



**HAL**  
open science

# Fonctionnement des populations de virus multipartites de plantes au cours des différentes étapes de leur cycle de vie

Anne Sicard

► **To cite this version:**

Anne Sicard. Fonctionnement des populations de virus multipartites de plantes au cours des différentes étapes de leur cycle de vie. Biologie végétale. 2014. Français. NNT : . tel-02801967

**HAL Id: tel-02801967**

**<https://hal.inrae.fr/tel-02801967>**

Submitted on 5 Jun 2020

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

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

# THÈSE

Pour obtenir le grade de  
**Docteur**

Délivré par le  
**Centre international d'études supérieures  
en sciences agronomiques  
Montpellier**

Préparée au sein de l'école doctorale SIBAGHE  
Et de l'unité de recherche UMR - BGPI

Spécialité : **Microbiologie/Parasitologie**

Présentée par **Anne SICARD**

**Fonctionnement des populations de virus  
multipartites de plantes au cours des  
différentes étapes de leur cycle de vie**

Soutenue le 12 décembre 2014 devant le jury composé de

Mme Isabelle JUPIN, Directrice de Recherche, CNRS, Paris	Examinatrice
Mme Eugénie HEBRARD, Chercheur, IRD, Montpellier	Examinatrice
M. Esteban DOMINGO, Professeur, Spanish Research Council	Rapporteur
M. Enrique MORIONES, Professeur, Consejo Superior de Investigaciones Cientificas	Rapporteur
M. Stéphane BLANC, Directeur de Recherche, INRA, Montpellier	Dir. de thèse



## Résumé :

Les virus multipartites sont caractérisés par un génome divisé en au moins deux segments, chacun étant encapsidé individuellement. Bien que ces virus représentent près de 30% des genres de virus de plantes connus, les coûts et avantages engendrés par cette organisation génomique restent mal connus. Basée sur l'étude d'un *Nanovirus*, le *Faba bean necrotic stunt virus* (FBNSV), ma thèse vise à mieux comprendre les dynamiques des populations d'un virus multipartite au cours de l'infection de son hôte et de sa transmission par pucerons.

Dans un premier temps, nous avons démontré que chacun des huit segments composant le génome du FBNSV s'accumulait de manière reproductible à une fréquence relative spécifique, certains segments représentant près de 30% de la population virale au sein de la fève et d'autres n'excédant pas 2%. En changeant de plante hôte, nous avons pu montrer que les fréquences relatives des différents segments étaient hôte-spécifiques. Nous avons par ailleurs observé des changements de fréquences relatives au sein des pucerons vecteurs, changements révélant des interactions intimes entre le FBNSV –précédemment perçu comme traversant simplement les barrières cellulaires des pucerons- et son vecteur.

Cette plasticité et l'ajustement correspondant de la composition génétique de la population virale au sein de l'hôte pourrait permettre aux virus multipartites de s'adapter rapidement à un nouvel environnement grâce à un changement spécifique du nombre de copies de ses gènes. Cependant, ces différences de fréquences relatives pourraient également constituer un coût additionnel pour le virus si, comme il est généralement admis, tous les segments ont besoin de pénétrer au sein de la même cellule pour garantir l'infection. En tenant compte des segments rares au sein des populations virales, ce coût apparaît énorme et nous avons donc décidé de vérifier l'existence d'un tel coût expérimentalement.

Une étude de la distribution des différents segments génomiques au sein des cellules végétales est donc en cours et ce, dans le but de regarder si les huit segments sont toujours ensemble au sein des cellules ou s'ils peuvent être physiquement séparés. Cette question est d'importance majeure non seulement en ce qu'elle permet de réévaluer le coût potentiel des virus multipartites mais aussi en ce qu'elle permet de tester le concept du cycle de réplication « cellule-autonome » chez les virus.

Enfin, le maintien de l'intégrité génomique des virus multipartites reposant sur de relativement faibles goulots d'étranglement tout au long du cycle viral, des outils ont été mis en place pour quantifier ces goulots d'étranglement.

**Mots clés :** *Nanoviridae*, *Faba bean necrotic stunt virus*, Nombre de copies de gène, Transmission circulante non-propagative, Multiplicité d'infection cellulaire.

## **Abstract :**

Multipartite viruses are characterized by a genome divided into two or more segments, each encapsidated individually into separated virus particles. Although they represent more than 30% of the known plant virus genera, little is known on how such viral systems actually work and on the advantage(s) and cost(s) of their peculiar genome organization. Based on the study of a *Nanovirus*, the *Faba bean necrotic stunt virus* (FBNSV), my thesis aimed at better understanding the dynamics of a multipartite virus population during within host disease development and during aphid transmission.

We have first demonstrated that each of the eight segments composing the FBNSV genome reproducibly accumulates at a specific relative frequency, some segments representing up to 30 % of the total viral DNA within an infected plant and others not exceeding 2%. By changing the host plant species, we could further show that the relative frequency of the different segments is host-specific. Besides, we observed changes in segment frequencies within the vectors, thus revealing intimate interactions between the FBNSV –which was previously thought to simply traverse cellular barriers- and its vectors.

This plasticity and the corresponding adjustment of the genetic composition of the within host viral population could enable the multipartite virus to rapidly adjust to a novel environment, by specifically changing the copy number of its different genes. However, it could also constitute an additional cost for the virus if, as generally assumed, all segments need to enter the same cell to secure infection. When considering the rare segments in FBNSV populations, the cost appear enormous and we have decided to experimentally address this question.

We are currently looking at the distribution of the different genome segments within individual cells of host plants in order to empirically assess whether all 8 segments are always together within individual cells or whether they can be physically separated. This question is of major importance, not only for re-evaluating the possible costs in multipartite viral systems, but also for challenging the concept of a cell-autonomous replication cycle in viruses.

Finally, as the genomic integrity preservation of multipartite viruses is based on rather small bottlenecks during the whole viral life cycle, tools have been developed to quantify these bottlenecks.

**Keywords :** *Nanoviridae*, *Faba bean necrotic stunt virus*, Gene copy number , Circulative non-propagative transmission, Multiplicity of Infection.

## Laboratoire d'accueil



Cette thèse a été réalisée au sein de l'équipe :

### Interactions Virus Insecte Plante (VIP)

#### Sous la Direction de :

Dr Stéphane BLANC, Directeur de Recherche INRA, Montpellier - Directeur de thèse

Dr Serafin Gutiérrez, Chercheur, CIRAD, Montpellier – Co-Directeur de thèse

#### UMR Biologie et Génétique des Interactions Plante-Parasite (BGPI)

INRA-CIRAD-SupAgro

Campus International de Baillarguet

34398 Montpellier cedex 5

France



## Remerciements

Je souhaite tout d'abord remercier les membres du jury d'avoir accepté de juger mon travail : Mme Eugénie Hébrard, Mme Isabelle Jupin et Messieurs Esteban Domingo, et Enrique Moriones.

Je souhaite également remercier les membres de mon comité de thèse pour leurs conseils : Mme Sylvie Dinant, M. Bruno Gronenborn, M. Michel Lebrun et M. Philippe Roumagnac.

Merci à Philippe Rott et à Claire Neema pour leur accueil au sein de l'UMR BGPI et leurs investissements pour que cette UMR soit un lieu propice au travail et aux échanges.

Je tiens tout particulièrement à remercier Stéphane Blanc d'avoir été un directeur de thèse très présent, à l'écoute et toujours très enthousiaste. Merci à toi Stéphane de m'avoir amenée à découvrir le monde fabuleux des virus multipartites et merci pour ton aide précieuse dans le rush final...

Merci également au Maître Jedi, Serafin Gutierrez, pour toutes les discussions que l'on a pu avoir et pour tous les conseils prodigués tout au long de cette thèse. Merci pour ta rigueur scientifique et tes idées judicieuses.

Je voudrais également remercier Dr T.L.German, mon professeur à l'université de Madison, WI de m'avoir initiée aux joies de la virologie et de m'y avoir fait prendre goût.

Merci à Yannis Michalakis pour son implication dans ce projet et pour toutes ces réunions très fructueuses tout au long de ma thèse. Mais pourquoi diable avoir choisi un nanovirus ?...

Un grand merci à tous les membres de l'équipe VIP, cette grande famille au sein de laquelle il est bon de grandir...

Je souhaite plus particulièrement remercier Michel Yvon pour son aide dans ce projet et pour sa constante bonne humeur, travailler à tes côtés allège tellement la tâche.

Merci à Marie-Stéphanie Vernerey pour toute son aide en microscopie, son investissement dans le projet, pour les joggings matinaux et sa gentillesse.

Merci à Daniel Gargani pour toutes les « belles images » émises.

Manue, cette partie est pour toi... Merci pour ton humour et ta joie de vivre.

Merci à Aurélie Bak, Déborah Conflon et Zineb Belabess pour tous ces bons moments passés ensemble.

Merci à Sophie Le Blaye et à Jean-Luc Macia pour leur aide et leur réactivité.

Merci à Jean-Louis Zeddami pour ses bonnes idées et à Maryline Uzest pour sa gentillesse et son aide.

Merci à « l'équipe gémini » pour tous les échanges techniques et conceptuels que nous avons pu avoir notamment au cours des réunions du lundi matin.

Je voudrais également remercier les membres de l'UMR BGPI pour tous les bons moments passés et tout particulièrement Katia Bonnemeyre.

Un grand merci à tous mes amis qui m'ont beaucoup soutenue tout au long de cette thèse et qui m'ont permis de passer des weekends formidables aux quatre coins de la France. Je vous attends tous en Californie !

Mes pensées se tournent évidemment vers ma famille qui m'a beaucoup encouragée et aidée à en arriver là. Je souhaite tout particulièrement remercier mes parents qui savent toujours trouver les mots justes et mes sœurs qui malgré la distance savent toujours se montrer présentes.

Enfin, merci à Marie-Steph et à Michel pour leur accueil chaleureux dans des endroits merveilleux lors de la rédaction de ce manuscrit.

Merci à Jérémie pour son aide dans l'édition de ce manuscrit et merci au comité de relecture, Hélène, David, Manuella et Déborah.



# Table des matières

Laboratoire d'accueil.....	4
Remerciements .....	5
Table des matières .....	7
Liste des figures et tableaux.....	9
Liste des abréviations.....	10
Introduction générale.....	12
Introduction bibliographique .....	17
1. Review : The strange lifestyle of multipartite viruses .....	18
2. Le <i>Faba bean necrotic stunt virus</i> , un modèle pour l'étude des virus multipartites.....	49
2.1. Les nanovirus, des virus multipartites au génome segmenté « à l'extrême ».....	49
2.2. Répartition géographique et importance agronomique des Nanoviridae.....	49
2.3. Le <i>Faba bean necrotic stunt virus</i> comme modèle d'étude .....	50
2.4. Transmission du FBNSV .....	51
2.5. Information génétique du FBNSV.....	52
2.6. Fonction des 8 protéines du FBNSV et importance de celles-ci dans l'infection de la fève .....	54
2.7. Réplication du FBNSV .....	55
2.8. Evolution du FBNSV .....	56
Chapitre I : Le nombre de copies de gènes est régulé différemment chez un virus multipartite..	59
1. Contexte et objectifs .....	59
2. Article 1 : Gene copy number is differentially regulated in a multipartite virus.....	62
3. Complément à l'article 1 .....	75
3.1. Essai de mise en place de l'évolution expérimentale du FBNSV.....	75
3.2. Protocole .....	76
3.3. Résultats et discussion .....	79
Chapitre II : La transmission circulante non-propagative des nanovirus : une vision simplificatrice .....	84
1. Contexte et objectif.....	84
2. Circulative non-propagative aphid-transmission of nanoviruses : an oversimplified view (Article 2).....	88
3. Conclusion .....	110
Chapitre III : Les virus multipartites : un fonctionnement décentralisé.....	113
1. Introduction.....	113
2. Matériel et méthodes.....	116
3. Résultats .....	124
4. Discussion.....	141

5. Perspectives à plus long terme.....	144
Chapitre IV : Insertion des marqueurs génétiques neutres dans le clone infectieux du FBNSV pour l'estimation des goulots d'étranglement au cours de son cycle de vie .....	147
1. Introduction.....	147
2. Matériel et méthodes.....	152
2.1. Construction des clones « segment-tige boucle » .....	152
2.2. Approche 2 : Construction de clones en tandem avec marqueurs génétiques .....	158
3. Résultats .....	163
4. Conclusion .....	165
Conclusion générale .....	168
Annexes .....	174
1. Annexe I : Purification du FBNSV.....	175
2. Annexe II : Calcul de la probabilité de perdre au moins un segment au cours de l'infection de la fève.....	176
3. Annexe III : Phloème.....	177
4. Annexe IV : Carte des plasmides utilisés au cours de cette thèse .....	178
Références bibliographiques.....	182

## Liste des figures et tableaux

Figure 1 : Schematic representation of plant virus families and genera.....	21
Figure 2: Genome formula of FBNSV in two different host plant species .....	27
Figure 3: Comparison of the genome or segment sizes in viruses.....	31
Figure 4: Relationships among putative 30K MP superfamily members determined by bootstrapped parsimony.....	34
Figure 5: Répartition géographique mondiale des nanovirus (jaune) et des babuvirus (rouge) .....	50
Figure 6: Symptômes causés par le Faba bean necrotic stunt virus sur la fève.....	51
Figure 7: Organisation génomique du FBNSV .....	53
Figure 8: Particules virales de Faba bean necrotic stunt virus .....	53
Figure 9: Réplication des virus à ADN simple brin selon le mécanisme du cercle roulant .....	56
Figure 10 : Evolution expérimentale de la formule génomique du FBNSV .....	77
Figure 11 : Mesure de l'intensité de fluorescence de chaque segment dans une cellule coïnfectée par les segments S (en rouge) et U4 (en vert), 29 jours post-infection avec le logiciel ZEN 2009 (ZEISS)	123
Figure 12 : Spécificité des sondes S-oligonucléotidique (ON-S, A) , S-random priming (RP-S, B), R-random priming (RP-R, C), U4-random priming (RP-U4, D), R1-random priming (RP-R1, E) et R2-random priming (RP-R2, F).....	125
Figure 13: Hybridation fluorescente in situ (FISH) sur faisceaux conducteurs de pétioles de Vicia faba .....	128
Figure 14: Effets de l'infection du FBNSV sur les tissus phloémiens de pétiole de fève 29 jours post-inoculation.....	130
Figure 15 : Localisation des segments S et U4 du FBNSV dans les cellules de pétioles de fève. ....	132
Figure 16: Localisation des segments R et S du FBNSV dans les cellules de pétioles de fève.....	134
Figure 17: Localisation des segments R et S dans les cellules de pétioles de fève après traitement RNase. ....	136
Figure 18 : Colocalisation des deux moitiés du segment R du FBNSV au sein du phloème des pétioles de fèves, 29 jours post-inoculation .....	138
Figure 19 : Localisation du segment R et de son produit d'expression dans des pétioles de fève infectée.....	140
Figure 20 : Etapes de construction des clones pCambia2300-segment-tige boucle.....	153
Figure 21 : Alignement des séquences non-codantes des huit segments du Faba bean necrotic stunt virus mettant en évidence la région commune de la tige-boucle (encadrée en noir).....	155
Figure 22 : Etapes de la construction de clones des segments tandem avec marqueurs génétiques	158
Figure 23 : Zones d'insertion des marqueurs au sein des segments S et N du FBNSV. ....	161
Figure 24 : Localisation du phloème au sein d'un pétiole de feuille de fève saine.....	177
Figure 25 : Schéma représentatif des éléments du phloème .....	177
Figure 26 : Représentation schématique des étapes de différenciation des tubes criblés .....	177
Figure 27 : Carte du plasmide pLitmus 28.....	178
Figure 28 : Carte du plasmide pCambia2300 .....	179
Figure 29 : Carte du plasmide pBin19 .....	180
Tableau 1 : Amorces utilisées en PCR pour amplifier les régions codantes de chaque segment .....	119
Tableau 2 : Séquences des amorces utilisées pour l'obtention des « megaprimers » .....	155
Tableau 3: Séquences des marqueurs insérés en différentes positions des segments S et N.....	160

