initial dataset comprising 1246 analyzed plants. Each sample included 8 frequencies, resulting in a total of 9534 samples. The average group size was approximately 25.

We compared genome formulas of parental genotypes between different batches in each host and observed variations between batches. Consequently, we incorporated the batch effect into the model. To investigate whether different segments accumulated at varying frequencies in different host plants and according to the infecting genotype, we modeled the logit of segment frequency. This model included the identity of the segment, the host plant, the infecting genotype, and all possible interactions between them as fixed effects. Additionally,

batch was included as a random effect. The lmer function from the lme4 package was used, specifying "sum-to-zero" contrasts.

The logit transformation was necessary to meet the assumptions of the mixed model. Significance of fixed effects and their interactions was evaluated using type-III Sum of Squares, with the Anova function from the car package (**Supplementary Table 8-A**). To further investigate the impact of reassortment on the initial genome formula, we conducted pairwise comparisons of back-transformed segment frequencies between the reassortant and the parental genotype within all possible genotype-host combinations. For this analysis, we used the contrast function from the emmeans package, specifying "sidak" as the method for adjustment of multiple comparisons and to calculate confidence intervals for estimated differences in frequencies (**Supplementary Table 8-B**). To assess the extent of differentiation between the compared genome formulas (e.g., reassortant versus one of the parental genotypes), we calculated distances by summing up the absolute differences in frequencies for each segment (denoted as $fi1$ and $fi2$ for the ith segment) using the following formula:

$$distance = \sum |f_i^1 - f_i^2| \text{ for } i = 1 \text{ to } 8.$$

We calculated distances for the following pairs: reassortant - major parent, reassortant - minor parent, and genome formulas in different hosts (**Supplementary Figure 1**).

9-6. Transmission rates

To compare the transmission rates between reassortants and their parental genotypes, we individually modeled the infectivity status (1 or 0) of recipient plants of the transmission assays as the response variable following a binomial distribution and genotype, host, and genotype-host interaction as predictors for each batch (glm function from lme4 package, binomial distribution). To take into account the significant host effect, individual model on subsets of each corresponding hosts was done with only genotype as fixed effect (**Supplementary Table 9**). This model uses a genotype of reference that was either **[AZ;15]** or **[AZ;10_12b]** to compare the corresponding associated reassortants.

RESULTATS

1. Viability of reassortants

Before delving further into the characterization of the viral phenotypic traits, a key initial finding during this experiment was that all single-segment reassortants were viable on all host plant species used during this experiment, including faba beans, vetches, and lentils (**Table 1**). Furthermore, all reassortants proved transmissible from faba beans to vetches and lentils (mandatory for step 3 in **Figure 1**), as well as from vetch to vetch (V>V) and lentil to lentil (L>L), as detailed later.

NAME	SEGMENT	MAJ PARENT	MIN PARENT	FABA BEAN	VETCH	LENTIL
C – 10	C	AZ;15	AZ;10_12b	+	+	+
C – 15	C	AZ;10_12b	AZ;15	+	+	+
M – 10	M	AZ;15	AZ;10_12b	+	+	+
M – 15	M	AZ;10_12b	AZ;15	+	+	+
N – 10	N	AZ;15	AZ;10_12b	+	+	+
N – 15	N	AZ;10_12b	AZ;15	+	+	+
R – 10	R	AZ;15	AZ;10_12b	+	+	+
R – 15	R	AZ;10_12b	AZ;15	+	+	+
S – 10	S	AZ;15	AZ;10_12b	+	+	+
S – 15	S	AZ;10_12b	AZ;15	+	+	+
U1 – 10	U1	AZ;15	AZ;10_12b	+	+	+
U1 – 15	U1	AZ;10_12b	AZ;15	+	+	+
U2 – 10	U2	AZ;15	AZ;10_12b	+	+	+
U2 – 15	U2	AZ;10_12b	AZ;15	+	+	+
U4 – 10	U4	AZ;15	AZ;10_12b	+	+	+
U4 – 15	U4	AZ;10_12b	AZ;15	+	+	+

Table 1 : Viability of reassortants on three host plant species. The first column lists the name of the reassortant genotype. The letter indicates the reassorted segment and the number the minor parent from which it originates. The remaining segments originate from the major parent. The "Maj" and "Min" parent columns specify the major and minor parents, respectively. A positive sign (+) indicates reassortant viability on each of the three hosts.

2. Plant height

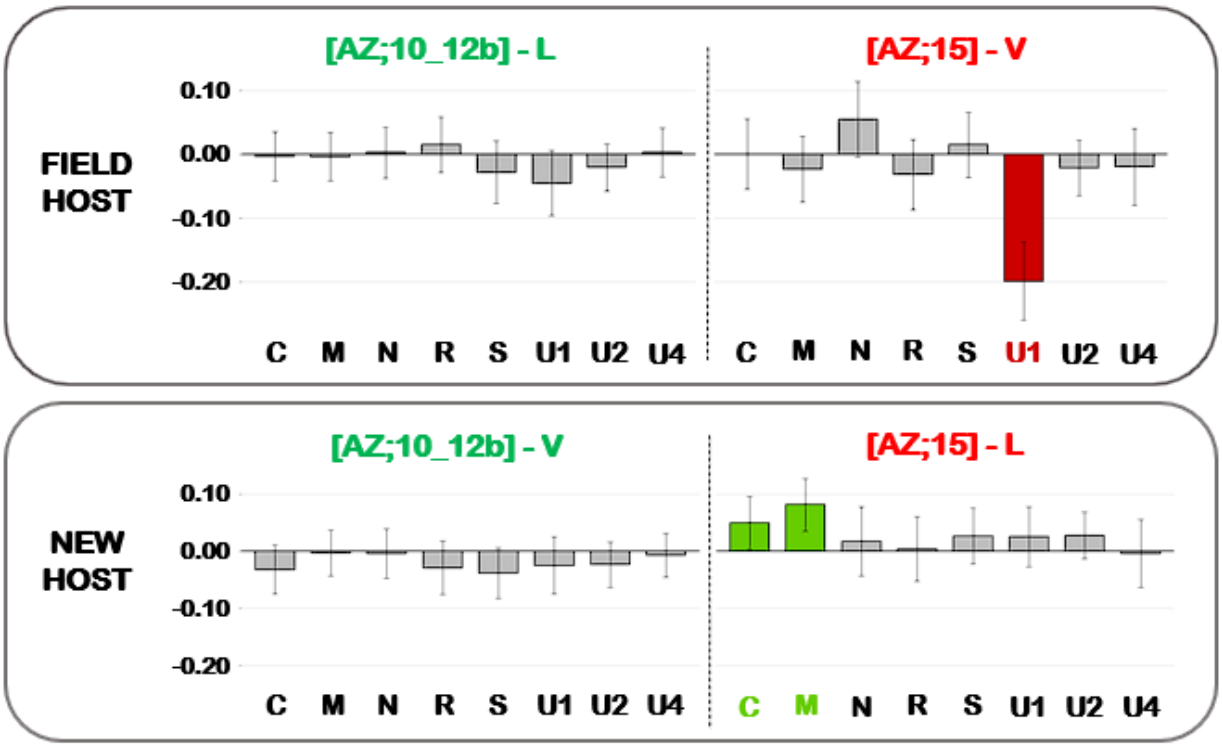


Figure 2 : Impact of each reassortant relative to its major parental genotype on plant height, in the field host of the major (top) and minor (bottom) parent. Y-axis represents the standardized plant height differences. The bars represent the difference in plant height reduction between the reassortants for a given segment, as indicated below the graphs, and its major parental genotype, as indicated in each panel. Red bars indicate statistically significant reductions, while green bars denote statistically significant increases. Grey bars are used when the variation is not statistically significant. The top row shows results on the field host of the major parent and the bottom row on the field host of the minor parent (L for lentil and V for vetch).

To begin with, it's important to highlight the contrasting phenotypes of the two parental genotypes, as determined in a prior experiment (**Torralba et al. 2024a – Chapter 2**), that might probably influence the outcome of reassortments between these two genotypes. In this regard, **[AZ;15]** had a less pronounced impact on the reduction of plant height compared to **[AZ;10_12b]** on both hosts.

Reassortants with **[AZ;10_12b]** as major parental genotype did not induce significant changes of the plant height when compared to **[AZ;10_12b]** isolate. This trait exhibits considerable robustness in the context of single-segment reassortments for **[AZ;10_12b]** (**Figure 2**). In contrast, three out of eight reassortants with **[AZ;15]** as major parental genotype differentially affected plant height in one of the two hosts. The exceptional robustness of **[AZ;10_12b]** is especially intriguing, considering that **[AZ;15]** has a statistically significantly reduce impact on

plant height in comparison. Additionally, it's worth highlighting the massive reduced effect associated with the U1-15 reassortant. This finding likely implies an important role of DNA-U1 as pathogenicity determinant for **[AZ;15]** in vetches as it was the case for **[JKI-2000]** in faba beans (**Grigoras et al. 2018**).

3. Symptom dynamics

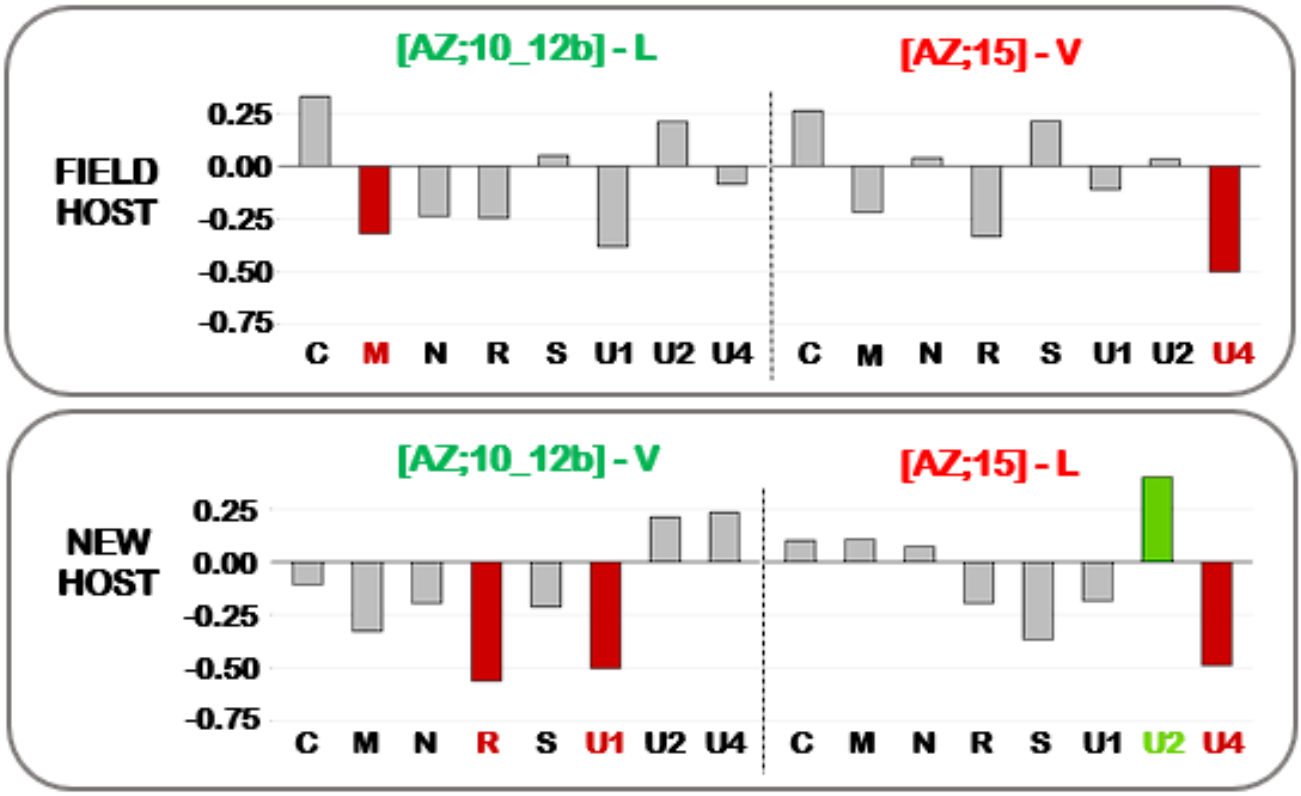


Figure 3 : Impact of each reassortant relative to its major parental genotype on the evolution of symptom severity, in the field host of the major (top) and minor (bottom) parent. Y-axis represent the RTE differences. The bars represent the RTE differences interpreted as the modification of symptom severity evolution between the reassortant of a specific segment, as indicated below the graphs, and its major parental genotype. Red bars indicate statistically significant reductions, while green bars denote statistically significant increases. Grey bars are used when the variation is not statistically significant. The top row shows results on the field host of the major parent and the bottom row on the field host of the minor parent (L for lentil and V for vetch).

Based on the previous characterization of the **[AZ;15]** and **[AZ;10_12b]** parental genotypes, there was no statistically significant difference in symptom dynamics between the two isolates in both hosts (**Torralba et al. 2024a – Chapter 2**).

Variations in the development of symptoms timing and severity were observed for both major parental genomic backgrounds and on both hosts (**Figure 3**). Those reassortants with

[AZ;10_12b] as major parental genotype whose effects statistically significantly differed from those of their major parent always showed milder symptoms than the latter. Regarding reassortants with [AZ;15] as major parental genotype, U4-10 caused milder symptoms in both hosts, suggesting that DNA-U4 is an important determinant of the pathogenicity for [AZ;15]. The reassortant U2-10 is the only one inducing more severe symptoms when compared to [AZ;15] in the new host. This might also indicate a significant role for U2 as a pathogenicity determinant in the new host lentils for [AZ;15].

4. Plant viral load

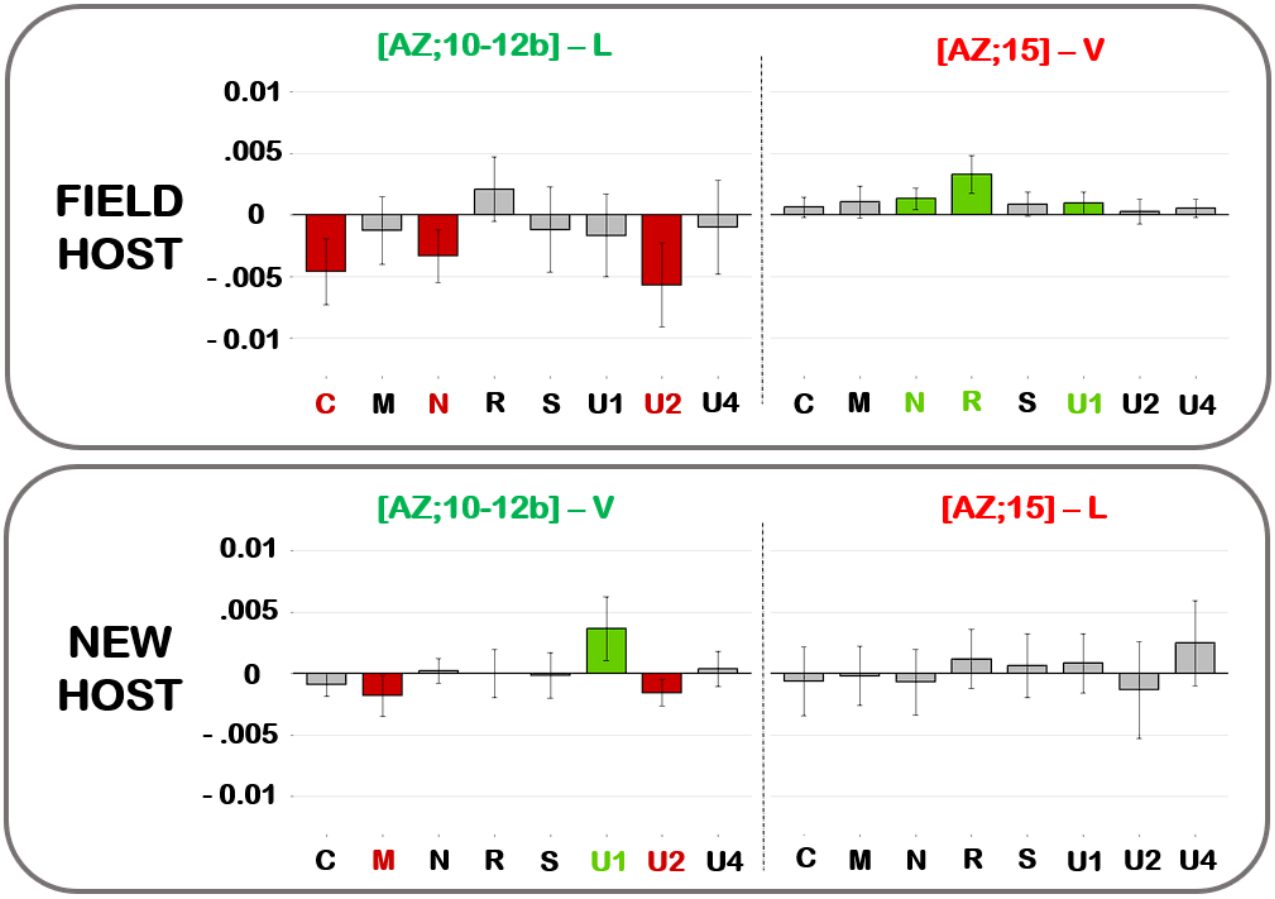


Figure 4 : Impact of each reassortant in comparison to its major parental genotype on within-plant viral accumulation in the field host of the major (top) and minor (bottom) parent. Y-axis represents the relative viral load difference**.** The bars indicate the difference in modification of plant viral load between the reassortant for a specific segment, as indicated below the graphs, and its major parental genotype. Red bars indicate statistically significant reductions, while green bars denote statistically significant increases. Grey bars are used when the variation is not statistically significant. The top row shows results on the field host of the major parent and the bottom row on the field host of the minor parent (L for lentil and V for vetch).

Based on the previous characterization of the **[AZ;15]** and **[AZ;10_12b]** parental genotypes, there was no statistically significant difference in viral load between the two isolates in lentils. However, in vetches, **[AZ;10_12b]** accumulated a significantly higher viral load compared to **[AZ;15]** (**Torralba et al. 2024a – Chapter 2**).

Reassortments induced variation of the viral accumulation in both major parental genomic backgrounds (**Figure 4**). A strong pattern that emerged was that all three reassortants with **[AZ;15]** as major parent which displayed statistically significant variations accumulated more in the major parent's field host while in contrast, all three reassortants with **[AZ;10_12b]** as major parent which exhibited statistically significant differences accumulated less in the corresponding field host. Considering the previously established differences in viral load between **[AZ;10_12b]** and **[AZ;15]** in vetches, these findings may not be unexpected, as **[AZ;10_12b]** appears to possess an advantage in viral accumulation over **[AZ;15]** at least in vetches.

Remarkably, DNA-N was significantly affected in both major parental contexts in the field host condition. This is noteworthy because it represents the only example we noted with a "reciprocal" effect: increased accumulation with DNA-N of **[AZ;10_12b]** in the background of **[AZ;15]**, and negative effect in the reciprocal reassortant.

No **[AZ;15]** reassortants displayed statistically significant differences in the new host. On the contrary, **[AZ;10_12b]** reassortants exhibited statistically significant differences as often in the new host as in the field host, although only DNA-U2 exchange had an effect in the two hosts, where this effect was negative. This might suggest that DNA-U2 is a key component regulating the viral load for **[AZ;10_12b]** in both lentils and vetches.

Overall, **[AZ;10_12b]** appears less robust against reassortments, regarding the viral load. Remarkably, as plant viral load is used as viral fitness proxy in our experimental setup, when significantly affected, the reassortants with **[AZ;15]** as major parent can do better, whereas those with **[AZ;10_12b]** as major parents most often do worse.

5. Genome formula

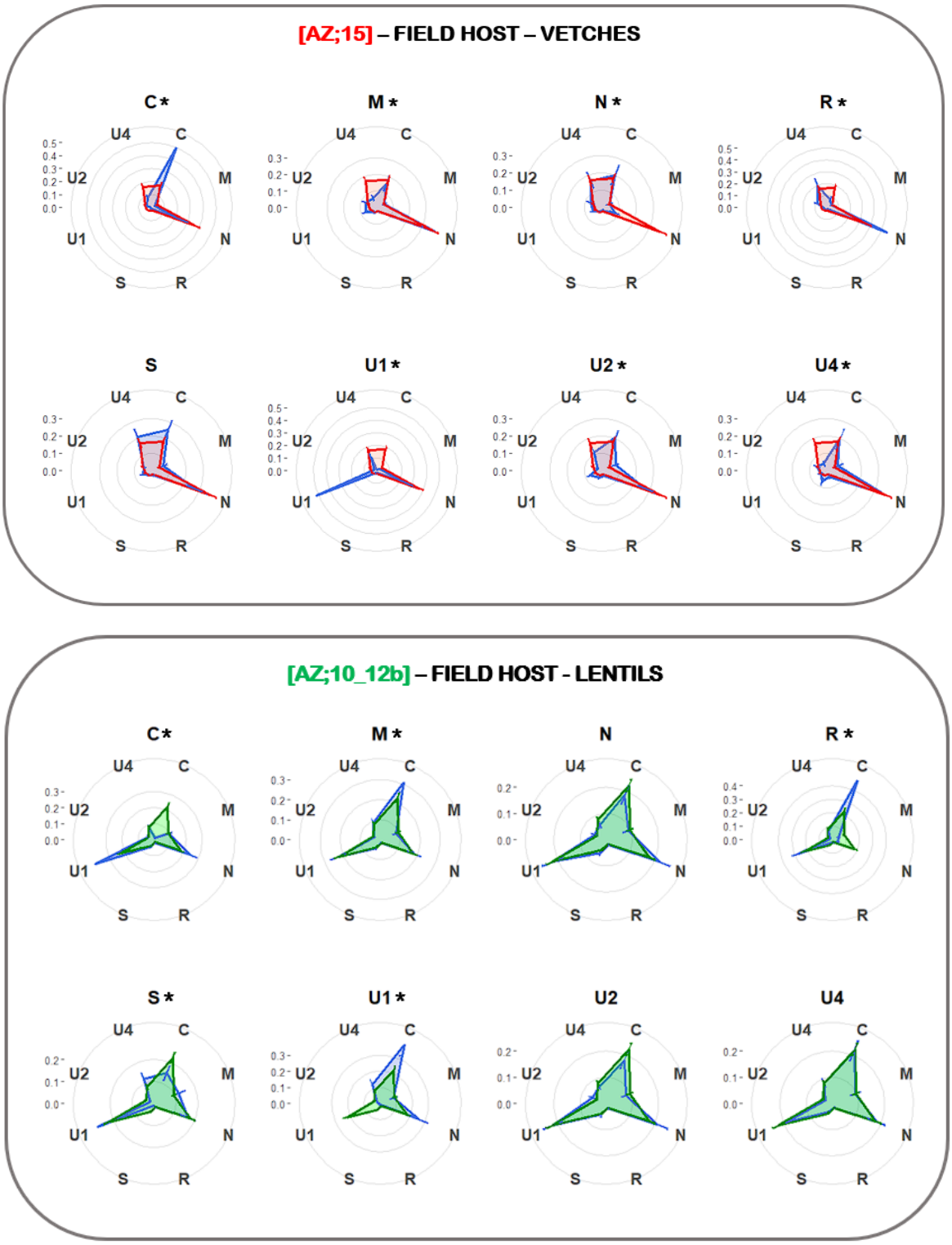


Figure 5 : Radar plots comparing the genome formula of each single-segment reassortant to that of its major parental genotype in its "field" host. Each circle on the radar plot corresponds to the relative frequency indicated on the left scale. The genome formula of the reassortant is depicted by the blue areas, while the genome formula of the major parental genotype is represented by the red and green areas for **[AZ;15]** and **[AZ;10_12b]**, respectively.

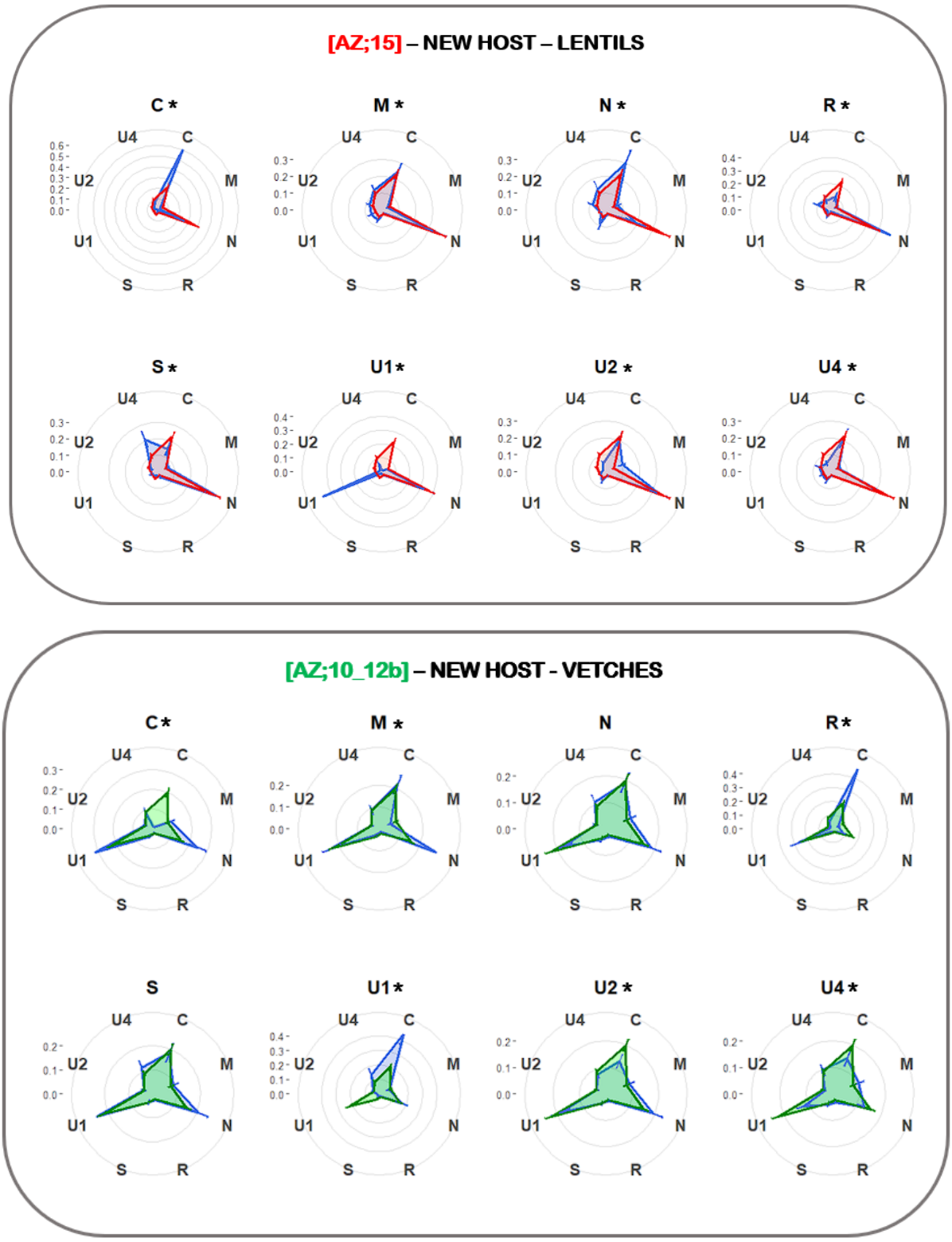


Figure 6 : Radar plots comparing the genome formula of each single-segment reassortant to that of its major parental genotype in its "new" host. Each circle on the radar plot corresponds to the relative frequency indicated on the left scale. The genome formula of the reassortant is depicted by the blue areas, while the genome formula of the major parental genotype is represented by the red and green areas for **[AZ;15]** and **[AZ;10_12b]**, respectively.

Variations in the genome formula were commonly observed after reassortments in both major parental genotypes (**Figure 5 and 6**). In the field host, only the S-10 reassortant did not exhibit a statistically significant change of the genome formula when compared to its major parent **[AZ;15]**, and solely N-15, U2-15, and U4-15 maintained the formula of their major parent **[AZ;10_12b]**. The genome formulas were even more often affected in the new host, with no reassortant preserving that of their respective major parent **[AZ;15],** except for the reassortant S-15 with **[AZ;10_12b]** as major parent. **[AZ;15]** appeared less robust than **[AZ;10_12b]** to variations of the genome formula due to reassortments on both hosts.