## Liste des abréviations

### Abréviations Virales

Acronyme	Nom	Genre	Famille
BBTV	Banana bunchy top virus	Babuvirus	Nanoviridae
BMV	Brome mosaic virus	Bromovirus	Bromoviridae
CaLCuV	Cabbage leaf curl virus	Begomovirus	Geminiviridae
CaMV	Cauliflower mosaic virus	Caulimovirus	Caulimoviridae
CCMV	Cowpea chlorotic mottle virus	Bromovirus	Bromoviridae
CMV	Cucumber mosaic virus	Cucumovirus	Bromoviridae
CPMV	Cowpea mosaic virus	Comovirus	Secoviridae
FBNSV	Faba bean necrotic stunt virus	Nanovirus	Nanoviridae
FBNYV	Faba bean necrotic yellows virus	Nanovirus	Nanoviridae
FLUV	Influenza virus	Influenzavirus A	Orthomyxoviridae
FMDV	Foot and mouse disease virus	Aphthovirus	Picornaviridae
PVY	Potato virus Y	Potyvirus	Potyviridae
RCNMV	Red clover necrotic mosaic virus	Dianthovirus	Tombusviridae
SBWMV	Soil-borne wheat mosaic virus	Furovirus	Virgaviridae
SSCSV	Subterranean clover stunt virus	Nanovirus	Nanoviridae
TAV	Tomato aspermy virus	Cucumovirus	Bromoviridae
TBSV	Tomato bushy stunt virus	Tombusvirus	Tombusviridae
TRV	Tobacco rattle virus	Tobravirus	Virgaviridae
TSWV	Tomato spotted wilt virus	Tospovirus	Bunyaviridae
TUMV	Turnip mosaic virus	Potyvirus	Potyviridae
TYLCV	Tomato yellow leaf curl virus	Begomovirus	Geminiviridae

### Abréviations courantes

CR-M	Major common region
CR-SL	Common Region-Stem Loop
FISH	Fluorescence in situ hybridization
HC	Helper component
GCN	Gene copy number
MOI	Multiplicity of infection
qPCR	quantitative PCR

# **INTRODUCTION GENERALE**

## Introduction générale

---

La nature, le fonctionnement et l'évolution des génomes viraux inspirent de nombreux biologistes depuis près de deux siècles (Zaitlin and Palukaitis, 2000). Assez tôt dans l'histoire de la découverte des virus, une architecture génomique particulièrement fascinante a été mise à jour ; il s'agit de celle qui caractérise les virus à composants multiples ou virus « multipartites ». Ces virus représentent un groupe polyphylétique dans lequel le génome est segmenté en deux à huit molécules d'acide nucléique, et dans lequel chacun des segments est encapsidé individuellement dans une particule virale. Les virus multipartites sont très fréquemment rencontrés chez les virus infectant les plantes –puisque'ils constituent plus d'un tiers des genres et famille de virus de plante connus- et sont responsables de maladies dévastatrices sur des cultures d'importance majeure dans le monde entier. Les virus multipartites ont été aussi reportés chez les mycovirus mais, pour des raisons qui restent à ce jour totalement énigmatiques, ils n'ont jamais été décrits chez les bactériophages et pas non plus chez les animaux, exceptés chez de très rares espèces virales bipartites infectant les insectes. Vu les nombreux amalgames trouvés dans la littérature, il est important de souligner que, contrairement aux virus multipartites, les virus segmentés encapsident leurs différents segments génomiques au sein de la même particule virale. Les virus segmentés peuvent infecter à la fois les bactéries, les champignons, les plantes et les animaux.

L'origine des virus multipartites intrigue depuis des décennies les évolutionnistes et virologistes. Deux scénarios principaux ont ainsi été proposés pour expliquer l'évolution de cette architecture génomique. Dans le premier scénario, les virus multipartites seraient issus de la compétition entre un virus monopartite « parental » (ou ancestral) et des virus défectifs interférents produits au cours de leur réplication, ces derniers pouvant être fonctionnels par complémentarité. Ils pourraient ainsi être sélectionnés comme un groupe de plusieurs segments interdépendants dans le contexte d'une forte multiplicité d'infection cellulaire (MOI) (Garcia-Arriaza et al., 2004; Iranzo and Manrubia, 2012; Nee, 1987; Simon-Loriere and Holmes, 2011). La production, le maintien et même la fixation dans la population virale de deux génomes défectifs complémentaires lors de passages en séries effectués à forte MOI en cultures cellulaires semblent corroborer ce scénario (Garcia-Arriaza et al., 2004; Ojosnegros et al., 2011). La co-infection d'une cellule par au moins deux virus distincts, qui auraient ensuite évolué pour fonctionner ensemble par complémentarité, constitue le second scénario proposé

(Roossinck, 1997; Simon-Loriere and Holmes, 2011). L'exemple du Pea enation mosaic virus, virus représentant une symbiose obligatoire entre un composant apparenté aux luteovirus et un composant assigné au genre Umbravirus, pourrait étayer cette hypothèse. Cependant, l'indépendance de réplication de ces deux composants ainsi que l'absence d'homologie de séquence entre les régions terminales 3' et 5' des ARN-2 et ARN-1 -une condition obligatoire pour la reconnaissance de la matrice dans la réplication de tous les virus multipartites à ARN- ne permettent pas de classer ce virus parmi les virus multipartites (de Zoeten and Skaf, 2001). Si chacun des deux scénarios pourrait être à l'origine des virus multipartites constitués d'un petit nombre de segments, l'origine des virus composés d'un nombre beaucoup plus élevé de segments (>4) semble plus complexe. Une étude théorique récente menée par Iranzo et Manrubia (Iranzo and Manrubia, 2012) a montré que le nombre de particules virales qui doivent pénétrer et infecter efficacement chaque cellule (multiplicité d'infection cellulaire ou MOI) pour qu'un virus multipartite composé de plus de quatre segments puisse être fixé, était près de dix fois supérieur aux valeurs de MOI reportées jusqu'à présent. Les virus pour lesquels le nombre de segments génomiques est supérieur à 4 pourraient peut-être résulter de la conjonction des deux scénarios précédents : la segmentation d'un virus monopartite à laquelle se serait ajoutée ultérieurement une capture de gènes lors de co-infections par d'autres espèces virales (Iranzo and Manrubia, 2012).

Le premier scénario, celui de la coopération des génomes défectifs, repose sur l'existence d'avantages sélectifs associés à la segmentation du génome. De tels avantages potentiels ont été imaginés et étudiés quasi exclusivement par des approches de modélisation mathématique. A ce jour, les avantages proposés sont : (i) une fidélité de réplication accrue puisque les segments d'acide nucléique courts représentent des cibles plus petites pour la mutation (Pressing and Reaney, 1984) ; (ii) une durée de synthèse plus courte donc une réplication plus rapide pour les segments courts (Nee, 1987; Spiegelman et al., 1975) ; (iii) une plus grande stabilité des particules virales qui encapsident des segments plus courts (Ojosnegros et al., 2011) ; (iv) une facilitation des échanges génétiques intra génome par un échange de modules d'information que représente chaque segment, sans nécessité de crossing-over entre les acides nucléiques (Chao, 1988, 1991, 1992) et enfin (v) une réplication et transcription indépendantes permettant aux virus à génome segmenté de maintenir un cycle de réplication au cours duquel les segments peuvent être exprimés en quantités différentes (Wagner 2008). Quels que soient les avantages proposés jusqu'à ce jour, et quels que soient les scénarios envisagés pour expliquer l'origine des virus multipartites, tous reposent sur une nécessité commune : la

multiplicité d'infection cellulaire (MOI) doit être très élevée pour permettre une complémentation fonctionnelle entre les divers segments génomiques. Cette nécessité de complémentation fonctionnelle et la forte MOI requise sont unanimement perçues et modélisées comme le coût majeur, pour ne pas dire unique, que doivent supporter ou surmonter ces systèmes viraux. En effet, une vision communément partagée dans tous les articles traitant de l'évolution de ces virus, et dans la communauté des virologistes en général, est que tous les segments génomiques composant une espèce donnée doivent pénétrer au sein de la même cellule pour qu'il y ait complémentation fonctionnelle et donc succès de l'infection, et/ou pour éviter la perte d'une partie de l'information.

Il est très surprenant de noter que l'ensemble des études qui s'intéressent spécifiquement au fait que ces virus soient multipartites, s'y intéressent pour comprendre le « pourquoi » de leur existence mais jamais le « comment » ces systèmes fonctionnent réellement. Les études citées ci-dessus cherchent toutes à confronter le ratio coût/bénéfice, la plupart au travers d'approches purement théoriques, et très peu voire aucune démarche expérimentale simple et sans *a priori* ne cherche à décrire le fonctionnement de ces systèmes viraux (hormis des études purement moléculaires qui n'informent pas sur les propriétés du système). Cet état de fait est analysé dans un article de revue en cours de préparation, dont la version actuelle est proposée dans la partie I de cette introduction.

Vu cette carence en études expérimentales s'adressant au fonctionnement et aux propriétés des systèmes viraux multipartites, ma thèse avait un objectif simple qui était d'élargir le champ de connaissances sur la biologie de ces virus. Pour ce faire, le *Faba bean necrotic stunt virus* (FBNSV), appartenant à la famille des *Nanoviridae*, a été choisi comme modèle d'étude. Ce virus représente un des cas les plus extrêmes de virus multipartite puisque son génome est constitué de huit segments génomiques, le nombre le plus élevé décrit à ce jour. Les différentes caractéristiques du FBNSV sont présentées dans la partie II de l'introduction.

Dans un premier temps, nous avons démontré que chaque segment génomique du FBNSV s'accumule de manière très reproductible à une fréquence relative spécifique (et différente pour chaque segment) au sein des populations virales *in planta*, certains segments étant très fréquents et d'autres très rares. Nous avons aussi pu observer que le pattern des fréquences des segments était spécifique à l'hôte. Un nouvel avantage tenant compte de ces données a alors été proposé pour la segmentation et la multi-encapsidation des génomes viraux: la possibilité de réguler différenciellement le nombre de copies des gènes (gene copy number, GCN), un



paramètre régulant l'expression génique et phénotypique chez tous les organismes où il a été étudié (Hastings et al., 2009; Stranger et al., 2007). Le changement de fréquences relatives des segments en fonction de l'environnement dans lequel le virus se trouve pourrait être à l'origine d'une plasticité de ces systèmes viraux, leur conférant un niveau d'ajustement supplémentaire à un environnement changeant, sans impliquer la sélection immédiate de mutation adaptative. Cette partie importante de mon travail de thèse est décrite sous la forme d'un article publié, dans le Chapitre I.

Par ailleurs, les différences de fréquences relatives de chacun des segments, et le fait que ces fréquences soient hôte-dépendantes, nous ont amenés à tester l'impact du passage à l'intérieur des pucerons vecteurs du FBNSV sur sa composition génétique. Nos résultats remettent en question une vision largement partagée quant au mode de transmission « circulant non-propagatif » des *Nanoviridae* (Chapitre II). En effet, ce mode de transmission invoque un simple passage de particules virales au travers des barrières cellulaires de leur insecte vecteur, sans décapsidation, réplication, ni même transcription. Les résultats obtenus montrent un changement des fréquences relatives des segments du FBNSV au sein des pucerons, qui ne peut que résulter d'une interaction intime plus complexe entre le *Faba bean necrotic stunt virus* et son insecte vecteur. Cette partie constitue le Chapitre II de ma thèse, et est présentée sous la forme d'un article qui est actuellement soumis.

Que ce soit dans les plantes hôtes ou dans les pucerons vecteurs, de telles différences de fréquences relatives entre les segments peuvent également augmenter le coût généralement associé à ces systèmes viraux. En effet, si toutes les particules virales se ressemblent quel que soit le segment qu'elles contiennent, l'entrée aléatoire des segments génomiques au sein des cellules pourrait être à l'origine de la perte des segments plus rares, sauf à imaginer une MOI qui devient totalement démesurée au cours de l'infection.

L'existence réelle de ce coût n'a, à notre connaissance, jamais été testée expérimentalement. Le chapitre III vise à pallier ce manque et ainsi, tout simplement, à vérifier si les différents segments ont toujours besoin de coexister au sein des cellules pour que le système viral puisse être fonctionnel.

Enfin, le maintien de l'intégrité génomique des virus multipartites est basé sur la nécessité que les goulots d'étranglement soient relativement faibles au cours de l'infection systémique de la plante, du passage de feuille à feuille, et aussi au moment de la transmission d'une plante à une autre par pucerons. Nous avons développé des outils destinés à quantifier ces goulots

d'étranglement, outils qui seront présentés au sein du chapitre IV bien que ce travail soit encore très préliminaire.

**REVIEW: THE STRANGE LIFESTYLE  
OF MULTIPARTITE VIRUSES**

# 1. Review : The strange lifestyle of multipartite viruses

## **Préambule**

Depuis la revue de Jaspars (1974) relatant la découverte des virus multipartites dans les années 1960, aucune autre synthèse n'a été publiée sur le sujet. Pourtant, de nombreux travaux ont été menés sur ces virus et peuvent être regroupés en trois grandes catégories (i) ceux sur les génomes, leurs caractéristiques, les fonctions des différents segments, et des différentes protéines dans le cycle viral, (ii) ceux traitant de l'évolution, et (iii) ceux plus théoriques discutant des avantages et coûts de ces systèmes viraux.

Nous avons donc décidé d'écrire une synthèse bibliographique (destinée à être publiée dans les mois à venir) confrontant ces différentes données. Cette synthèse vise à mettre en exergue les principales caractéristiques qui pourraient être spécifiques aux virus multipartites et, qui pourraient *in fine* représenter une base biologique pour expliquer les avantages et les inconvénients de cette architecture génomique particulière.

Cette synthèse regroupe et discute les résultats de différentes équipes dans le monde, y compris ceux obtenus lors de ma thèse, mais elle est aussi influencée par les nombreuses discussions, interprétations et questions que nous avons menées au laboratoire durant ces trois années. En conséquence, elle introduit d'inévitables redondances avec d'autres chapitres de ce mémoire. Nous avons néanmoins choisi de l'insérer en guise d'introduction bibliographique.

La version présentée peut être considérée comme une première version comportant les éléments qui nous paraissent importants ou pertinents. Elle sera *in fine* certainement modifiée et peut être formatée sous forme de revue plus succincte. Quoiqu'il en soit, elle reflète notre vision actuelle des virus multipartites et vise à dégager les principaux axes de recherche qui permettront de mieux appréhender leur fonctionnement.

## Introduction

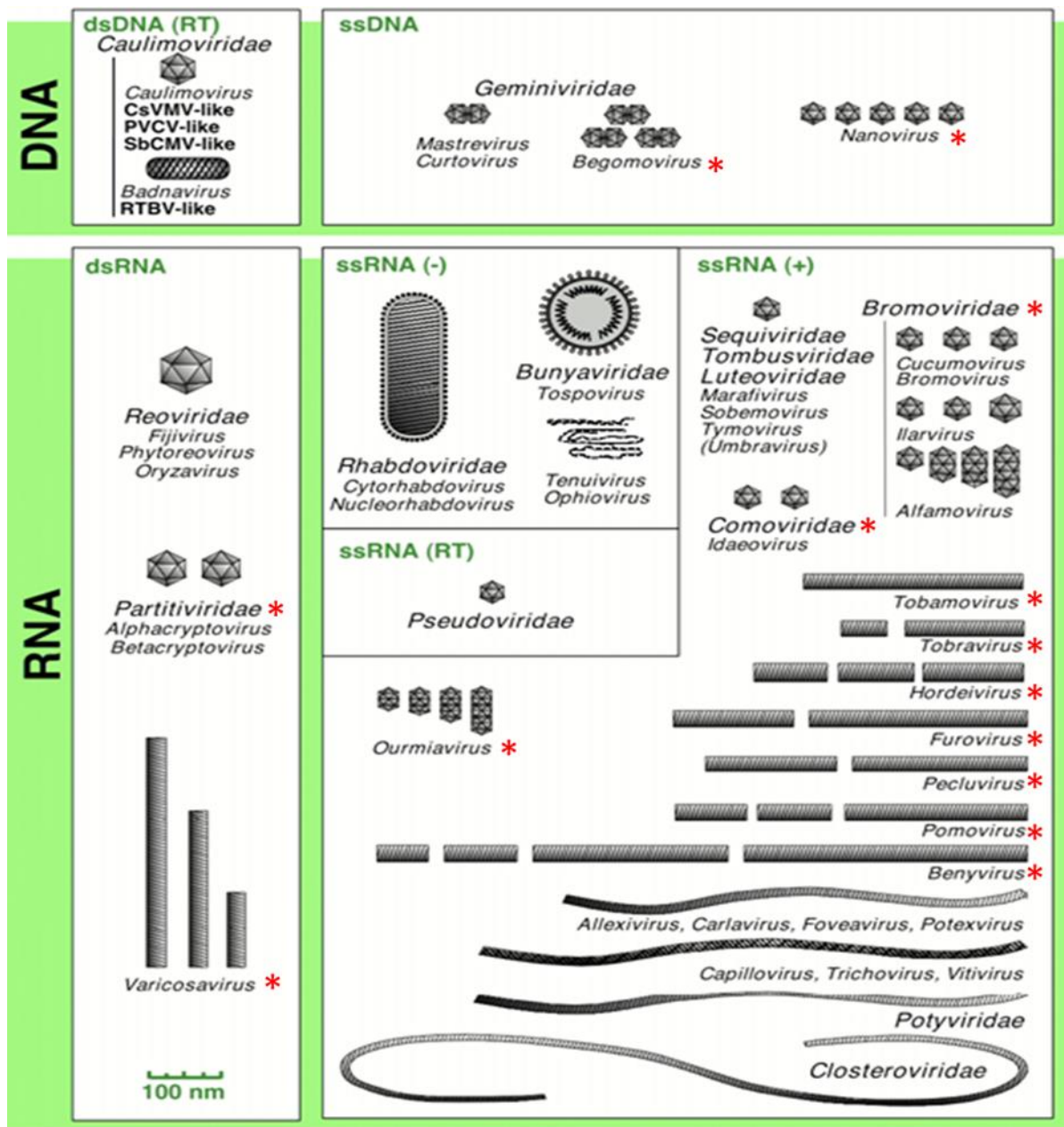
The general architecture, organization, and packaging of viral genetic information can be divided into three distinct categories, here distinguished under the terms “monopartite”, “segmented” and “multipartite” viruses. The monopartite type is a virus with all genetic information encoded by a single nucleic acid molecule protected in a shell made of proteins (and sometimes also lipids), forming the virus particle. The segmented type here refers to cases where the viral genome is divided into two or more nucleic acid segments that are all encapsidated together in a single virus particle. Finally, the multipartite viruses (the terms multicomponent viruses or coviruses are also used in the literature) are species where the genome is divided into two or more nucleic acid segments, just as the segmented type, but where each of these segments is individually packaged into its own virus particle. This latter peculiar organization is the only one resulting in virus particles that do not contain the entirety of the viral genetic information, and where the co-entry of several virus particles in a new cell or a new host is assumed as mandatory to maintain the integrity and functionality of the viral genome. As demonstrated below, the biology and evolution of multipartite viruses challenge some basic concepts of virology. In fact, at this point, it remains hard (if possible at all) to conceive how multipartite viruses have evolved and can actually be functional. This review is intended to briefly review the successive steps of the life cycle of a virus, trying to extract and highlight empirical or theoretical data specifically relevant for the biology of multipartite viruses. It also tries to identify major gaps and thus future research lines which would allow a better comprehension of the functioning of these peculiar biological systems.

The report of the called “multicomponent viruses” marked an important step in the history of the discovery of viruses. When analytical centrifugation techniques emerged, in particular the centrifugation through density gradient (Brakke, 1951), it was rapidly noted that some viral-like diseases appeared associated to two or more components of different density made of proteins and nucleic acid. The first reported cases were the purification of *Tobacco rattle virus* (Lister, 1968) and of *Cowpea mosaic virus* (Van Kammen, 1968), but at this time the multicomponent nature of such disease-inducing agent could not be understood. Together with earlier dose-related infectivity studies (Bald, 1937, 1950; Fulton, 1962), further development of biochemistry, electron microscopy, and later molecular biology and sequencing, definitely evidenced that many viruses are composed of two or more physically separated particles, each containing a complementary portion of the genetic information (Jaspars, 1975; Hull, 2014; Zaitlin and Palukaitis, 2000). While monopartite and segmented viruses infect all possible

living organisms, multipartite viruses appear extremely successful in plants (and fungi), but not in animals where very rare cases have only been reported in insects (Hayakawa et al., 2000; Ribière et al., 2010; Wang et al., 2007). The success of multipartite viruses in plants, where 30 to 40 % of virus genera and families are multipartite (Hull, 2014), is a long-standing mystery for which a few tentative and questionable answers have been proposed. For example, it had early been speculated that the multipartite architecture of the genome could be related to its RNA nature, and logically proposed that multipartite viruses are so frequent in plants because most plant viruses are RNA viruses (Pressing and Reanney, 1984). However, it is now clear that multipartite viruses are also frequent among plant viruses with a DNA genome. In fact, multipartite viruses can be (+)ssRNA, (-)ssRNA, dsRNA, and ssDNA viruses, their genome size is highly variable, they can form icosahedral, rod-like, or filamentous virus particles, and none of these features demarcates them from monopartite and segmented viruses (Figure 1).

Soon after their discovery, both virologists and evolutionary biologists begun to question and investigate the possible benefits and the counterpart costs in these complex biological systems. Putative benefits proposed and sometimes disputed are mostly related to the smaller size of the genome segments, as opposed to a larger single molecule, which may induce a better tolerance to high mutation rates (Pressing and Reanney, 1984), a faster replication (Nee, 1987, 1988), a facilitated genetic exchange between information modules born by the different genome segments (Chao, 1988, 1991, 1992), or a higher stability of viral particles (Ojosnegros et al., 2011). The counterpart cost opposed to these benefits is the reduced chance to infect new cells (and new hosts) with all components required for functionality and integrity of the viral genome (Chao, 1991; Iranzo and Manrubia, 2012; Nee, 1987; Pressing and Reanney, 1984). This cost is equally assumed in all studies striving to define conditions allowing multipartite viruses to evolve from monopartite ancestors. Contrary to the multiple proposed benefits, the reduced chance of success at infection steps appears so intuitively logical, that it is considered an evidence and has never been questioned. The quantitative importance of the cost has been suggested to be related directly to the number of genome segments, thus the number of virus particles, required to gather the entire genome information when entering a cell or host (Iranzo and Manrubia, 2012).

A main concern in this field of research is that most of the above cited studies are theoretical, and that empirical supports for the various benefits proposed and for the actual existence of the cost are rare, if available at all. When considering the common wisdom on viruses, not only in



**Figure 1 : Schematic representation of plant virus families and genera**

All drawings are approximately scaled, providing a comparative view of the sizes of the virus particles. Only genera most representative of the virus structure are indicated. Families and genera with a red asterisk are those whose member species are multipartite viruses. (N.B.: The genus *Begomovirus* is composed of both monopartite and bipartite viruses).

Reproduced from *M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 30, copyright 2000*

a general public, but also in the wide majority of the scientific community, it appears clearly that a conceptual frame derived from the understanding of monopartite viruses largely dominates. As a consequence, it is quite possible that this frame biases the conception of theoretical models intended to explain the evolution of multipartite viral systems, and hampers the design of experiments that would relevantly address biological processes specifically adapted to the way of life of multipartite viruses.

Here below, we will illustrate step by step the fact that experimental investigations targeting the specifics of the biology of multipartite viruses are extremely scarce, and emphasize what we consider major gaps to be addressed in the near future.

## **1 - Replication**

Multipartite viruses, just as monopartite and segmented viruses, are replicated through a diversity of mechanisms depending partly on their genome nature. Whatever the molecular details or the cellular location of viral replication, two specific benefits have been proposed and modeled to explain the evolution of genome segmentation, thus of segmented and multipartite viruses. The first putative replication-related benefit is that splitting a genome into several smaller segments should engender faster replication (Nee, 1987; Spiegelman et al., 1975). This is based on the trivial fact that, when the speed of the replicase is constant and when this replicase is not a limiting factor, the time required for the duplication of a genome of 10kb is twice longer than that for two segments of 5kb each. Experimental studies looking at the replication kinetics as a general function of genome length seem to corroborate this hypothesis (Mills et al., 1967; Sakai et al., 1999). However, the only study directly comparing a monopartite genome to its bipartite derivative obtained in infected cell cultures could not confirm the expected faster replication of the bipartite variant (Ojosnegros et al., 2011).

This contrasting report calls for more studies comparing near-isogenic monopartite and multipartite viruses, in order to empirically confirm that faster replication can indeed benefit segmented genomes. In addition, while faster replication can be a competitive advantage in simple experimental designs or simple theoretical models, its effect in more realistic host and ecological context calls for possible nuances (see for example Goldhill and Turner, 2014).

The second replication-benefit proposed for segmented genome, is that smaller segments are smaller targets for mutations, and thus result in a higher replication-fidelity (Nee, 1987; Pressing and Reaney, 1984). Initially invoked for RNA viruses replicated by error-prone



RNA-dependent RNA-polymerases (Drake, 1993), this argument could now be extended to ssDNA viruses of plants (Gemini- and nanoviruses) and animals (circo-, denso-, and parvoviruses), where mutation rates comparable to that of RNA viruses have been repeatedly documented (Duffy and Holmes, 2009; Grigoras et al., 2010; López-Bueno et al., 2006; Ren et al., 2013; Sarker et al., 2014; van der Walt et al., 2008; Yang et al., 2014). Different studies comparing the mutation rate in distinct families of RNA viruses have suggested a possible negative correlation between the mutation rate and the genome length (Drake, 1993; Sanjuan et al., 2010), suggesting that smaller genomes or segments can bear higher mutation frequencies. In contrast, very recently, a comprehensive study has compiled a large dataset of 118 mutation rates, from 91 genes of 28 viral species (Hicks and Duffy, 2014). By testing the relationship between mutation rates, several viral genome properties and ecological factors, the authors concluded that the nature of the target cells is the only good predictor of the viral mutation rates, whereas genome length and genome segmentation are not.

Altogether, it seems reasonable to conclude that the data currently available do not corroborate the hypotheses explaining the evolution of genome segmentation by a faster replication and/or an easier mutational escape of smaller genome segments. In addition, whatever the future outcome of this debate, it cannot specifically explain the separate encapsidation of segments in multipartite viruses.

As opposed to the above arguable benefits, one replication-associated constraint for segmented genomes is the need to bear similar origins of replication and regulatory elements on all segments, in order to be efficiently recognized and processed by the replication complex. The existence of conserved origins of replication among genome segments has been demonstrated for a number of segmented and multipartite viruses, as for example bromoviruses (Ahlquist et al., 1984; Sivakumaran et al., 2000), begomoviruses (Argüello-Astorga et al., 1994; Fontes et al., 1994; Gladfelter et al., 1997) and nanoviruses (Horser et al., 2001; Timchenko et al., 2000).

A deletion/mutation in this region drastically affects the replication efficiency of the corresponding segments and thus, in order to coordinate the various functions distributed onto distinct segments, its concerted evolution is mandatory (Hughes, 2004; Savory and Ramakrishnan, 2014). While the conservation of these regulatory sequences may here appear has a burden for segmented and multipartite viral genomes, it might also promote genetic exchanges, as discussed in the next section.

## 2 - Genetic exchanges

The genetic exchanges in between viral genomes can depend directly on the replication mechanisms. However, it can also be replication-independent, through genome break and repair processes and, most of all, through reassortments in segmented and multipartite viruses.

A reassortment can be defined as an exchange of homologous segments between two related virus isolate or species, thus a form of genetic exchange not involving intra-molecular cross-overs (Chao, 1991). Reassortment had been suggested to substitute for intramolecular recombination in segmented RNA genomes at a time when RNA cross-overs were believed to be rare, if not impossible (Chao, 1988, 1991; Pressing and Reaney, 1984; Van Vloten-Doting and Jaspars, 1977). Thus, for RNA viruses, reassortment was perceived as the major mean to promote genetic exchange and to both purge deleterious mutations and create new advantageous combinations (Chao, 1991).

When empirically evaluating whether reassortants are favored in the real world, the published results appear contradictory. Numerous studies describe the existence of reassortants from the analysis of sequence data sets, and postulate that they play an important role in the evolutionary history of the corresponding viruses. However, investigations explicitly demonstrating a selective advantage or cost in reassortants are few.

Specific reassortants were shown or suggested to have a higher fitness in segmented and multipartite viruses of animals and plants, such as for example *Influenza virus*, *Tomato spotted wilt virus*, *Cucumber mosaic virus* and several nanoviruses (Grigoras et al., 2014; Ince et al., 2013; Kuroda et al., 2005; Lin et al., 2004; Medina and García-Sastre, 2011; Neumann et al., 2009; Pierrugues et al., 2007; Qiu and Moyer, 1999; Roossinck, 2002). In contrast, in plant samples from fields where several related strains of *Cucumber mosaic virus* (CMV) co-circulate, two independent studies concluded that reassortants are much fewer than expected, and thus mainly counter selected (Bonnet et al., 2005; Nouri et al., 2014). Interestingly, even for *Influenza virus* or the plant nanoviruses, which are respectively segmented and multipartite viruses with an elevated number of segments (6 to 8), reassortments seem to be constrained to one or two segments only, suggesting that many other possible combinations do rarely emerge (Brown et al., 1998; Grigoras et al., 2014).

In 1986, Bujarski and colleagues (Bujarski and Kaesberg, 1986) demonstrated that *Brome mosaic virus* (BMV), a monopartite (+)ssRNA plant virus could recombine by intramolecular RNA cross-overs. More recently it became evident that most (+)ssRNA virus have

recombination rates as elevated as DNA viruses (Froissart et al., 2005; Grigoras et al., 2014; Tromas et al., 2014; Urbanowicz et al., 2005), and that intra-segment recombinants are as frequent as inter-segment reassortants in multipartite viruses (Bonnet et al., 2005; Grigoras et al., 2014; Nouri et al., 2014; Stainton et al., 2012; Urbanowicz et al., 2005). From this wealth of new data, it is now evident that both recombination and reassortment promote frequent genetic exchanges in multipartite viruses. How this is affecting the hypothesis that genome segmentation has evolved to allow sex through reassortment is unclear, but the demonstrated ease of intra-molecular recombination may alleviate some of its authority.

The mechanisms of intra-molecular recombination have a priori no reason to differ between monopartite, segmented and multipartite viruses, and are thus not detailed further. In contrast, as mentioned in section 1, the role of the conserved replication origin in distinct segments of the same viral genome deserves mention in this section. During the replication process, this conserved region promotes homology-driven inter-segment recombination.