General quantification of the alterations in the relative frequencies of the distinct segments varied greatly among the reassortants themselves, with these effects being however relatively consistent in both hosts for each reassortant (**Supplementary Figure 1** and **Figure 9**). Cand U1 reassortants exhibited the most significant overall differences from the genome formulas of both major parental genotypes, while R-15 also displayed a notable difference with its major parent **[AZ;10_12b]** (**Supplementary Figure 1**).

The most remarkable pattern observed (**Figures 5, 6 and 8**) was the dramatic decrease in the relative frequency of the reassorted DNA-C in **[AZ;10_12b]** genomic background (C-15 reassortant), almost reaching 0%, while the relative frequency of DNA-U1 significantly increased. Symmetrically, the U1-15 reassortant exhibited a similar pattern, with a significant drop in DNA-U1 relative frequency, nearly reaching 0%, and a substantial increase in DNA-C frequency. Moreover, this pattern was reversed for reassortants in **[AZ;15]** background, where DNA-C relative frequency increased while DNA-U1 dropped significantly for C-10, and the opposite was observed for U1-10. This consistent observation suggests a potential interaction between DNA-C and DNA-U1 in both **[AZ;10_12b]** and **[AZ;15]**.

6. Transmission rates

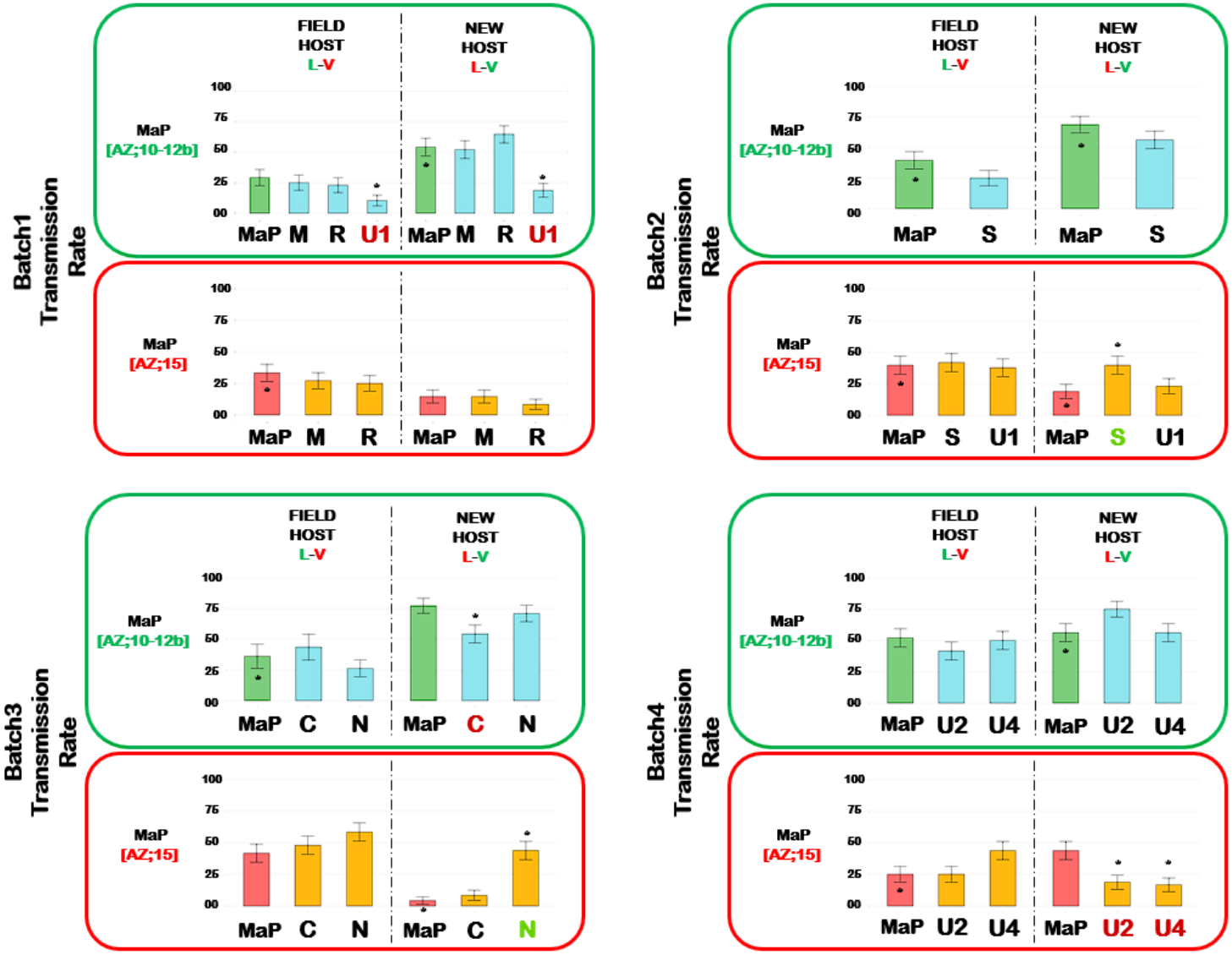


Figure 7 : Evaluation of the transmission rates (infected plants/inoculated plants) of reassortants and their major parental genotypes on both field (left side of each panel) and new hosts (right side of each panel). For practical reasons, this large experiment was divided in 4 batches that present the V>V and L>L transmission rates for each reassortant relative to its major parental genotype, both in the field and new host. The Y-axis represents the transmission rate as a percentage (infected plants/inoculated plants). The colors used to represent reassortants are consistent across batches and correspond to the major parents (cyan for reassortants with [AZ;10_12b] as major parent and orange for reassortants with [AZ;15] as major parent)., Red and green bars always represent [AZ;15] and [AZ;10_12b] parental genotypes transmission rate in all batches. Asterisks (*) placed on reassortants indicate a statistically significant variation from their respective major parental genotypes (MaP), while asterisks (*) on the controls indicate statistically significant variations between both parental genotypes in the same host (MaP in diagonal). Corresponding field and new hosts are displayed with same color as the major parental genome.

Based on the previous characterization of the **[AZ;15]** and **[AZ;10_12b]** parental genotypes, **[AZ;10_12b]** was transmitted statistically significantly better than **[AZ;15]** on vetches (**Torralba et al. 2024a – Chapter 2**).

We conducted four separate transmission batches, each assessing specific reassortants alongside both parental genotypes serving as controls (**Figure 7**). One general trend we observed was a higher transmission rate from vetch to vetch (V>V) compared to lentil to lentil (L>L) except in the batch 4, with a consistent pattern of higher transmission rates for **[AZ;10_12b]** compared to **[AZ;15]**. These findings align with our earlier parental genotypes characterization (**Torralba et al. 2024a**).

Overall, only few reassortants were statistically significantly affected in their transmission rates, with only four out of the results of the 4 out 16 for **[AZ;15]** and 3 out of 16 for **[AZ;10_12b]**. Moreover, only reassortants with a statistically significant reduced transmission rates were observed for **[AZ;10_12b]** as major parent, in both plant hosts. In contrast, reassortants with both statistically significant increased and decreased transmission rates were observed with **[AZ;15]** as major parent. For both parental genotypes, only one reassortant U1-15, exhibited statistically significant difference when present in the field host of its major parent, indicating that transmission rate in this context seems highly robust. Considering the generally superior transmission rate of **[AZ;10_12b]** compared to **[AZ;15]**, it is coherent that more reassortants would be negatively affected compared to **[AZ;15]**, if not all of them.

Interestingly, U1-15 was consistently associated with a reduced transmission rate in both field and new hosts in comparison to its major parent **[AZ;10_12b],** suggesting that this segment, despite its unknown function, potentially influences the transmission efficiency of its major genotype. Noteworthy due to their apparent substantial effect on transmission rate, although only concerning either one of the evaluated hosts, N-10 appeared to greatly enhance the transmission rate in the new host in comparison to its major parent **[AZ;15]**. This observation implies an optimized role of the helper component in the genomic context of **[AZ;15]** within this particular host. Furthermore, U2-10 and U4-10 also notably reduced the transmission rate within the new host environment, which hints at potential roles in either transmission or infection establishment in this host.

7. Biological traits assessment

FIELD HOST

NAME	SEGMENT	MaP	HEIGHT	SYMPT	PLANT VL	TR	C GF	M GF	N GF	R GF	S GF	U1 GF	U2 GF	U4 GF
C-10	C	AZ;15	=	=	=	=	63.75	-82.06	-48.58	-81.90	-133.95	-624.36	=	-111.1
M-10	M	AZ;15	=	=	=	=	=	=	=	=	=	=	=	-194.76
N-10	N	AZ;15	=	=	+	=	=	=	-340.58	=	=	=	=	=
R-10	R	AZ;15	=	=	+	=	-147.40	=	34.26	=	=	-78.03	=	=
S-10	S	AZ;15	=	=	=	=	=	=	=	=	=	=	=	=
U1-10	U1	AZ;15	-	=	+	=	-1860.33	-81.33	=	=	=	92.22	-186.00	=
U2-10	U2	AZ;15	=	=	=	=	=	44.39	=	=	=	44.67	-119.14	-55.19
U4-10	U4	AZ;15	=	-	=	=	=	=	=	=	=	=	=	-298.37
C-15	C	AZ;10_12b	=	=	-	=	-2504.22	=	=	=	=	44.14	=	=
M-15	M	AZ;10_12b	=	-	=	=	27.46	=	=	=	=	=	=	=
N-15	N	AZ;10_12b	=	=	-	=	=	=	=	=	=	=	=	=
R-15	R	AZ;10_12b	=	=	=	=	53.52	-87.71	-286.04	=	-75.24	=	-121.49	=
S-15	S	AZ;10_12b	=	=	=	=	-49.80	=	=	=	-326.17	=	-90.073	=
U1-15	U1	AZ;10_12b	=	=	=	-	43.28	=	=	-73.94	-253.27	-9324.9	-110.58	=
U2-15	U2	AZ;10_12b	=	=	-	=	=	=	=	=	=	=	=	=
U4-15	U4	AZ;10_12b	=	=	=	=	=	=	=	=	=	=	=	=

NEW HOST

NAME	SEGMENT	MaP	HEIGHT	SYMPT	PLANT VL	TR	C GF	M GF	N GF	R GF	S GF	U1 GF	U2 GF	U4 GF
C-10	C	AZ;15	+	=	=	=	62.37	-156.86	-74.98	-178.35	-92.51	-2113.08	-57.98	=
M-10	M	AZ;15	+	=	=	=	=	=	=	=	=	43.26	=	=
N-10	N	AZ;15	=	=	=	+	=	=	-94.20	=	=	=	=	=
R-10	R	AZ;15	=	=	=	=	-109.86	=	=	-90.20	=	-153.96	=	=
S-10	S	AZ;15	=	=	=	+	-52.21	=	=	=	-99.64	=	=	50.24
U1-10	U1	AZ;15	=	=	=	=	-2909.85	=	=	=	-224.32	91.33	-199.98	-73.979
U2-10	U2	AZ;15	=	+	=	-	=	52.11	=	=	=	=	-185.044	-103.61
U4-10	U4	AZ;15	=	-	=	-	=	=	=	=	=	=	=	-99.43
C-15	C	AZ;10_12b	=	=	=	-	-1792.92	=	37.46	=	=	29.48	=	=
M-15	M	AZ;10_12b	=	=	-	=	=	-47.35	44.73	=	=	=	=	=
N-15	N	AZ;10_12b	=	=	=	=	=	=	=	=	=	=	=	=
R-15	R	AZ;10_12b	=	-	=	=	58.22	-80.80	-90.74	=	-63.16	=	-107.11	=
S-15	S	AZ;10_12b	=	=	=	=	=	=	=	=	=	=	=	=
U1-15	U1	AZ;10_12b	=	-	+	-	55.75	=	=	-63.28	-80.67	-2583.73	=	=
U2-15	U2	AZ;10_12b	=	=	-	=	-48.80	=	=	=	=	=	=	=
U4-15	U4	AZ;10_12b	=	=	=	=	=	=	=	=	=	-85.14	=	=

Figure 8 : Summary tables depicting the phenotypic effects assessed in reassortants compared to their major parental genotypes in both field and new hosts. The first column presents the reassorted segment associated with its minor parental genotype. The "Maj Parent" column displays the major parental genotype. For each evaluated biological trait (height = impact on viral plant height, sympt = symptom severity evolution, plant vl = plant viral load, TR = transmission rate, X GF = segment X relative frequency in genome formula), case colors signify a statistical increase (+ and green) or decrease (- and red) in comparison to the major parental genotype. If there are no statistically significant variations, an equal sign is used instead. Finally, in the genome formula columns, the percentage change is indicated instead of (+) or (-) signs if it is significant.

The effects of the different traits are summarized in (**Figure 8**) for each reassortant in both the field and new host. The outcomes are categorized into three groups. The first one focuses on the effects on the host plant, namely impact on plant height and symptom severity. The second pertains to proxies of viral fitness, viral load and transmission rates. The final category addresses alterations in the genome formula, highlighting specific variations for each genomic segment.

Based on the outcomes of the viral fitness proxies, namely plant viral load and transmission rates, we identified, 3 out of 8 reassortants that exhibited improved fitness for **[AZ;15]** in its vetch field host, whereas 4 out of 8 reassortants displayed reduced fitness for **[AZ;10_12b]** in its lentil field host. In contrast, when evaluating **[AZ;15]** in its new lentil host, we observed 2 out of 8 reassortants with a negative impact and 2 out of 8 with a positive impact, while for **[AZ;10_12b]** in its new vetch host, 3 out of 8 reassortants were negatively impacted. Thus **[AZ;10_12b]** tended to experience more negative effects on its fitness following reassortments than **[AZ;15]** in both host environments, which aligns with its overall superior fitness in both hosts (**Torralba et al. 2024a**).

DISCUSSION

In our experimental setup, diverse reassortant profiles emerged across all evaluated biological traits. While the majority of these profiles showed no statistical deviations from the major parental genotypes, a notable subset of reassortants displayed significant positive or negative effects on some of the monitored traits. One intriguing finding pertains to the substantial alterations observed in the genome formula following reassortments. Some of these changes were quite dramatic, resulting in significantly reduced accumulation of reassorted segments. These alterations could be attributed to either intrinsic dynamics within the genome formula of the virus or processes affecting replication or degradation of these segments. We observed that most reassortants provoked significant alterations of the viral genome formula. Among pronounced changes in genome formula the cases of U1-10 in its new lentil host or M-10 in its vetch field host are intriguing as these reassortants did not display alterations in other evaluated biological traits. To account for this phenomenon, one hypothesis is that the observed variations in their genome formula served as compensation for the intra-genomic disruptions which in turn did not manifest as phenotype variations. This finding might suggest the possibility of a role played by the genome formula in adapting to the effect of genetic exchanges and genomic disruptions.

Of particular interest was the observation that when either DNA-C or -U1 was reassorted, there were significant opposed variations of their relative frequencies, and this trend was in opposite directions in the two major parental genotypes. This suggests an unforeseen interaction between these two segments or a common function mediated by their respective proteins. DNA-C is suspected to play a role in regulating host cell-cycle based on conserved protein domains (**Lageix et al. 2007**). DNA-U1 function remains completely unknown and formulating sound hypotheses about their putative physical or functional interaction remains a challenge.

Upon investigating and comparing the phenotypes of the 16 reassortants, several insights could be revealed. For instance, there seem to be no apparent correlation observed between our assessment of general symptom severity and the impact on plant height for a given genotype, although no correlations have been statistically assessed yet (**Figure 8**). While this might be surprising at first glance, it likely suggests that plant dwarfing is one of the earliest symptoms to manifest, whereas subsequent effects predominantly affect lateral bushing and branching, with limited influence on overall plant height. Furthermore, we did not observe a general link (general statistically significant alteration for the whole genotype) between general symptom development and plant viral load either, except for one genotype, U1-15, in its corresponding field host. And neither between our viral fitness proxies (**Figure 8**). For instance, reassortants that showed a significant reduction in transmission rates did not necessarily exhibit corresponding changes in plant viral load, and vice versa. However, it's important to note that these two biological traits were not assessed simultaneously and not on the same set of plants with viral load determined on 30 days old plants and the transmission rate using 18 days old plant sources, limiting their direct comparison. Again, these correlations will undergo further evaluation when analyzing individual plants in the near future.

Focusing on genomic segments, we found that certain segments were linked to reassortants affected either in both environmental contexts (field and new host) for a specific genotype, or that influenced both genotypes within the same host. These observations could offer valuable insights into the functions they are associated with. For instance, we found significant variations in the total viral load when DNA-N was reassorted in both major parental genotypes within their respective field hosts. This observation is surprising as it raises the possibility that DNA-N, whose only known function is in aphid transmission, may somehow also affect virus accumulation through, thus far, unreported processes. Delving further into the individual phenotypic effects of these N-reassortants may inform on additional functions of DNA-N, which could also be indirectly related to transmission. Another reassortant that exhibited an impact on its viral load was U2-15, which showed a negative effect in both field and new hosts

when compared to its major parent, **[AZ;10_12b]**. While DNA-U2 has previously been linked to reduced symptoms in faba beans for the **[JKI-2000]** isolate (**Grigoras et al. 2018**), these findings suggest a potentially new role in viral accumulation for the **[AZ;10_12b]** isolate.

Concerning transmission rates, we observed the reduce impact of transmission rate on both field and new host for the U1-15 reassortant that might indicate a role of DNA-U1 in either infection establishment or direct transmission efficiency for the [AZ;10_12b] isolate. Finally, we consistently observed a reduction in symptom severity for the U4-10 reassortant in both field and new hosts. This finding suggests that DNA-U4 could be a key segment influencing pathogenicity for **[AZ;15]** in both lentils and vetches. This result is particularly intriguing because DNA-U4, along with other segments, is frequently lost under laboratory conditions and prolonged maintenance on laboratory hosts (**Knierim et al. 2019**). This observation might confirm the pivotal role of DNA-U4 as a pathogenicity determinant in different host environments and might underscore its significance in the nanovirus infection cycle, which encompasses a wide range of host plant species.

These findings open up intriguing possibilities, as they reveal distinct effects stemming from reassortants linked to genomic segments of unknown functions. A more comprehensive exploration of the phenotypic traits of these reassortants could provide valuable insights into the functions of these enigmatic genomic segments.

A major result from this study is that all single-segment reassortants examined exhibited the capacity to infect all evaluated hosts, including faba beans, lentils, and vetches, as well as to be efficiently transmitted by the aphid vector *Aphis craccivora*. This observation indicates a propensity for reassortment between isolates of the same nanovirus species, probably facilitating genome reconstitutions and thus impacting on ecological and evolutionary dynamics of nanoviruses in regions where multiple isolates coexist. This finding is consistent with the observations in natural populations where the majority of identified reassortment events involve intra-specific single-segment reassortments. Nonetheless, it does not align with the observation that DNA-C is never found as a reassortant in natural population, as the corresponding genotypes in our experimental setup did not appear drastically affected in fitness traits, when compared to any others. Therefore, most likely, the absence of DNA-C reassortants in field samples analyzed and reported thus far can be attributed to the limited extent of nanovirus genomic studies (**Grigoras et al. 2014; Kraberger et al. 2017; Torralba, Blanc, and Michalakis 2023**).

We observed that at least 50% of reassortants did not appear to be significantly affected in terms of viral fitness. Those that were affected typically showed changes in either viral load or transmission rate with only U1-15 in the new host affected in both. However, the biological traits that we used as viral fitness proxies are associated with several limitations, including likely variations in plant viral load over the course of infection, the unknown optimal timeframe and dynamics for aphid vector infestation, their viral acquisition and subsequent transmission efficiency, or even the correlation of our evaluated biological traits with plants infected at later stages of development, and might thus not reflect an effective link between viral load and transmission rate in natural conditions.

When assessing viral fitness, we nevertheless observed a greater number of reassortants that displayed enhanced fitness for **[AZ;15]** major parent in its field host, and an equal number of reassortants with diminished fitness in its new host. This suggests a general advantage for reassortments between the two genotypes for **[AZ;15]** genomic background. In contrast, most of the reassortants with affected viral fitness in **[AZ;10_12b]** background appeared less fit than the major parental genotype in both the field and new host. This observation aligns with the previously reported overall better fitness of **[AZ;10_12b]** compared to **[AZ;15]** in the two evaluated hosts (**Torralba et al. 2024a**), a trend that is further validated in this experiment. It also suggests that the fitness cost resulting from reassortment events between different genotypes is primarily influenced by the dominant genotype in situations involving mixed infections and their overall fitness. While we can evaluate the relative fitness of reassortants, we currently do not possess an estimate of how frequently they emerge.

With multipartite ssDNA viruses not requiring mixed infection to produce reassortants (**Torralba, Blanc, Michalakis 2023**), the fact that many reassortants appear less fit than the parental genotypes but are still viable may be considered. Indeed, in such viral systems, the reassortants may often appear even in a single infection context. As such, our findings can be consistent again with a reduction of the cost of maintaining genome integrity, through genome reconstitution. However, co-infections are possible and may even be frequent, urging additional studies where the two alleles of the same segments and even full genotypes could compete. In this context, most of the viable reassortants characterized in this study may be rapidly outcompeted in natural conditions and thus go extinct. In any case, this shifts the focus from investigating viability to directly evaluate their competitivity in heterogenous viral populations.

In conclusion, these findings affirm the capacity for intra-specific complementation between divergent isolates of the same nanovirus species, which could hold profound implications for

their ecological dynamics, genetic exchanges, the emergence of hybrid genotypes, and their evolution, potentially leading to a reduction in the cost associated with the maintenance of genomic integrity. However, the limitations of our experiments underscore the need for further validation of the viability of these reassortants in a competitive environment, involving both homologous segments and other complete genotypes. This validation would help determine whether these reassortants not only possess the potential for emergence but also indeed arise naturally and maintain themselves within nanovirus populations. Additionally, future investigation should expand towards multiple-segments and inter-specific reassortments. These two factors are associated with significant disparities between experimental data predictions and their very low prevalence in natural populations observations (**Torralba, Blanc, and Michalakis 2023**). It would be particularly interesting to evaluate the threshold value at which the number of segments concomitantly exchanged in a reassorting event for a given couple of parental genotypes induces near systematic failure and to subsequently evaluate the number of feasible combinations in multiple-segment reassortants. This exploration would provide valuable insights into both the constraints and advantages of the reassortment process in nanoviruses that appears already consequential.

Bibliography

Chakraborty, S., Vanitharani, R., Chattopadhyay, B., & Fauquet, C. M. (2008). 'Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India', *The Journal of General Virology*, 89/Pt 3: 818–28. DOI: 10.1099/vir.0.82873-0

Chao, L., Tran, T., & Matthews, C. (1992). 'MULLER'S RATCHET AND THE ADVANTAGE OF SEX IN THE RNA VIRUS ϕ6', *Evolution; International Journal of Organic Evolution*, 46/2: 289–99. DOI: 10.1111/j.1558-5646.1992.tb02038.x

Di Mattia, J., Torralba, B., Yvon, M., Zeddam, J.-L., Blanc, S., & Michalakis, Y. (2022). 'Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution', *Proceedings of the National Academy of Sciences of the United States of America*, 119/32: e2201453119. DOI: 10.1073/pnas.2201453119

Di Mattia, J., Vernerey, M.-S., Yvon, M., Pirolles, E., Villegas, M., Gaafar, Y., Ziebell, H., et al. (2020). 'Route of a Multipartite Nanovirus across the Body of Its Aphid Vector', *Journal of Virology*, 94/9: e01998-19. DOI: 10.1128/JVI.01998-19

Escriu, F., Fraile, A., & García-Arenal, F. (2007). 'Constraints to genetic exchange support gene coadaptation in a tripartite RNA virus', *PLoS pathogens*, 3/1: e8. DOI: 10.1371/journal.ppat.0030008

Fraile, A., Alonso-Prados, J. L., Aranda, M. A., Bernal, J. J., Malpica, J. M., & García-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: 934–40. DOI: 10.1128/JVI.71.2.934-940.1997

Gallet, R., Di Mattia, J., Ravel, S., Zeddam, J.-L., Vitalis, R., Michalakis, Y., & Blanc, S. (2022). 'Gene copy number variations at the within-host population level modulate gene expression in a multipartite virus', *Virus Evolution*, 8/2: veac058. DOI: 10.1093/ve/veac058

Grigoras, I., Ginzo, A. I. D. C., Martin, D. P., Varsani, A., Romero, J., Mammadov, A. C., Huseynova, I. M., et al. (2014). 'Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe', *The Journal of General Virology*, 95/Pt 5: 1178–91. DOI: 10.1099/vir.0.063115-0

Gronenborn, B. (2004). 'Nanoviruses: genome organisation and protein function', *Veterinary Microbiology*, 98/2: 103–9. DOI: 10.1016/j.vetmic.2003.10.015

Holland, J., & Domingo, E. (1998). 'Origin and Evolution of Viruses', *Virus Genes*, 16/1: 13–21. DOI: 10.1023/A:1007989407305

Idris, A. M., Mills-Lujan, K., Martin, K., & Brown, J. K. (2008). 'Melon chlorotic leaf curl virus: characterization and differential reassortment with closest relatives reveal adaptive virulence in the squash leaf curl virus clade and host shifting by the host-restricted bean calico mosaic virus', *Journal of Virology*, 82/4: 1959–67. DOI: 10.1128/JVI.01992-07

Iranzo, J., & Manrubia, S. C. (2012). 'Evolutionary dynamics of genome segmentation in multipartite viruses', *Proceedings. Biological Sciences*, 279/1743: 3812–9. DOI: 10.1098/rspb.2012.1086

Kraberger, S., Kumari, S. G., Najar, A., Stainton, D., Martin, D. P., & Varsani, A. (2018). 'Molecular characterization of faba bean necrotic yellows viruses in Tunisia', *Archives of Virology*, 163/3: 687–94. DOI: 10.1007/s00705-017-3651-3

Lageix, S., Catrice, O., Deragon, J.-M., Gronenborn, B., Pélissier, T., & Ramírez, B. C. (2007). 'The nanovirus-encoded Clink protein affects plant cell cycle regulation through interaction with the retinoblastoma-related protein', *Journal of Virology*, 81/8: 4177–85. DOI: 10.1128/JVI.02103-06

Lal, A., Vo, T. T. B., Sanjaya, I. G. N. P. W., Ho, P. T., Kim, J.-K., Kil, E.-J., & Lee, S. (2020). 'Nanovirus Disease Complexes: An Emerging Threat in the Modern Era', *Frontiers in Plant Science*, 11: 558403. DOI: 10.3389/fpls.2020.558403

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., & Varsani, A. (2011). 'Recombination in Eukaryotic Single Stranded DNA Viruses', *Viruses*, 3/9: 1699–738. DOI: 10.3390/v3091699

Martin, D. P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.-Y., Dijoux, B., Vincent, C., et al. (2011). 'Complex Recombination Patterns Arising during Geminivirus Coinfections Preserve and Demarcate Biologically Important Intra-Genome Interaction Networks', *PLoS Pathogens*, 7/9: e1002203. DOI: 10.1371/journal.ppat.1002203

Matsuzaki, Y., Mizuta, K., Sugawara, K., Tsuchiya, E., Muraki, Y., Hongo, S., Suzuki, H., et al. (2003). 'Frequent Reassortment among Influenza C Viruses', *Journal of Virology*, 77/2: 871–81. DOI: 10.1128/JVI.77.2.871-881.2003

McDonald, S. M., Nelson, M. I., Turner, P. E., & Patton, J. T. (2016). 'Reassortment in segmented RNA viruses: mechanisms and outcomes', *Nature Reviews. Microbiology*, 14/7: 448–60. DOI: 10.1038/nrmicro.2016.46

Ohshima, K., Matsumoto, K., Yasaka, R., Nishiyama, M., Soejima, K., Korkmaz, S., Ho, S. Y. W., et al. (2016). 'Temporal analysis of reassortment and molecular evolution of Cucumber mosaic virus: Extra clues from its segmented genome', *Virology*, 487: 188–97. DOI: 10.1016/j.virol.2015.09.024

Roossinck, M. J. (1997). 'Mechanisms of plant virus evolution', *Annual Review of Phytopathology*, 35: 191–209. DOI: 10.1146/annurev.phyto.35.1.191

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den Hoff, M. J. B., & Moorman, A. F. M. (2009). 'Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data', *Nucleic Acids Research*, 37/6: e45. DOI: 10.1093/nar/gkp045

Ryckebusch, F., Peterschmitt, M., Granier, M., & Sauvion, N. (2021). 'Alfalfa leaf curl virus is efficiently acquired by its aphid vector Aphis craccivora but inefficiently transmitted', *The Journal of General Virology*, 102/2: 001516. DOI: 10.1099/jgv.0.001516

Sicard, A., Pirolles, E., Gallet, R., Vernerey, M.-S., Yvon, M., Urbino, C., Peterschmitt, M., et al. (2019). 'A multicellular way of life for a multipartite virus', *eLife*, 8: e43599. DOI: 10.7554/eLife.43599

Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakis, Y., Gutierrez, S., & Blanc, S. (2013). 'Gene copy number is differentially regulated in a multipartite virus', *Nature Communications*, 4: 2248. DOI: 10.1038/ncomms3248

Tentchev, D., Verdin, E., Marchal, C., Jacquet, M., Aguilar, J. M., & Moury, B. (2011). 'Evolution and structure of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes', *The Journal of General Virology*, 92/Pt 4: 961–73. DOI: 10.1099/vir.0.029082-0

Varsani, A., Lefeuvre, P., Roumagnac, P., & Martin, D. (2018). 'Notes on recombination and reassortment in multipartite/segmented viruses', *Current Opinion in Virology*, 33: 156–66. DOI: 10.1016/j.coviro.2018.08.013

Villa, M., & Lässig, M. (2017). 'Fitness cost of reassortment in human influenza', *PLoS pathogens*, 13/11: e1006685. DOI: 10.1371/journal.ppat.1006685

Supplementary Information for

Systematic phenotypic characterization of three isolates of the faba bean necrotic stunt virus (FBNSV) in multiple host species and common vector

Babil Torralba, Aleksandra Greshnova, Andy Brousse, Prune Lacôte-Popovic, Stéphane Blanc and Yannis Michalakis

Species	Isolate	Segment	Number
FBNSV	AZ;10	C	KC978976.1
FBNSV	AZ;10	M	KC978977.1
FBNSV	AZ;10	N	KC978978.1
FBNSV	AZ;10	R	KC978974.1
FBNSV	AZ;10	S	KC978975.1
FBNSV	AZ;10	U1	KC978979.1
FBNSV	AZ;12b	U2	KC978988.1
FBNSV	AZ;10	U4	KC978981.1
FBNSV	AZ;15	C	KC978994.1
FBNSV	AZ;15	M	KC978995.1
FBNSV	AZ;15	N	KC978996.1
FBNSV	AZ;15	R	KC978992.1
FBNSV	AZ;15	S	KC978993.1
FBNSV	AZ;15	U1	KC978997.1
FBNSV	AZ;15	U2	KC978998.1
FBNSV	AZ;15	U4	KC978999.1

Supplementary Table 1 : Summary table containing the GenBank accession numbers for all the sequences of genomic segments used in this study.