This has been experimentally forced under high selection pressure for *Brome mosaic virus* (Bujarski and Kaesberg, 1986), and also observed in natural populations for *Nanoviridae* (Grigoras et al., 2014; Stainton et al., 2012). Because these replication origins are sometimes conserved in different species of the same family, homologous-recombination process can paradoxically occur in between heterologous segments originating from distinct species, as previously observed in nanovirus samples collected from the field (Sivakumaran et al., 2000; Timchenko et al., 2000). One could state that such recombination between segments encoding distinct functions is a non-sense. So, the only sound “raison d’être” of recombination events driven by the replication origin appears to be its facilitated exchange between segments of related genomes (Hughes, 2004; Savory and Ramakrishnan, 2014). This might help reassorting segments from distinct strains or species by rapidly matching their replication origin to the viral system into which they incorporate.

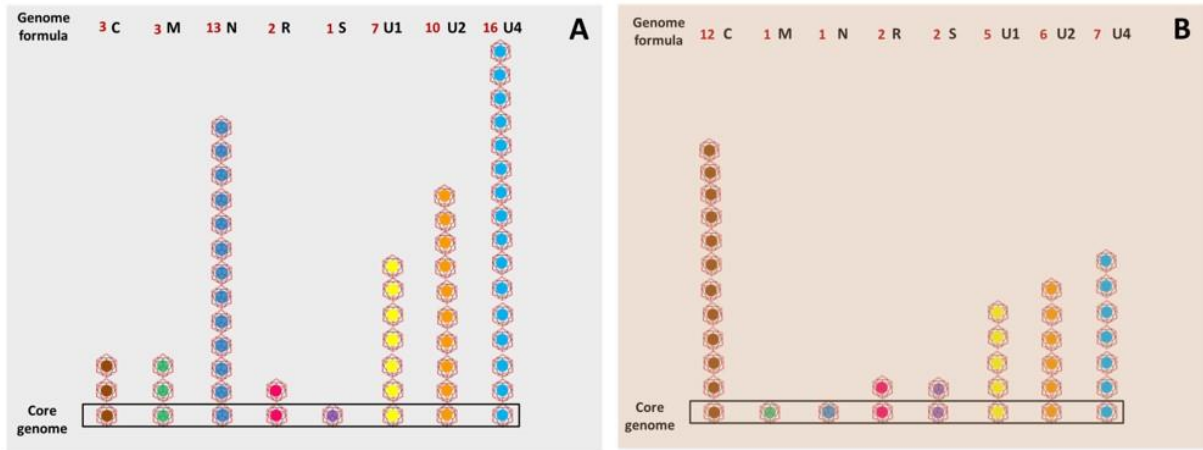
### **3 - Gene expression**

Viruses have been extremely inventive in order to encode all required genes in very small genomes, and to regulate their coordinated expression. In particular, viruses have to deal with cell machinery, which generally limits the translation of mRNAs to only one ORF. Over ten distinct strategies have been described for gene expression of monopartite viruses, and most are also used by segmented and multipartite viruses (Hull, 2014). For example, viral genomes or

segments can encode a single or several genes, these genes can be expressed from a single mRNA or from subgenomic (or subsegment) mRNAs, as a single protein which can sometimes be subsequently cleaved in several products with specific functions. In other cases, the genomes or segments can have internal ribosome entry sites (IRES) allowing several ORF to be translated, leaky scanning sequences allowing polycistronic RNA to produce more than one protein, or strategies to promote reinitiation of translation of several ORFs encoded on a single mRNA.

In front of this complexity, the viral genome segmentation could provide an extreme simplification where each segment would encode a single gene. A long-foreseen advantage of segmented and multipartite viral systems would be that each segment could possess its own specific regulatory sequence (Godefroy-Colburn et al., 1985; Kwon and Chung, 2000; Shirasawa-Seo, 2005; Zagorski, 1978). Surprisingly enough, only two virus groups have evolved this ultimate simplification: the multipartite virus families *Nanoviridae* and *Partitiviridae*. In all other segmented or multipartite virus species, the one-segment/one-gene strategy is either not found or combined with other segments encoding multiple genes. In conclusion, that the genome segmentation could be a simple way to match the eukaryotic gene expression and RNA translation machinery does not seem totally satisfactory, because in most cases the control of gene expression in these viruses is observed to be as complex as in monopartite viruses.

One additional way to regulate gene expression, which is seemingly possible in segmented and multipartite, but not in monopartite viruses, is the differential regulation of gene copy numbers (Sicard et al., 2013). Despite, an important literature on the significant impact of gene copy number variations on gene expression in all cellular organisms (Hollox and Hoh, 2014; Katju and Bergthorsson, 2013; Mileyko et al., 2008), this idea has thus far hardly made its ways into our scientific community (Elde et al., 2012; Filée, 2013). This concept has been so ignored in virology that the relative frequency of the different genome segments had not been explicitly investigated until recently, neither in segmented nor in multipartite viruses. We directly addressed this question in the nanovirus *Faba bean necrotic stunt virus* (FBNSV) (Sicard et al., 2013), and found that each ssDNA segment (so each gene) accumulates reproducibly with a specific relative copy number in a given environment. We proposed that these copy numbers, each associated to a specific segment, define the “genome formula” (Figure 2A), which proved to be specific to the host plant species (Figure 2B). Because viral populations closer to the steady state or setpoint genome formula accumulated to higher levels, we hypothesized that the



**Figure 2: Genome formula of FBNSV in two different host plant species**

The genome formulae presented 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, and the median copy number of each segment is represented here as relative to the less abundant one. The core genome corresponds to the classical conception of the viral FBNSV genome (rectangle). *Adapted from Sicard et al., 2013*

differential regulation of gene copy number might be adaptive, and could stand as an unforeseen benefit for segmented and multipartite viruses. Several earlier hints indicating that multipartite viruses other than nanoviruses could also regulate their gene (or segment) copy number are discussed in (Sicard et al., 2013), and a direct demonstration has been published recently for the tripartite (+)ssRNA *Alfalfa mosaic virus* (Sanchez-Navarro et al., 2013). The genome formula has not been investigated upon infection of a host by a segmented virus, where it could play a role at the intracellular level. While multipartite viruses might control a differential copy number of their different genes (or segments) at cell or host entry, segmented viruses appear constrained at these specific steps by the fact that in most cases one copy of each segment is encapsidated in every single virus particles (McDonald and Patton, 2011; Poranen and Tuma, 2004; Rager et al., 2002; Reguera et al., 2013; Wichgers Schreur et al., 2014).

This discovery might represent a significant step forward in the understanding of the biology of multipartite viruses for several reasons: (i) it represents an unprecedented putative advantage for the regulation of gene expression in segmented viral genomes, (ii) this advantage applies best to multipartite viruses because GCN can be regulated at all viral step, whereas it could mainly be regulated at the level of individual cells for segmented virus, (iii) as discussed later, this advantage is the only one described thus far that imposes a constraint on the relative frequency of the segments, and that can explain that these viral systems have not evolved to the situation of the minimum cost where all segments would ideally accumulate at equal frequency.

#### **4-Encapsidation**

This step of the virus life cycle is really key in this review because it markedly distinguishes the segmented from the multipartite viruses. These are seemingly opposite strategies where in the former the virus obviously tries to have its whole genetic information travelling together within plant and from plant to plant, whereas in the latter it apparently tries not.

Segmented viruses have a very specific constraint at encapsidation, which is the sorting of distinct segments to ensure that at least one of each is present in every single particle. The molecular means by which this process is successfully accomplished are partly understood. Best-documented cases are those of *Influenza virus* or the phage Phi6. The segments bear different and complementary packaging signal sequences, which induce specific secondary RNA folding and a timely concerted interaction with the structural protein ensuring the sorting of one of each per virion (Chou et al., 2012; Mindich, 2004; Noda and Kawaoka, 2012).

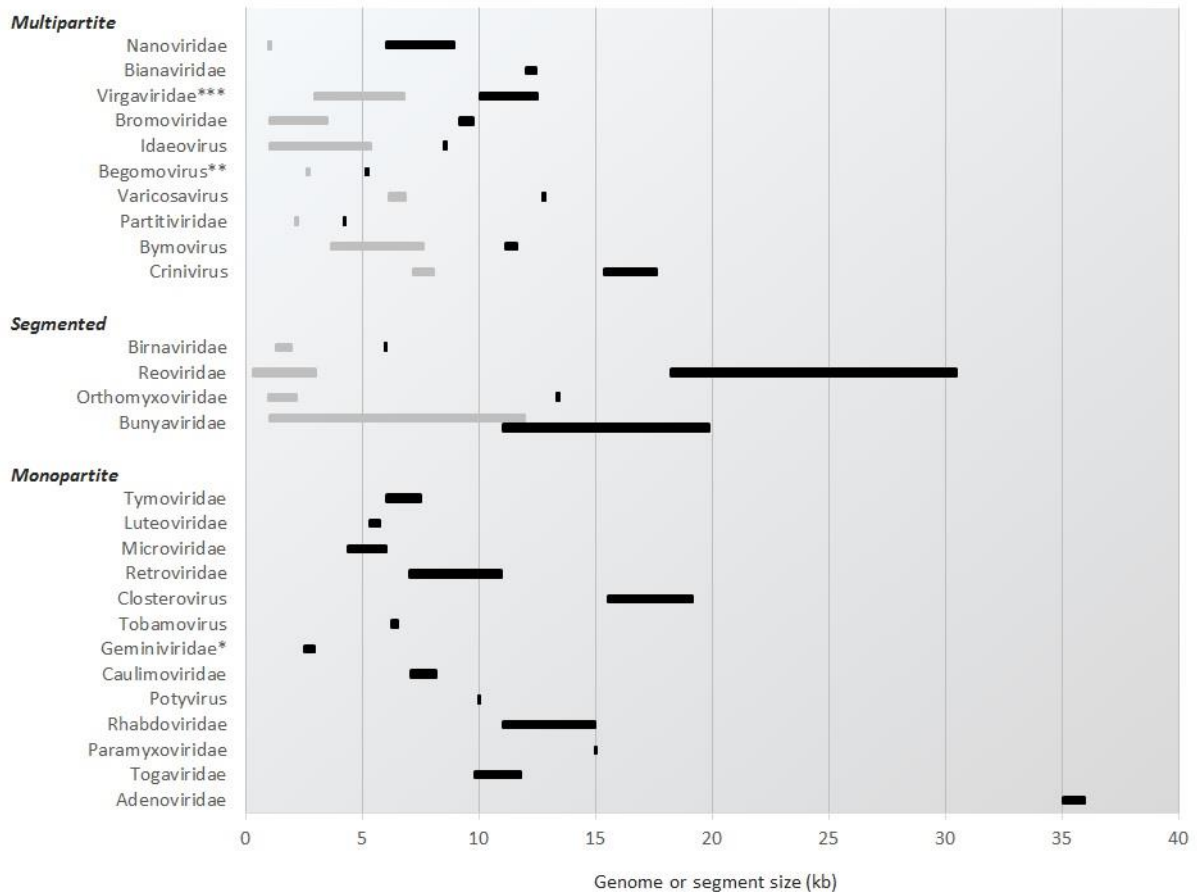
A priori, multipartite viruses do not have to sort segments at encapsidation and they could encapsidate in ways similar to monopartite viruses, the frequency of encapsidated segments directly depending of the frequency of these segments within producing cells. As for monopartite viruses, the specific packaging of viral segments of multipartite genomes relies on the presence of assembly signal sequences (Basnayake et al., 2009; Choi and Rao, 2003). A noticeable difference with monopartite viruses, however, is that the same coat protein(s) have to accommodate the packaging of segments of different sizes, sequences, and secondary/tertiary structures. Multipartite virus particles can be either spherical (*Partitiviridae*, *Nanoviridae*, *Begomovirus*, *Secoviridae*, *Idaeovirus*...), bacilliform (*Ourmiavirus*, *Alfamovirus*...), rod-shaped (*Virgaviridae*, *Varicosavirus*, *Benyvirus*...) or filamentous (*Closteroviridae*, *Capiloviridae*...). While rod-shaped, bacilliform and filamentous viruses can easily accommodate segments of different sizes, by accordingly adjusting the length of the virus particles, physical constraints exist for viruses with an icosahedral structure. A nice illustration of this constraint are found in *in vitro* studies of particles made with the coat protein of *Cowpea chlorotic mottle virus* (CCMV) with a symmetry T=3 (Comas-Garcia et al., 2012). The particles can encapsidate segments from 100 to 12000 base-long RNAs, but they preferentially package one or more RNA segments with a total size around 3.2kb, consistently resulting in an optimal protein/RNA ratio of 6/1 that corresponds to the natural situation for this virus.

In some icosahedral viral species, all segments have a comparable size, indicating that this size may be optimal for particle stability. This is particularly striking for ssDNA multipartite nanoviruses, where the 8 genome segments vary between 920 and 1022 nt in all described species. The ssDNA bipartite geminiviruses of plants and bidensoviruses of insects also have very constrained genome segment sizes of around 2.7 kb and 6 kb, respectively. In other cases however the size of segments encapsidated in icosahedral particles can widely vary. In many species of the family *Bromoviridae* the RNAs, 1, 2, 3 and 4 are approximately of 3, 3, 2 and 1 kb, respectively. This is somehow intriguing because it could theoretically allow the encapsidation of two or more short segments in a single particle. A case study illustrating this is that of *Brome mosaic virus* (BMV) where some particles contain either one copy of a RNA-1 or one copy of RNA-2 (around 3kb each) whereas others contain one copy of RNA-3 and one of RNA-4, together also summing up to approximately 3kb (Annamalai and Rao, 2008; Choi and Rao, 2000). It was recently shown that the situation in BMV is even more complex, because some smaller virus particles have been found and might also accommodate a single copy of either RNA-3 or RNA-4 (Ni et al., 2014).

Once encapsidated, the different interactions occurring between the distinct segments and the capsid, resulting from distinct 3D structure of the packaged RNA, can engender different virion stability (Vaughan et al., 2014). This difference could actually constitute an advantage of multi-encapsidation since it may regulate a differential timing of RNA release and thus the kinetics of gene expression, as suggested for BMV (Ni et al., 2014; Vaughan et al., 2014). Likewise, in rod- or filamentous shape particles, the possible time shift associated with decapsidation of particles of different size could participate in the timely regulation of gene expression. Unfortunately, this possibility has been thus far proposed for BMV only and further work will be required to decide whether this might be a general feature in multipartite viral systems.

Perhaps related to similar questions of encapsidation constraints is the very elegant work recently published on the experimental evolution of *Foot and mouse disease virus* (FMDV) (Ojosnegros et al., 2011). Through repeated passages in cell cultures at elevated multiplicity of infection, these authors observed that two defective molecules complementing each other could outcompete their ancestral monopartite FMDV. The selective advantage of this bipartite derivative was demonstrated to be associated to a higher stability and infectivity of the virus particles packaging smaller RNA segments, and not to a faster replication. The authors thus proposed that genome segmentation and more specifically the multi-encapsidation could in some cases result from a «trade-off between segment length and particle stability». While it has recently been shown that particle size was correlated to genome length (Zandi and van der Schoot, 2009), there is no clear general correlation between genome length and particle stability. For this reason, we believe that the case of segmented FMDV variant taking over the monopartite parental genome might be a specific example, related to experimental conditions, hardly expandable as a general rule to explain the evolution of multipartite viruses. Perhaps consistently, it should be noted that this FMDV experiment was conducted with a virus of vertebrates, where no multipartite natural systems have ever been described. Finally, and this might be a most definitive argument pleading against the genome multi-encapsidation as a way to preserve particle stability for oversized genomes, (i) multipartite viruses do not necessarily have longer genomes than monopartite viruses and (ii) they sometimes encapsidate segments which are longer than the whole genome of some monopartite viruses (Figure 3).





**Figure 3: Comparison of the genome or segment sizes in viruses**

The listed families and genera are those with significant differences in genome architectures. The size-range of whole genomes and that of individual segments are illustrated by black and grey lines, respectively. All size data come from Viralzone.

\*All genera of the family *Geminiviridae* are composed of monopartite virus species, except for the genus *Begomovirus*

\*\*The genus *Begomovirus* is composed of both monopartite and bipartite virus species.

\*\*\* All genera of the family *Virgaviridae* except the genus *Tobamovirus* are composed of multipartite virus species.

## 5 - Within host movement

When viruses move from cell to cell or long distance to colonize systemically their host, the question of transferring all genome information becomes a real issue distinguishing multipartite viruses from both segmented and monopartite. Indeed, for the two latter, the whole genetic information may move as a whole, packaged as complete information units within individual virions, whereas multipartite viruses scatter away fragmentary genetic information in distinct virus particles, which must later reunite to resume infection. The question in this section is thus to see whether some striking specific features emerge in the known mechanisms of within host spread for multipartite viruses.

The discussion on this point is confused by the fact that molecular and cellular data available are showing that viral forms actually trafficking within their host plant is multifarious (reviewed in Hipper et al., 2013; Scholthof, 2005). Some plant virus species move both cell-to-cell and long distance as mature virus particles. Others can move cell-to-cell as nucleoprotein complexes not assembled into mature virus particles, which are only required for long distance movement. Finally, in rare cases, some viral species can spread both cell-to-cell and in the vasculature as nucleoprotein complexes that do not even contain the coat protein. All three cases have been suggested in both monopartite and multipartite viruses (Scholthof, 2005). For example monopartite *Cauliflower mosaic virus* (Carluccio et al., 2014) and multipartite *Cowpea mosaic virus* (Pouwels, 2003) move both cell to cell and long distance as mature virions. On the opposite, both the monopartite *Tomato bushy stunt virus* (Scholthof et al., 1995) and the bipartite *Cabbage leaf curl virus* (Pooma et al., 1996) move cell-to-cell and long distance within their host in the absence of the coat protein.

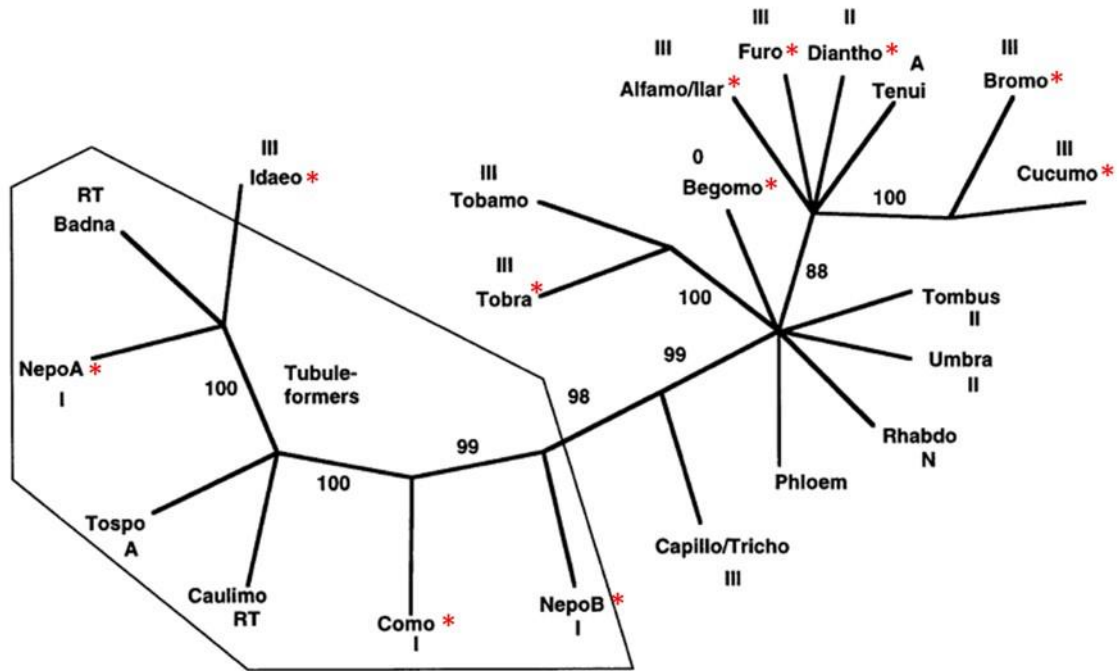
Whatever the viral form that is actually transported, with regard to limitations linked to the size exclusion limit of plasmodesmata, viruses have developed specific different mechanisms to insure their passage to the adjacent cells and into the vasculature, all depending on one or more movement proteins. Through intricate interactions with multiples host factors, some of these movement proteins simply enlarge the size of plasmodesmata, allowing passage of infectious material, whereas others polymerize into a tubular structure, which serves as a syringe to inject the virus into the next cells (Ritzenthaler, 2011; Tilsner and Oparka, 2012). These different mechanisms to pass through plasmodesmata are also shared by multipartite, segmented and monopartite viruses (Hipper et al., 2013; Scholthof, 2005), and thus the present data do not distinguish multipartite virus movement mechanisms from that of other viruses. One further illustrative example of this conclusion is the phylogeny of the viral movement proteins of the

30K super family, established by Melcher in 2000 (Melcher, 2000). In Figure 4, it is apparent that the relatedness between these 30K movement proteins does not reflect any commonalities within multi- or multipartite viruses.

If no distinct molecular or cellular features appear in the mechanisms of within host movement of multipartite viruses, one may try to look at it with a different angle, that of the number of virus particles or genomes that actually enter and infect individual cells (MOI, Multiplicity Of cellular Infection). For multipartite viruses the “independent action hypothesis” (IAH) stipulates that infection of a cell and/or host can be initiated by a single infectious unit, thus a single virus particle, and that each infectious unit can act independently (Druett, 1952). Theoretical models predict a deviation from IAH when one infecting segment depends on the presence of one or more others. This could be empirically verified in several instances (Sanchez-Navarro et al., 2013; Zwart et al., 2011, 2013) and indicate that the number of viral particles efficiently entering a cell (here simplified as MOI) must be much higher in multipartite viruses than it is in multipartite or segmented ones. The MOI is even predicted to be directly related to the number of genome segments and should reach very high values when multipartite viruses have more than 3 or 4 segments. In this case, Iranzo and Manrubia (Iranzo and Manrubia, 2012) have estimated that this number should reach hundreds in order to possibly explain the evolution of such viruses.

MOI values have been experimentally estimated in a number of multipartite viruses infecting bacteria (Turner et al., 1999), insects (Bull et al., 2001), vertebrates (Josefsson et al., 2011; Jung et al., 2002) and plants (Bergua et al., 2014; Gonzalez-Jara et al., 2009; Gutiérrez et al., 2010), and repeatedly found to be relatively small (from 1-13) most often below 5 (reviewed in Gutiérrez et al., 2012). Unfortunately, the only estimate of the MOI of a multipartite virus is that of the bipartite *Soil-borne wheat mosaic virus*, where solely RNA-2 segment was analyzed (Miyashita and Kishino, 2010). The estimated value is similar to that in multipartite viruses, in the order of 5, not clearly supportive of the prediction that multipartite viruses infect cells with higher numbers of virus particles or genome segments.

Due to this paucity of data, and considering the importance of this question in the biology of multipartite viruses, more investigation are wanted and should focus on multipartite viruses with a high number of segments, such as for example member species of the family *Bromoviridae* (3 ssRNA segments + 1 sgRNA), or *Nanoviridae* (6-8 ssDNA segments).



**Figure 4: Relationships among putative 30K MP superfamily members determined by bootstrapped parsimony**

Branches with less than 80% support (100 bootstrap replicates) were collapsed. 0, RT, N, A, I, II and III represent the type of polymerase encoded by the viruses: none, RNA-dependent DNA polymerase, negative-strand virus, ambisense-strand virus and positive-strand virus, supergroups I, II, III RNA-dependent DNA polymerases, respectively. The thin-lined polygon encloses those MPs known to form virion-bearing tubules. Genera with a red asterisk are those whose member species are multipartite viruses (N.B.: The genus Begomovirus is composed of both monopartite and bipartite viruses).

Reproduced from *U. Melcher, The '30K' superfamily of viral movement proteins. Journal of General Virology (2000), 81, 263.*

## 6 - Inter-host movement

The problem of maintaining the integrity of multipartite viral genomes not only questions steps of within host cell-to-cell or long-distance transfers, but also that of host-to-host transmission. In this section, we will focus on horizontal transmission by vectors because this is the major and best-documented mode of transmission of plant viruses. Although data are still lacking in some cases, viruses are all generally assumed to be acquired from infected plants, and inoculated into healthy ones, under the form of virus particles. This extends the risk of scattering and thus losing parts of the genetic information while the virus is spread in the host population, in the environment. That vector-transmission can induce severe bottlenecks into virus populations has been described both in monopartite viruses of animals and plants (reviewed in Gutiérrez et al., 2012). However, cases where this bottleneck was actually quantified are rare in plants and not yet available in animal viruses. The transmission of *Potato virus Y* (PVY) was shown to result from the transfer of as few as 1 or 2 genomes per aphid vector (Moury et al., 2007), and similar figures were obtained for the whitefly-transmission of *Tomato yellow leaf curl virus* (TYLCV) (Péréfarres et al., 2014). Higher numbers are intuitively expected for multipartite viruses, in order to increase the chances of co-transmitting a set of particles containing at least one copy of each segment. Surprisingly, when quantifying the effective number of founders that initiate an infection after aphid-transmission of the tripartite *Cucumber mosaic virus* (CMV), Betancourt and colleagues (Betancourt et al., 2008) found numbers similar to that estimated for the monopartite PVY and TYLCV. It should be noted however, that the genetic diversity of CMV might not be distributed homogeneously in infected plants. The aphids may have acquired and transmitted similar CMV genotypes, resulting in the observed genetic bottleneck even if the actual number of particles transmitted is high (discussed previously in Gutiérrez et al., 2012). In other words, while the transmission-related genetic bottlenecks suffered by viral populations is clearly comparable in mono- and multipartite viruses, the actual number of viral particles that needs to be acquired and inoculated might be different and remains to be evaluated and compared.

These results open a number of questions, which are thus far a mystery in the transmission of multipartite viruses, and which represents appealing lines of research in the next future. In particular, that distinct genome segments accumulate at very different frequencies within host plants (Sanchez-Navarro et al., 2013; Sicard et al., 2013) poses the question of the rescue of the rare segments upon transmission process. Likewise, because virus particles containing distinct genome segments can have a different stability (Vaughan et al., 2014), they might well be

differentially degraded during the passage within the insect vectors, again questioning whether the relative frequency of the segments may change within vectors and how the most labile particles can be transmitted as efficiently as the others.

Finally, the bottlenecks induced by the transmission of viral populations by a single insect vector, as calculated in the above-cited studies, might be alleviated by the fact that plants could be visited by numerous viruliferous insects and thus inoculated repeatedly. Unfortunately, as discussed in the next section, the ecology of multipartite viruses is poorly known, and this possibility is not sufficiently documented.

## **7 - Ecology of multipartite viruses**

As far as we are aware, there is no study available on specific ecological features that may be associated with multipartite viruses. One investigation on possible viral intrinsic features that could affect the breadth of the viral host range is suggesting that mono- and bipartite viruses have a narrower host range than tripartite viruses, a trend that is seemingly reverting when the number of segments is further increasing (Moury et al In press). The authors could not provide any explanation for this observation and so its significance and/or relevance is not clear.

A huge literature is providing information on the geographical variations in genetic diversity of viruses in general, and of multipartite viruses in particular. In the latter case however, apart for confirming that reassortment and recombination are frequent (discussed in section 2), and thus likely play a key role in the evolution of multipartite viruses, few (if at all) investigation of the ecology of these viruses have been reported. The studies on the frequency of reassortants in natural population of CMV (Bonnet et al., 2005; Nouri et al., 2014), of nanoviruses (Grigoras et al., 2014; Savory and Ramakrishnan, 2014), and of bipartite geminiviruses (De Bruyn et al., 2012) are certainly a strong step in this direction, particularly the fact that different segments appear to have distinct evolutionary histories and distinct phylogeographies (De Bruyn et al., 2012).

A detailed analysis of the species community network into which multipartite virus actually circulate (both host plants and vectors) has never been reported. It is noteworthy that in certain viral species, specific genome segments are dispensable for infection of plants in laboratory conditions, indicating a function at other levels. Because several related species or strains are often found in the same fields or even plants, (i) how the distinct segments actually circulate in nature, (ii) whether they are always together or whether they can be lost and retrieved at distinct

nodes of the host and vectors community network, (iii) how the homologous segments of related species circulate and compete in this network, (iv) what are the fitness values of the different possible reassortants, and (v) how these various reassortants actually do in different nodes of the network, are all questions that appear fascinating, specific to multipartite viruses, and yet totally unexplored.

### **Concluding remarks: future prospects**

The perusal of the literature presented above convincingly illustrates the current ignorance on the reasons why multipartite viruses are so successful. The simple fact that they can be ssRNA, dsRNA, or ssDNA, likely indicates that multi-encapsulation has evolved more than once independently, but we nevertheless remain unable to identify and confirm any associated beneficial aspects. Most putative benefits proposed are possibly controversial because of the cruel lack of data on the way of life of these viruses. We believe that this is the key: investigating the specifics of the biology of multipartite viruses would allow confirming or detracting the existing hypotheses, and certainly proposing new ones perhaps sometimes more documented and thus more pertinent or unforeseen. As concluding remarks, we here select and highlight a few research directions that could clarify or assign specific properties to multipartite viruses that we judge immediately critical. Structural, physical and biochemical properties of the virus particles, depending on the contained segment(s), should be investigated in more details. The bipartite *Red clover necrotic mosaic virus* (RCNMV) can form two distinct populations of particles, one containing both RNA1 and RN2 and the other packaging multiple copies of RNA2 (Basnayake et al., 2006, 2009). Likewise, the detailed biochemical studies on BMV (section 4) have revealed that the nature and the number of encapsidated segments can be more versatile than generally thought. The most important message from these series of investigations is that distinct physical and chemical properties of particles containing distinct segments reflect an adaptive process involved in the timely regulation gene expression. This far-reaching possibility could extend to the regulation of host-specific (Ni et al., 2014) or even tissue-specific functions. Moreover because these variable properties of virus particles depend on the RNA folding structure, they might also be important for ssDNA multipartite viruses where secondary folding structures of various segments appears to have unknown biological function (Muhire et al., 2014) and may confer distinct properties to their respective particles.

Still related to the regulation of gene expression, the discovery of the genome formula in the nanovirus FBNSV is promising (Sicard et al., 2013, section 3).

This is contributing to the consideration of a putative important role of gene copy number variations in the biology of viruses. Prominent questions are whether the genome formula is also controlled in other multipartite viruses, whether it actually regulates gene and phenotype expression, and whether it is an adaptive and evolvable trait. One argument in favor of the adaptive regulation of gene copy number in multipartite viruses is the observation of rare segments or genes (Feng et al., 2008, 2006; Kwon et al., 2012; Sanchez-Navarro et al., 2013; Sicard et al., 2013). Would the gene copy number be irrelevant in the biology of these viruses, then all segments would be expected to accumulate at equal ratios, a situation where the probability of a segment loss is minimal.