Host plant	Score	Associated symptoms
Vicia sativa	0	No apparent symptoms, leaves can be slightly curved
Vicia sativa	0.5	Suspected curvatures of the leaves or small leaves development
Vicia sativa	1	Apparent symptoms, apex leaflets are curving on themselves
Vicia sativa	1.5	Multiple foliar levels are curving and start to twist
Vicia sativa	2	Pronounced leaves deformations. New branches start to develop at the nodes
Vicia sativa	2.5	New branches are developing and are symptomatic (small and/or twisted leaves)
Vicia sativa	3	Multiple branches are well developed and clearly symptomatic, bushy appearance starting
Vicia sativa	3.5	Bushing development
Vicia sativa	4	Extreme case of bushing
Lens culinaris	0	No apparent symptoms, leaves can be slightly curved
Lens culinaris	0.5	Suspected curvatures of the leaves or small leaves development
Lens culinaris	1	Apparent symptoms, apex leaflets are curving on themselves
Lens culinaris	1.5	Multiple foliar levels are curving and start to twist
Lens culinaris	2	New branches start to develop at the nodes
Lens culinaris	2.5	New branches are developing and are symptomatic (small and/or curled leaves)
Lens culinaris	3	Multiple branches are well developed and clearly symptomatic, bushy appearance starting
Lens culinaris	3.5	Bushing development
Lens culinaris	4	Extreme case of bushing

<u>Supplementary Table 2</u> : Symptom severity evaluation grid for the assessment of the progression of disease index scoring on *Vicia sativa* (vetch) and *Lens culinaris* (lentil). Bottom panel display extreme cases of bushing for both lentils and vetches with plants usually >30 dpi. Blue darts indicates the developments of multiple branches in stark contrast to the usual 2 to 3 stems of healthy plants (**Torralba et al. 2024a – Chapter 2**).

ISOLATE	SEGMENT	SEQUENCE	SENSE	LENGTH	[ ] (µM)
AZ;10_12b	C - For	ACCCATCAACTGAAGATTTG	F	20	1
AZ;10_12b	C - Rev	ACCATACTTAACTTGGAAGGG	R	21	
AZ;10_12b	M - For	AGGAACGATGTCGGATT	F	17	0.8
AZ;10_12b	M - Rev	ACCTATCATGTACAGAGCCT	R	20	
AZ;10_12b	N - For	AGTGTTAGCTTCAACGGGA	F	19	0.5
AZ;10_12b	N - Rev	ACGAACGTTTCTGTTTCCTC	R	20	
AZ;10_12b	R - For	GGAGTATGGAGAGTTCATTGTT	F	22	0.8
AZ;10_12b	R - Rev	GAACTCTCTTAATGTGCCTGAA	R	22	
AZ;10_12b	S - For	GTTTACAGCACCGGCTT	F	17	0.8
AZ;10_12b	S - Rev	ATACAGCCTTAACAGCCTCA	R	20	
AZ;10_12b	U1 - For	AGTTAGATGCCGTTAGTCG	F	19	0.8
AZ;10_12b	U1 - Rev	CAGACAACAGATCTGCGT	R	18	
AZ;10_12b	U2 - For	CTGGTTTTAATGGCTACCG	F	19	0.5
AZ;10_12b	U2 - Rev	GGTGCGTAACTGGGTAAC	R	18	
AZ;10_12b	U4 - For	TTCGAGGATGTTGATGCAG	F	19	0.5
AZ;10_12b	U4 - Rev	TCGATTCCCATAGGCGTT	R	18	
AZ;15	C - For	CAAGATGCTTAGTCAAGTGGAA	F	22	1.5
AZ;15	C - Rev	CAGTGAATAACTTATCTAATTCACTGC	R	27	
AZ;15	M - For	ATCAAGACGACGGAGATG	F	18	1
AZ;15	M - Rev	CTTCTTGACAAATCCAGGTAG	R	21	
AZ;15	N - For	GTTAAATGATTCAAGGAAGGTGTTAC	F	26	1.5
AZ;15	N - Rev	GACCACGAACATTTCTATTTCCG	R	23	
AZ;15	R - For	ATGGTGAGTTCGTTACGTC	F	19	1
AZ;15	R - Rev	CCTCTAAATTCTCTCAACGTACTA	R	24	
AZ;15	S - For	ACGACGATATTCAAGGCCA	F	19	1
AZ;15	S - Rev	CCTGCAACACCATTAGATTCC	R	21	
AZ;15	U1 - For	CTCTGGTTAGTCGAGCTTC	F	19	1
AZ;15	U1 - Rev	CTGAGAATCATCTTCACAACATGG	R	21	
AZ;15	U2 - For	CTTCATTTCATCTTGTTCTATTCGC	F	25	1
AZ;15	U2 - Rev	CCAAACGAGTAGGAGCATATG	R	21	
AZ;15	U4 - For	GTTGTTAAATCCTTCTTTGGTGG	F	23	1
AZ;15	U4 - Rev	CGTCTTCTTTCTTCTTACCCATG	R	23	

<u>Supplementary Table 3</u> : Summary table presenting the pairs of primers employed for the targeted amplification of the genomic segments from [AZ;15] and [AZ;10_12b] isolates. The "SENSE" column specifies whether the corresponding primer is in the forward (F) or reverse (R) orientation. The final column [ ] indicates the recommended primer concentration (µm) for use in the qPCR mix

ISOLATE	SEGMENT	HOST	CT	n
AZ;10_12b	C	FB	30.007	19
AZ;10_12b	M	FB	32.717	19
AZ;10_12b	N	FB	31.176	19
AZ;10_12b	R	FB	27.590	19
AZ;10_12b	S	FB	30.761	19
AZ;10_12b	U1	FB	30.540	19
AZ;10_12b	U2	FB	33.620	16
AZ;10_12b	U4	FB	33.196	9
AZ;10_12b	C	L	30.486	38
AZ;10_12b	M	L	31.953	26
AZ;10_12b	N	L	27.729	34
AZ;10_12b	R	L	30.952	31
AZ;10_12b	S	L	32.014	29
AZ;10_12b	U1	L	27.937	23
AZ;10_12b	U2	L	28.169	43
AZ;10_12b	U4	L	32.421	32
AZ;10_12b	C	V	27.087	35
AZ;10_12b	M	V	31.097	31
AZ;10_12b	N	V	30.566	31
AZ;10_12b	R	V	32.508	24
AZ;10_12b	S	V	27.592	30
AZ;10_12b	U1	V	25.242	23
AZ;10_12b	U2	V	34.351	28
AZ;10_12b	U4	V	30.888	39
AZ;15	C	FB	30.767	19
AZ;15	M	FB	31.371	20
AZ;15	N	FB	28.497	19
AZ;15	R	FB	30.548	19
AZ;15	S	FB	29.450	10
AZ;15	U1	FB	32.723	19
AZ;15	U2	FB	27.674	18
AZ;15	U4	FB	27.620	18
AZ;15	C	L	29.299	27
AZ;15	M	L	29.376	25
AZ;15	N	L	33.095	18
AZ;15	R	L	32.960	21
AZ;15	S	L	32.131	11
AZ;15	U1	L	33.706	27
AZ;15	U2	L	27.543	39
AZ;15	U4	L	27.946	21
AZ;15	C	V	28.357	18
AZ;15	M	V	31.371	20
AZ;15	N	V	25.356	10
AZ;15	R	V	32.084	16
AZ;15	S	V	32.569	22
AZ;15	U1	V	32.036	14
AZ;15	U2	V	26.780	32
AZ;15	U4	V	31.434	14

Supplementary Table 4 : Threshold Ct values that were statistically determined for each genomic segment of the respective isolates across all assessed plant host species.
The "Host" column specifies the host plant species (FB = faba bean, L = lentil, V = vetch) for which the thresholds (CT) were statistically determined. The column labeled 'n' indicates the number of plants used to establish these thresholds.

A

Source	Chisq	Df	Pr..Chisq.
(Intercept)	3700.412	1	<0.001
id	232.022	17	<0.001
host	205.460	1	<0.001
id:host	107.800	17	<0.001

B

Reass	MaP	Host	Diff	Lower	Upper	P.value	SE	df
M	AZ15	L	-0.081	-0.127	-0.035	<0.001	0.016	1127
R	AZ15	L	-0.004	-0.060	0.052	1.000	0.019	1137
S	AZ15	L	-0.027	-0.076	0.023	0.850	0.017	1117
U1	AZ15	L	-0.025	-0.077	0.027	0.934	0.018	1124
C	AZ15	L	-0.049	-0.096	-0.003	0.029	0.016	1091
N	AZ15	L	-0.017	-0.078	0.043	1.000	0.021	1128
U2	AZ15	L	-0.027	-0.069	0.014	0.550	0.014	1119
U4	AZ15	L	0.004	-0.056	0.063	1.000	0.020	1138
M	AZ10_12b	L	0.004	-0.034	0.042	1.000	0.013	1121
R	AZ10_12b	L	-0.015	-0.059	0.028	0.996	0.015	1131
S	AZ10_12b	L	0.028	-0.021	0.077	0.780	0.017	1053
U1	AZ10_12b	L	0.045	-0.005	0.096	0.128	0.017	1065
C	AZ10_12b	L	0.003	-0.035	0.042	1.000	0.013	1096
N	AZ10_12b	L	-0.002	-0.043	0.038	1.000	0.014	1104
U2	AZ10_12b	L	0.021	-0.016	0.058	0.805	0.012	1111
U4	AZ10_12b	L	-0.003	-0.042	0.036	1.000	0.013	1117
M	AZ15	V	0.023	-0.028	0.075	0.960	0.017	1137
R	AZ15	V	0.032	-0.023	0.087	0.778	0.019	1138
S	AZ15	V	-0.015	-0.066	0.037	1.000	0.017	1101
U1	AZ15	V	0.199	0.137	0.260	<0.001	0.021	1126
C	AZ15	V	0.000	-0.055	0.055	1.000	0.019	1122
N	AZ15	V	-0.055	-0.114	0.005	0.101	0.020	1129
U2	AZ15	V	0.022	-0.022	0.065	0.915	0.015	1132
U4	AZ15	V	0.020	-0.040	0.080	0.998	0.020	1140
M	AZ10_12b	V	0.003	-0.037	0.043	1.000	0.014	1133
R	AZ10_12b	V	0.029	-0.018	0.076	0.687	0.016	1138
S	AZ10_12b	V	0.039	-0.006	0.083	0.150	0.015	1022
U1	AZ10_12b	V	0.025	-0.025	0.075	0.905	0.017	1071
C	AZ10_12b	V	0.032	-0.011	0.075	0.356	0.014	1099
N	AZ10_12b	V	0.004	-0.039	0.047	1.000	0.015	1101
U2	AZ10_12b	V	0.024	-0.016	0.064	0.731	0.013	1118
U4	AZ10_12b	V	0.007	-0.031	0.046	1.000	0.013	1114

Supplementary Table 5 : [A] Type III test of fixed effects for standardized height ratio model with host, and genotype as fixed effect and batch as random effect. Significative p.values are colored in green (<0.05). **[B] Pairwise comparisons of standardized height ratio (reassortant – major parent) using Emmeans.** Negative "Diff" values mean a more pronounced effect on plant height (more reduction) of the reassortant, compared to the major parent (control) in the same host. Significative p.values are colored in green (<0.05).

A

Model F2.LD.F1

[AZ;15]

Source	Chisq	Df1	Df2	P.Value
Host	15.762	1.000	80.428	<0.001
Batch	5.577	2.513	80.428	0.003
Host:Batch	1.737	2.513	80.428	0.174

Model F2.LD.F1

[AZ;10-12b]

Source	Chisq	Df1	Df2	P.Value
Host	65.871	1.000	192.429	<0.001
Batch	32.910	2.951	192.429	<0.001
Host:Batch	7.162	2.951	192.429	<0.001

B

Full Model F2.LD.F1

Source	Chisq	Df1	Df2	P.Value
Host	79.638	1.000	167.479	<0.001
Genotype	15.208	7.952	167.479	<0.001
Host:Genotype	3.369	7.952	167.479	0.001

C

Models F1.LD.F1

Batch	Host	Chisq	Df1	Df2	P.Value
1	V	0.689	4.259	87.385	0.610
1	L	0.900	4.129	82.403	0.470
2	V	2.119	3.422	91.781	0.095
2	L	7.321	3.410	46.409	<0.001
3	V	6.739	3.661	54.122	<0.001
3	L	0.927	2.380	32.910	0.420
4	V	7.658	4.150	92.296	<0.001
4	L	1.649	4.764	116.375	0.156

D

Batch	Reass	MaP	Host	Diff	P.Value
1	M	AZ15	L	0.105	1.099
1	M	AZ10/12b	L	-0.319	0.040
1	R	AZ15	L	-0.195	0.533
1	R	AZ10/12b	L	-0.247	0.178
1	M	AZ15	V	-0.218	0.526
1	M	AZ10/12b	V	-0.328	0.157
1	R	AZ15	V	-0.334	0.151
1	R	AZ10/12b	V	-0.562	0.005
2	S	AZ15	L	-0.368	0.051
2	S	AZ10/12b	L	0.050	1.563
2	U1	AZ15	L	-0.184	0.499
2	U1	AZ10/12b	L	-0.380	0.069
2	S	AZ15	V	0.212	0.493
2	S	AZ10/12b	V	-0.213	0.449
2	U1	AZ15	V	-0.110	1.047
2	U1	AZ10/12b	V	-0.503	0.013
3	C	AZ10/12b	L	0.329	0.063
3	C	AZ15	L	0.100	1.186
3	N	AZ10/12b	L	-0.239	0.301
3	N	AZ15	L	0.072	1.388
3	C	AZ10/12b	V	-0.107	1.109
3	C	AZ15	V	0.262	0.300
3	N	AZ10/12b	V	-0.197	0.587
3	N	AZ15	V	0.038	1.656
4	U2	AZ10/12b	L	0.209	0.262
4	U2	AZ15	L	0.398	0.019
4	U4	AZ10/12b	L	-0.085	1.110
4	U4	AZ15	L	-0.487	0.004
4	U4	AZ15	V	-0.500	0.004
4	U2	AZ10/12b	V	0.210	0.342
4	U2	AZ15	V	0.032	1.683
4	U4	AZ10/12b	V	0.232	0.256

<u>Supplementary Table 6</u> : [A] ANOVA-type statistic with Box-Cox transformation results reported for the between-subject factors (batch and host) and their interaction for both parental genotypes. Significative p.values are colored in green (<0.05). **[B] ANOVA-type statistic with Box-Cox transformation results reported for the between-subject factors (genotype and host) and their interaction.** Significative p.values are colored in green (<0.05). **[C] Symptom dynamics F1.LD.F1 models per batch, with genotype as between-subject factor and time as within-subject factor.** Significative p.values are colored in green (<0.05). **[D] Differences in RTE values for symptom dynamics between reassortants and their major parents**. Positive 'Diff' values mean a more pronounced effect than major parent. Significative p.values are colored in green (<0.05).

A

Batch	Source	Sum.Sq	Df	F.Value	Pr(>F)
1	(Intercept)	271.364	1.000	105276.900	>0.001
	id	0.125	5.000	9.724	>0.001
	host	0.206	1.000	79.884	>0.001
	id:host	0.034	5.000	2.648	0.023
	Residuals	0.776	301.000	NA	NA
2	(Intercept)	215.830	1.000	22536.092	>0.001
	id	0.350	5.000	7.308	>0.001
	host	0.834	1.000	87.033	>0.001
	id:host	0.336	5.000	7.018	>0.001
	Residuals	2.145	224.000	NA	NA
3	(Intercept)	225.452	1.000	111122.400	>0.001
	id	0.059	5.000	5.768	>0.001
	host	0.572	1.000	281.985	>0.001
	id:host	0.036	5.000	3.565	>0.001
	Residuals	0.562	277.000	NA	NA
4	(Intercept)	286.447	1.000	93902.650	>0.001
	id	0.222	5.000	14.534	>0.001
	host	1.538	1.000	504.223	>0.001
	id:host	0.012	5.000	0.787	>0.001
	Residuals	1.019	334.000	NA	NA

B

Batch	Reass	MaP	Host	Diff	Lower	Upper	P.Value	Df
1	M	AZ;15	L	0.000	-0.002	0.003	0.999	301
1	R	AZ;15	L	-0.001	-0.004	0.001	0.605	301
1	M	AZ;10_12b	L	0.001	-0.002	0.004	0.676	301
1	R	AZ;10_12b	L	-0.002	-0.005	0.001	0.163	301
1	M	AZ;15	V	-0.001	-0.002	0.000	0.170	301
1	R	AZ;15	V	-0.003	-0.005	-0.002	>0.001	301
1	M	AZ;10_12b	V	0.002	0.000	0.004	0.031	301
1	R	AZ;10_12b	V	0.000	-0.002	0.002	1.000	301
2	S	AZ;15	L	-0.001	-0.003	0.002	0.940	224
2	U1	AZ;15	L	-0.001	-0.003	0.002	0.827	224
2	S	AZ;10_12b	L	0.001	-0.002	0.005	0.835	224
2	U1	AZ;10_12b	L	0.002	-0.002	0.005	0.596	224
2	S	AZ;15	V	-0.001	-0.002	0.000	0.097	224
2	U1	AZ;15	V	-0.001	-0.002	0.000	0.026	224
2	S	AZ;10_12b	V	0.000	-0.002	0.002	1.000	224
2	U1	AZ;10_12b	V	-0.004	-0.006	-0.001	0.002	224
3	C	AZ;15	L	0.001	-0.002	0.003	0.955	277
3	N	AZ;15	L	0.001	-0.002	0.003	0.928	277
3	C	AZ;10_12b	L	0.005	0.002	0.007	>0.001	277
3	N	AZ;10_12b	L	0.003	0.001	0.006	0.001	277
3	C	AZ;15	V	-0.001	-0.002	0.000	0.221	277
3	N	AZ;15	V	-0.001	-0.002	0.000	0.001	277
3	C	AZ;10_12b	V	0.001	0.000	0.002	0.058	277
3	N	AZ;10_12b	V	0.000	-0.001	0.001	0.964	277
4	U2	AZ;15	L	0.001	-0.003	0.005	0.850	334
4	U4	AZ;15	L	-0.003	-0.006	0.001	0.252	334
4	U2	AZ;10_12b	L	0.006	0.002	0.009	>0.001	334
4	U4	AZ;10_12b	L	0.001	-0.003	0.005	0.929	334
4	U2	AZ;15	V	0.000	-0.001	0.001	0.909	334
4	U4	AZ;15	V	-0.001	-0.001	0.000	0.241	334
4	U2	AZ;10_12b	V	0.002	0.001	0.003	0.002	334
4	U4	AZ;10_12b	V	0.000	-0.002	0.001	0.919	334

Supplementary Table 7 : [A] Type III test of fixed effect for relative viral load models per batch with genotype, host as fixed effects. Significative p.values are colored in green (<0.05). **[B] Pairwise comparisons of relative viral load (reassortant – major parent) using Emmeans.** Negative 'Diff' values mean higher viral load of a reassortant, compared to the major parent (control). Significative p.values are colored in green (<0.05).

A

Source	ChiSq	Df	Pr(>ChiSq)
(Intercept)	4767,642	1,000	<0,001
id	575,386	17,000	<0,001
host	2,455	1,000	0,117
segment	10043,781	7,000	<0,001
id:host	84,633	17,000	<0,001
id:segment	8905,487	119,000	<0,001
host:segment	70,872	7,000	<0,001
id:host:segment	471,893	119,000	<0,001

B

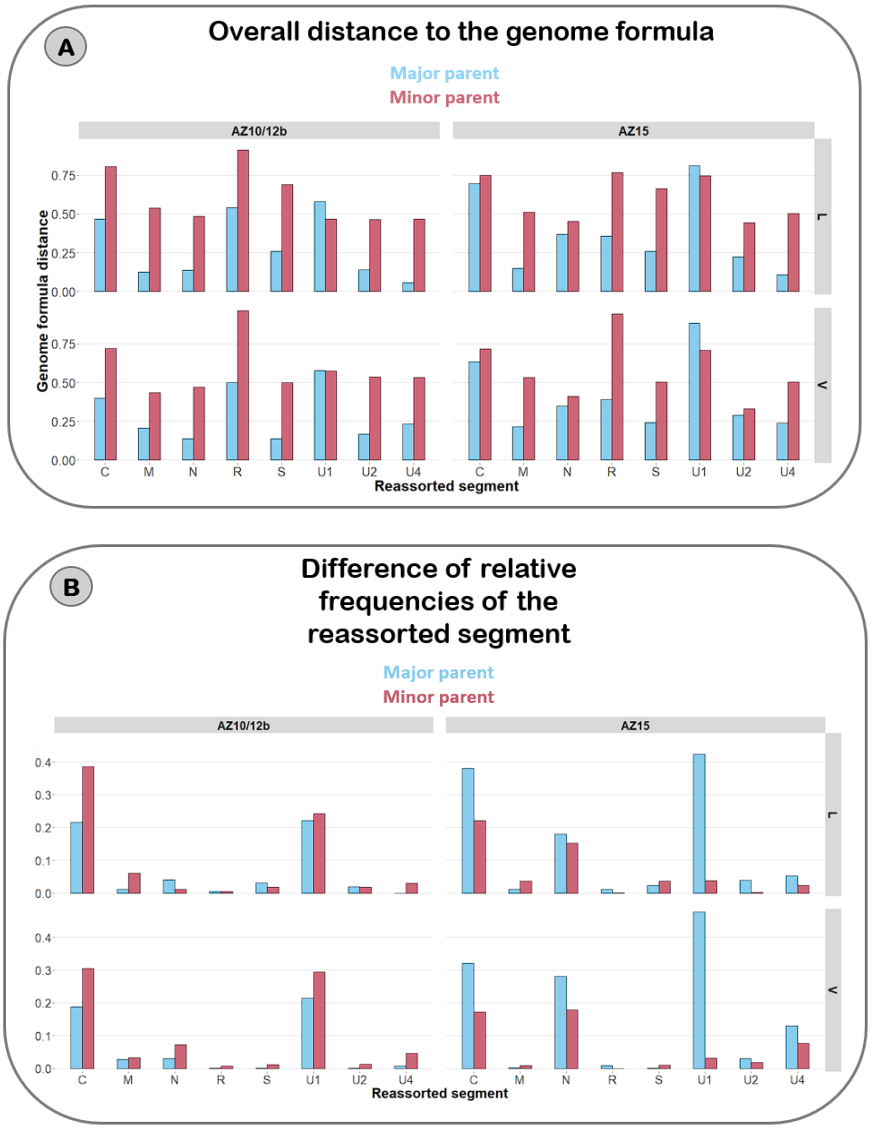
Host	MaP	id	segment	diff	SE	change	pval	upr.ci	lwr.ci
	AZ;10_12b	M-15	C	0,085	0,028	27,464	0,045	0,169	0,001
	AZ;10_12b	M-15	M	-0,012	0,010	-13,436	0,994	0,022	-0,045
	AZ;10_12b	M-15	N	0,005	0,020	2,866	1,000	0,068	-0,057
L	AZ;10_12b	M-15	R	-0,002	0,002	-10,516	1,000	0,005	-0,008
	AZ;10_12b	M-15	S	0,005	0,006	9,819	1,000	0,022	-0,013
	AZ;10_12b	M-15	U1	0,003	0,023	1,506	1,000	0,075	-0,068
	AZ;10_12b	M-15	U2	-0,006	0,004	-16,684	0,976	0,008	-0,019
	AZ;10_12b	M-15	U4	0,008	0,011	9,244	1,000	0,042	-0,025
	AZ;10_12b	R-15	C	0,259	0,034	53,519	0,000	0,368	0,150
	AZ;10_12b	R-15	M	-0,046	0,008	-87,713	0,000	-0,019	-0,072
	AZ;10_12b	R-15	N	-0,134	0,011	-286,037	0,000	-0,099	-0,169
L	AZ;10_12b	R-15	R	0,005	0,003	21,677	0,915	0,015	-0,006
	AZ;10_12b	R-15	S	-0,018	0,004	-75,238	0,001	-0,005	-0,030
	AZ;10_12b	R-15	U1	0,047	0,028	17,268	0,816	0,134	-0,041
	AZ;10_12b	R-15	U2	-0,021	0,003	-121,489	0,000	-0,011	-0,031
	AZ;10_12b	R-15	U4	-0,013	0,010	-18,674	0,964	0,018	-0,044
	AZ;10_12b	S-15	C	-0,075	0,020	-49,799	0,015	-0,009	-0,140
	AZ;10_12b	S-15	M	0,025	0,016	20,237	0,884	0,075	-0,026
	AZ;10_12b	S-15	N	-0,014	0,020	-8,422	1,000	0,051	-0,079
L	AZ;10_12b	S-15	R	0,003	0,003	12,962	1,000	0,012	-0,007
	AZ;10_12b	S-15	S	-0,031	0,003	-326,175	0,000	-0,022	-0,041
	AZ;10_12b	S-15	U1	0,055	0,029	19,768	0,679	0,148	-0,038
	AZ;10_12b	S-15	U2	-0,018	0,004	-90,074	0,000	-0,007	-0,030
	AZ;10_12b	S-15	U4	0,037	0,015	31,194	0,257	0,085	-0,011
	AZ;10_12b	U1-15	C	0,171	0,039	43,281	0,002	0,297	0,046
	AZ;10_12b	U1-15	M	-0,008	0,013	-9,077	1,000	0,034	-0,051
	AZ;10_12b	U1-15	N	0,079	0,031	30,475	0,224	0,178	-0,020
L	AZ;10_12b	U1-15	R	-0,008	0,002	-73,938	0,005	-0,002	-0,013
	AZ;10_12b	U1-15	S	-0,029	0,003	-253,274	0,000	-0,020	-0,039
	AZ;10_12b	U1-15	U1	-0,221	0,012	-9324,902	0,000	-0,183	-0,258
	AZ;10_12b	U1-15	U2	-0,020	0,004	-110,580	0,000	-0,009	-0,032
	AZ;10_12b	U1-15	U4	0,042	0,018	33,913	0,290	0,098	-0,013
	AZ;10_12b	C-15	C	-0,216	0,012	-2504,217	0,000	-0,179	-0,253
	AZ;10_12b	C-15	M	0,002	0,012	1,635	1,000	0,039	-0,036
	AZ;10_12b	C-15	N	0,065	0,023	26,386	0,138	0,139	-0,010
L	AZ;10_12b	C-15	R	0,002	0,002	8,119	1,000	0,009	-0,006
	AZ;10_12b	C-15	S	0,004	0,005	8,606	1,000	0,021	-0,013
	AZ;10_12b	C-15	U1	0,178	0,030	44,139	0,000	0,270	0,082
	AZ;10_12b	C-15	U2	0,002	0,005	3,637	1,000	0,017	-0,014
	AZ;10_12b	C-15	U4	-0,001	0,010	-1,749	1,000	0,029	-0,032
	AZ;10_12b	N-15	C	-0,040	0,021	-21,799	0,653	0,027	-0,107
	AZ;10_12b	N-15	M	-0,007	0,011	-7,808	1,000	0,029	-0,043
	AZ;10_12b	N-15	N	0,040	0,023	18,264	0,775	0,114	-0,033
L	AZ;10_12b	N-15	R	0,003	0,003	16,859	0,980	0,012	-0,005
	AZ;10_12b	N-15	S	0,013	0,007	23,427	0,649	0,033	-0,008
	AZ;10_12b	N-15	U1	-0,002	0,023	-0,740	1,000	0,072	-0,075
	AZ;10_12b	N-15	U2	0,011	0,006	21,628	0,799	0,030	-0,009
	AZ;10_12b	N-15	U4	-0,021	0,009	-34,149	0,284	0,007	-0,049
	AZ;10_12b	U2-15	C	-0,046	0,019	-26,029	0,275	0,014	-0,107
	AZ;10_12b	U2-15	M	-0,018	0,010	-21,791	0,722	0,013	-0,048
	AZ;10_12b	U2-15	N	0,033	0,021	16,392	0,884	0,100	-0,034
L	AZ;10_12b	U2-15	R	0,003	0,002	12,555	0,996	0,010	-0,005
	AZ;10_12b	U2-15	S	0,006	0,005	11,944	0,998	0,023	-0,012
	AZ;10_12b	U2-15	U1	-0,002	0,021	-0,811	1,000	0,066	-0,070
	AZ;10_12b	U2-15	U2	0,019	0,006	33,229	0,083	0,040	-0,001
	AZ;10_12b	U2-15	U4	-0,013	0,008	-19,331	0,873	0,013	-0,040
	AZ;10_12b	U4-15	C	-0,007	0,023	-3,228	1,000	0,065	-0,079
	AZ;10_12b	U4-15	M	-0,002	0,011	-1,940	1,000	0,034	-0,037
	AZ;10_12b	U4-15	N	0,001	0,019	0,445	1,000	0,062	-0,061
L	AZ;10_12b	U4-15	R	0,001	0,002	3,740	1,000	0,008	-0,007
	AZ;10_12b	U4-15	S	-0,004	0,005	-11,496	1,000	0,011	-0,020
	AZ;10_12b	U4-15	U1	-0,033	0,020	-17,407	0,857	0,032	-0,098
	AZ;10_12b	U4-15	U2	-0,010	0,004	-35,174	0,201	0,002	-0,022
	AZ;10_12b	U4-15	U4	0,000	0,010	0,009	1,000	0,031	-0,031

Host	MaP	id	segment	diff	SE	change	pval	upr.ci	lwr.ci
	AZ;10_12b	M-15	C	0,021	0,024	9,387	1,000	0,096	-0,055
	AZ;10_12b	M-15	M	-0,027	0,008	-47,351	0,032	-0,001	-0,053
	AZ;10_12b	M-15	N	0,126	0,026	44,734	0,000	0,208	0,043
V	AZ;10_12b	M-15	R	-0,007	0,003	-42,408	0,111	0,001	-0,015
	AZ;10_12b	M-15	S	-0,006	0,004	-19,366	0,948	0,008	-0,020
	AZ;10_12b	M-15	U1	0,013	0,024	5,551	1,000	0,090	-0,064
	AZ;10_12b	M-15	U2	-0,002	0,005	-4,031	1,000	0,015	-0,018
	AZ;10_12b	M-15	U4	0,004	0,012	3,788	1,000	0,040	-0,033
	AZ;10_12b	R-15	C	0,277	0,040	58,219	0,000	0,401	0,152
	AZ;10_12b	R-15	M	-0,038	0,008	-80,804	0,001	-0,012	-0,064
	AZ;10_12b	R-15	N	-0,074	0,015	-90,738	0,000	-0,028	-0,119
V	AZ;10_12b	R-15	R	0,001	0,004	5,110	1,000	0,014	-0,011
	AZ;10_12b	R-15	S	-0,015	0,004	-63,162	0,018	-0,002	-0,028
	AZ;10_12b	R-15	U1	0,048	0,032	17,187	0,934	0,147	-0,054
	AZ;10_12b	R-15	U2	-0,020	0,004	-107,113	0,000	-0,009	-0,031
	AZ;10_12b	R-15	U4	-0,031	0,010	-52,566	0,063	0,001	-0,063
	AZ;10_12b	S-15	C	-0,013	0,022	-7,108	1,000	0,056	-0,083
	AZ;10_12b	S-15	M	0,010	0,012	10,733	1,000	0,049	-0,029
	AZ;10_12b	S-15	N	0,051	0,023	24,814	0,405	0,123	-0,021
V	AZ;10_12b	S-15	R	0,004	0,004	14,494	0,996	0,016	-0,008
	AZ;10_12b	S-15	S	0,001	0,006	2,735	1,000	0,018	-0,016
	AZ;10_12b	S-15	U1	0,028	0,027	11,011	0,997	0,112	-0,057
	AZ;10_12b	S-15	U2	0,001	0,006	2,896	1,000	0,018	-0,016
	AZ;10_12b	S-15	U4	0,028	0,015	24,035	0,622	0,074	-0,017
	AZ;10_12b	U1-15	C	0,250	0,038	55,752	0,000	0,369	0,131
	AZ;10_12b	U1-15	M	-0,010	0,012	-12,738	1,000	0,027	-0,046
	AZ;10_12b	U1-15	N	0,014	0,022	8,063	1,000	0,083	-0,056
V	AZ;10_12b	U1-15	R	-0,010	0,003	-63,283	0,030	-0,001	-0,018
	AZ;10_12b	U1-15	S	-0,017	0,004	-80,672	0,002	-0,005	-0,030
	AZ;10_12b	U1-15	U1	-0,215	0,012	-2583,734	0,000	-0,176	-0,254
	AZ;10_12b	U1-15	U2	0,009	0,007	18,466	0,982	0,031	-0,013
	AZ;10_12b	U1-15	U4	0,055	0,019	37,978	0,109	0,116	-0,006
	AZ;10_12b	C-15	C	-0,188	0,011	-1792,920	0,000	-0,153	-0,224
	AZ;10_12b	C-15	M	0,015	0,012	14,738	0,986	0,052	-0,023
	AZ;10_12b	C-15	N	0,093	0,024	37,459	0,007	0,169	0,017
V	AZ;10_12b	C-15	R	0,001	0,004	5,099	1,000	0,012	-0,010
	AZ;10_12b	C-15	S	0,003	0,006	8,104	1,000	0,021	-0,014
	AZ;10_12b	C-15	U1	0,093	0,029	29,482	0,036	0,183	0,004
	AZ;10_12b	C-15	U2	0,003	0,006	7,312	1,000	0,020	-0,014
	AZ;10_12b	C-15	U4	0,003	0,012	3,013	1,000	0,039	-0,033
	AZ;10_12b	N-15	C	-0,009	0,022	-4,930	1,000	0,060	-0,079
	AZ;10_12b	N-15	M	0,016	0,013	15,737	0,983	0,055	-0,024
	AZ;10_12b	N-15	N	0,030	0,021	16,260	0,932	0,095	-0,035
V	AZ;10_12b	N-15	R	0,007	0,004	22,040	0,818	0,020	-0,006
	AZ;10_12b	N-15	S	0,005	0,006	11,708	1,000	0,024	-0,014
	AZ;10_12b	N-15	U1	-0,047	0,021	-26,950	0,399	0,019	-0,114
	AZ;10_12b	N-15	U2	0,002	0,006	4,642	1,000	0,019	-0,016
	AZ;10_12b	N-15	U4	0,021	0,014	19,060	0,902	0,065	-0,022
	AZ;10_12b	U2-15	C	-0,065	0,017	-48,801	0,007	-0,012	-0,118
	AZ;10_12b	U2-15	M	0,012	0,011	12,573	0,996	0,048	-0,024
	AZ;10_12b	U2-15	N	0,039	0,020	19,945	0,651	0,102	-0,025
V	AZ;10_12b	U2-15	R	0,003	0,004	11,609	1,000	0,015	-0,008
	AZ;10_12b	U2-15	S	-0,001	0,005	-1,764	1,000	0,015	-0,017
	AZ;10_12b	U2-15	U1	-0,036	0,021	-19,132	0,800	0,030	-0,102
	AZ;10_12b	U2-15	U2	-0,002	0,005	-4,486	1,000	0,014	-0,017
	AZ;10_12b	U2-15	U4	-0,010	0,010	-12,433	0,999	0,022	-0,042
	AZ;10_12b	U4-15	C	-0,054	0,018	-36,968	0,069	0,002	-0,109
	AZ;10_12b	U4-15	M	0,023	0,012	21,159	0,659	0,060	-0,015
	AZ;10_12b	U4-15	N	-0,027	0,016	-20,767	0,788	0,022	-0,075
V	AZ;10_12b	U4-15	R	0,009	0,004	27,840	0,358	0,022	-0,004
	AZ;10_12b	U4-15	S	0,006	0,006	12,683	0,998	0,023	-0,012
	AZ;10_12b	U4-15	U1	-0,103	0,017	-85,138	0,000	-0,050	-0,155
	AZ;10_12b	U4-15	U2	-0,007	0,004	-20,876	0,905	0,007	-0,020
	AZ;10_12b	U4-15	U4	0,007	0,011	7,438	1,000	0,043	-0,028

Host	MaP	id	segment	diff	SE	change	pval	upr.ci	lwr.ci
L	AZ:15	M-10	C	0,018	0,029	7,269	1,000	0,106	-0,070
	AZ:15	M-10	M	-0,012	0,006	-32,074	0,588	0,007	-0,031
	AZ:15	M-10	N	-0,027	0,037	-7,712	1,000	0,086	-0,139
	AZ:15	M-10	R	-0,003	0,003	-12,787	1,000	0,008	-0,013
	AZ:15	M-10	S	0,018	0,009	28,128	0,618	0,047	-0,011
	AZ:15	M-10	U1	0,031	0,010	43,257	0,039	0,061	0,001
	AZ:15	M-10	U2	0,020	0,011	25,198	0,686	0,054	-0,013
	AZ:15	M-10	U4	0,021	0,017	16,672	0,983	0,074	-0,032
L	AZ:15	R-10	C	-0,120	0,022	-109,864	0,000	-0,054	-0,187
	AZ:15	R-10	M	0,003	0,010	5,151	1,000	0,032	-0,027
	AZ:15	R-10	N	0,121	0,049	24,553	0,225	0,271	-0,029
	AZ:15	R-10	R	-0,011	0,003	-90,201	0,002	-0,003	-0,020
	AZ:15	R-10	S	0,009	0,010	15,557	1,000	0,038	-0,021
	AZ:15	R-10	U1	-0,024	0,004	-153,963	0,000	-0,012	-0,037
	AZ:15	R-10	U2	0,045	0,018	42,720	0,245	0,101	-0,011
	AZ:15	R-10	U4	-0,025	0,015	-31,259	0,821	0,021	-0,072
L	AZ:15	S-10	C	-0,079	0,024	-52,207	0,026	-0,005	-0,152
	AZ:15	S-10	M	0,011	0,010	17,546	0,994	0,040	-0,019
	AZ:15	S-10	N	0,012	0,040	3,156	1,000	0,135	-0,111
	AZ:15	S-10	R	0,009	0,005	27,941	0,770	0,026	-0,007
	AZ:15	S-10	S	-0,023	0,005	-99,642	0,000	-0,008	-0,038
	AZ:15	S-10	U1	0,003	0,007	7,804	1,000	0,026	-0,019
	AZ:15	S-10	U2	-0,017	0,008	-39,042	0,473	0,008	-0,041
	AZ:15	S-10	U4	0,106	0,027	50,238	0,003	0,188	0,024
L	AZ:15	U1-10	C	-0,222	0,014	-2909,853	0,000	-0,178	-0,265
	AZ:15	U1-10	M	-0,010	0,007	-26,132	0,922	0,011	-0,032
	AZ:15	U1-10	N	-0,036	0,039	-10,853	0,999	0,083	-0,156
	AZ:15	U1-10	R	-0,001	0,004	-2,135	1,000	0,012	-0,013
	AZ:15	U1-10	S	-0,032	0,004	-224,321	0,000	-0,019	-0,044
	AZ:15	U1-10	U1	0,425	0,041	91,331	0,000	0,551	0,298
	AZ:15	U1-10	U2	-0,040	0,005	-199,977	0,000	-0,024	-0,056
	AZ:15	U1-10	U4	-0,045	0,012	-73,979	0,005	-0,009	-0,081
L	AZ:15	C-10	C	0,380	0,035	62,368	0,000	0,489	0,272
	AZ:15	C-10	M	-0,030	0,005	-156,858	0,000	-0,017	-0,044
	AZ:15	C-10	N	-0,159	0,030	-74,982	0,000	-0,069	-0,250
	AZ:15	C-10	R	-0,015	0,002	-178,349	0,000	-0,009	-0,022
	AZ:15	C-10	S	-0,022	0,005	-92,511	0,000	-0,008	-0,036
	AZ:15	C-10	U1	-0,039	0,003	-2113,079	0,000	-0,029	-0,048
	AZ:15	C-10	U2	-0,022	0,007	-57,975	0,021	-0,002	-0,042
	AZ:15	C-10	U4	-0,028	0,012	-36,182	0,309	0,009	-0,065
L	AZ:15	N-10	C	0,069	0,045	23,013	0,898	0,207	-0,070
	AZ:15	N-10	M	0,019	0,014	28,100	0,941	0,061	-0,023
	AZ:15	N-10	N	-0,180	0,035	-94,200	0,000	-0,073	-0,288
	AZ:15	N-10	R	0,000	0,005	-1,154	1,000	0,015	-0,016
	AZ:15	N-10	S	0,044	0,017	48,906	0,150	0,095	-0,007
	AZ:15	N-10	U1	-0,004	0,008	-9,730	1,000	0,020	-0,027
	AZ:15	N-10	U2	0,024	0,016	28,877	0,890	0,073	-0,024
	AZ:15	N-10	U4	0,029	0,024	21,662	0,985	0,103	-0,045
L	AZ:15	U2-10	C	-0,023	0,023	-11,082	0,998	0,048	-0,093
	AZ:15	U2-10	M	0,054	0,012	52,105	0,000	0,089	0,019
	AZ:15	U2-10	N	-0,034	0,032	-10,008	0,996	0,064	-0,131
	AZ:15	U2-10	R	0,002	0,004	8,868	1,000	0,013	-0,008
	AZ:15	U2-10	S	0,014	0,008	23,300	0,665	0,037	-0,009
	AZ:15	U2-10	U1	-0,004	0,005	-10,002	1,000	0,012	-0,019
	AZ:15	U2-10	U2	-0,039	0,005	-185,044	0,000	-0,024	-0,054
	AZ:15	U2-10	U4	-0,054	0,009	-103,612	0,000	-0,026	-0,081
L	AZ:15	U4-10	C	-0,016	0,033	-7,566	1,000	0,086	-0,118
	AZ:15	U4-10	M	0,004	0,010	6,702	1,000	0,035	-0,028
	AZ:15	U4-10	N	-0,002	0,053	-0,636	1,000	0,161	-0,165
	AZ:15	U4-10	R	0,002	0,005	7,739	1,000	0,017	-0,013
	AZ:15	U4-10	S	0,011	0,010	18,908	0,997	0,042	-0,021
	AZ:15	U4-10	U1	0,011	0,009	22,108	0,984	0,040	-0,017
	AZ:15	U4-10	U2	0,010	0,013	14,412	1,000	0,049	-0,029
	AZ:15	U4-10	U4	-0,053	0,012	-99,428	0,001	-0,016	-0,089

Host	MaP	id	segment	diff	SE	change	pval	upr.ci	lwr.ci
V	AZ:15	M-10	C	-0,027	0,025	-17,656	0,995	0,049	-0,104
	AZ:15	M-10	M	0,002	0,009	4,742	1,000	0,029	-0,024
	AZ:15	M-10	N	0,020	0,044	5,269	1,000	0,155	-0,115
	AZ:15	M-10	R	0,002	0,005	5,448	1,000	0,016	-0,013
	AZ:15	M-10	S	0,002	0,006	7,260	1,000	0,019	-0,015
	AZ:15	M-10	U1	0,031	0,011	43,245	0,108	0,064	-0,003
	AZ:15	M-10	U2	0,017	0,012	23,967	0,924	0,053	-0,019
	AZ:15	M-10	U4	-0,115	0,015	-194,764	0,000	-0,070	-0,159
V	AZ:15	R-10	C	-0,109	0,017	-147,404	0,000	-0,058	-0,159
	AZ:15	R-10	M	-0,003	0,008	-6,629	1,000	0,022	-0,028
	AZ:15	R-10	N	0,190	0,046	34,261	0,001	0,330	0,050
	AZ:15	R-10	R	-0,009	0,004	-53,096	0,173	0,002	-0,020
	AZ:15	R-10	S	-0,007	0,004	-31,376	0,851	0,006	-0,020
	AZ:15	R-10	U1	-0,018	0,005	-78,030	0,009	-0,003	-0,033
	AZ:15	R-10	U2	0,027	0,015	33,137	0,694	0,072	-0,018
	AZ:15	R-10	U4	0,029	0,031	14,364	0,999	0,122	-0,064
V	AZ:15	S-10	C	0,071	0,031	28,148	0,310	0,165	-0,022
	AZ:15	S-10	M	0,027	0,012	35,501	0,322	0,062	-0,009
	AZ:15	S-10	N	-0,082	0,036	-29,134	0,341	0,028	-0,192
	AZ:15	S-10	R	0,004	0,005	13,002	1,000	0,019	-0,011
	AZ:15	S-10	S	-0,002	0,005	-5,398	1,000	0,013	-0,015
	AZ:15	S-10	U1	0,014	0,008	26,225	0,781	0,040	-0,011
	AZ:15	S-10	U2	-0,007	0,008	-15,529	0,999	0,017	-0,031
	AZ:15	S-10	U4	0,036	0,027	17,078	0,969	0,119	-0,047
V	AZ:15	U1-10	C	-0,173	0,013	-1880,332	0,000	-0,134	-0,212
	AZ:15	U1-10	M	-0,022	0,007	-81,331	0,022	-0,002	-0,042
	AZ:15	U1-10	N	-0,107	0,042	-41,380	0,191	0,021	-0,234
	AZ:15	U1-10	R	-0,003	0,005	-15,259	1,000	0,012	-0,018
	AZ:15	U1-10	S	-0,007	0,005	-32,037	0,921	0,008	-0,022
	AZ:15	U1-10	U1	0,478	0,050	92,224	0,000	0,630	0,326
	AZ:15	U1-10	U2	-0,036	0,006	-186,995	0,000	-0,018	-0,053
	AZ:15	U1-10	U4	-0,057	0,024	-49,442	0,267	0,016	-0,130
V	AZ:15	C-10	C	0,321	0,041	63,748	0,000	0,446	0,195
	AZ:15	C-10	M	-0,022	0,006	-82,059	0,003	-0,005	-0,039
	AZ:15	C-10	N	-0,119	0,035	-48,579	0,014	-0,014	-0,224
	AZ:15	C-10	R	-0,012	0,003	-81,905	0,004	-0,002	-0,021
	AZ:15	C-10	S	-0,017	0,003	-133,954	0,000	-0,007	-0,026
	AZ:15	C-10	U1	-0,035	0,003	-624,364	0,000	-0,024	-0,045
	AZ:15	C-10	U2	-0,019	0,007	-54,705	0,085	0,001	-0,040
	AZ:15	C-10	U4	-0,091	0,017	-111,100	0,000	-0,041	-0,142
V	AZ:15	N-10	C	0,020	0,032	9,894	1,000	0,117	-0,077
	AZ:15	N-10	M	0,001	0,010	1,778	1,000	0,030	-0,029
	AZ:15	N-10	N	-0,281	0,024	-340,581	0,000	-0,210	-0,353
	AZ:15	N-10	R	-0,009	0,004	-51,577	0,345	0,003	-0,021
	AZ:15	N-10	S	-0,006	0,005	-25,661	0,983	0,009	-0,021
	AZ:15	N-10	U1	0,014	0,010	25,182	0,951	0,043	-0,016
	AZ:15	N-10	U2	0,013	0,013	18,927	0,998	0,051	-0,026
	AZ:15	N-10	U4	-0,005	0,028	-2,801	1,000	0,081	-0,090
V	AZ:15	U2-10	C	0,022	0,024	10,830	0,999	0,094	-0,049
	AZ:15	U2-10	M	0,039	0,011	44,385	0,008	0,071	0,006
	AZ:15	U2-10	N	-0,085	0,032	-30,475	0,148	0,013	-0,183
	AZ:15	U2-10	R	0,007	0,005	20,087	0,933	0,020	-0,007
	AZ:15	U2-10	S	0,013	0,006	30,847	0,382	0,030	-0,005
	AZ:15	U2-10	U1	0,033	0,009	44,697	0,009	0,060	0,005
	AZ:15	U2-10	U2	-0,030	0,005	-119,139	0,000	-0,014	-0,046
	AZ:15	U2-10	U4	-0,062	0,017	-55,188	0,010	-0,009	-0,114
V	AZ:15	U4-10	C	0,012	0,033	6,294	1,000	0,113	-0,088
	AZ:15	U4-10	M	0,022	0,016	31,010	0,960	0,071	-0,027
	AZ:15	U4-10	N	-0,017	0,048	-4,814	1,000	0,130	-0,163
	AZ:15	U4-10	R	0,011	0,008	30,543	0,913	0,035	-0,012
	AZ:15	U4-10	S	0,038	0,013	56,929	0,051	0,077	0,000
	AZ:15	U4-10	U1	-0,001	0,008	-2,372	1,000	0,024	-0,026
	AZ:15	U4-10	U2	0,007	0,012	11,147	1,000	0,044	-0,030
	AZ:15	U4-10	U4	-0,130	0,014	-298,375	0,000	-0,086	-0,174

Supplementary Table 8 : [A] Type III test of fixed effects for relative frequency of segments model with genotype, host and segment as fixed effects and batch as random effect. Significative p.values are colored in green (<0.05). **[B] Pairwise comparisons of segment frequencies in reassortants in comparison to their major parent (MaP) in both vetches (V) and lentils (L).** "Change" column represents the percentage changes such as [diff/(response)*100] with response as the corresponding segment mean frequency. Significative p.values are colored in green (<0.05).



<u>Supplementary Figure 1</u> : [A] Visual summary illustrating the overall difference in genome formula for each reassortant relative to both its major parental and minor parental genotypes. The Y-axis represents the percentage change in relation to the major parental genotype, depicted in blue, and the minor parental genotype, shown in red. X-axis display the corresponding single-segment reassortant with the associated major parent. **[B] Visual summary representing the absolute differences in relative frequencies of the reassorted segments compared to the frequencies in the major parental genotype and minor parental genotypes in all host plant species (lentil L and vetch V).** The Y-axis indicates the absolute differences in the relative frequency of the reassorted segment compared to the segment relative frequency in the genome formula of the major parent shown in blue, and in the genome formula of the minor parent, depicted in red. X-axis display the corresponding single-segment reassortant with the associated major parent.

Batch	Host	Reass	Estimate	Std.Err	Z.Value	Pr(>|Z|)
1	V	(Intercept) AZ;15	-0,693	0,306	-2,264	0,024
1	V	AZ;10_12b	0,860	0,422	2,041	0,041
1	V	M-10	-0,297	0,446	-0,666	0,505
1	V	R-10	-0,405	0,453	-0,896	0,370
1	V	(Intercept) AZ;10_12b	0,167	0,290	0,577	0,564
1	V	AZ;15	-0,860	0,422	-2,041	0,041
1	V	M-15	-0,084	0,409	-0,205	0,838
1	V	R-15	0,434	0,418	1,037	0,300
1	V	U1-15	-1,633	0,470	-3,477	0,001
1	L	(Intercept) AZ;15	-1,768	0,409	-4,322	<0,001
1	L	AZ;10_12b	0,880	0,518	1,700	0,089
1	L	M-10	0,000	0,578	0,000	1,000
1	L	R-10	-0,630	0,663	-0,950	0,342
1	L	(Intercept) AZ;10_12b	-0,887	0,318	-2,794	0,005
1	L	AZ;15	-0,880	0,518	-1,700	0,089
1	L	M-15	-0,211	0,460	-0,459	0,646
1	L	R-15	-0,326	0,468	-0,696	0,486
1	L	U1-15	-1,264	0,569	-2,221	0,026
2	V	(Intercept) AZ;15	-0,423	0,295	-1,433	0,152
2	V	AZ;10_12b	1,211	0,429	2,823	0,005
2	V	S-10	0,086	0,416	0,208	0,835
2	V	U1-10	-0,088	0,420	-0,210	0,834
2	V	(Intercept) AZ;10_12b	0,788	0,311	2,532	0,011
2	V	AZ;15	-1,211	0,429	-2,823	0,005
2	V	S-15	-0,537	0,426	-1,260	0,208
2	L	(Intercept) AZ;15	-1,466	0,370	-3,965	<0,001
2	L	AZ;10_12b	1,043	0,473	2,205	0,027
2	L	S-10	1,043	0,473	2,205	0,027
2	L	U1-10	0,253	0,505	0,502	0,616
2	L	(Intercept) AZ;10_12b	-0,423	0,295	-1,433	0,152
2	L	AZ;15	-1,043	0,473	-2,205	0,027
2	L	S-15	-0,676	0,445	-1,518	0,129

Batch	Host	Reass	Estimate	Std.Err	Z.Value	Pr(>|Z|)
3	V	(Intercept) AZ;15	-0,336	0,293	-1,149	0,250
3	V	C-10	0,253	0,411	0,615	0,538
3	V	N-10	0,673	0,414	1,625	0,104
3	V	AZ;10_12b	1,549	0,451	3,434	0,001
3	V	(Intercept) AZ;10_12b	1,213	0,343	3,532	<0,001
3	V	AZ;15	-1,549	0,451	-3,434	0,001
3	V	C-15	-1,046	0,449	-2,328	0,020
3	V	N-15	-0,326	0,468	-0,696	0,486
3	L	(Intercept) AZ;15	-3,135	0,722	-4,342	<0,001
3	L	AZ;10_12b	2,560	0,834	3,071	0,002
3	L	C-10	0.738	0.891	0.828	0.408
3	L	N-10	2.884	0.779	3.704	<0,001
3	L	(Intercept) AZ;10_12b	-0,575	0,417	-1,381	0,167
3	L	AZ;15	-2,560	0,834	-3,071	0,002
3	L	C-15	0,313	0,592	0,529	0,597
3	L	N-15	-0,461	0,545	-0,846	0,398
4	V	(Intercept) AZ;15	-1,099	0,333	-3,296	0,001
4	V	U2-10	0,000	0,471	0,000	1,000
4	V	U4-10	0,847	0,442	1,915	0,055
4	V	AZ;10_12b	1,350	0,442	3,051	0,002
4	V	(Intercept) AZ;10_12b	0,251	0,291	0,864	0,388
4	V	AZ;15	-1,350	0,442	-3,051	0,002
4	V	U2-15	0,847	0,442	1,915	0,055
4	V	U4-15	0,000	0,411	0,000	1,000
4	L	(Intercept) AZ;15	-0,251	0,291	-0,864	0,388
4	L	AZ;10_12b	0,335	0,410	0,816	0,414
4	L	U2-10	-1,215	0,471	-2,582	0,010
4	L	U4-10	-1,358	0,484	-2,804	0,005
4	L	(Intercept) AZ;10_12b	0,083	0,289	0,289	0,773
4	L	AZ;15	-0,335	0,410	-0,816	0,414
4	L	U2-15	-0,420	0,411	-1,021	0,307
4	L	U4-15	-0,083	0,408	-0,204	0,838

<u>Supplementary Table 9</u> : Coefficients output of generalized linear models per batch and hosts with infectivity status of recipient plants as the binary response variable and the genotype as fixed effect. Significative p.values are colored in green (<0.05).