Then comes the question of the MOI, here meaning the number of virus particles efficiently penetrating individual cells, and releasing segments that will actually replicate. When complying with the belief that all segments should co-penetrate a cell in order to be functional, there are two undecided possibilities. First, particles could penetrate cells with an equal chance, whatever the contained segment. Second, it is imaginable (though hardly) that multipartite viruses have developed sophisticated mechanisms sorting particles containing specific segments, and promoting the cell entry as packages with complete genetic information (at least a copy of each). In the first scenario, the observed MOI values for the different segments should be high (to make sure that at least one rare segment makes it through), and directly related to their frequency within the viral population. In the second scenario, the MOI values might be low and non-correlated to the segment frequencies within the population. This understudied aspect would be very informative on the way of life of the multipartite viruses. For example, the long standing question of why multipartite viruses do not exist in animals has always been related to the existence of a cytoplasmic continuum in between cells of plants and fungi, that do not exist in animals, and that could allow massive transfers of viral particles from cell to cell. As discussed in section 5, MOI values calculated thus far with monopartite viruses of plants and animals do not differ so much, and thus their values for multipartite viruses is highly wanted.

Finally, the most important view to be incontrovertibly verified is that the distinct segments of a multipartite viruses need be together within cells for the system to be functional. This assertion is at the basis of most if not all the upheaval concerning these viruses, because the scientific community is striving to both estimate how costly it is and what benefits could compensate for this cost.



We believe that experimentally assessing whether multipartite viruses could function with their different segments scattered in different cells is not a heretic proposal. Indeed, many studies are reporting the capacity of proteins (Gallagher et al., 2014; Solovyev and Savenkov, 2014; Turgeon and Wolf, 2009; Ueki and Citovsky, 2011) and RNA (Ding, 2010; Flores et al., 2012; Gopinath and Kao, 2007; Palukaitis, 2014; Sarkies and Miska, 2014; Spiegelman et al., 2013) to traffic autonomously from cell-to-cell or long distance within host plants. Assuming that some RNA and/or proteins of multipartite viruses might traffic on their own opens the possibility for a viral function to efficiently act within cells devoid of the gene encoding it. In this view, the reduced chances to infect individual cells with all segments together would no longer be such an acute problem, and the angle with which the biology of multipartite viruses could best be conceptualized would be dramatically different.

## References

- Ahlquist, P., Dasgupta, R., and Kaesberg, P. (1984). Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. *J. Mol. Biol.* *172*, 369–383.
- Annamalai, P., and Rao, A.L.N. (2008). RNA encapsidation assay. *Methods Mol. Biol. Clifton NJ* *451*, 251–264.
- Argüello-Astorga, G.R., Guevara-González, R.G., Herrera-Estrella, L.R., and Rivera-Bustamante, R.F. (1994). Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology* *203*, 90–100.
- Bald, J.G. (1937). The use of numbers of infections for comparing the concentration of plant virus suspensions: dilution experiments with purified suspensions. *Ann. Appl. Biol.* *24*, 33–55.
- Bald, J.G. (1950). Measurement of concentration of plant virus suspensions. *Viruses* *1950*, 17–29.
- Basnayake, V.R., Sit, T.L., and Lommel, S.A. (2006). The genomic RNA packaging scheme of Red clover necrotic mosaic virus. *Virology* *345*, 532–539.
- Basnayake, V.R., Sit, T.L., and Lommel, S.A. (2009). The Red clover necrotic mosaic virus origin of assembly is delimited to the RNA-2 trans-activator. *Virology* *384*, 169–178.
- Bergua, M., Zwart, M.P., El-Mohtar, C., Shilts, T., Elena, S.F., and Folimonova, S.Y. (2014). A Viral Protein Mediates Superinfection Exclusion at the Whole-Organism Level but Is Not Required for Exclusion at the Cellular Level. *J. Virol.* *88*, 11327–11338.
- Betancourt, M., Fereres, A., Fraile, A., and Garcia-Arenal, F. (2008). Estimation of the Effective Number of Founders That Initiate an Infection after Aphid Transmission of a Multipartite Plant Virus. *J. Virol.* *82*, 12416–12421.
- Bonnet, J., Fraile, A., Sacristán, S., Malpica, J.M., and García-Arenal, F. (2005). Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus. *Virology* *332*, 359–368.
- Brown, I.H., Harris, P.A., McCauley, J.W., and Alexander, D.J. (1998). Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. *J. Gen. Virol.* *79*, 2947–2955.
- De Bruyn, A., Villemot, J., Lefeuvre, P., Villar, E., Hoareau, M., Harimalala, M., Abdoul-Karime, A.L., Abdou-Chakour, C., Reynaud, B., Harkins, G.W., 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 Evol. Biol.* *12*, 228.
- Bujarski, J.J., and Kaesberg, P. (1986). Genetic recombination between RNA components of a multipartite plant virus. *Nature* *321*, 528–531.
- Bull, J.C., Godfray, H.C.J., and O'Reilly, D.R. (2001). Persistence of an Occlusion-Negative Recombinant Nucleopolyhedrovirus in *Trichoplusia ni* Indicates High Multiplicity of Cellular Infection. *Appl. Environ. Microbiol.* *67*, 5204–5209.
- Carluccio, A.V., Zicca, S., and Stavolone, L. (2014). Hitching a ride on vesicles: cauliflower mosaic virus movement protein trafficking in the endomembrane system. *Plant Physiol.* *164*, 1261–1270.
- Chao, L. (1988). Evolution of sex in RNA viruses. *J. Theor. Biol.* *133*, 99–112.

- Chao, L. (1991). Levels of selection, evolution of sex in RNA viruses, and the origin of life. *J. Theor. Biol.* *153*, 229–246.
- Chao, L. (1992). Evolution of sex in RNA viruses. *Trends Ecol. Evol.* *7*, 147–151.
- Choi, Y.G., and Rao, A.L. (2000). Molecular studies on bromovirus capsid protein. VII. Selective packaging on BMV RNA4 by specific N-terminal arginine residuals. *Virology* *275*, 207–217.
- Choi, Y.G., and Rao, A.L.N. (2003). Packaging of Brome Mosaic Virus RNA3 Is Mediated through a Bipartite Signal. *J. Virol.* *77*, 9750–9757.
- Chou, Y.-y., Vafabakhsh, R., Doganay, S., Gao, Q., Ha, T., and Palese, P. (2012). One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl. Acad. Sci.* *109*, 9101–9106.
- Comas-Garcia, M., Cadena-Nava, R.D., Rao, A.L.N., Knobler, C.M., and Gelbart, W.M. (2012). In Vitro Quantification of the Relative Packaging Efficiencies of Single-Stranded RNA Molecules by Viral Capsid Protein. *J. Virol.* *86*, 12271–12282.
- Ding, B. (2010). Viroids: self-replicating, mobile, and fast-evolving noncoding regulatory RNAs. *Wiley Interdiscip. Rev. RNA* *1*, 362–375.
- Drake, J.W. (1993). Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci.* *90*, 4171–4175.
- Druett, H.A. (1952). Bacterial invasion. *Nature* *170*, 288.
- Duffy, S., and Holmes, E.C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J. Gen. Virol.* *90*, 1539–1547.
- Elde, N.C., Child, S.J., Eickbush, M.T., Kitzman, J.O., Rogers, K.S., Shendure, J., Geballe, A.P., and Malik, H.S. (2012). Poxviruses Deploy Genomic Accordions to Adapt Rapidly against Host Antiviral Defenses. *Cell* *150*, 831–841.
- Feng, J., Zeng, R., and Chen, J. (2008). Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. *BioTechniques* *44*, 901–912.
- Feng, J.-L., Chen, S.-N., Tang, X.-S., Ding, X.-F., Du, Z.-Y., and Chen, J.-S. (2006). Quantitative Determination of Cucumber Mosaic Virus Genome RNAs in Virions by Real-Time Reverse Transcription-Polymerase Chain Reaction. *Acta Biochim. Biophys. Sin.* *38*, 669–676.
- Filée, J. (2013). Route of NCLDV evolution: the genomic accordion. *Curr. Opin. Virol.* *3*, 595–599.
- Flores, R., Serra, P., Minoia, S., Di Serio, F., and Navarro, B. (2012). Viroids: From Genotype to Phenotype Just Relying on RNA Sequence and Structural Motifs. *Front. Microbiol.* *3*.
- Fontes, E.P., Gladfelter, H.J., Schaffer, R.L., Petty, I.T., and Hanley-Bowdoin, L. (1994). Geminivirus replication origins have a modular organization. *Plant Cell* *6*, 405–416.
- Froissart, R., Roze, D., Uzest, M., Galibert, L., Blanc, S., and Michalakis, Y. (2005). Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. *PLoS Biol.* *3*, e89.
- Fulton, R.W. (1962). The effect of dilution on necrotic ringspot virus infectivity and the enhancement of infectivity by noninfective virus. *Virology* *18*, 477–485.

- Gallagher, K.L., Sozzani, R., and Lee, C.-M. (2014). Intercellular Protein Movement: Deciphering the Language of Development. *Annu. Rev. Cell Dev. Biol.* 30, 207–233.
- Gladfelter, H.J., Eagle, P.A., Fontes, E.P.B., Batts, L., and Hanley-Bowdoin, L. (1997). Two Domains of the AL1 Protein Mediate Geminivirus Origin Recognition. *Virology* 239, 186–197.
- Godefroy-Colburn, T., Thivent, C., and Pinck, L. (1985). Translational discrimination between the four RNAs of alfalfa mosaic virus. *Eur. J. Biochem.* 147, 541–548.
- Goldhill, D.H., and Turner, P.E. (2014). The evolution of life history trade-offs in viruses. *Curr. Opin. Virol.* 8C, 79–84.
- Gonzalez-Jara, P., Fraile, A., Canto, T., and Garcia-Arenal, F. (2009). The Multiplicity of Infection of a Plant Virus Varies during Colonization of Its Eukaryotic Host. *J. Virol.* 83, 7487–7494.
- Gopinath, K., and Kao, C.C. (2007). Replication-independent long-distance trafficking by viral RNAs in *Nicotiana benthamiana*. *Plant Cell* 19, 1179–1191.
- Grigoras, I., Timchenko, T., Grande-Perez, A., Katul, L., Vetten, H.-J., and Gronenborn, B. (2010). High Variability and Rapid Evolution of a Nanovirus. *J. Virol.* 84, 9105–9117.
- Grigoras, I., Ginzo, A.I. d. C., Martin, D.P., Varsani, A., Romero, J., Mammadov, A.C., Huseynova, I.M., Aliyev, J.A., Kheyr-Pour, A., Huss, H., et al. (2014). Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *J. Gen. Virol.* 95, 1178–1191.
- Gutiérrez, S., Yvon, M., Thébaud, G., Monsion, B., Michalakis, Y., and Blanc, S. (2010). Dynamics of the Multiplicity of Cellular Infection in a Plant Virus. *PLoS Pathog.* 6, e1001113.
- Gutiérrez, S., Michalakis, Y., and Blanc, S. (2012). Virus population bottlenecks during within-host progression and host-to-host transmission. *Curr. Opin. Virol.* 2, 546–555.
- Hayakawa, T., Kojima, K., Nonaka, K., Nakagaki, M., Sahara, K., Asano, S., Iizuka, T., and Bando, H. (2000). Analysis of proteins encoded in the bipartite genome of a new type of parvo-like virus isolated from silkworm—structural protein with DNA polymerase motif. *Virus Res.* 66, 101–108.
- Hicks, A.L., and Duffy, S. (2014). Cell Tropism Predicts Long-term Nucleotide Substitution Rates of Mammalian RNA Viruses. *PLoS Pathog.* 10, e1003838.
- Hipper, C., Brault, V., Ziegler-Graff, V., and Revers, F. (2013). Viral and Cellular Factors Involved in Phloem Transport of Plant Viruses. *Front. Plant Sci.* 4.
- Hollox, E.J., and Hoh, B.-P. (2014). Human gene copy number variation and infectious disease. *Hum. Genet.* 133, 1217–1233.
- Horser, C.L., Karan, M., Harding, R.M., and Dale, J.L. (2001). Additional Rep-encoding DNAs associated with banana bunchy top virus. *Arch. Virol.* 146, 71–86.
- Hughes, A.L. (2004). Birth-and-death evolution of protein-coding regions and concerted evolution of non-coding regions in the multi-component genomes of nanoviruses. *Mol. Phylogenet. Evol.* 30, 287–294.
- (Hull, 2014). *Plant virology* (Amsterdam ; Boston: Elsevier/AP).

- Ince, W.L., Gueye-Mbaye, A., Bennink, J.R., and Yewdell, J.W. (2013). Reassortment Complements Spontaneous Mutation in Influenza A Virus NP and M1 Genes To Accelerate Adaptation to a New Host. *J. Virol.* *87*, 4330–4338.
- Iranzo, J., and Manrubia, S.C. (2012). Evolutionary dynamics of genome segmentation in multipartite viruses. *Proc. R. Soc. B Biol. Sci.* *279*, 3812–3819.
- Jaspars, E.M.J. (1975). Plant viruses with a multipartite genome. *Adv. Virus Res.* *19*, 37–149.
- Josefsson, L., King, M.S., Makitalo, B., Brannstrom, J., Shao, W., Maldarelli, F., Kearney, M.F., Hu, W.-S., Chen, J., Gaines, H., et al. (2011). Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proc. Natl. Acad. Sci.* *108*, 11199–11204.
- Jung, A., Maier, R., Vartanian, J.-P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S., and Meyerhans, A. (2002). Recombination: Multiply infected spleen cells in HIV patients. *Nature* *418*, 144–144.
- Van Kammen, A. (1968). The relationship between the components of cowpea mosaic virus. *Virology* *34*, 312–318.
- Katju, V., and Bergthorsson, U. (2013). Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Front. Genet.* *4*.
- Kuroda, T., Natsuaki, T., Okuda, S., and Teranaka, M. (2005). Analysis of natural populations and possible natural reassortment of Cucumber mosaic virus. *J. Gen. Plant Pathol.* *71*, 221–229.
- Kwon, C.S., and Chung, W.-I. (2000). Differential roles of the 5' untranslated regions of cucumber mosaic virus RNAs 1, 2, 3 and 4 in translational competition. *Virus Res.* *66*, 175–185.
- Kwon, S.-J., Seo, J.-K., and Rao, A.L.N. (2012). Detection and Quantification of Viral and Satellite RNAs in Plant Hosts. In *RNA Abundance Analysis*, H. Jin, and W. Gassmann, eds. (Totowa, NJ: Humana Press), pp. 131–141.
- Lin, H.-X., Rubio, L., Smythe, A.B., and Falk, B.W. (2004). Molecular Population Genetics of Cucumber Mosaic Virus in California: Evidence for Founder Effects and Reassortment. *J. Virol.* *78*, 6666–6675.
- Lister, R.M. (1968). Functional Relationships between Virus-specific Products of Infection by Viruses of the Tobacco Rattle Type. *J. Gen. Virol.* *2*, 43–58.
- López-Bueno, A., Villarreal, L.P., and Almendral, J.M. (2006). Parvovirus variation for disease: a difference with RNA viruses? *Curr. Top. Microbiol. Immunol.* *299*, 349–370.
- Luque, D., Rivas, G., Alfonso, C., Carrascosa, J.L., Rodríguez, J.F., and Castón, J.R. (2009). Infectious bursal disease virus is an icosahedral polypliod dsRNA virus. *Proc. Natl. Acad. Sci.* *106*, 2148–2152.
- McDonald, S.M., and Patton, J.T. (2011). Assortment and packaging of the segmented rotavirus genome. *Trends Microbiol.* *19*, 136–144.
- Medina, R.A., and García-Sastre, A. (2011). Influenza A viruses: new research developments. *Nat. Rev. Microbiol.* *9*, 590–603.
- Melcher, U. (2000). The “30K” superfamily of viral movement proteins. *J. Gen. Virol.* *81*, 257–266.