GENERAL DISCUSSION

While this doctoral research primarily focused on investigating nanovirus reassortments, its broader aim was to make significant contributions to unraveling the mysteries surrounding multipartite genomic architecture and advancing our understanding of nanovirus biology, an emerging worldwide group of phytopathogens. The overarching findings, their implications and perspectives are exposed in the following sections.

1. Genome reconstitution through non-concomitant transmission of genomic segments

In the **Introduction**, we explored the hypothesis that the multipartite genomic segments of FBNSV could be transmitted separately leading to the reconstitution of complete genomes when they eventually all come together again (**Di Mattia et al. 2022**). This reconstitution process was investigated through two distinct scenarios (i) sequential horizontal transmission where aphids acquired two incomplete sets of segments sequentially from two different plants. In this case, genome reconstitution occurred directly within the aphid vector. And (ii) parallel horizontal transmission where two aphids, each carrying an incomplete set of segments acquired on recipient plants, inoculated them on the same new host plant. Genome reconstitution occurred within the recipient plant in both scenarios. Subsequently, the reconstituted genomes were capable of full transmission, functioning as an integral complete genome (**Di Mattia et al. 2022**).

The occurrence of genome reconstitutions in both scenarios implies that these reconstitutions can occur between genotypes that do not simultaneously infect the same host and/or that are not transmitted by the same vector. This phenomenon significantly reduces the inter-host cost associated with the maintenance of genomic integrity, as discussed in the introduction of this manuscript, that was hypothetically attributed to an assumed mandatory concomitant transmission of all segments and thus on a high MOI to establish an infection in the newly infected host if the viral particles are randomly disseminated with no mechanism to sort or specifically group them.

Furthermore, since reconstitution relies on the complementation between genomic segments, it is conceivable that the complementation of incomplete sets of segments from different genotypes would lead to the production of reconstituted hybrid genomes through reassortments. This characteristic holds significant implications, as it greatly expands the spatial scale at which multipartite viruses, particularly nanoviruses, can undergo reassortment

when compared to segmented viruses, which rely on co-infections of individual host cells due to the co-packaging of genomic segments. In the appropriate ecological context, the viability of reassortants becomes key to the success of hybrid genome reconstitution and thus to the cost associated of maintaining genomic integrity.

2. Reassortments in ssDNA multipartite viruses

To explore and substantiate this assertion, the review (**Torralba, Blanc and Michalakis 2023**), developed in **Chapter 1**, focused on identifying the molecular constraints that impact the reassorted segments following reassortments in single-stranded DNA multipartite viruses. These segments must undergo processes such as replication, encapsidation, viral movement, and transmission that are partially performed by their genomic background. As such, they need to be complemented by the specific genomic background with which they are associated. The review explores a combination of experimental reports and data from natural populations to determine if the predicted molecular/cellular constraints align with the actual frequency and nature of reassortants found in natural populations. Given the limited data for other viral genera, the review primarily deals with the *Begomovirus* genus and the *Nanoviridae* family.

The primary molecular constraint that emerged was related to the trans-replication of the reassorted segments. Since the majority of known ssDNA multipartite viruses are CRESS DNAs, which mostly rely on rolling-circle replication based on iteron recognition (**Torralba, Blanc and Michalakis 2023 – Chapter 1**), the review pointed out limitations regarding iteron identification and recognition process with the Rep proteins, despite its critical role in the viability of reassortments. Mostly, the compatibilities between iterons and Rep sequences are not fully understood and a comprehensive listing of known associations between iterons and Rep sequences should be beneficial. Interestingly, although involving different genera, aspects such as packaging, viral movement, and horizontal transmission did not appear to inherently impose important severe constraints on a newly reassorted segment as they seemed to be less dependent on direct sequences recognition.

Additionally, it was notably evident that reassortments played a substantial role in the evolution of these viruses. Indeed, despite the likelihood of underestimating the detection of reassortments, they were still associated with a considerable number of surveyed genomes. This could suggest that reassortments are either a common occurrence or that they are often

selected. These findings are consistent with the concept of facilitated reassortment in single-stranded DNA multipartite viruses, at least nanoviruses.

When comparing experimental data with observations from natural populations, a perplexing inconsistency emerged regarding nanovirid reassortants. Population surveys revealed a strikingly lower frequency of inter-specific reassortants when compared to inter-specific recombinants. Although the data from natural population surveys of nanovirids is limited, this observation most likely reflects a biological reality. Considering previous findings that reassortment can occur without the need for co-infections of the same host and cell, unlike recombination which depends on cell co-infection, the opportunity for inter-specific reassortment should theoretically be more prevalent than inter-specific recombinations. This inconsistency becomes even more intriguing when considering experimental data which suggests relatively few constraints on inter-specific reassortments, as multiple species are capable of complementing essential steps of the infection cycle such as replication (**Timchenko et al. 2000**), encapsidation (**Timchenko et al. 2006**), and transmission (**Grigoras et al. 2018**). it is likely that the discrepancy between the observation of interspecific complementation and the relative rarity of interspecific reassortments results from more subtle deleterious effects accompanying complementation which have not yet been characterized. This possibility calls for further investigation and comparison of several proxies of the fitness of intra-specific and inter-specific reassortants versus recombinants, as further discussed in future prospects.

Another intriguing inconsistency lies in the prevalence of single-segment reassortment events with very few events involving multiple segments detected despite repeated experimental observations of missing genomic segments such as DNA-C, -N, -U1, -U2, and -U4 which *a priori* seem compatible with the complementation of multiple segments at once. This observation likely suggests severe limitations to multiple-segment reassortment that probably lies in severe intra-genomic disruptions caused by additional reassorted segments that might also limit genome reconstitution through reassortment. Here again a quantitative analysis of fitness proxies of multiple segment reassortants would be highly informative.

3. Production and phenotypic characterization of two distinct isolates of the FBNSV

To initiate the examination of intra-specific single-segment reassortant fitness in FBNSV, we needed infectious clones for at least two distinct isolates. At the beginning of my PhD work, only one infectious clone **[JKI-2000]** (**Grigoras et al. 2009**) was available. Consequently, we opted to choose two new isolates for the development of bitmer infectious clones (**Urbino et al. 2008; Yepes et al. 2018**). Our preference was to select isolates that represented different phylogenetic clades within the species, were sampled from different hosts, and coexisted geographically.

Using the available genomic data, we delved into the phylogenetic structure of nanoviruses, specifically focusing on FBNSV. This analysis confirmed the organization in three distinct clades, genetically equi-distant from each other. We employed additional metadata (**Grigoras et al. 2014**) to inform our selection of discriminable isolates from distinct clades, sampled on different field hosts for further study of host adaptation and co-circulating in the same geographical area, thus making natural encounters plausible. Ultimately, we chose isolates **[AZ;15]** and **[AZ;10_12b]** as they shared close geographical proximity in Azerbaijan. Furthermore, **[AZ;15]** represented the "Red" clade and was collected from vetches, while **[AZ;10_12b]** belonged to the "Green" clade and was isolated from lentils.

We successfully generated two additional infectious clones of FBNSV, namely **[AZ;15]** and **[AZ;10_12b]** (**Torralba et al. 2024a – Chapter 2**). These infectious clones, in conjunction with the previously existing **[JKI-2000]**, together represent the three distinct phylogenetic clades observed within FBNSV species. These genetic resources open up avenues for further research, including investigations into mixed infections, genetic exchange phenomena (as elaborated upon in future prospects), and the exploration of phenotypic diversity within the species.

All three isolates demonstrated their ability to infect vetches, lentils, and faba beans, and they were all transmitted by the same aphid vector, *A. craccivora*. This is ecologically and evolutionary significant, particularly given the coexistence of **[AZ;10_12b]** and **[AZ;15]** within the same geographic area. Previously, different nanovirus species have been found to infect the same host plant species (including peas, faba beans, vetches, lentils, and beans) and be successfully transmitted by the same aphid vector species (*A. craccivora* and *A. pisum*). This suggest that nanovirus mixed-infections might be common with consequences on their genetic exchange dynamics and the emergence of hybrid genotypes.

The three isolates were thoroughly characterized in the three distinct hosts, revealing varying phenotypic outcomes. Interestingly, no discernible correlations were found between the severity of symptoms and our indicators of viral fitness (viral load and transmission rate). **[AZ;10_12b]** consistently demonstrated the highest transmission rate, while **[JKI-2000]** consistently exhibited the lowest transmission rates, as well as lower viral loads in both plants and aphid vectors. An important finding from this study was the divergence in genome formula among FBNSV isolates when tested in various plant hosts and aphid vectors. This research also confirmed that the loss of genomic segments was observed following both agroinoculation and aphid inoculation. Furthermore, this phenomenon was observed across all isolates and host plant species in our study. These observations suggest that that the loss of genomic segments following aphid inoculation can probably happen in natural conditions although this phenomenon is probably selected against or very uncommon due to aphid vector ecology in natural conditions.

4. Systematic phenotypic characterization of single-segment reassortants between two FBNSV isolates

To begin unraveling the fitness implications of reassortments, we chose to systematically characterize intra-specific single-segment reassortants between **[AZ;15]** and **[AZ;10_12b]** on both vetches and lentils (**Torralba et al. 2024b – Chapter 3**). The objective of this approach was to gain insights on the frequency of fitness alterations associated with reassortments – how frequently they result in reduced, enhanced, or unaltered fitness – and which specific segments are linked to these effects. Additionally, we examined how these changes in fitness played out in different host plant species.

One of the key findings of this study is that all generated reassortants were capable of establishing infections in both host plants and subsequently exhibited efficient transmission. This observation aligns with existing data indicating that the majority of viable natural reassortment events are intra-specific single-segment reassortments (**Torralba, Blanc and Michalakis – Chapter 1**). However, it also raises question regarding the absence or low occurrence of natural nanovirus reassortants linked to certain genomic segments like DNA-C. In fact, the DNA-C reassortant derived from the primary parent **[AZ;15]** exhibited no apparent reduction in fitness in both hosts, while the counterpart reassortant originating from the **[AZ;10_12b]** major parent was linked to decreased fitness in both hosts. These findings suggest, however, that this inconsistency likely results from limited data and

research on nanovirus genomics. Our study will serve as a crucial foundation for delving deeper into the dynamics of natural reassortments, a topic we will explore in the future prospects section.

With the evaluation of several biological traits of each reassortant, we observed important variations in the genome formula following reassortment. In fact, this was the biological trait that appeared to be the least robust, with 26 out of 32 reassortants with significant alterations of the major parent genome formula. Interestingly, some reassortants displayed important changes in their genome formula without any other variations of evaluated phenotype. This observation might suggest a role of the genome formula in adapting to intra-genomic disruptions following genetical exchanges. However, our study did not include any analysis of the transcription profiles nor the protein expression that would enable us to link the genome formula changes to other phenotypical variations. Conducting such investigations and comparing the reassortant transcription and expression profiles with those of the major parents in future research will be essential for gaining a comprehensive understanding of the factors influencing their associated fitness and phenotypes. It will also allow a comparison between the genome formula changes in reassortants and gene expression, to evaluate a possible function of the formula variation in adjusting gene copy numbers to regulate gene expression (**Gallet et al. 2022**).

Analyzing the reassortant phenotypes highlighted several of them, and their associated genomic segments, which exhibited consistent effects on both genotypes or host species. Four specific segments, DNA-N, -U1, -U2, and -U4, appeared to have potential new roles in transmission, viral accumulation, or as pathogenicity determinants. This is especially intriguing since most of these segments have unknown functions, and DNA-N only known function is related to transmission.

Assessing the fitness of reassortants, using viral load and transmission rate as indicators, revealed that 50% of them were not significantly affected. Variations in fitness between reassortants associated with different major parents underscored the differential effects of reassortments on parental genotypes, probably directly linked to overall fitness imbalances between them. However, the generation of viable single-segment reassortants encompassing all possible combinations between both isolates, which frequently exhibited minimal changes in viral fitness, implies that reassortments might alleviate the burden of maintaining genomic integrity through genome reconstitutions in compatible ecological conditions that support the co-circulation of viral diversity in shared hosts and vectors. Moreover, this process enables the generation of reassortants without competition from parental genotypes or other reassortants.

This appears particularly relevant, considering that nanoviruses seem to infect a broad range of hosts within the *Fabaceae* plant family, as also demonstrated in this study.

4. Conclusion

In summary, my doctoral research has significantly advanced our understanding of the profound implications of reassortments in the evolution, genome integrity maintenance, and biology of single-stranded DNA multipartite viruses, particularly nanoviruses. It has provided valuable insights into the multipartite genomic architecture with the demonstration of non-concomitant genome reconstitutions that reduces inter-host maintenance cost and expands the scale of potential reassortments.

Reviewing the current state of the art on single-stranded DNA multipartite reassortment and the associated constraints, trans-replication emerged as the main constraint for reassorted segment compatibility. This investigation also shed light on the significant prevalence of reassortment in begomovirus and nanovirid natural populations while revealing intriguing disparities concerning nanovirus inter-specific and multi-segment reassortments.

Furthermore, we developed two new infectious clones of FBNSV, namely **[AZ;15]** and **[AZ;10_12b]**, enabling us to partially explore phenotypic diversity. Specifically, we assessed the reaction norms of several FBNSV genotypes and provided a glimpse into the phenotypic diversity within the species. In addition, our research deepened our understanding of the viability of intra-specific single-segment reassortments in FBNSV, elucidating their phenotypic effects and associated fitness. We also observed indications of the genome formula potentially playing a role in maintaining a form of homeostasis following reassortments.

In a broader context, this study represents the first systematic examination of the phenotypic consequences of single-segment reassortment in a virus with a segmented genome. It also suggests that reassortments in nanoviruses probably participate in mitigating the cost of preserving genomic integrity under favorable ecological conditions.

Looking ahead, the development of infectious clones for two new FBNSV isolates opens the door to further research avenues, some of which are outlined as prospective directions in the following section.

FUTURE PROSPECTS

1. Genome reconstitution and rescue hypothesis

Demonstrating genome reconstitutions between complementary incomplete sets of genomic segments is undeniably significant, as mentioned earlier. Nevertheless, there are limitations to our study, primarily that these sets of segments, while complementary, can independently initiate an infection. It was feasible because certain nanovirus genomic segments are non-essential for establishing an infection namely DNA-C, -N, -U1, -U2, and -U4.

As a secondary and incomplete aspect of this part of the PhD research, the ectopic expression of an essential nanovirus protein, such as M-Rep, CP, or MP, using a viral vector able to co-infect a common nanovirus host has been attempted by another PhD student of the team (Mélia BONNAMY) and myself. The intention behind this effort was to produce incomplete sets of segments incapable of establishing an infection on their own. This would be facilitated by the complementation provided by the protein produced via the viral vector, for example the M-Rep or the coat protein, of an inoculation of an FBNSV infectious clone lacking either DNA-R or -S, respectively. Similar to the prior research on non-concomitant transmission of FBNSV segments (**Di Mattia et al. 2022**), this approach would extend genome reconstitution to sets of segments that are unable to initiate an infection on their own but can fully reconstitute a complete genome if functional complementation is accessible within the corresponding host and/or aphid vector. In this context, preliminary work on two viral vectors has been initiated (see annex).

If indeed reconstitutions can occur from incomplete sets of genomic segments incapable of initiating an infection on their own, then another question emerges regarding the potential time-window for the reconstitutions. Is it possible for sets of segments that cannot establish an infection to remain "latent" in plant hosts or vectors before being complemented by a second set of segments and if yes for how long (**Blanc and Michalakis 2020**)?

The question for a latent phase can be even more general with a segment possibly being rescued by a secondary infection (incomplete or not). In this scenario, each genomic segment would function as an individual entity, dispersing in the ecosystem until conditions for complementation are met, somewhere. It would also further expand the spatial scale of possible reassortments. To test this hypothesis, an experiment was devised using two distinct isolates to distinguish between the latent segment(s) and the rescuing infection. The idea to introduce latent segments in a plant relies on the assumption that the plants that do not display symptoms following aphid inoculation could not establish an infection due to the absence of essential segments. Thus, these plants may contain some segments that have been

inoculated by aphids but not a set of segments that could initiate a systemic infection. This possibility is supported by our aphid analysis in (**Torralba et al. 2024a – Chapter 2**), which indicates that aphid transmission could generate plants lacking segments. Thus, among the uninfected plants, many lack at least one essential segment while containing other segments in inoculated plant tissues. The use of a different genotype genetically distinguishable from the first isolate for a second inoculation of these plants may generate a recognizable infection that could rescue the latent segments from the first inoculation. If successful this experiment may generate infections with two alleles for some of the segments, or reassortants between both genotypes. In this context, the findings from (**Torralba et al. 2024b – Chapter 3**) are promising, as all single-segment reassortants between **[AZ;15]** and **[AZ;10_12b]** were found to be viable in faba beans, lentils, and vetches and that the majority of viable reassortants in natural populations are intra-specific single-segment reassortants (**Torralba, Blanc and Michalakis – Chapter 1**). However, this will probably reflect limitations in the possible complementations and primo-infection rescue.

The experimental design of this 'rescue' experiment, which relies on the production of reassortants and/or mixed infections at least for some of the segments, may impose significant limitations despite the fact that we have earlier successfully generated mixed infections of all 16 segments of **[AZ;15]** and **[AZ;10_12b]** (see annex).

An alternative approach to investigate the latent segments and rescue hypothesis using only one isolate could involve leveraging the very low infection rate observed for sets of segments lacking DNA-C (**Di Mattia et al. 2022**) during the agroinoculation step or using the extremely low aphid transmission rate of sets of segments lacking DNA-U1 (**Grigoras et al. 2018**) as the primo-inoculation. These segments, C or U1, while non-essential for the establishment of infections (**Grigoras et al. 2018**), and not affecting virion production, likely play a critical role, as they are rarely lost under experimental conditions (**Torralba et al. 2024a – Chapter 2**) and may act as surrogate essential segments. A comparison of the total infected plants with both primo-inoculation and second inoculation controls, where each control represents separate parallel inoculations conducted under the same conditions, would indirectly demonstrate the occurrence of delayed complementation if we observe statistical differences of the presence/absence of genomic segments with the controls (see annex).

2. Expanding genomic analysis and experimental validation

As emphasized in the review (**Torralba, Blanc and Michalakis 2023**), our understanding of nanovirus iteron and of trans-replication compatibilities when a segment is introduced in a new genomic background is limited. Therefore, preliminary results from a nanovirus genomic study will hopefully lay the foundation for *in silico* identification of candidate nanovirus iteron sequences through sequence alignments and further experimental validation. This would enable a listing of nanovirus iterons for every isolate from nanovirus species.

The detection of reassortments is constrained by the dataset sequence counts and their pairwise similarity. By centralizing nanovirus genomic data, we aim to update the known nanovirus diversity, identify new natural reassortants, and gain further insights into compatibilities of parental genotypes previously identified with reassortants. This knowledge might potentially be used to retrace nanovirus evolution history and understand their natural population dynamics. This would also enable us to investigate the observed bias between intra- and inter-specific reassortants or even the absence of reassorting DNA-C in nanovirus natural populations.

Expanding and advancing the analysis of nanovirus genomics is a project of significant importance to me because I firmly believe that easy access to genomic sequences and segment phylogenetics can greatly enhance our understanding of nanovirus biology and evolution, serving as a valuable source of information on genetic and genomic patterns. With the rapid discovery of new isolates, species, and hosts, updating the natural population survey to detect reassortments in nanoviruses will enable us to identify new reassortment events with the appropriate methodologies. To achieve this, I intend to use the RDP5 software (**Martin et al. 2020**) and other methods on the updated nanovirus genomic data for reassortment detection.

Additionally, employing motif recognition software such as MEME Suite (**Bailey et al. 2015**) on this dataset will offer additional validation for *in silico* iteron sequence candidates, which could be subject to functional validation via direct mutagenesis, similar to the approach used for babuviruses (**Herrera-Valencia et al. 2006**). This dual approach of exploring both the diversity of iterons within the nanovirus genus and their compatibility with M-Rep sequences in documented reassortment events will provide us with enhanced insights into the ecological dynamics of nanoviruses and potential emerging associations. Furthermore, it may assist us in identifying the putative nanovirus iteron-recognition domain on M-Rep, thereby enhancing our understanding of iteron-M-Rep interactions.

3. Investigating the genome formula

Due to limited time, I was unable to conduct an in-depth statistical analysis of the extensive datasets generated in both the (**Torralba et al. 2024a – Chapter 2**) and (**Torralba et al. 2024b – Chapter 3**) studies. However, the individual phenotyping of a large number of plants for distinct isolates and reassortants that is now available most likely contains precious information that will help answer the intriguing question of whether the genome formula is an adaptive trait, which could be a significant advantage of the multipartite architecture (**Gutiérrez and Zwart 2018; Zwart and Elena 2020; Gallet et al. 2022**). If the set point genome formula is indeed established through an equilibrium that represents the virus's best-adapted state, this might be observable through statistical analyses of correlations between our evaluated traits.

4. Exploring reassortment discrepancies, inter-specificity and multiple-segment events

When reviewing the natural population survey of nanoviruses, two intriguing discrepancies emerged: (i) the prevalence of single-segment reassortment events and (ii) the rarity of inter-specific reassortment events compared to inter-specific recombinations. To address these puzzling observations, it would be valuable to characterize multiple-segment reassortments, beginning with the direct reassortment of two segments. However, due to logistical constraints, as this would involve 56 reciprocal combinations using two isolates, it is way more practical to select combinations of segments associated with single-segment reassortants that either showed no fitness effect compared to the major parent or exhibit improved fitness like N-10 which appeared to have an increased fitness in both field hosts, S-10 or R-10. This approach could provide insights into the limitations affecting multiple-segment reassortants and shed light on the interactions between genomic segments.

It appears evident to me that the next significant step in investigating nanovirus reassortments involves generating single-segment inter-specific reassortants. This experiment is feasible since there are available infectious clones for three other species besides FBNSV: FBNYV, BMLRV, and PNYDV (**Grigoras et al. 2014**). Furthermore, FBNYV and PNYDV are routinely manipulated (**Krenz et al. 2017; Mansourpour et al. 2022**) and share host plant species and aphid vectors with FBNSV (**Grigoras et al. 2014**). Based on phylogenetics (**Figure 6**), FBNYV is genetically much closer to FBNSV than PNYDV, making it more likely for them to complement

each other. Therefore, a priority should be to characterize single-segment reassortants in only one host to reduce logistical challenges between FBNSV and FBNYV and assess the viability of corresponding reassortants. If a significant proportion of them proves viable, it might be interesting to compare them with FBNSV and PNYD to further assess how genetic diversity affects reassortment viability.

5. Adaptation and competition of reassortants

The results of (**Torralba et al. 2024b – Chapter 3**) have identified multiple segments associated with a better fit reassortant than the major parent. However, the context in which we produced these reassortants did not consider the competition between the heterologous and homologous segments from the minor and major parent respectively. Therefore, conducting competition experiments will be interesting to determine whether these reassortants are reproduced in a competitive context and if they take over. This approach may provide insights into processes that prioritize the integrity of homologous genomes, potentially explaining why inter-specific reassortments are less common in natural environments.

Exploring how reassortant genomes adapt to their new genomic context and the rate at which this adaptation occurs is a captivating area of study. In our reassortant characterization (**Torralba et al. 2024b – Chapter 3**), we identified several reassortants that demonstrated reduced fitness compared to their major parent in both host environments. To delve deeper into these inquiries, one potential avenue for exploration is to conduct experimental evolution experiments under controlled conditions.

6. Controlled field experiments for ecological data on nanoviruses

Ultimately, what we crucially lack is experimental field data, and this need has been underscored by the significant disparities observed between laboratory-derived data and observations in natural populations (**Torralba, Blanc and Michalakis 2023 – Chapter 1**). These disparities include the loss of certain segments, the overrepresentation of inter-specific and single-segment reassortants, among others. As plant viruses, nanoviruses rely on their insect vectors to spread in the environment. Gaining insight into the ecological dynamics of various factors such as host plant species, insect vectors, viral species, isolates, reassortant dissemination and interactions will make a substantial contribution to our understanding of

the underlying evolutionary dynamics and nanovirus ecological properties. This is especially relevant if multipartite viruses like nanoviruses have indeed adapted well to changing environments, using mechanisms such as the genome formula, within-host supra-cellular complementation, and inter-host non-concomitant genome reconstitutions.

ANNEX

1. Infectious clone characterization

1-1. Infectious clone validation and agro-inoculation protocol optimization

We designed the **[AZ;15]** and **[AZ;10_12b]** infectious clones as bitmers (**Urbino et al. 2008; Yepes et al. 2018**), where each segment is constructed as a partial tandem repeat encompassing the entire segment flanked by two conserved regions CR-SL. This design ensures that the conserved region CR-SL, associated with the origin of replication, is present at both ends of the whole segment, allowing for its liberation and circularization by M-Rep or homologous recombination.

We outsourced the production of all clones to Agate Bioservices company (Mialet, France) who synthesized the insert sequences and cloned them in a binary plasmid pCAMBIA2300. Two sets of eight plasmids, each containing a genomic segment bit-mer, respectively constitute the FBNSV **[AZ;15]** and **[AZ;10_12b]** infectious clones. All the plasmids were delivered and individually transformed by electroporation in distinct clones of *Agrobacterium tumefaciens* STR COR308.

The conventional agroinoculation technique, which involves making punctures in the main stems of faba beans, a commonly used experimental host within our research team, was employed with both infectious clones. We observed the appearance of symptomatic plants for both isolates. Preliminary results showed an agroinoculation success rate of 30 % for **[AZ;10_12b]** and 5% for **[AZ;15]**.

In an effort to increase the agroinoculation success rate, we transformed the pCAMBIA-segment clones of both isolates into multiple strains of *A. tumefaciens*. Specifically, we selected *A. tumefaciens* STR COR308, C58C1, and LBA4404 and subsequently compared their performance by conducting agroinoculation experiments on faba beans (**Table 2 - A**). The results revealed that there was no significant improvement in the infection rate when using the C58C1 and LBA4404 strains. Consequently, we made the decision to continue working with the transformed *A. tumefaciens* COR308 for further manipulations of the new infectious clones **[AZ;15]** and **[AZ;10_12b]**.

A

ISOLATE	AGRO STRAIN	INOC MET	INOC PLANTS	INF PLANTS	INF RATE
AZ;15	COR308	Piercing	73	3	4,11 %
AZ;15	C58C1	Piercing	105	2	1,90 %
AZ;15	LBA4404	Piercing	91	0	0,00 %
AZ;10_12b	COR308	Piercing	66	10	15,15 %
AZ;10_12b	C58C1	Piercing	65	11	16,92 %
AZ;10_12b	LBA4404	Piercing	65	2	3,08 %

B

ISOLATE	AGRO STRAIN	INOC MET	INOC PLANTS	INF PLANTS	INF RATE
AZ;15	COR308	Infiltration	37	11	29,73 %
AZ;10_12b	COR308	Infiltration	43	22	51,16 %

Table 2 : [A] Agroinoculation table of infectious clones [AZ;15] and [AZ;10/12b] evaluating multiple *A. tumefaciens strains*. Multiple agrobacterium strains and inoculation methods were compared to evaluate the viability of the two infectious clones and optimize their infection rates. "Inf plants" column indicates the number of infected plants, detected through qPCR, with the corresponding isolate. **[B] Agroinfiltration table of infectious clones [AZ;15] and [AZ;10_12b] using *A. tumefaciens* STR COR308 clones.**

ISOLATE	SEGMENT	SEQUENCE	SENSE	LENGTH	[ ] (µM)
AZ;10_12b	C - For	ACCCATCAACTGAAGATTTG	F	20	1
AZ;10_12b	C - Rev	ACCATACTTAACTTGGAAGGG	R	21	
AZ;10_12b	M - For	AGGAACGATGTCGGATT	F	17	0,8
AZ;10_12b	M - Rev	ACCTATCATGTACAGAGCCT	R	20	
AZ;10_12b	N - For	AGTGTTAGCTTCAACGGGA	F	19	0,5
AZ;10_12b	N - Rev	ACGAACGTTTCTGTTTCCTC	R	20	
AZ;10_12b	R - For	GGAGTATGGAGAGTTCATTGTT	F	22	0,8
AZ;10_12b	R - Rev	GAACTCTCTTAATGTGCCTGAA	R	22	
AZ;10_12b	S - For	GTTTACAGCACCGGCTT	F	17	0,8
AZ;10_12b	S - Rev	ATACAGCCTTAACAGCCTCA	R	20	
AZ;10_12b	U1 - For	AGTTAGATGCCGTTAGTCG	F	19	0,8
AZ;10_12b	U1 - Rev	CAGACAACAGATCTGCGT	R	18	
AZ;10_12b	U2 - For	CTGGTTTTAATGGCTACCG	F	19	0,5
AZ;10_12b	U2 - Rev	GGTGCGTAACTGGGTAAC	R	18	
AZ;10_12b	U4 - For	TTCGAGGATGTTGATGCAG	F	19	0,5
AZ;10_12b	U4 - Rev	TCGATTCCCATAGGCGTT	R	18	
AZ;15	C - For	CAAGATGCTTAGTCAAGTGGAA	F	22	1,5
AZ;15	C - Rev	CAGTGAATAACTTATCTAATTCACTGC	R	27	
AZ;15	M - For	ATCAAGACGACGGAGATG	F	18	1
AZ;15	M - Rev	CTTCTTGACAAATCCAGGTAG	R	21	
AZ;15	N - For	GTTAAATGATTCAAGGAAGGTGTTAC	F	26	1,5
AZ;15	N - Rev	GACCACGAACATTTCTATTTCCG	R	23	
AZ;15	R - For	ATGGTGAGTTCGTTACGTC	F	19	1
AZ;15	R - Rev	CCTCTAAATTCTCTCAACGTACTA	R	24	
AZ;15	S - For	ACGACGATATTCAAGGCCA	F	19	1
AZ;15	S - Rev	CCTGCAACACCATTAGATTCC	R	21	
AZ;15	U1 - For	CTCTGGTTAGTCGAGCTTC	F	19	1
AZ;15	U1 - Rev	CTGAGAATCATCTTCACAACATGG	R	21	
AZ;15	U2 - For	CTTCATTTCATCTTGTTCTATTCGC	F	25	1
AZ;15	U2 - Rev	CCAAACGAGTAGGAGCATATG	R	21	
AZ;15	U4 - For	GTTGTTAAATCCTTCTTTGGTGG	F	23	1
AZ;15	U4 - Rev	CGTCTTCTTTCTTCTTACCCATG	R	23	

Table 3 : Table overview of the pairs of primers used for the specific amplification of the segments of [AZ;15] and [AZ;10_12b] isolates. The final column [ ] indicates the recommended primer concentration for use in the qPCR mix. "Sense" column indicates if the primer sequence concerns the forward or reverse sense for amplification.