- Mileyko, Y., Joh, R.I., and Weitz, J.S. (2008). Small-scale copy number variation and large-scale changes in gene expression. *Proc. Natl. Acad. Sci.* *105*, 16659–16664.
- Mills, D.R., Peterson, R.L., and Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U. S. A.* *58*, 217–224.
- Mindich, L. (2004). Packaging, replication and recombination of the segmented genomes of bacteriophage  $\Phi 6$  and its relatives. *Virus Res.* *101*, 83–92.
- Miyashita, S., and 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. *J. Virol.* *84*, 1828–1837.
- Moury, B., Fabre, F., and Senoussi, R. (2007). Estimation of the number of virus particles transmitted by an insect vector. *Proc. Natl. Acad. Sci.* *104*, 17891–17896.
- Muhire, B.M., Golden, M., Murrell, B., Lefevre, P., Lett, J.-M., Gray, A., Poon, A.Y.F., Ngandu, N.K., Semegni, Y., Tanov, E.P., et al. (2014). Evidence of Pervasive Biologically Functional Secondary Structures within the Genomes of Eukaryotic Single-Stranded DNA Viruses. *J. Virol.* *88*, 1972–1989.
- Nee, S. (1987). The evolution of multicompartmental genomes in viruses. *J. Mol. Evol.* *25*, 277–281.
- Nee, S. (1988). Deleterious mutation and genetic recombination. *Nature* *331*, 308.
- Neumann, G., Noda, T., and Kawaoka, Y. (2009). Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* *459*, 931–939.
- Ni, P., Vaughan, R.C., Tragesser, B., Hoover, H., and Kao, C.C. (2014). The Plant Host Can Affect the Encapsidation of Brome Mosaic Virus (BMV) RNA: BMV Virions Are Surprisingly Heterogeneous. *J. Mol. Biol.* *426*, 1061–1076.
- Noda, T., and Kawaoka, Y. (2012). Packaging of influenza virus genome: Robustness of selection. *Proc. Natl. Acad. Sci.* *109*, 8797–8798.
- Nouri, S., Arevalo, R., Falk, B.W., and Groves, R.L. (2014). Genetic Structure and Molecular Variability of Cucumber mosaic virus Isolates in the United States. *PLoS ONE* *9*, e96582.
- Ojosnegros, S., García-Arriaza, J., Escarmís, C., Manrubia, S.C., Perales, C., Arias, A., Mateu, M.G., and Domingo, E. (2011). Viral Genome Segmentation Can Result from a Trade-Off between Genetic Content and Particle Stability. *PLoS Genet.* *7*, e1001344.
- Palukaitis, P. (2014). What has been happening with viroids? *Virus Genes* *49*, 175–184.
- Péréfarres, F., Thébaud, G., Lefevre, P., Chiroleu, F., Rimbaud, L., Hoareau, M., Reynaud, B., and Lett, J.-M. (2014). Frequency-dependent assistance as a way out of competitive exclusion between two strains of an emerging virus. *Proc. Biol. Sci.* *281*, 20133374.
- Pierrugues, O., Guilbaud, L., Fernandez-Delmond, I., Fabre, F., Tepfer, M., and Jacquemond, M. (2007). Biological properties and relative fitness of inter-subgroup cucumber mosaic virus RNA 3 recombinants produced in vitro. *J. Gen. Virol.* *88*, 2852–2861.
- Pooma, W., Gillette, W.K., Jeffrey, J.L., and Petty, I.T. (1996). Host and viral factors determine the dispensability of coat protein for bipartite geminivirus systemic movement. *Virology* *218*, 264–268.

- Poranen, M.M., and Tuma, R. (2004). Self-assembly of double-stranded RNA bacteriophages. *Virus Res.* *101*, 93–100.
- Pouwels, J. (2003). Identification of distinct steps during tubule formation by the movement protein of Cowpea mosaic virus. *J. Gen. Virol.* *84*, 3485–3494.
- Pressing, J., and Reaney, D.C. (1984). Divided genomes and intrinsic noise. *J. Mol. Evol.* *20*, 135–146.
- Qiu, W., and Moyer, J.W. (1999). Tomato Spotted Wilt Tosspovirus Adapts to the TSWV N Gene-Derived Resistance by Genome Reassortment. *Phytopathology* *89*, 575–582.
- Rager, M., Vongpunsawad, S., Duprex, W.P., and Cattaneo, R. (2002). Polyploid measles virus with hexameric genome length. *EMBO J.* *21*, 2364–2372.
- Reguera, J., Malet, H., Weber, F., and Cusack, S. (2013). Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc. Natl. Acad. Sci.* *110*, 7246–7251.
- Ren, X., Tao, Y., Cui, J., Suo, S., Cong, Y., and Tijssen, P. (2013). Phylogeny and evolution of porcine parvovirus. *Virus Res.* *178*, 392–397.
- Ribi re, M., Olivier, V., and Blanchard, P. (2010). Chronic bee paralysis: A disease and a virus like no other? *J. Invertebr. Pathol.* *103*, S120–S131.
- Ritzenthaler, C. (2011). Parallels and distinctions in the direct cell-to-cell spread of the plant and animal viruses. *Curr. Opin. Virol.* *1*, 403–409.
- Roossinck, M.J. (2002). Evolutionary History of Cucumber Mosaic Virus Deduced by Phylogenetic Analyses. *J. Virol.* *76*, 3382–3387.
- Sakai, Y., Kiyotani, K., Fukumura, M., Asakawa, M., Kato, A., Shioda, T., Yoshida, T., Tanaka, A., Hasegawa, M., and Nagai, Y. (1999). Accommodation of foreign genes into the Sendai virus genome: sizes of inserted genes and viral replication. *FEBS Lett.* *456*, 221–226.
- Sanchez-Navarro, J.A., Zwart, M.P., and Elena, S.F. (2013). Effects of the Number of Genome Segments on Primary and Systemic Infections with a multipartite Plant RNA Virus. *J. Virol.* *87*, 10805–10815.
- Sanjuan, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral Mutation Rates. *J. Virol.* *84*, 9733–9748.
- Sarker, S., Patterson, E.I., Peters, A., Baker, G.B., Forwood, J.K., Ghorashi, S.A., Holdsworth, M., Baker, R., Murray, N., and Raidal, S.R. (2014). Mutability Dynamics of an Emergent Single Stranded DNA Virus in a Na ve Host. *PLoS ONE* *9*, e85370.
- Sarkies, P., and Miska, E.A. (2014). Small RNAs break out: the molecular cell biology of mobile small RNAs. *Nat. Rev. Mol. Cell Biol.* *15*, 525–535.
- Savory, F.R., and Ramakrishnan, U. (2014). Asymmetric patterns of reassortment and concerted evolution in Cardamom bushy dwarf virus. *Infect. Genet. Evol.* *24*, 15–24.
- Scholthof, H.B. (2005). Plant virus transport: motions of functional equivalence. *Trends Plant Sci.* *10*, 376–382.

- Scholthof, H.B., Scholthof, K.-B.G., Kikkert, M., and Jackson, A.O. (1995). Tomato bushy stunt virus spread is regulated by two nested genes that function in cell-to-cell movement and host-dependent systemic invasion. *Virology* 213, 425–438.
- Shirasawa-Seo, N. (2005). Characteristics of the promoters derived from the single-stranded DNA components of Milk vetch dwarf virus in transgenic tobacco. *J. Gen. Virol.* 86, 1851–1860.
- Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakakis, Y., Gutierrez, S., and Blanc, S. (2013). Gene copy number is differentially regulated in a multipartite virus. *Nat. Commun.* 4.
- Sivakumaran, K., Bao, Y., Roossinck, M.J., and Kao, C.C. (2000). Recognition of the Core RNA Promoter for Minus-Strand RNA Synthesis by the Replicases of Brome Mosaic Virus and Cucumber Mosaic Virus. *J. Virol.* 74, 10323–10331.
- Solovyev, A.G., and Savenkov, E.I. (2014). Factors involved in the systemic transport of plant RNA viruses: the emerging role of the nucleus. *J. Exp. Bot.* 65, 1689–1697.
- Spiegelman, S., Mills, D.R., and Kramer, F.R. (1975). The extracellular evolution of structure in replicating RNA molecules. *Stab. Orig. Biol. Inf.*
- Spiegelman, Z., Golan, G., and Wolf, S. (2013). Don't kill the messenger: Long-distance trafficking of mRNA molecules. *Plant Sci.* 213, 1–8.
- Stainton, D., Kraberger, S., Walters, M., Wiltshire, E.J., Rosario, K., Halafihi, M., Lolohea, S., Katoa, I., Faitua, T.H., Aholelei, W., et al. (2012). Evidence of inter-component recombination, intra-component recombination and reassortment in banana bunchy top virus. *J. Gen. Virol.* 93, 1103–1119.
- Tilsner, J., and Oparka, K.J. (2012). Missing links? - The connection between replication and movement of plant RNA viruses. *Curr. Opin. Virol.* 2, 705–711.
- Timchenko, T., Katul, L., Sano, Y., de Kouchkovsky, F., Vetten, H.J., and Gronenborn, B. (2000). The Master Rep Concept in Nanovirus Replication: Identification of Missing Genome Components and Potential for Natural Genetic Reassortment. *Virology* 274, 189–195.
- Tromas, N., Zwart, M.P., Lafforgue, G., and Elena, S.F. (2014). Within-Host Spatiotemporal Dynamics of Plant Virus Infection at the Cellular Level. *PLoS Genet.* 10, e1004186.
- Turgeon, R., and Wolf, S. (2009). Phloem Transport: Cellular Pathways and Molecular Trafficking. *Annu. Rev. Plant Biol.* 60, 207–221.
- Turner, P.E., Burch, C.L., Hanley, K.A., and Chao, L. (1999). Hybrid frequencies confirm limit to coinfection in the RNA bacteriophage phi6. *J. Virol.* 73, 2420–2424.
- Ueki, S., and Citovsky, V. (2011). To Gate, or Not to Gate: Regulatory Mechanisms for Intercellular Protein Transport and Virus Movement in Plants. *Mol. Plant* 4, 782–793.
- Urbanowicz, A., Alejska, M., Formanowicz, P., Blazewicz, J., Figlerowicz, M., and Bujarski, J.J. (2005). Homologous Crossovers among Molecules of Brome Mosaic Bromovirus RNA1 or RNA2 Segments In Vivo. *J. Virol.* 79, 5732–5742.
- Vaughan, R., Tragesser, B., Ni, P., Ma, X., Dragnea, B., and Kao, C.C. (2014). The Tripartite Virions of the Brome Mosaic Virus Have Distinct Physical Properties That Affect the Timing of the Infection Process. *J. Virol.* 88, 6483–6491.



- Van Vloten-Doting, L., and Jaspars, E.M.J. (1977). Plant covirus system: three component systems. *Compr. Virol.* 11, 1–53.
- Van der Walt, E., Martin, D.P., Varsani, A., Polston, J.E., and Rybicki, E.P. (2008). Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. *Virol. J.* 5, 104.
- Wang, Y.J., Yao, Q., Chen, K.P., Wang, Y., Lu, J., and Han, X. (2007). Characterization of the genome structure of *Bombyx mori* densovirus (China isolate). *Virus Genes* 35, 103–108.
- Wichgers Schreur, P.J., Oreshkova, N., Moormann, R.J.M., and Kortekaas, J. (2014). Creation of Rift Valley Fever Viruses with Four-Segmented Genomes Reveals Flexibility in Bunyavirus Genome Packaging. *J. Virol.* 88, 10883–10893.
- Yang, X., Zhou, M., Qian, Y., Xie, Y., and Zhou, X. (2014). Molecular variability and evolution of a natural population of tomato yellow leaf curl virus in Shanghai, China. *J. Zhejiang Univ. Sci. B* 15, 133–142.
- Zagórski, W. (1978). Translational Regulation of Expression of the Brome-Mosaic-Virus RNA Genome in vitro. *Eur. J. Biochem.* 86, 465–472.
- Zaitlin, M., and Palukaitis, P. (2000). Advances in understanding plant viruses and virus diseases. *Annu. Rev. Phytopathol.* 38, 117–143.
- Zandi, R., and van der Schoot, P. (2009). Size Regulation of ss-RNA Viruses. *Biophys. J.* 96, 9–20.
- Zwart, M.P., Daròs, J.-A., and Elena, S.F. (2011). One Is Enough: In Vivo Effective Population Size Is Dose-Dependent for a Plant RNA Virus. *PLoS Pathog.* 7, e1002122.
- Zwart, M.P., Tromas, N., and Elena, S.F. (2013). Model-Selection-Based Approach for Calculating Cellular Multiplicity of Infection during Virus Colonization of Multi-Cellular Hosts. *PLoS ONE* 8, e64657.

**LE *FABA BEAN NECROTIC STUNT VIRUS*, UN MODELE POUR  
L'ETUDE DES VIRUS MULTIPARTITES**

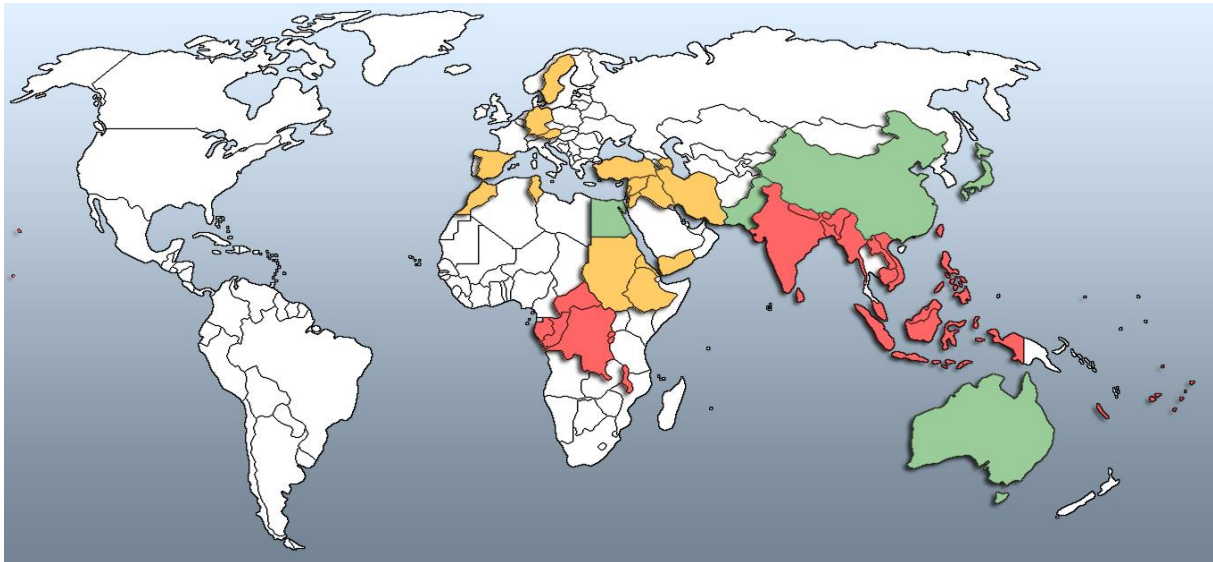
## 2. Le *Faba bean necrotic stunt virus*, un modèle pour l'étude des virus multipartites

### 2.1. *Les nanovirus, des virus multipartites au génome segmenté « à l'extrême »*

Notre modèle d'étude a été choisi parmi les *Nanoviridae*. Cette famille de virus de plantes est composée des deux genres *Babuvirus* et *Nanovirus* et se caractérise par des génomes composés de six à huit molécules d'ADN circulaire simple brin (Grigoras et al., 2014). La plupart des virus multipartites étant bi- et tripartites, les membres de la famille *Nanoviridae* apparaissent comme les cas les plus extrêmes d'encapsidation multiple. Ainsi, on peut supposer que tous les traits biologiques spécifiquement reliés à la multiencapsidation, qu'ils soient associés à des avantages ou à des inconvénients, seront maximisés chez ces virus. Par ailleurs, une grande majorité des virus multipartites connus ont un ou plusieurs de leurs segments qui codent pour plusieurs protéines (via l'utilisation d'ARN subgénomiques, de polyprotéines...). *A contrario*, à l'exception du segment R du *Banana bunchy top virus*, tous les segments des nanoviridae décrits jusqu'à maintenant sont monocistroniques (Timchenko and Bernardi, 2007). Cette correspondance « un segment - un gène » apporte donc une certaine clarté au modèle, notamment en termes de stratégie d'expression et de régulation des gènes.