Concurrently, our research team devised a novel agroinfiltration protocol that demonstrated a noteworthy enhancement in the infection rate for **[JKI-2000]**. Instead of using a needle to puncture the stems of 10-day-old seedlings, we infiltrate with an open syringe our bacteria mixture into the first leaflets of 9-day-old seedlings, at a very early stage of their growth before their first leaves fully expanded. This new agroinoculation method substantially increased the infection rate for both **[AZ;10_12b]** and **[AZ;15]** infectious clones, bringing it within an acceptable range for subsequent experiments (**Table 2 - B**).

1-2. Design of qPCR primers to specifically detect new infectious clone genomic segments

I devised pairs of qPCR primers designed to be segment-specific for both **[AZ;15]** and **[AZ;10_12b]**, taking into account the genetic characteristics of both isolates as well as **[JKI-2000]**. Subsequent to several modifications and adjustments, two sets of primer pairs capable of specifically amplifying each segment in an isolate-specific manner were successfully developed (**Table 3**). To validate the efficacy and specificity of these primer sets, we conducted cross-amplification tests using pCAMBIA-plasmids of **[AZ;10_12b]** and **[AZ;15]** and extracts from infected plants of the three isolates, namely **[JKI-2000]** (not shown in this manuscript), **[AZ;10_12b]** and **[AZ;15** (**Figure 17**).

Because **[JKI-2000]** primers were already available at the start of the PhD, they were not designed to be specific only to its genomic segments in a mixed-infection with **[AZ;10_12b]** or **[AZ;15]**. In this regard, they probably amplify segments from either **[AZ;10_12b]** and **[AZ;15]**, especially concerning very conserved segments like DNA-N, -R and -M, but this assumption has not been verified. To use **[JKI-2000]** in this context would probably need new primers.

As previously mentioned for **[JKI-2000]** isolate, our laboratory conditions sometimes result in infected plants missing viral segments. This implies that relying solely on observed symptoms is insufficient to confirm the presence of all genomic segments. Analysis of samples extracted from symptomatic plants of the previous agroinoculation experiment detected the presence of all 8 segments, at least in some cases. These findings provide confirmation that all segments from both isolates are able to replicate and systematically spread within the plant during the course of disease development (**Figure 17**).

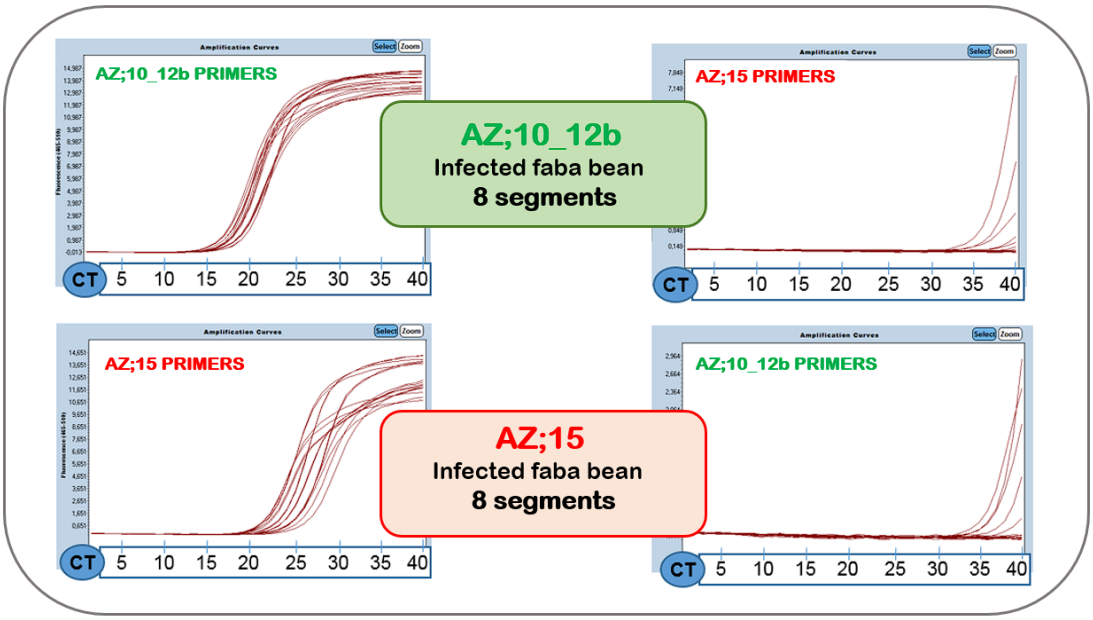


Figure 17 : Amplification curves during qPCR detection of all genomic segments of both infectious clones [AZ;15] and [AZ;10_12b] using primers described in Table 3. The two top panels correspond the amplification of a faba bean [AZ;10_12b] complete infection using sets of primers designed for the detection of either [AZ;10_12b] or [AZ;15] segments. The two bottom panels correspond to the amplification of a faba bean [AZ;15] complete infection using sets of primers designed for the detection of either [AZ;10_12b] or [AZ;15] segments. All segments are detected and in the same cycles range observed for [JKI-2000] usual detection, between 10 and 25 CT. Each couple of primers designed for the detection of a specific segment amplify only this corresponding segment among all of either [AZ;10_12b], [AZ;15] or [JKI-2000] infectious clones.

PLANT	C - 10_12b	C - 15	M - 10_12b	M - 15	N - 10_12b	N - 15	R - 10_12b	R - 15	S - 10_12b	S - 15	U1 - 10_12b	U1 - 15	U2 - 10_12b	U2 - 15	U4 - 10_12b	U4 - 15
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																

Figure 18 : **[AZ;15] and [AZ;10_12b] Table of genomic segment presence and absence in co-inoculated plants extracted at 30-days post-inoculation**. Plant number (name) is displayed in the first column. Subsequent columns (e.g., C - 15) represent the segment amplified from the respective genotypes (15 corresponds to [AZ;15] and 10_12b represents [AZ;10_12b]). Green cases correspond to the presence of the segment whereas red case corresponds to its absence. Thresholds of amplification are set with reassorted thresholds from (**Torralba et al. 2024b – Chapter 3**).

1-3. Mixed-infection in faba beans

To explore the interactions between the two isolates in preparation for future experiments requiring both genotypes, mixed infections were intentionally induced in faba bean plants. Two groups of five *Aphis craccivora* individuals, which had previously fed on **[AZ;10_12b]** or **[AZ;15]** infected faba beans with the 8 segments, were subsequently transferred to young faba bean seedlings. Each set of aphids was deposited and caged onto a distinct leaflet of the same recipient plant. Plants were extracted and their segments quantified at 30 dpi and filtered with the thresholds determined in (**Torralba et al. 2024b – Chapter 3**). The findings demonstrate that all sixteen segments can coexist within the same host plant, as illustrated by 3 plants out of 17 (**Figure 18**). No additional assessment was conducted on these co-infected plants to determine transmission rates or to evaluate their symptom severity.

We observed 15 out of 16 plants fully infected by **[AZ;10_12b]** and one plant only missing DNA-R. On contrast, 5 out of 16 plants were fully infected by **[AZ;15]** with 3 out of 16 not infected at all. The rest of the plants have at least one segment of **[AZ;15]**. These results might imply that **[AZ;10_12b]** develops more rapidly than **[AZ;15]** in faba beans. Alternatively, this observation might indicate a competitive advantage or a variation in transmission rates between the two isolates, even when using five aphids for inoculation.

Plants 6 to 14 exhibited incomplete infections caused by **[AZ;15]**, and these infections excluded essential segments such as DNA-R, which is crucial for replication. Plant 5 exhibited incomplete infection of **[AZ;10_12b]** missing the essential DNA-R that appeared complemented by **[AZ;15]**. In a standalone infection scenario, these incomplete infections would not have led to disease development. This observation confirms the complementation of essential **[AZ;15]** functions by **[AZ;10_12b]** and reciprocally, at least for replication. In some instances, segments from **[AZ;15]** appeared to be maintained and could potentially give rise to reassortants, as exemplified by DNA-M in plant 8, -N in plants 8 and 11, -U1 in plants 8, 9, 10, 12, 13, 14 and -U2 segments in plants 8, 9 and 10.

2. Viral-Vector strategies for the heterologous production of an essential nanovirus protein

2-1. Heterologous expression of nanovirus protein in susceptible host approaches

To pursue our demonstration of genome reconstitution, that has thus far been addressed through non-concomitant transmission of incomplete but viable infections (**Di Mattia et al. 2022**), we were interested in exploring methods to investigate non-concomitant host-to-host transmission of non-viable incomplete sets of segments and possible reconstitution. These incomplete sets are characterized by the absence of essential genomic segments such as DNA-M, -R, or -S.

Two possible approaches were considered. Firstly, the stable transformation of a host plant susceptible to FBNSV, with the objective of producing either M-Rep, MP, or CP proteins. The feasibility of this approach seems questionable as the transformation of legume hosts is not a routine practice while *Arabidopsis thaliana* infection by FBNSV is not easily achieved (**Vega-Arreguín, Gronenborn and Ramírez 2007**).

The second approach involves co-infection by FBNSV and a modified viral vector designed to express one of the essential FBNSV proteins. This expression is intended to complement the corresponding essential components of the viral genome, as the virus would be unable to establish an infection without this complementation. The same transmission methods, both parallel and sequential, as in (**Di Mattia et al. 2022**) would be employed to potentially demonstrate genome reconstitution from sets of segments incapable of establishing an infection alone. This would serve as confirmation that individual genomic segments, whatever their identity, can independently spread and initiate infections under the right complementary conditions in a susceptible host.

The second approach was selected. The envisioned viral vectors should possess the capability to infect the phloem of a *Fabaceae* host plant susceptible to FBNSV, ideally faba beans. They should be transmitted by a different insect vector than aphids or through a different transmission mode. This is essential to enable the exclusive transmission of the FBNSV segments from co-infected plants without co-transmitting the complementing viral-vector. Furthermore, these viral vectors should be available as infectious clones, which would enable easier genetic engineering for the expression of a nanovirus protein.

NAME	ISOLATES	SEGMENT	SENSE	SEQUENCE
CPMV JKI-2000 C – For	FBNSV JKI-2000	C	Forward	CACGCA GGGCCC ATGGGTCTGAAATAT
CPMV JKI-2000 C – Rev	FBNSV JKI-2000	C	Reverse	TCGGTC AGGCCT TTAATTAATTACAAT
CPMV JKI-2000 M – For	FBNSV JKI-2000	M	Forward	CACGCA GGGCCC ATGGCTTCGTCTAGC
CPMV JKI-2000 M – Rev	FBNSV JKI-2000	M	Reverse	TCGGTC AGGCCT TTAAAATACAGATGT
CPMV JKI-2000 N – For	FBNSV JKI-2000	N	Forward	CACGCA GGGCCC ATGGCAGATTGGTTT
CPMV JKI-2000 N – Rev	FBNSV JKI-2000	N	Reverse	TCGGTC AGGCCT TTACACTTTGATTCT
CPMV JKI-2000 R – For	FBNSV JKI-2000	R	Forward	CACGCA GGGCCC ATGGCTCGGCAAGTT
CPMV JKI-2000 R – Rev	FBNSV JKI-2000	R	Reverse	TCGGTC AGGCCT TCATGCGTATACAAA
CPMV JKI-2000 S – For	FBNSV JKI-2000	S	Forward	CACGCA GGGCCC ATGGTGAGCAATTGG
CPMV JKI-2000 S – Rev	FBNSV JKI-2000	S	Reverse	TCGGTC AGGCCT TTAAACTTCAATATA
CPMV JKI-2000 U1 – For	FBNSV JKI-2000	U1	Forward	CACGCA GGGCCC ATGCCTTTGGTCGAT
CPMV JKI-2000 U1 – Rev	FBNSV JKI-2000	U1	Reverse	TCGGTC AGGCCT TTAAGTATTATTAAA
CPMV JKI-2000 U2 – For	FBNSV JKI-2000	U2	Forward	CACGCA GGGCCC ATGGATGCCGGCTTT
CPMV JKI-2000 U2 – Rev	FBNSV JKI-2000	U2	Reverse	TCGGTC AGGCCT TTAATTACATGAAGT
CPMV JKI-2000 U4 – For	FBNSV JKI-2000	U4	Forward	CACGCA GGGCCC ATGGAATGTAGGTTC
CPMV JKI-2000 U4 – Rev	FBNSV JKI-2000	U4	Reverse	TCGGTC AGGCCT TTATTTCTTACTACT

Table 4 : Table overview of primer pairs designed to amplify corresponding FBNSV coding sequences. Restriction site enabling the insertion in CPMV corresponding plasmid is integrated in each primer as the red sequence. Blue sequence corresponds to the matching sequence on the nanovirus ORF for hybridization.

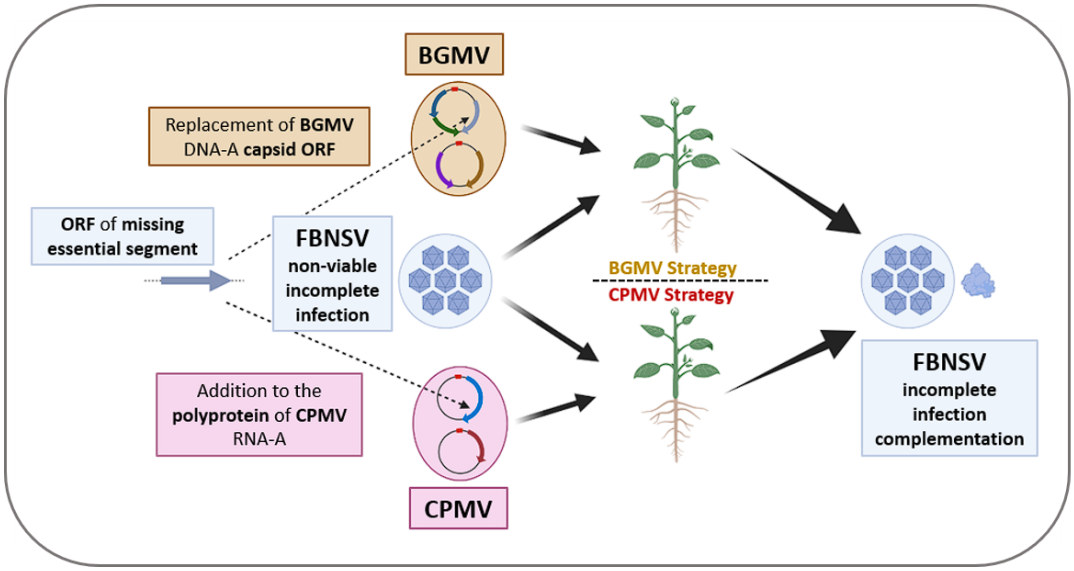


Figure 19 : Schematic illustration of both viral vector strategies designed for the production and complementation of a nanovirus essential function. ORF of an essential associated segment (M, R or S) will be genetically engineered in either BGMV viral vector by the replacement of its capsid ORF, or CPMV viral vector in addition to its polyprotein. Co-infection of one of the modified viral-vectors and agroinoculation of sets of FBNSV segments omitting the corresponding essential segment will hopefully result in the complementation and maintenance of a FBNSV population where an essential segment is missing. A viral segment population that would only survive in co-infected plants. Figure was created using **BioRender.com**.

2-2. Viral vector strategies for heterologous production of a nanovirus protein

I chose to try to use two distinct viral vectors. The first strategy uses a modified infectious clone of the bipartite *Cowpea mosaic virus* (CPMV), which belongs to the *Comovirus* genus from the *Secoviridae* family, provided by George LOMONOSSOFF. The CPMV strategy involves the direct replacement of the GFP ORF within a modified CPMV RNA-A viral vector with the nanovirus gene intended for expression (**Figure 19**). This viral vector was originally designed for the production of heterologous proteins in plants. The plasmids have been received and subsequently transformed into both *Escherichia coli* DH5α and *A. tumefaciens* COR308. We are presently in the process of testing agroinoculation and systemic infection on plant hosts, specifically *Vigna unguiculata* and *Vicia faba*. Mosaic-like symptoms were observed following the agroinoculation of the *Vigna unguiculata* plant hosts. These plants were co-infected with FBNSV and exhibited both curling and mosaic symptoms. However, the confirmation of co-infection by CPMV through qPCR has not been established as primers for RT-qPCR were not designed yet. I have nevertheless designed primers that require testing for both the detection of CPMV genomic segments and for the amplification of the FBNSV ORF. These latter primers have been designed to include the essential restriction sites required for their incorporation into the CPMV plasmids (**Table 4**).

The second approach involves the *Bean golden mosaic virus* (BGMV), a bipartite begomovirus, from the *Geminiviridae* family. This infectious clone was provided by Murilo ZERBINI. This approach is based on the observations that some bipartite begomoviruses, such as BGMV in bean, can establish systemic infections in specific plant hosts even when their capsid gene is absent. Therefore, the BGMV strategy involves replacing a portion of the BGMV DNA-A capsid ORF with the selected nanovirus protein ORF (MP, M-Rep, or CP) while maintaining the overall size of the BGMV DNA-A unaltered in order to prevent length constraints that could favor the emergence of revertants with an optimal length (**Rojas et al. 1998; Gilbertson et al. 2003**) (**Figure 19**).

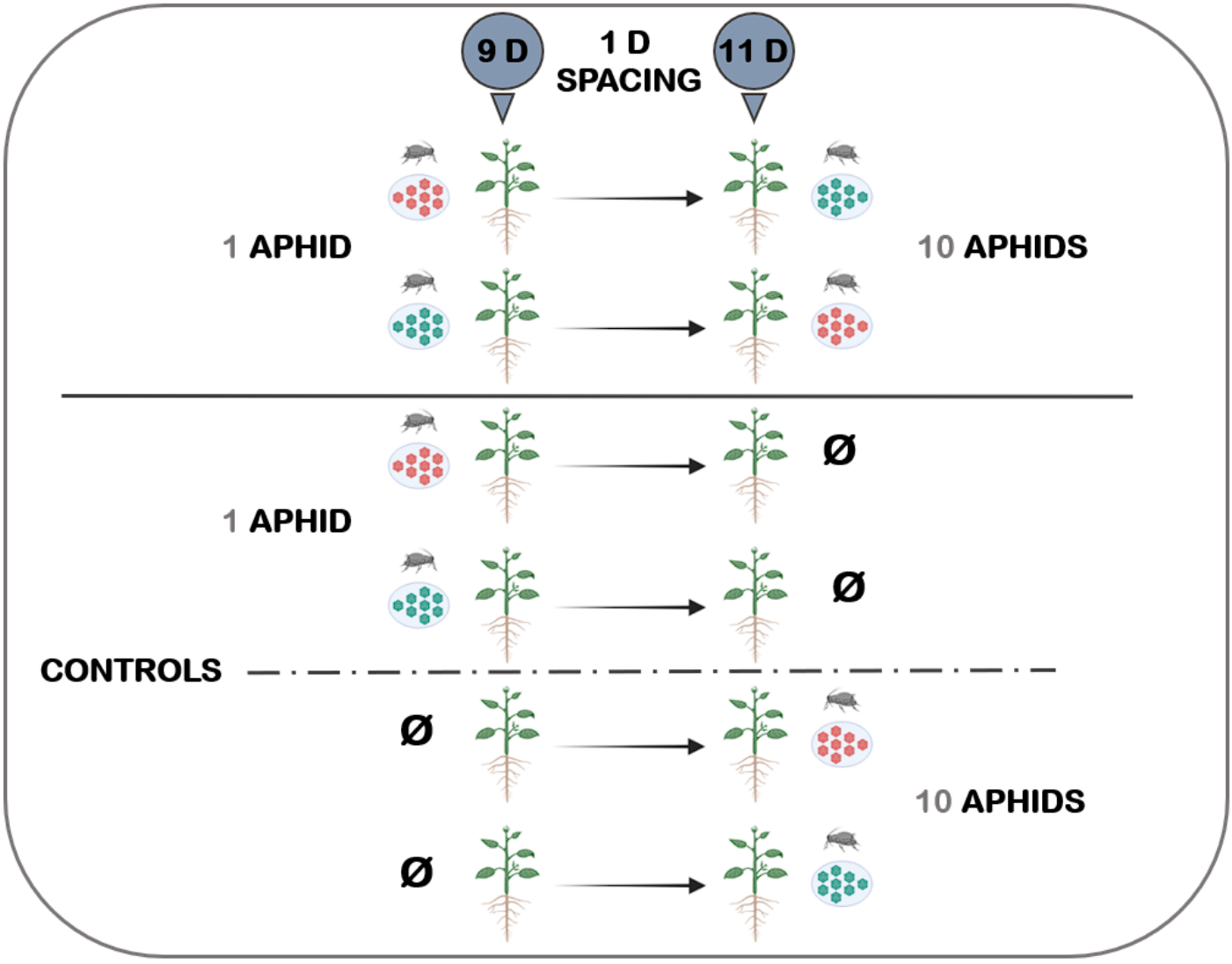


Figure 20 : Graphical abstract of the rescue experiment. One aphid will be used to inoculate with one isolate 9 days-old (D) plants thus starting an infection between 20 and 40 % of the plants. After a spacing day (D) to avoid the formation of reassortants foci that could start infections by themselves, 10 aphids will be used to inoculate those same plants in order to assure the infection of 100 % plants with the second isolate. If we detect primo-inoculations lacking some essential segments it will gather proofs to a potential rescue of this infection by the second inoculation. As controls we will inoculate 9 days-old and 11 days-old (D) plants in the same conditions to compare the transmission rates. Figure was created using **BioRender.com**.

3. Rescue experiment design

The hypothesis that genomic segments (alone or a group lacking essential ones) can stay latent in insect vectors or susceptible hosts, awaiting future rescue or complementation, would greatly reduce the inter-host cost to maintain multipartite genomic integrity. To test for this possible latent phase of one or several segments, an experiment was designed.

In order to identify one or a group of latent segments originating from a failed primo-infection among segments of a secondary infection which could have rescued them, we needed to consider certain criteria:

- An essential segment must be absent from the primo-infection, while other segments may be present.
- The corresponding essential segment from the secondary infection should be detectable.
- It should not be possible to ascribe the absence of the essential segment from the primo-infection to competition with the segments associated with the secondary infection. An essential segment that is consistently maintained in a competitive environment would be a good candidate for the primo-infection.

Two separate experiments would be possibly conducted simultaneously in order to assess both the competition effect between genomic segments of both genotypes and the actual rescue experiment itself (**Figure 20**). The only distinction between the two experiments lies in the number of aphids employed for the inoculation of the primo-infection, which determines whether or not all receiving plants become infected. In the rescue experiment, one aphid will be used to inoculate 9-day-old plants for a single day, resulting in a transmission rate expected to be significantly less than 100 %. Following a one-day gap to induce latency, a duration that would be increased if we obtain positive results, a set of 10 viruliferous aphids of the second genotype will then inoculate the same plants, aiming for a transmission rate as close as possible to 100 %. Concurrently, control groups for both primary and second transmissions under the same conditions will provide a basis for comparing their respective transmission rates, without the influence of rescue or competition effects. Conversely, for the competition experiment, 10 aphids will be used to achieve an almost 100 % infection rate from the primo-infection and the same number of aphids for the secondo-infection. This would ensure the presence of essential segments required for this infection. The absence of these segments may suggest that they were outcompeted.

GENERAL BIBLIOGRAPHY

Abraham, A. D., Bencharki, B., Torok, V., Katul, L., Varrelmann, M., & Josef Vetten, H. (2010). 'Two distinct nanovirus species infecting faba bean in Morocco', *Archives of Virology*, 155/1: 37–46. DOI: 10.1007/s00705-009-0548-9

Abraham, A. D., Varrelmann, M., & Josef Vetten, H. (2012). 'Three Distinct Nanoviruses, One of Which Represents a New Species, Infect Faba Bean in Ethiopia', *Plant Disease*, 96/7: 1045–53. DOI: 10.1094/PDIS-09-11-0734-RE

Ågren, J. A., & Clark, A. G. (2018). 'Selfish genetic elements', *PLoS genetics*, 14/11: e1007700. DOI: 10.1371/journal.pgen.1007700

Amin, I., Ilyas, M., Qazi, J., Bashir, R., Yadav, J. S., Mansoor, S., Fauquet, C. M., et al. (2011). 'Identification of a major pathogenicity determinant and suppressors of RNA silencing encoded by a South Pacific isolate of Banana bunchy top virus originating from Pakistan', *Virus Genes*, 42/2: 272–81. DOI: 10.1007/s11262-010-0559-3

Argüello-Astorga, G. R., & Ruiz-Medrano, R. (2001). 'An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach', *Archives of Virology*, 146/8: 1465–85. DOI: 10.1007/s007050170072

Avalos-Calleros, J. A., Pastor-Palacios, G., Bolaños-Martínez, O. C., Mauricio-Castillo, A., Gregorio-Jorge, J., Martínez-Marrero, N., Bañuelos-Hernández, B., et al. (2021). 'Two strains of a novel begomovirus encoding Rep proteins with identical β1 strands but different β5 strands are not compatible in replication', *Archives of Virology*, 166/6: 1691–709. DOI: 10.1007/s00705-021-05066-2

Babin, M., Ortíz, V., Castro, S., & Romero, J. (2000). 'First Detection of Faba bean necrotic yellow virus in Spain', *Plant Disease*, 84/6: 707. DOI: 10.1094/PDIS.2000.84.6.707B

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). 'The MEME Suite', *Nucleic Acids Research*, 43/W1: W39-49. DOI: 10.1093/nar/gkv416

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: 12416–21. DOI: 10.1128/JVI.01542-08

Chakraborty, S., Vanitharani, R., Chattopadhyay, B., & Fauquet, C. M. (2008). 'Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India', *The Journal of General Virology*, 89/Pt 3: 818–28. DOI: 10.1099/vir.0.82873-0

Chao, L., Tran, T., & Matthews, C. (1992). 'MULLER'S RATCHET AND THE ADVANTAGE OF SEX IN THE RNA VIRUS ϕ6', *Evolution; International Journal of Organic Evolution*, 46/2: 289–99. DOI: 10.1111/j.1558-5646.1992.tb02038.x

Chu, P. W., Keese, P., Qiu, B. S., Waterhouse, P. M., & Gerlach, W. L. (1993). 'Putative full-length clones of the genomic DNA segments of subterranean clover stunt virus and identification of the segment coding for the viral coat protein', *Virus Research*, 27/2: 161–71. DOI: 10.1016/0168-1702(93)90079-3

Dale, J. L. (1987). 'Banana Bunchy Top: An Economically Important Tropical Plant Virus Disease'. Maramorosch K., Murphy F. A., & Shatkin A. J. (eds) *Advances in Virus Research*, Vol. 33, pp. 301–25. Academic Press. DOI: 10.1016/S0065-3527(08)60321-8

Di Mattia, J., Ryckebusch, F., Vernerey, M.-S., Pirolles, E., Sauvion, N., Peterschmitt, M., Zeddam, J.-L., et al. (2020). 'Co-Acquired Nanovirus and Geminivirus Exhibit a Contrasted Localization within Their Common Aphid Vector', *Viruses*, 12/3: 299. DOI: 10.3390/v12030299

Di Mattia, J., Torralba, B., Yvon, M., Zeddam, J.-L., Blanc, S., & Michalakis, Y. (2022). 'Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution', *Proceedings of the National Academy of Sciences of the United States of America*, 119/32: e2201453119. DOI: 10.1073/pnas.2201453119

Di Mattia, J., Vernerey, M.-S., Yvon, M., Pirolles, E., Villegas, M., Gaafar, Y., Ziebell, H., et al. (2020). 'Route of a Multipartite Nanovirus across the Body of Its Aphid Vector', *Journal of Virology*, 94/9: e01998-19. DOI: 10.1128/JVI.01998-19

Escriu, F., Fraile, A., & García-Arenal, F. (2007). 'Constraints to genetic exchange support gene coadaptation in a tripartite RNA virus', *PLoS pathogens*, 3/1: e8. DOI: 10.1371/journal.ppat.0030008

Fulton, R. W. (1980). 'Biological Significance of Multicomponent Viruses', *Annual Review of Phytopathology*, 18/1: 131–46. DOI: 10.1146/annurev.py.18.090180.001023

Gaafar, Y., Grausgruber-Gröger, S., & Ziebell, H. (2016). 'Vicia faba, V. sativa and Lens culinaris as new hosts for Pea necrotic yellow dwarf virus in Germany and Austria', *New Disease Reports*, 34/1: 28–28. DOI: 10.5197/j.2044-0588.2016.034.028

Gallet, R., Di Mattia, J., Ravel, S., Zeddam, J.-L., Vitalis, R., Michalakis, Y., & Blanc, S. (2022). 'Gene copy number variations at the within-host population level modulate gene expression in a multipartite virus', *Virus Evolution*, 8/2: veac058. DOI: 10.1093/ve/veac058

Gallet, R., Fabre, F., Thébaud, G., Sofonea, M. T., Sicard, A., Blanc, S., & Michalakis, Y. (2018). 'Small Bottleneck Size in a Highly Multipartite Virus during a Complete Infection Cycle', *Journal of Virology*, 92/14: e00139-18. DOI: 10.1128/JVI.00139-18

Gallet, R., Kraberger, S., Filloux, D., Galzi, S., Fontes, H., Martin, D. P., Varsani, A., et al. (2018). 'Nanovirus-alphasatellite complex identified in Vicia cracca in the Rhône delta region of France', *Archives of Virology*, 163/3: 695–700. DOI: 10.1007/s00705-017-3634-4

García-Arriaza, J., Manrubia, S. C., Toja, M., Domingo, E., & Escarmís, C. (2004). 'Evolutionary transition toward defective RNAs that are infectious by complementation', *Journal of Virology*, 78/21: 11678–85. DOI: 10.1128/JVI.78.21.11678-11685.2004

González-Jara, P., Fraile, A., Canto, T., & García-Arenal, F. (2009). 'The multiplicity of infection of a plant virus varies during colonization of its eukaryotic host', *Journal of Virology*, 83/15: 7487–94. DOI: 10.1128/JVI.00636-09

Grigoras, I., Gronenborn, B., & Vetten, H. J. (2010). 'First Report of a Nanovirus Disease of Pea in Germany', *Plant Disease*, 94/5: 642. DOI: 10.1094/PDIS-94-5-0642C

Grigoras, Ioana, Ginzo, A. I. D. C., Martin, D. P., Varsani, A., Romero, J., Mammadov, A. C., Huseynova, I. M., et al. (2014). 'Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe', *The Journal of General Virology*, 95/Pt 5: 1178–91. DOI: 10.1099/vir.0.063115-0

Grigoras, Ioana, Timchenko, T., Grande-Pérez, A., Katul, L., Vetten, H.-J., & Gronenborn, B. (2010). 'High variability and rapid evolution of a nanovirus', *Journal of Virology*, 84/18: 9105–17. DOI: 10.1128/JVI.00607-10

Grigoras, Ioana, Timchenko, T., Katul, L., Grande-Pérez, A., Vetten, H.-J., & Gronenborn, B. (2009). 'Reconstitution of authentic nanovirus from multiple cloned DNAs', *Journal of Virology*, 83/20: 10778–87. DOI: 10.1128/JVI.01212-09

Grigoras, Ioana, Vetten, H.-J., Commandeur, U., Ziebell, H., Gronenborn, B., & Timchenko, T. (2018). 'Nanovirus DNA-N encodes a protein mandatory for aphid transmission', *Virology*, 522: 281–91. DOI: 10.1016/j.virol.2018.07.001

Gutiérrez, S., Michalakis, Y., & Blanc, S. (2012). 'Virus population bottlenecks during within-host progression and host-to-host transmission', *Current Opinion in Virology*, 2/5: 546–55. DOI: 10.1016/j.coviro.2012.08.001

Gutiérrez, S., & Zwart, M. P. (2018). 'Population bottlenecks in multicomponent viruses: first forays into the uncharted territory of genome-formula drift', *Current Opinion in Virology*, 33: 184–90. DOI: 10.1016/j.coviro.2018.09.001

Guyot, V., Rajeswaran, R., Chu, H. C., Karthikeyan, C., Laboureau, N., Galzi, S., Mukwa, L. F. T., et al. (2022). 'A newly emerging alphasatellite affects banana bunchy top virus replication, transcription, siRNA production and transmission by aphids', *PLoS pathogens*, 18/4: e1010448. DOI: 10.1371/journal.ppat.1010448

Hasanvand, V., Heydanejad, J., Massumi, H., Kleinow, T., Jeske, H., Fontenele, R. S., Kraberger, S., et al. (2021). 'Genome characterization of parsley severe stunt-associated virus in Iran', *Virus Genes*, 57/3: 293–301. DOI: 10.1007/s11262-021-01835-y

Hassan-Sheikhi, P., Heydarnejad, J., Massumi, H., Kraberger, S., & Varsani, A. (2020). 'Novel nanovirus and associated alphasatellites identified in milk vetch plants with chlorotic dwarf disease in Iran', *Virus Research*, 276: 197830. DOI: 10.1016/j.virusres.2019.197830

Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E., & Hatfull, G. F. (1999). 'Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage', *Proceedings of the National Academy of Sciences*, 96/5: 2192–7. Proceedings of the National Academy of Sciences. DOI: 10.1073/pnas.96.5.2192

Herrera-Valencia, V. A., Dugdale, B., Harding, R. M., & Dale, J. L. (2006). 'An iterated sequence in the genome of Banana bunchy top virus is essential for efficient replication', *The Journal of General Virology*, 87/Pt 11: 3409–12. DOI: 10.1099/vir.0.82166-0

Heydarnejad, J., Kamali, M., Massumi, H., Kvarnheden, A., Male, M. F., Kraberger, S., Stainton, D., et al. (2017). 'Identification of a Nanovirus-Alphasatellite Complex in Sophora alopecuroides', *Virus Research*, 235: 24–32. DOI: 10.1016/j.virusres.2017.03.023

Holland, J., & Domingo, E. (1998). 'Origin and Evolution of Viruses', *Virus Genes*, 16/1: 13–21. DOI: 10.1023/A:1007989407305

Horser, C. L., Harding, R. M., & Dale, J. L. (2001). 'Banana bunchy top nanovirus DNA-1 encodes the "master" replication initiation protein', *The Journal of General Virology*, 82/Pt 2: 459–64. DOI: 10.1099/0022-1317-82-2-459

Hu, ZhaoYang, Li, G., Li, G., Yao, Q., & Chen, K. (2013). 'Bombyx mori bidensovirus: The type species of the new genus Bidensovirus in the new family Bidnaviridae', *Chinese Science Bulletin*, 58/36: 4528–32. DOI: 10.1007/s11434-013-5876-1

Hu, Zhaoyang, Zhang, X., Liu, W., Zhou, Q., Zhang, Q., Li, G., & Yao, Q. (2016). 'Genome segments accumulate with different frequencies in Bombyx mori bidensovirus', *Journal of Basic Microbiology*, 56/12: 1338–43. DOI: 10.1002/jobm.201600120

Idris, A. M., Mills-Lujan, K., Martin, K., & Brown, J. K. (2008). 'Melon chlorotic leaf curl virus: characterization and differential reassortment with closest relatives reveal adaptive virulence in the squash leaf curl virus clade and host shifting by the host-restricted bean calico mosaic virus', *Journal of Virology*, 82/4: 1959–67. DOI: 10.1128/JVI.01992-07

Iranzo, J., & Manrubia, S. C. (2012). 'Evolutionary dynamics of genome segmentation in multipartite viruses', *Proceedings. Biological Sciences*, 279/1743: 3812–9. DOI: 10.1098/rspb.2012.1086

Iranzo, J., Puigbò, P., Lobkovsky, A. E., Wolf, Y. I., & Koonin, E. V. (2016). 'Inevitability of Genetic Parasites', *Genome Biology and Evolution*, 8/9: 2856–69. DOI: 10.1093/gbe/evw193

Kamran, A., Hou, H., Xie, Y., Zhao, C., Wei, X., Zhang, C., Yu, X., et al. (2019). 'Full genome sequence analysis and putative host-shifting of Milk vetch dwarf virus infecting tobacco (Nicotiana tabacum) in China', *Virology Journal*, 16/1: 38. DOI: 10.1186/s12985-019-1129-5

Knierim, D., Barrière, Q., Grigoras, I., Winter, S., Vetten, H.-J., Schwinghamer, M., Thomas, J., et al. (2019). 'Subterranean Clover Stunt Virus Revisited: Detection of Two Missing Genome Components', *Viruses*, 11/2: 138. DOI: 10.3390/v11020138

Koonin, E. V. (2014). 'The origins of cellular life', *Antonie Van Leeuwenhoek*, 106/1: 27–41. DOI: 10.1007/s10482-014-0169-5

Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., Zerbini, F. M., et al. (2020). 'Global Organization and Proposed Megataxonomy of the Virus World', *Microbiology and molecular biology reviews: MMBR*, 84/2: e00061-19. DOI: 10.1128/MMBR.00061-19

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: MMBR*, 85/3: e0005321. DOI: 10.1128/MMBR.00053-21

Koonin, E. V., Krupovic, M., & Dolja, V. V. (2023). 'The global virome: How much diversity and how many independent origins?', *Environmental Microbiology*, 25/1: 40–4. DOI: 10.1111/1462-2920.16207

Krapp, S., Greiner, E., Amin, B., Sonnewald, U., & Krenz, B. (2017). 'The stress granule component G3BP is a novel interaction partner for the nuclear shuttle proteins of the nanovirus pea

necrotic yellow dwarf virus and geminivirus abutilon mosaic virus', *Virus Research*, 227: 6–14. DOI: 10.1016/j.virusres.2016.09.021

Krenz, B., Schießl, I., Greiner, E., & Krapp, S. (2017). 'Analyses of pea necrotic yellow dwarf virus-encoded proteins', *Virus Genes*, 53/3: 454–63. DOI: 10.1007/s11262-017-1439-x

Krupovic, M., Dolja, V. V., & Koonin, E. V. (2019). 'Origin of viruses: primordial replicators recruiting capsids from hosts', *Nature Reviews. Microbiology*, 17/7: 449–58. DOI: 10.1038/s41579-019-0205-6

——. (2020). 'The LUCA and its complex virome', *Nature Reviews. Microbiology*, 18/11: 661–70. DOI: 10.1038/s41579-020-0408-x

Krupovic, M., & Koonin, E. V. (2017). 'Multiple origins of viral capsid proteins from cellular ancestors', *Proceedings of the National Academy of Sciences of the United States of America*, 114/12: E2401–10. DOI: 10.1073/pnas.1621061114

Kumari, S. G., Attar, N., Mustafayev, E., & Akparov, Z. (2009). 'First Report of Faba bean necrotic yellows virus Affecting Legume Crops in Azerbaijan', *Plant Disease*, 93/11: 1220. DOI: 10.1094/PDIS-93-11-1220C

Kumari, S. G., Makkouk, K. M., Attar, N., Ghulam, W., & Lesemann, D.-E. (2004). 'First Report of Chickpea Chlorotic Dwarf Virus Infecting Spring Chickpea in Syria', *Plant Disease*, 88/4: 424. DOI: 10.1094/PDIS.2004.88.4.424C

Lageix, S., Catrice, O., Deragon, J.-M., Gronenborn, B., Pélissier, T., & Ramírez, B. C. (2007). 'The nanovirus-encoded Clink protein affects plant cell cycle regulation through interaction with the retinoblastoma-related protein', *Journal of Virology*, 81/8: 4177–85. DOI: 10.1128/JVI.02103-06

Lal, A., Vo, T. T. B., Sanjaya, I. G. N. P. W., Ho, P. T., Kim, J.-K., Kil, E.-J., & Lee, S. (2020). 'Nanovirus Disease Complexes: An Emerging Threat in the Modern Era', *Frontiers in Plant Science*, 11: 558403. DOI: 10.3389/fpls.2020.558403

Leeks, A., Young, P. G., Turner, P. E., Wild, G., & West, S. A. (2023). 'Cheating leads to the evolution of multipartite viruses', *PLoS biology*, 21/4: e3002092. DOI: 10.1371/journal.pbio.3002092

Li, P., Wang, S., Zhang, L., Qiu, D., Zhou, X., & Guo, L. (2020). 'A tripartite ssDNA mycovirus from a plant pathogenic fungus is infectious as cloned DNA and purified virions', *Science Advances*, 6/14: eaay9634. DOI: 10.1126/sciadv.aay9634

Londoño, A., Riego-Ruiz, L., & Argüello-Astorga, G. R. (2010). 'DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs', *Archives of Virology*, 155/7: 1033–46. DOI: 10.1007/s00705-010-0674-4

Lotfipour, M., Behjatnia, S. A. A., Dall'Ara, M., & Ratti, C. (2020). 'The full-length genome characterization and diversity of faba bean necrotic stunt virus in Iran', *European Journal of Plant Pathology*, 157/2: 239–50. DOI: 10.1007/s10658-020-01973-3

Lucía-Sanz, A., Aguirre, J., & Manrubia, S. (2018). 'Theoretical approaches to disclosing the emergence and adaptive advantages of multipartite viruses', *Current Opinion in Virology*, 33: 89–95. DOI: 10.1016/j.coviro.2018.07.018

Lucía-Sanz, A., & Manrubia, S. (2017). 'Multipartite viruses: adaptive trick or evolutionary treat?', *NPJ systems biology and applications*, 3: 34. DOI: 10.1038/s41540-017-0035-y

Mandal, B., Shilpi, S., Barman, A. R., Mandal, S., & Varma, A. (2013). 'Nine novel DNA components associated with the foorkey disease of large cardamom: evidence of a distinct babuvirus species in Nanoviridae', *Virus Research*, 178/2: 297–305. DOI: 10.1016/j.virusres.2013.09.027

Mansourpour, M., Gallet, R., Abbasi, A., Blanc, S., Dizadji, A., & Zeddam, J.-L. (2022). 'Effects of an Alphasatellite on the Life Cycle of the Nanovirus Faba Bean Necrotic Yellows Virus', *Journal of Virology*, 96/3: e0138821. DOI: 10.1128/JVI.01388-21

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., & Varsani, A. (2011). 'Recombination in Eukaryotic Single Stranded DNA Viruses', *Viruses*, 3/9: 1699–738. DOI: 10.3390/v3091699

Martin, D. P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.-Y., Dijoux, B., Vincent, C., et al. (2011). 'Complex Recombination Patterns Arising during Geminivirus Coinfections Preserve and Demarcate Biologically Important Intra-Genome Interaction Networks', *PLoS Pathogens*, 7/9: e1002203. DOI: 10.1371/journal.ppat.1002203

Matsuzaki, Y., Mizuta, K., Sugawara, K., Tsuchiya, E., Muraki, Y., Hongo, S., Suzuki, H., et al. (2003). 'Frequent Reassortment among Influenza C Viruses', *Journal of Virology*, 77/2: 871–81. DOI: 10.1128/JVI.77.2.871-881.2003

McDonald, S. M., Nelson, M. I., Turner, P. E., & Patton, J. T. (2016). 'Reassortment in segmented RNA viruses: mechanisms and outcomes', *Nature Reviews. Microbiology*, 14/7: 448–60. DOI: 10.1038/nrmicro.2016.46

Michalakis, Y., & Blanc, S. (2020). 'The Curious Strategy of Multipartite Viruses', *Annual Review of Virology*, 7/1: 203–18. DOI: 10.1146/annurev-virology-010220-063346

Miyashita, S., & Kishino, H. (2010). 'Estimation of the size of genetic bottlenecks in cell-to-cell movement of soil-borne wheat mosaic virus and the possible role of the bottlenecks in speeding up selection of variations in trans-acting genes or elements', *Journal of Virology*, 84/4: 1828–37. DOI: 10.1128/JVI.01890-09

Moury, B., Fabre, F., & Senoussi, R. (2007). 'Estimation of the number of virus particles transmitted by an insect vector', *Proceedings of the National Academy of Sciences of the United States of America*, 104/45: 17891–6. DOI: 10.1073/pnas.0702739104

Mushegian, A. R. (2020). 'Are There 1031 Virus Particles on Earth, or More, or Fewer?', *Journal of Bacteriology*, 202/9: e00052-20. DOI: 10.1128/JB.00052-20

Ohshima, K., Matsumoto, K., Yasaka, R., Nishiyama, M., Soejima, K., Korkmaz, S., Ho, S. Y. W., et al. (2016). 'Temporal analysis of reassortment and molecular evolution of Cucumber mosaic virus: Extra clues from its segmented genome', *Virology*, 487: 188–97. DOI: 10.1016/j.virol.2015.09.024

Ojosnegros, S., García-Arriaza, J., Escarmís, C., Manrubia, S. C., Perales, C., Arias, A., Mateu, M. G., et al. (2011). 'Viral genome segmentation can result from a trade-off between genetic content and particle stability', *PLoS genetics*, 7/3: e1001344. DOI: 10.1371/journal.pgen.1001344

Pénzes, J. J., Pham, H. T., Chipman, P., Smith, E. W., McKenna, R., & Tijssen, P. (2023). 'Bipartite genome and structural organization of the parvovirus Acheta domesticus segmented densovirus', *Nature Communications*, 14/1: 3515. DOI: 10.1038/s41467-023-38875-x

Ponsen, M. B. (1972). *The site of potato leafroll virus multiplication in its vector, Myzus persicae : an anatomical study* (phd). Veenman, Wageningen. Retrieved September 21, 2023, from <https://library.wur.nl/WebQuery/wurpubs/421730>

Qazi, J. (2016). 'Banana bunchy top virus and the bunchy top disease', *Journal of General Plant Pathology*, 82/1: 2–11. DOI: 10.1007/s10327-015-0642-7

Ren, R., Zheng, L., Han, J., Perdoncini Carvalho, C., Miyashita, S., Zhang, D., & Qu, F. (2023). 'Intracellular bottlenecking permits no more than three tomato yellow leaf curl virus genomes to initiate replication in a single cell', *PLoS pathogens*, 19/5: e1011365. DOI: 10.1371/journal.ppat.1011365

Roossinck, M. J. (1997). 'Mechanisms of plant virus evolution', *Annual Review of Phytopathology*, 35: 191–209. DOI: 10.1146/annurev.phyto.35.1.191

Safari Murhububa, I., Tougeron, K., Bragard, C., Fauconnier, M.-L., Bisimwa Basengere, E., Walangululu Masamba, J., & Hance, T. (2021). 'Banana Tree Infected with Banana Bunchy Top Virus Attracts Pentalonia nigronervosa Aphids Through Increased Volatile Organic Compounds Emission', *Journal of Chemical Ecology*, 47/8–9: 755–67. DOI: 10.1007/s10886-021-01298-3

Sánchez-Navarro, J. A., Zwart, M. P., & Elena, S. F. (2013). 'Effects of the number of genome segments on primary and systemic infections with a multipartite plant RNA virus', *Journal of Virology*, 87/19: 10805–15. DOI: 10.1128/JVI.01402-13

Savory, F. R., & Ramakrishnan, U. (2014). 'Asymmetric patterns of reassortment and concerted evolution in Cardamom bushy dwarf virus', *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 24: 15–24. DOI: 10.1016/j.meegid.2014.02.012

Sharman, M., Thomas, J. E., Skabo, S., & Holton, T. A. (2008). 'Abacá bunchy top virus, a new member of the genus Babuvirus (family Nanoviridae)', *Archives of Virology*, 153/1: 135–47. DOI: 10.1007/s00705-007-1077-z

Sicard, A., Michalakis, Y., Gutiérrez, S., & Blanc, S. (2016). 'The Strange Lifestyle of Multipartite Viruses', *PLoS pathogens*, 12/11: e1005819. DOI: 10.1371/journal.ppat.1005819

Sicard, A., Pirolles, E., Gallet, R., Vernerey, M.-S., Yvon, M., Urbino, C., Peterschmitt, M., et al. (2019). 'A multicellular way of life for a multipartite virus', *eLife*, 8: e43599. DOI: 10.7554/eLife.43599

Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakis, Y., Gutierrez, S., & Blanc, S. (2013). 'Gene copy number is differentially regulated in a multipartite virus', *Nature Communications*, 4: 2248. DOI: 10.1038/ncomms3248

Sicard, A., Zeddam, J.-L., Yvon, M., Michalakis, Y., Gutiérrez, S., & Blanc, S. (2015). 'Circulative Nonpropagative Aphid Transmission of Nanoviruses: an Oversimplified View', *Journal of Virology*, 89/19: 9719–26. DOI: 10.1128/JVI.00780-15

Stainton, D., Martin, D. P., Collings, D. A., & Varsani, A. (2017). 'Comparative analysis of common regions found in babuviruses and alphasatellite molecules', *Archives of Virology*, 162/3: 849–55. DOI: 10.1007/s00705-016-3168-1

Stainton, D., Martin, D. P., Muhire, B. M., Lolohea, S., Halafihi, M., Lepoint, P., Blomme, G., et al. (2015). 'The global distribution of Banana bunchy top virus reveals little evidence for frequent recent, human-mediated long distance dispersal events', *Virus Evolution*, 1/1: vev009. DOI: 10.1093/ve/vev009

Sun, X., Xiao, L., Cheng, C., Shi, Z., Sun, Z., Ma, Y., & Zhu, X. (2022). 'Genome-wide analysis of milk vetch dwarf virus and related alphasatellites from garlic and aphids in Shandong Province, China', *Archives of Virology*, 167/10: 2071–7. DOI: 10.1007/s00705-022-05490-y

Tadesse, N., Ali, K., Gorfu, D., Abraham, A., Lencho, A., Ayalew, M., Yusuf, A., et al. (1999). 'First Report of Soybean Dwarf Virus Infecting Lentil and Beet Western Yellows Virus Infecting Lentil and Chickpea Crops in Ethiopia', *Plant Disease*, 83/6: 589. DOI: 10.1094/PDIS.1999.83.6.589B

Tentchev, D., Verdin, E., Marchal, C., Jacquet, M., Aguilar, J. M., & Moury, B. (2011). 'Evolution and structure of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes', *The Journal of General Virology*, 92/Pt 4: 961–73. DOI: 10.1099/vir.0.029082-0

Timchenko, T., Katul, L., Aronson, M., Vega-Arreguín, J. C., Ramirez, B. C., Vetten, H. J., & Gronenborn, B. (2006). 'Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus', *The Journal of General Virology*, 87/Pt 6: 1735–43. DOI: 10.1099/vir.0.81753-0

Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H. J., & Gronenborn, B. (2000). 'The master rep concept in nanovirus replication: identification of missing genome components and potential for natural genetic reassortment', *Virology*, 274/1: 189–95. DOI: 10.1006/viro.2000.0439

Trapani, S., Bhat, E. A., Yvon, M., Lai-Kee-Him, J., Hoh, F., Vernerey, M.-S., Pirolles, E., et al. (2023). 'Structure-guided mutagenesis of the capsid protein indicates that a nanovirus requires assembled viral particles for systemic infection', *PLOS Pathogens*, 19/1: e1011086. DOI: 10.1371/journal.ppat.1011086

Urbino, C., Thébaud, G., Granier, M., Blanc, S., & Peterschmitt, M. (2008). 'A novel cloning strategy for isolating, genotyping and phenotyping genetic variants of geminiviruses', *Virology Journal*, 5: 135. DOI: 10.1186/1743-422X-5-135

Varsani, A., Lefeuvre, P., Roumagnac, P., & Martin, D. (2018). 'Notes on recombination and reassortment in multipartite/segmented viruses', *Current Opinion in Virology*, 33: 156–66. DOI: 10.1016/j.coviro.2018.08.013

Vega-Arreguín, J. C., Gronenborn, B., & Ramírez, B. C. (2007). 'Arabidopsis thaliana is a host of the legume nanovirus Faba bean necrotic yellows virus', *Virus Research*, 128/1–2: 81–7. DOI: 10.1016/j.virusres.2007.04.011

Vetten, H. J., Knierim, D., Rakoski, M. S., Menzel, W., Maiss, E., Gronenborn, B., Winter, S., et al. (2019). 'Identification of a novel nanovirus in parsley', *Archives of Virology*, 164/7: 1883–7. DOI: 10.1007/s00705-019-04280-3

Villa, M., & Lässig, M. (2017). 'Fitness cost of reassortment in human influenza', *PLoS pathogens*, 13/11: e1006685. DOI: 10.1371/journal.ppat.1006685

Wang, X.-W., & Blanc, S. (2021). 'Insect Transmission of Plant Single-Stranded DNA Viruses', *Annual Review of Entomology*, 66: 389–405. DOI: 10.1146/annurev-ento-060920-094531

Wanitchakorn, R., Hafner, G. J., Harding, R. M., & Dale, J. L. (2000). 'Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6', *The Journal of General Virology*, 81/Pt 1: 299–306. DOI: 10.1099/0022-1317-81-1-299

Wanitchakorn, R., Harding, R. M., & Dale, J. L. (1997). 'Banana bunchy top virus DNA-3 encodes the viral coat protein', *Archives of Virology*, 142/8: 1673–80. DOI: 10.1007/s007050050188

Werren, J. H., Nur, U., & Wu, C.-I. (1988). 'Selfish genetic elements', *Trends in Ecology & Evolution*, 3/11: 297–302. DOI: 10.1016/0169-5347(88)90105-X

Yu, N.-T., Xie, H.-M., Zhang, Y.-L., Wang, J.-H., Xiong, Z., & Liu, Z.-X. (2019). 'Independent modulation of individual genomic component transcription and a cis-acting element related to high transcriptional activity in a multipartite DNA virus', *BMC genomics*, 20/1: 573. DOI: 10.1186/s12864-019-5901-0

Zhang, C., Zheng, H., Yan, D., Han, K., Song, X., Liu, Y., Zhang, D., et al. (2017). 'Complete genomic characterization of milk vetch dwarf virus isolates from cowpea and broad bean in Anhui province, China', *Archives of Virology*, 162/8: 2437–40. DOI: 10.1007/s00705-017-3348-7

Zwart, M. P., Blanc, S., Johnson, M., Manrubia, S., Michalakis, Y., & Sofonea, M. T. (2021). 'Unresolved advantages of multipartitism in spatially structured environments', *Virus Evolution*, 7/1: veab004. DOI: 10.1093/ve/veab004

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*, 2/1: 161–79. DOI: 10.1146/annurev-virology-100114-055135

Zwart, M. P., & Elena, S. F. (2020). 'Modeling multipartite virus evolution: the genome formula facilitates rapid adaptation to heterogeneous environments†', *Virus Evolution*, 6/1: veaa022. DOI: 10.1093/ve/veaa022

RÉSUMÉ FRANÇAIS

1. Introduction

Les virus ont développé trois stratégies d'encapsidation et d'architecture génomique : monopartite, lorsque toute l'information génétique est portée par une seule molécule d'acide nucléique encapsidée dans une particule virale ; segmenté, lorsque le génome est divisé en plusieurs segments co-encapsidés dans la même particule virale ; et multipartite, lorsque le génome est également divisé mais que les segments sont encapsidés séparément.

L'architecture génomique multipartite représente un mystère évolutif du fait de l'important coût théorique associé à la perte potentielle d'un segment et donc de l'intégrité de l'information génétique essentielle pour le déroulement du cycle infectieux. Ces coûts se traduisent par une augmentation importante du nombre de particules virales *a priori* nécessaire pour établir une infection chez un organisme ou une cellule hôtes (multiplicity of infection, MOI) sans perdre de segments, cette augmentation est d'autant plus importante que le génome est fragmenté en de nombreux segments. À partir de trois segments, les valeurs de MOI nécessaires sont disproportionnées en comparaison des mesures réelles de MOI observées au sein des vecteurs et organismes hôtes.

Parmi les avantages hypothétiques associés à la segmentation du génome, on retrouve la possibilité d'échanger des segments génomiques entiers, une forme d'échange génétique appelée réassortiment. Les réassortiments concernent à la fois les virus segmentés et multipartites, et le bénéfice éventuel qu'ils pourraient apporter ne semble donc pas expliquer l'émergence évolutive de l'encapsidation séparée des segments, qui est la caractéristique distinctive du « multipartitisme ». Les réassortiments peuvent être associés à des effets phénotypiques très importants, mais qui ne sont pas nécessairement bénéfiques. Une évaluation systématique des effets phénotypiques associés aux réassortiments manque actuellement dans la littérature scientifique et pourrait nous en apprendre plus sur l'apport réel de cet échange génétique au sein des populations virales.

Afin d'étudier le multipartitisme, notre groupe de recherche travaille sur le genre *Nanovirus* qui comporte les virus multipartites les plus fragmentés avec un génome divisé en huit segments génomiques. Ces virus à ADN simple brin infectent majoritairement les plantes de la famille *Fabaceae* (légumineuses) et sont transmis horizontalement par des pucerons vecteurs. Comme les autres virus multipartites, ils accumulent leurs segments génomiques différentiellement, créant un pattern de fréquence relative des segments appelé « formule génomique » (**Sicard et al. 2013**). À l'intérieur d'un individu hôte ces nanovirus accumulent leurs segments dans des cellules différentes et sont capables de complémenter

fonctionnellement leur génome à un niveau supra-cellulaire, via l'échange des produits des gènes entre cellules (**Sicard et al. 2019**).

2. Thématique de la thèse

Cette thèse étudie l'architecture génomique multipartite, en se concentrant sur les processus d'échange génétique appelés réassortiments. L'objectif est d'analyser leur influence sur les coûts liés à la préservation de l'intégrité génomique, ainsi que d'évaluer dans quelle mesure ces réassortiments affectent la fitness des génotypes hybrides engendrés.

3. Méthodologie

Notre groupe de recherche se consacre à l'étude du multipartitisme en utilisant principalement le *Faba bean necrotic stunt virus* (**FBNSV**) comme organisme modèle, espèce faisant partie du genre *Nanovirus*. Les recherches menées au cours de cette thèse reposent sur diverses techniques expérimentales nous permettant d'induire des infections de nanovirus sur une plante hôte dans un contexte expérimental par agro-inoculation ou inoculation par puceron.

L'agro-inoculation repose sur la capacité naturelle des bactéries *Agrobacterium tumefaciens* à transformer les cellules végétales compatibles en leur injectant une partie des séquences présentes sur leur plasmide pTi. Cette capacité est utilisée en virologie végétale grâce à la création de plasmides binaires permettant l'injection de séquences virales d'intérêt au sein de la plante hôte. Le génome viral ainsi introduit peut se répliquer et démarrer un cycle infectieux si le reste de son génome est présent au sein d'un hôte susceptible. Ces "clones infectieux" se composent, dans le cas des nanovirus, de huit clones distincts d'*A. tumefaciens*, chacun contenant l'un des huit segments génomiques constituant le génome complet des nanovirus. En injectant cette population de clones dans les tissus d'une plante hôte susceptible, nous pouvons ainsi induire des infections systémiques complètes ou partielles (qui peuvent ne pas inclure certains segments génomiques non essentiels) de nanovirus. La deuxième méthode d'inoculation implique l'utilisation de populations de pucerons capables de transmettre les nanovirus. Ces pucerons sont isolés et placés sur des plantes infectées, ce qui les rend « virulifères », leur permettant ainsi de transmettre le virus après quelques jours passés à se nourrir sur des plantes réceptrices. Cette thèse utilise différentes plantes en tant qu'hôtes susceptibles (fèves, lentilles et vesces) et différentes espèces de pucerons en tant que vecteurs (*Acyrthosiphon pisum* et *Aphis craccivora*).

Les acides nucléiques provenant des plantes hôtes et des pucerons vecteurs peuvent être extraits puis soumis à une amplification et à une quantification à l'aide de paires d'amorces, spécifiques pour chaque segment génomique, par le biais de la PCR quantitative (qPCR). Cette quantification nous permet de déterminer la présence de chaque segment génomique, la charge virale associée, ainsi que la formule génomique pour chaque échantillon.

Enfin, pour observer la présence des segments en question au sein des pucerons, préalablement disséqués et chimiquement fixés, nous utilisons des techniques d'observation en microscopie confocale avec marquage des segments génomiques par hybridation in-situ en fluorescence (FISH).

4. Objectifs de la thèse

L'objectif principal de cette thèse consiste à mener une caractérisation phénotypique systématique des 16 combinaisons possibles de réassortiments impliquant un unique segment génomique entre deux isolats différents du FBNSV. Au démarrage de ce travail, un seul isolat était disponible sous forme de clone infectieux, **[JKI-2000]** (**Grigoras et al. 2009**). Avant d'entreprendre la caractérisation, il a été décidé de sélectionner puis produire deux nouveaux clones infectieux représentant deux isolats génétiquement distincts, **[AZ;15]** et **[AZ;10_12b]**. La sélection de ces isolats s'est fondée sur plusieurs critères, notamment leur présence dans la même zone géographique, leur appartenance à des clades phylogénétiques distinctes, et enfin le fait qu'ils aient été échantillonnés à partir d'hôtes différents. Nous avons utilisé une analyse phylogénétique du genre *Nanovirus* ainsi que les informations associées à l'échantillonnage et au séquençage de divers isolats du FBNSV (**Grigoras et al., 2014)** pour faire notre choix. Deux clones infectieux ont par la suite été générés à partir des séquences génomiques correspondantes obtenues dans les bases de données. Ces clones ont été soumis dans un premier temps à une caractérisation phénotypique avec le clone **[JKI-2000]** afin d'établir les fondements de l'étude des réassortiments. En parallèle de ces investigations, nous avons examiné comment les contraintes moléculaires associées aux réassortiments chez les virus multipartites à ADN simple brin pouvaient influer sur la gestion des coûts associés au maintien de l'intégrité génomique lors de la transmission inter-hôte par un travail approfondi de synthèse bibliographique.

5. Organisation du manuscrit

Ce manuscrit de thèse est structuré en trois parties principales. En complément de l'introduction initiale du contexte scientifique et de la thématique générale de la thèse, une première étude réalisée dans le cadre de ce travail de thèse a mis en évidence la capacité du FBNSV à reconstituer son génome par la transmission non-concomitante d'infections partielles possédant des ensembles de segments génomiques complémentaires. Cette aptitude, en plus de réduire considérablement les rsiques de pertes de segments associés à la transmission horizontale inter-hôte, accroît les potentialités de réassortiment de ces virus. En complément de résultats précédents obtenus au sein de l'équipe (**Sicard et al. 2019**), cette démonstration indique que ces virus peuvent réassortir sans nécessairement que les génotypes parentaux n'infectent la même cellule, ni même la même plante hôte, ou soient transmis par le même individu d'insecte vecteur. Cela constitue une différence majeure en comparaison des virus segmentés, qui exigent la co-infection de la même cellule par deux génotypes distincts pour pouvoir réassortir, puisque tous les segments sont encapsidés dans la même particule virale.

La première section de ce manuscrit est centrée sur la rédaction d'une revue exhaustive des contraintes moléculaires liées aux réassortiments chez les virus multipartites à ADN simple brin afin d'évaluer si cette faculté étendue de réassortir démontrée par notre équipe de recherche est cohérente avec l'ensemble des données disponibles dans la littérature. À la suite de l'identification de la trans-réplication des segments réassortants comme la contrainte majeure chez ces virus, une étude génomique du genre *Nanovirus* a été initiée et développée en complément à cette partie du manuscrit. L'objectif final de cette analyse génomique est d'identifier *in silico* des séquences itératives candidates, contrôlant la réplication virale et déterminantes pour la trans-réplication des divers segments génomiques. Les prédictions *in silico* pourront ensuite être validées expérimentalement, comme cela a été fait pour le genre *Babuvirus*, second genre composant la famille *Nanoviridae*. Un attendu supplémentaire de cette analyse génomique est de détecter de nouveaux événements de réassortiments à partir de données plus récemment déposées dans les bases. Ceci permettra de compléter les précédentes études de détection de réassortiments dans les populations naturelles, qui datent de 2014 et 2017 (**Grigoras et al. en 2014; Kraberger et al. en 2017**).

La deuxième section du manuscrit de thèse décrit en détail le choix des deux isolats du FBNSV, **[AZ;15]** et **[AZ;10_12b]**, la création des deux nouveaux clones infectieux correspondants et leur caractérisation phénotypique qui inclue aussi le clone **[JKI-2000]** préalablement disponible. Cette caractérisation a été effectuée sur leurs trois hôtes de terrain respectifs, à

savoir la vesce (*Vicia sativa*), la lentille (*Lens culinaris*) et la fève (*Vicia faba*), ainsi que sur le puceron vecteur *Aphis craccivora*. Cette caractérisation phénotypique approfondie a servi de fondement à l'étude suivante sur les réassortiments.

La troisième partie de ce manuscrit se concentre sur la production expérimentale des 16 combinaisons possibles de réassortants impliquant un unique segment génomique entre les isolats **[AZ;15]** et **[AZ;10_12b]**, suivi de leur caractérisation phénotypique au sein des trois espèces hôtes mentionnées précédemment (la fève, la lentille et la vesce).

Pour conclure, une section de discussion générale synthétise les principales conclusions de ces recherches, suivie par l'exploration des perspectives futures en fonction des résultats obtenus et de travaux en cours qui n'ont pas été finalisés. Le manuscrit se termine par une section contenant des documents annexes supplémentaires liés à des travaux préliminaires ou inachevés ainsi que par une bibliographie générale des références du manuscrit.

6. Résultats principaux et discussion générale

6-1. Reconstitution du génome par transmission non-concomitante

Une première collaboration avec un autre étudiant en thèse, investiguant la transmission horizontale du FBNSV par puceron, a permis de mettre en évidence la capacité des nanovirus à reconstituer un génome complet grâce à la transmission non concomitante des segments génomiques par l'insecte vecteur (**Di Mattia et al. 2022 – Introduction**). Cette démonstration s'est appuyée sur la possibilité de générer des infections partielles qui ne contiennent pas un segment génomique donné (DNA-C, -N ou -U4), mais qui sont néanmoins capables de déclencher des infections systémiques (**Grigoras et al., 2018**) par agro-inoculation.

Deux modes de transmission, séquentielle et parallèle, ont permis la reconstitution des génomes. Un mode de transmission séquentielle : la reconstitution du génome se fait au sein d'un même puceron vecteur ayant acquis séquentiellement des groupes de segments incomplets et complémentaires sur deux plantes infectées. Un mode de transmission parallèle : la reconstitution du génome se fait au sein d'une nouvelle plante receveuse inoculée par deux pucerons vecteurs, chacun apportant son groupe de segments génomiques complémentaires. Cette aptitude diminue considérablement le coût lié à la préservation de l'intégrité génomique lors de la transmission horizontale, car les segments génomiques peuvent être transmis de manière indépendante et ont la capacité de reconstituer un génome complet lorsque les conditions de complémentation le permettent.

Ainsi les nanovirus possèdent la capacité de complémenter les fonctions de leurs différents segments au niveau supra-cellulaire (**Sicard et al., 2019**) et de transmettre leurs segments génomiques de manière non non-concomitante, conduisant à la reconstitution d'un génome complet. Ces deux caractéristiques augmentent considérablement les possibilités de réassortiment pour ces virus. Deux génotypes parentaux sont ainsi en mesure de réaliser des échanges de segments sans avoir besoin d'infecter la même cellule, le même hôte ou le même insecte vecteur. En outre, en fonction de la fréquence et de la viabilité des réassortiments en question, ces échanges génétiques pourraient avoir un impact sur la gestion des coûts associés au maintien de l'intégrité génomique lors de la transmission inter-hôte.

6-2. Réassortiment chez les virus multipartites à ADN simple brin

La revue bibliographique (**Torralba, Blanc and Michalakis 2023 – Chapitre 1**), consacrée à l'étude des contraintes associées aux réassortiments chez les virus multipartites à ADN simple brin et à leur comparaison avec les observations dans les populations naturelles, visait à déterminer si cette amélioration des capacités de réassortiment était en accord avec les observations disponibles. Malgré des données limitées et une probable sous-estimation, la prévalence d'isolats résultant d'événements de réassortiments semble être grande (entre 20 et 70 % chez les begomovirus et 40-50 % chez les nanovirus), ce qui pourrait être cohérent avec nos hypothèses. Néanmoins, il convient de nuancer cette information car, comme pour tout variant génétique, cette fréquence dépend non seulement de la fréquence d'apparition mais aussi de l'effet sur la fitness.

L'évaluation des données expérimentales semblent suggérer que les nanovirus ont une facilité à réaliser des réassortiments entre différentes espèces. Cependant, il est important de noter que très peu de réassortants inter-espèces sont observés dans les populations naturelles, malgré la fréquence élevée contrastante des recombinaisons inter-spécifiques. Cette observation suggère qu'il existe des limitations significatives aux réassortiments inter-espèces, lesquelles méritent une étude plus approfondie pour en comprendre les mécanismes.

Une autre observation concerne le fait que les évènements de réassortiment impliquant directement plus d'un segment génomique sont rares chez les nanovirus. Pourtant, la capacité à générer en laboratoire des infections incomplètes avec l'absence de trois ou quatre segments pourrait suggérer le contraire. Cette observation semble indiquer qu'il existe une perturbation significative des interactions génomiques lorsque plus d'un segment est impliqué dans le réassortiment.

La revue a mis en évidence que la trans-réplication des segments réassortants constitue une contrainte majeure dans la complémentation fonctionnelle du segment par son nouveau contexte génomique chez les begomovirus bipartites et chez les nanovirus. En comparaison, l'encapsidation et le mouvement intra- et inter-hôtes semblent être moins restrictifs. La plupart des virus multipartites à ADN simple brin se répliquent par un mécanisme de Rolling-circle (RCR) qui dépend de la liaison de la protéine Rep à l'origine de réplication de chaque segment génomique, grâce à la reconnaissance de séquences itératives appelées itérons. Cette revue scientifique a également souligné le manque d'identification et de compréhension des compatibilités entre itérons et Rep chez les nanovirus, et dans une moindre mesure chez les begomovirus. Suite à ces observations, l'étude génomique préliminaire du genre *Nanovirus* a conduit à la proposition d'au moins trois séquences itératives candidates, similaires à ce qui a été identifié chez les babuvirus. Cependant, ce travail nécessitera un soutien supplémentaire en analyses *in silico* et en validation expérimentale afin d'être véritablement finalisé.

6-3. Production et caractérisation phénotypique de deux nouveaux clones infectieux du FBNSV

Nos travaux de recherche ont conduit à la création de deux clones infectieux viables à partir de différents isolats du FBNSV, **[AZ;15]** et **[AZ;10_12b]**, portant à trois le nombre de clones infectieux disponibles pour le FBNSV, avec **[JKI-2000]**. L'évaluation phénotypique de ces trois isolats a révélé leur capacité à infecter la fève, la lentille et la vesce, ainsi qu'à être transmis par les pucerons *Aphis craccivora* et *Acyrthosiphon pisum*. Bien que **[JKI-2000]** provienne d'un isolat initialement échantillonné en Éthiopie, contrairement aux isolats **[AZ;15]** et **[AZ;10_12b]** qui coexistent en Azerbaïdjan, cette observation suggère que le spectre d'hôte au sein d'une même espèce, voir au sein du genre nanovirus, est partagé, ce qui devrait favoriser considérablement les échanges génétiques entre les isolats et les espèces partageant une aire géographique commune.

Cette étude (**Torralba et al. 2024a – Chapitre 2**) a révélé des variations de virulence entre les divers isolats, cependant, elle n'a pas établi de corrélation entre la gravité des symptômes et la fitness virale. En effet, les isolats **[AZ;10_12b]** et **[JKI-2000]** ont montré respectivement les meilleures et les plus faibles capacités de transmission, indépendamment de l'hôte étudié et de la sévérité des symptômes associés à leur infection. Ces résultats ont aussi permis de mettre en évidence une partie de la diversité phénotypique au sein de l'espèce FBNSV.

Enfin, la formule génomique associée à chaque isolat a été déterminée dans les trois plantes hôtes ainsi que chez le puceron vecteur *A. craccivora*. Nos recherches mettent en lumière pour la première fois de nombreuses différences significatives entre les formules génomiques des différents isolats, à la fois au sein des plantes hôtes et chez le puceron vecteur. Ainsi, la formule génomique s'avère spécifique de l'hôte, du virus et de son l'isolat.

6-4. Caractérisation phénotypique systématique des réassortants impliquant un seul segment génomique entre deux isolats du FBNSV

En conclusion, la caractérisation phénotypique systématique des 16 réassortants impliquant un seul segment génomique entre **[AZ;15]** et **[AZ;10_12b]** (**Torralba et al. 2024b – Chapitre 3**) a mis en évidence la viabilité de chaque reassortant sur les trois hôtes étudiés. Ce résultat majeur indique que les réassortiments pourraient fréquemment aboutir à la reconstitution de génomes réassortants impliquant un seul segment génomique, ce qui amplifierait la réduction du coût de maintien de l'intégrité génomique observée précédemment lors de la reconstitution génomique par transmission non-concomitante des segments génomiques (**Di Mattia et al. 2022 – Introduction**).

Ces résultats sont en accord avec les observations antérieures faites dans les populations naturelles et la prédominance des réassortants intra-spécifiques impliquant un seul segment qui ne semblent pas être associés à d'importantes contraintes. Toutefois, les segments génomiques ne sont pas également impliqués dans des réassortiments viables détectés au sein des populations naturelles, en particulier le segment C qui n'a encore jamais été observé dans un réassortiment au sein des deux études de détection de réassortants chez les nanovirus. Une analyse approfondie de la compétitivité associée à ces 16 réassortants viables dans un environnement naturel contrôlé pourrait fournir davantage d'informations sur cette disparité. La capacité de survie de ces réassortants, même en environnement exempt de compétition particulière, souligne également la facilité avec laquelle les échanges génétiques peuvent se produire entre deux isolats d'une même espèce, ce qui pourrait avoir des implications majeures sur l'évolution, l'adaptation et la formation de génotypes hybrides.

En outre, cette étude a également révélé la diversité des effets phénotypiques associés aux réassortiments et leur influence parfois importante sur la formule génomique chez le FBNSV. Ces résultats appellent à des analyses statistiques plus approfondies afin d'évaluer un effet adaptatif de la formule génomique et s'il existe des interactions spécifiques entre différents

segments génomiques en examinant les modifications de la formule génomique consécutives aux réassortiments.

7. Perspectives

Ces travaux de thèse ouvrent de nombreuses perspectives d'études des réassortiments mais aussi de la biologie des nanovirus et de leur gestion des couts associés au multipartitisme.

La nouvelle disponibilité de trois isolats génétiquement discriminables du FBNSV permet désormais d'explorer les échanges génétiques tels que la recombinaison et les réassortiments au sein d'isolats d'une même espèce, mais aussi la diversité des interactions biologiques liées au cycle infectieux de ces virus.

La mise en évidence de la capacité des nanovirus à reconstituer un génome complet via la transmission non-concomitante de particules virales par insecte vecteur peut être approfondie afin de déterminer si cette capacité concerne aussi des lots de segments incapables d'établir une infection autonome. Cette hypothèse soulève la possibilité d'une phase « latente » des particules virales qui pourraient ainsi être complémentées *a posteriori* par des infections successives. Cette éventualité aurait des conséquences très importantes sur la réduction du coût associé au maintien de l'intégrité génomique chez ces virus multipartites. Répondre à cette question est désormais possible via la disponibilité de clones infectieux de deux isolats différents et capables de produire de nombreux réassortants viables. Le design d'une expérience correspondante dénommée expérience de « rescue » est décrit au sein du manuscrit (**Annex**).

Poursuivre l'étude génomique en vue de valider l'identification des itérons chez les nanovirus et de répertorier cette diversité au sein du genre serait d'une importance cruciale pour une meilleure compréhension des compatibilités en matière de réassortiments entre espèces et isolats. Cette démarche pourrait renforcer notre aptitude à prédire quelles combinaisons sont compatibles et sont susceptibles de générer des réassortiments viables ou émergents au sein du genre *Nanovirus*.

Pour approfondir notre compréhension des réassortiments chez les nanovirus, il serait pertinent d'explorer les réassortiments inter-espèces impliquant un seul segment. Des clones infectieux sont déjà disponibles pour les espèces *Faba bean necrotic yellows virus* (FBNYV) et *Pea necrotic yellow dwarf virus* (PNYDV), qui partagent des hôtes et pucerons vecteurs communs. De plus, il serait intéressant d'étudier certains réassortants impliquant plusieurs

segments afin de déterminer si des perturbations majeures des interactions génomiques sont véritablement responsables de leur faible prévalence observée dans les populations naturelles.

Afin de mieux simuler un contexte naturel, il serait pertinent de confronter les réassortants qui présentent une meilleure fitness avec le segment homologue correspondant du génotype parental majeur. Cette démarche nous permettrait d'examiner s'il existe des mécanismes favorisant la préservation de l'intégrité génomique, même lorsque des génotypes de réassortiment plus performants sont produits par rapport au génotype parental dominant.

Pour mieux comprendre les interactions entre segments réassortants et contexte génomique, il serait instructif de soumettre les réassortants présentant une réduction de fitness, comparativement au génotype parental majeur, à une évolution expérimentale. Cette démarche nous permettrait de mieux comprendre les mécanismes par lesquels les segments réassortants s'adaptent à leur nouvel environnement génomique. Les détails de cette expérience sont décrits au sein du manuscrit.

Enfin une analyse statistique approfondie des données de phénotypage obtenues au cours de la thèse pourrait nous fournir des informations clefs sur la biologie des nanovirus et sur l'aspect adaptatif de la formule génomique chez les virus multipartites.

LIST OF COMMUNICATIONS AND PUBLICATIONS

September **2021** – Oral communication at international congress - **18ème Rencontres de Virologie Végétale (RVV)** in Aussois, France – Babil TORRALBA, Stéphane BLANC and Yannis MICHALAKIS - **Reassortments in an octopartite plant virus model**.

September **2022** – Oral communication at international congress –**10th International symposium on ssDNA Viruses (IS^{3}DV)** 2021 in Sète, France – Babil TORRALBA, Stéphane BLANC and Yannis MICHALAKIS – **Intra-Specific variations of the genome formula**.

January **2023** – Oral communication at international congress – **19ème Rencontres de Virologie Végétale (RVV)** in Aussois, France – Babil TORRALBA, Stéphane BLANC and Yannis MICHALAKIS - **Intra-Specific variations of the genome formula**.

(Published) Jérémy DI MATTIA, Babil TORRALBA, Michel YVON, Jean-Louis ZEDDAM, Stéphane BLANC and Yannis MICHALAKIS – 2022 - **Nonconcomitant host-to-host transmission of multipartite virus genome segments may lead to complete genome reconstitution**. *Proceedings of the National Academy of Sciences of the United States of America*, DOI: 10.1073/pnas.2201453119

(Published) Babil TORRALBA, Stéphane BLANC and Yannis MICHALAKIS – 2024 **– Reassortments in Single-Stranded DNA Multipartite Viruses: Confronting expectations based on molecular constraints with field observations.** *Virus Evolution*, DOI: 10.1093/ve/veae010

(In preparation) Babil TORRALBA, Andy BROUSSE, Prune LACOTE-POPOVIC, Stéphane BLANC and Yannis MICHALAKIS – 2024 – **Phenotypic characterization of three isolates of the *faba bean necrotic stunt virus* (FBNSV, *Nanovirus*) in multiple host species.**

(In preparation) Babil TORRALBA, Andy BROUSSE, Aleksandra GRESHNOVA, Prune LACOTE-POPOVIC, Stéphane BLANC and Yannis MICHALAKIS – 2024 – **Systematic phenotypic characterization of single-segment reassortants of the *faba bean necrotic stunt virus* (FBNSV, *Nanovirus*) in multiple host species**.

Exploring Reassortments in an Octopartite Phytovirus: A Study of the *Faba Bean Necrotic Stunt Virus* (FBNSV) from the *Nanovirus* Genus

The viral multipartite genomic architecture and packaging strategy, in which the genome is divided in multiple segments separately encapsidated, stands as an evolutionary conundrum associated with theoretical high fitness costs associated to the maintenance of genome integrity. This cost becomes especially acute as the number of genome segments composing the genome increases. Theoretical models, which operate on the assumption that all genomic segments must infect the same cells/hosts to establish an infection and that genomic segments are randomly disseminated, predict an exceedingly high multiplicity of infection (MOI) *i.e.* the number of viral particles necessary to infect a host or a cell, for genomes divided into more than three segments. However, these theoretical predictions do not align with empirical *in vivo* data, which demonstrate significantly lower MOI values in both hosts and vectors. Earlier studies within the research team showed the capacity of the *Faba bean necrotic stunt virus* (**FBNSV**; *Nanovirus* genus, genome composed of 8 segments) to accumulate each segment in distinct host cells and to complement the system at the supra-cellular level by exchanging gene products among interconnected cells, thereby greatly reducing the theorized intra-host necessary MOI.

The aim of this doctoral research was to delve into the mechanisms through which nanoviruses manage to maintain their genome integrity, with a specific focus on reassortments and on their effects on the fitness of resulting hybrid genotypes. A reassortment corresponds to the replacement of one or more segments by homologous ones from a distinct parental genotype. The primary objective was to conduct a systematic comprehensive phenotypic characterization of all 16 possible single-segment reassortants involving two distinct FBNSV isolates.

The objectives were addressed through the utilization of experimental techniques such as agro-inoculation, aphid inoculation, qPCR, phenotyping of infected plants, and confocal microscopy. Our findings revealed the FBNSV capacity to transmit its distinct segments from host-to-host non-concomitantly and to eventually reconstitute a complete genome, thereby significantly diminishing the costs related to the maintenance of their genomic integrity.

These findings, along with the capacity of nanovirus genomic segments to effectively complement each other at the supra-cellular level, significantly expand the spatial scale at which reassortment can occur. A thorough literature review on reassortments in single-stranded DNA multipartite viruses underscored that the replication of reassorted segments appears as the primary constraint for the success of the new generated genotypes. Additionally, it highlighted two notable disparities within nanoviruses: the rarity of inter-specific reassortments in comparison to inter-specific recombination, and the limited prevalence of multiple-segment reassortments relative to single-segment reassortments in natural populations. A follow up genomics investigation was initiated to identify potential replication determinants within the sequences of the segments of species of the *Nanovirus* genus using *in silico* methods, with the aim to identify the required future experimental validation.

This doctoral research has led to the production of two new infectious clones, **[AZ;15]** and **[AZ;10_12b]**, each representing distinct FBNSV isolates. Together with the isolate **[JKI-2000]**, which was already available at the start of the PhD, we now possess three infectious clones representing isolates from the three main FBNSV phylogenetic clades. These clones will facilitate additional investigations of genetic exchanges, mixed infections, and ecological dynamics.

The phenotypic characterization of isolates was conducted across three host plants (faba beans, lentils, and vetches) and one aphid vector (*Aphis craccivora*), who appeared to be infected by all three viral isolates holding significant implications on ecological dynamics of nanovirus natural populations. This analysis revealed phenotypic diversity among isolates, notably regarding their genome formula, across all host plants and aphid vectors.

Finally, the systematic phenotypic characterization of single-segment reassortants between **[AZ;15]** and **[AZ;10_12b]** isolates demonstrated the viability of all 16 possible combinations across three host plants. While these findings hold potential ecological and evolutionary significance for nanoviruses as they confirm that reassortments might reduce the cost to maintain genomic integrity through genome reconstitutions, it is important to emphasize that they were obtained under controlled environment with no competition from other genotypes.

The results from this doctoral research significantly expand the genetic resources and opportunities available for exploring nanovirus biology, reassortment dynamics, and the strategies employed by ssDNA multipartite viruses in general, at least nanoviruses, to mitigate the costs associated with the maintenance of genome integrity.

Key words : Multipartitism, ssDNA, multipartite virus, *Nanovirus*, FBNSV, reassortments, genomics, infectious clones, phenotypic characterization, fitness, recombination, genetic exchange