### 2.2. *Répartition géographique et importance agronomique des Nanoviridae*

Le membre le plus connu des *Babuvirus*, le *Banana bunchy top virus*, est le virus causant le plus de dommages sur les cultures de bananiers en Afrique, en Asie du Sud-Est et en Australie (Blomme et al., 2013). L'espèce type du genre *Nanovirus* est le *Subterranean clover stunt virus*, qui infecte de nombreuses légumineuses (Grylls and Butler, 1959), tout comme la plupart sinon la totalité des autres espèces du genre (Grigoras et al., 2014) et, donc tout comme le *Faba bean necrotic stunt virus* (FBNSV), notre modèle d'étude décrit en détail ci-après. De récentes prospections en Afrique (Abraham et al., 2010, 2012), en Europe (Grigoras et al., 2010a, 2014) et en Asie (Kumari et al., 2010) ont révélé de nombreuses nouvelles souches ou espèces de nanovirus soulignant ainsi le caractère émergent de ce genre. La répartition géographique actuelle des nanoviridae est représentée en figure 5.



**Figure 5: Répartition géographique mondiale des nanovirus (jaune) et des babuvirus (rouge)**

Les pays au sein desquels à la fois des espèces appartenant aux genres *Nanovirus* et *Babuvirus* ont été décrites sont représentés en vert.

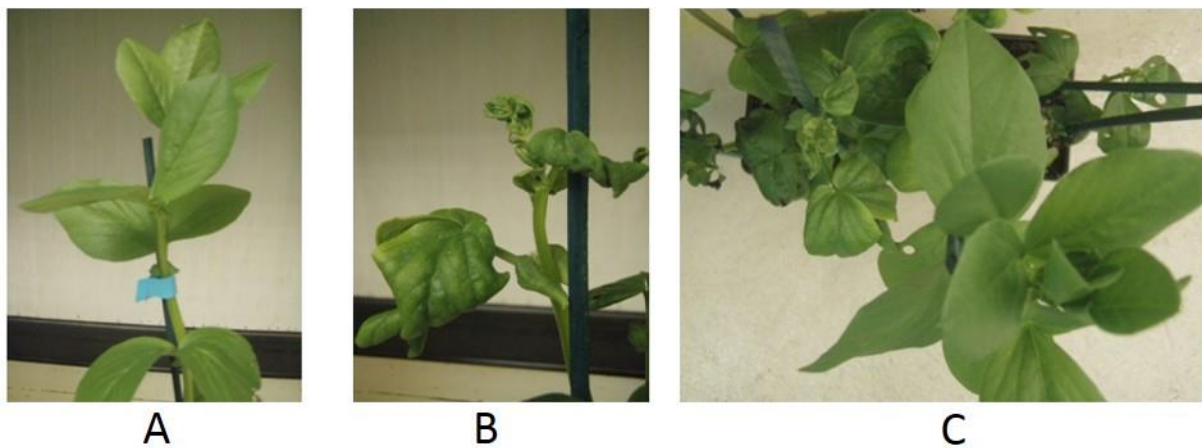
### *2.3. Le Faba bean necrotic stunt virus comme modèle d'étude*

Le *Faba bean necrotic stunt virus* ou FBNSV a été choisi comme modèle d'étude des virus multipartites dans la mesure où ce virus était le seul nanovirus pour lequel des clones infectieux transmissibles étaient disponibles au début du projet de thèse (Grigoras et al., 2009).

Le FBNSV a été isolé pour la première fois sur *Vicia faba* en Ethiopie en 2009 (Grigoras et al., 2009). Comme les autres membres connus de son genre, il s'agit d'un virus émergent causant des pertes de rendement chez les légumineuses. L'étendue de sa distribution géographique ainsi que celle de son spectre d'hôte sont encore inconnues. A ce jour, sans même y avoir été recherchés de manière systématique, des isolats de ce virus ont été découverts en Afrique (Abraham et al., 2010, 2012; Grigoras et al., 2009) et en Europe (Grigoras et al., 2014) sur des plants de fève (*Vicia faba*), de lentille (*Lens culinaris*) et de vesce (*Vicia sativa*).

Le spectre d'hôte et les pucerons vecteurs de ce virus pourraient être en partie similaires à ceux du *Faba bean necrotic yellows virus* (FBNYV), avec lequel le FBNSV partage près de 75% d'homologie globale (Grigoras et al., 2009). Suite à son inoculation à une plante saine par un puceron, le FBNSV peut se multiplier et envahir le système phloémien, provoquant après 10 à 20 jours un arrêt du développement chez la fève, ainsi que l'enroulement et le jaunissement de

ses feuilles (Figure 6). L'arrêt de croissance de la plante n'est pas un symptôme généralisable, tout du moins en laboratoire, à toutes les espèces végétales infectées par le FBNSV ; par exemple, des plantes comme le pois (*Pisum sativum*) et la luzerne (*Medicago truncatula*) continuent de croître après infection. Par contre, le FBNSV provoque chez toutes les espèces de légumineuses que nous avons testées (pois, lentille, fève, luzerne, haricot et cornille) un enroulement puis un jaunissement des feuilles apicales.



**Figure 6: Symptômes causés par le *Faba bean necrotic stunt virus* sur la fève**

(a) fève saine et (b) fève infectée par le *Faba bean necrotic stunt virus*, 25 jours après inoculation par pucerons. La photo (c) montre la différence de développement entre ces deux plantes vues d'en haut.

#### *2.4. Transmission du FBNSV*

A l'instar des autres membres de sa famille, le *Faba bean necrotic stunt virus* est transmis de façon circulante non-propagative par pucerons (Grigoras et al., 2009). Suivant ce modèle, lorsqu'un puceron vecteur s'alimente sur une plante infectée, il ingurgite des particules virales à partir du phloème.

Une fois ingérées, les particules traversent la paroi intestinale, puis diffusent dans l'hémolymphe jusqu'aux glandes salivaires, où elles transitent et parfois s'accumulent avant d'être inoculées avec la salive à de nouvelles plantes hôtes (Blanc et al., 2014). Ce qui distingue la transmission circulante non-propagative de la transmission circulante propagative, c'est que le virus traverse les barrières cellulaires de son vecteur (intestin et glande salivaire) sous la forme de particules virales, sans jamais s'y répliquer (ni même être décapsidé et transcrit).

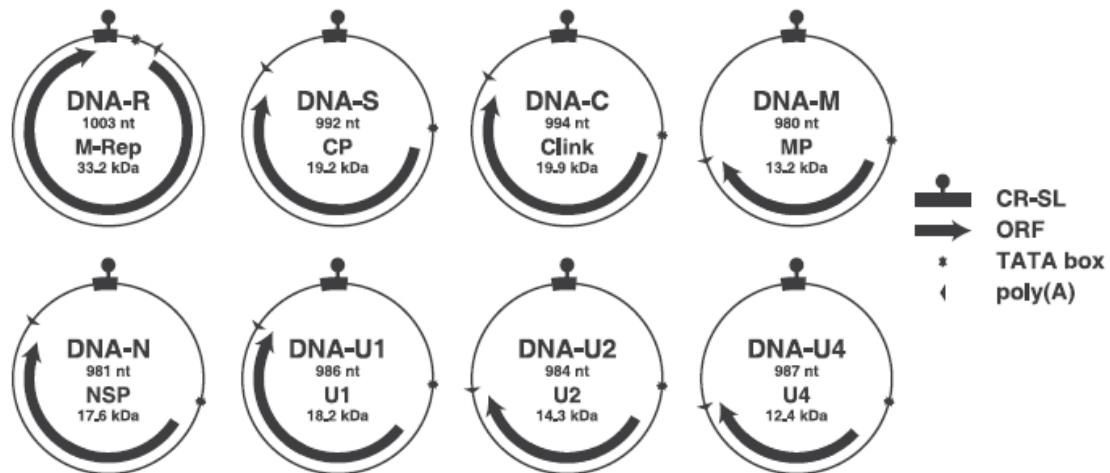
D'après une étude menée par Franz et collaborateurs (Franz et al., 1999), le FBNYV nécessiterait, outre la protéine de capsid, un facteur assistant viro-codé (helper component, HC) pour sa transmission. La découverte de ce HC a tiré parti d'un isolat non-transmissible du FBNYV, qui pouvait malgré tout être efficacement transmis par pucerons lorsque ceux-ci étaient nourris sur des plantes co-infectées par un isolat transmissible (Franz et al., 1999). La complémentation de la transmission a même pu être obtenue séquentiellement, lorsque les pucerons étaient d'abord nourris sur des plantes infectées par l'isolat transmissible, puis ensuite sur d'autres infectées par l'isolat non-transmissible. Dans la même étude, les auteurs ont montré que le virus purifié à partir de plantes infectées par l'isolat transmissible ne pouvait pas être efficacement acquis puis transmis par pucerons, qu'il soit ingéré au travers d'une membrane de parafilm ou injecté directement dans l'hémocèle des pucerons. Ceci semble démontrer deux choses : i) le HC est perdu durant la purification du virus, et ii) le HC joue un rôle dans le contrôle du passage du FBNYV depuis l'hémocèle (où le virus est injecté) vers les glandes salivaires et la salive.

Depuis cette étude, et donc depuis plus de quinze ans, le segment du génome viral codant pour ce facteur n'a pas encore été déterminé. Il est intéressant de signaler que les HC sont très fréquents dans la transmission des virus non-circulants, comme par exemple les caulimovirus et les potyvirus (Blanc et al., 2014), mais que leur description dans le contexte de la transmission circulante est unique. En quoi ce phénomène spécifique aux nanovirus indique une déviation par rapport au modèle canonique de la transmission circulante non-propagative est une question qui n'a pas été abordée.

## *2.5. Information génétique du FBNSV*

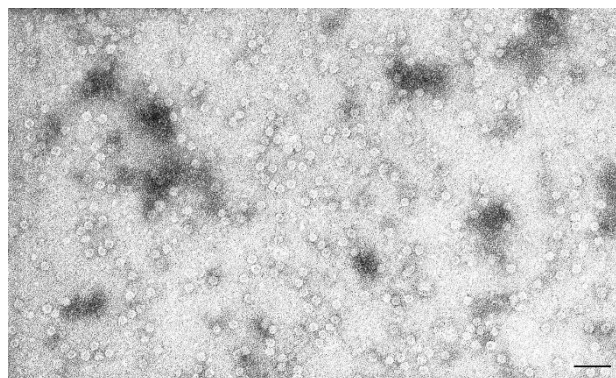
L'information génétique de ce nanovirus est répartie sur huit molécules d'ADN circulaire simple brin, toutes d'une taille comprise entre 980 et 1003 nucléotides (Figure 7). Chacune de ces molécules comporte deux régions communes : la région contenant l'origine de réplication ou région de la tige-boucle (CR-SL) et une région commune majeure (CR-M). Ces molécules codent chacune pour un seul gène et donc pour une seule protéine. Elles sont chacune encapsidées individuellement dans une particule icosaédrique de 18–20 nm de diamètre de symétrie T=1 et, toutes les particules étant composées de soixante copies de la même protéine de capsid, elles sont *a priori* identiques quel que soit le segment génomique contenu (cf discussion du deuxième article). Une image des particules virales est disponible en figure 8 (le

protocole de purification du virus est disponible en annexe I). La protéine de capsid étant requise pour l'infection systémique des plantes, on pense que le mouvement de l'information génétique à longue distance s'effectue sous forme encapsidée.



**Figure 7: Organisation génomique du FBNSV**

Les huit segments génomiques sont représentés sous forme circulaire et désignés selon la nomenclature actuelle. Les flèches représentent les cadres ouverts de lecture. Les tailles des segments et de leurs protéines correspondantes sont indiquées respectivement en nucléotides (nt) et en kDa. CP : protéine de capsid, MP : protéine de mouvement, NSP: protéine de navette noyau-cytoplasme, M-Rep : Master Rep, U1-U2-U4 : protéines aux fonctions inconnues, CR-SL : région commune de la tige-boucle. Reproduit de (Grigoras et al., 2009)



**Figure 8: Particules virales de *Faba bean necrotic stunt virus***

Image de microscopie électronique à transmission prise par D. Gargani. L'échelle correspond à 50 nm.

## 2.6. Fonction des 8 protéines du FBNSV et importance de celles-ci dans l'infection de la fève

Les protéines des nanovirus ainsi que leurs fonctions ont été déterminées en grande partie par homologie avec celles des géminivirus (Timchenko and Bernardi, 2007). Cinq des huit segments ont des fonctions au moins partiellement connues : le segment C code pour la protéine Clink qui interfère avec la régulation du cycle cellulaire, notamment en interagissant avec la protéine cellulaire du rétinoblastome, pouvant ainsi enclencher la phase de synthèse d'ADN. Cette interaction serait à l'origine d'une augmentation du taux de réplication des molécules virales (Timchenko and Bernardi, 2007); le segment M code pour la protéine de mouvement essentielle pour le mouvement du virus de cellule-à-cellule et à longue distance; le segment N code pour une protéine qui, à l'instar de la protéine NSP des géminivirus, permettrait le transport de l'ADN nouvellement synthétisé vers le cytoplasme (nuclear shuttle protein); le segment R code pour la Master Rep, protéine qui initie la réplication de tous les segments; et enfin le segment S code pour la protéine de capsid. Trois segments supplémentaires, U<sub>1</sub>, U<sub>2</sub> et U<sub>4</sub>, codent pour des protéines aux fonctions qui restent à ce jour totalement inconnues (Grigoras et al., 2009) (Figure 7).

D'après une étude menée par Timchenko (2006) sur le *Faba bean necrotic yellows virus*, un minimum de cinq segments comprenant les molécules M, R, S, U<sub>1</sub> et U<sub>2</sub> est nécessaire pour obtenir une plante symptomatique dans les conditions de laboratoire. Cependant, quand tous les autres segments sont présents, les composants U<sub>1</sub> ou U<sub>2</sub> ne sont pas indispensables pour induire la maladie. Il semblerait donc qu'il y ait plusieurs compositions au moins partiellement fonctionnelles pour ce génome, ouvrant la voie à des régulations complexes qui restent entièrement à découvrir. En outre, si le segment N n'est pas indispensable à l'infection de la fève, il est toutefois nécessaire lors de la transmission par vecteurs et pourrait ainsi représenter le HC décrit dans la section 2.3 (B. Gronenborn, communication personnelle). Par ailleurs, bien que le segment C ne soit pas non plus nécessaire dans le développement de symptômes chez la fève, il semble être néanmoins positivement sélectionné au cours de l'infection (Timchenko, 2006).

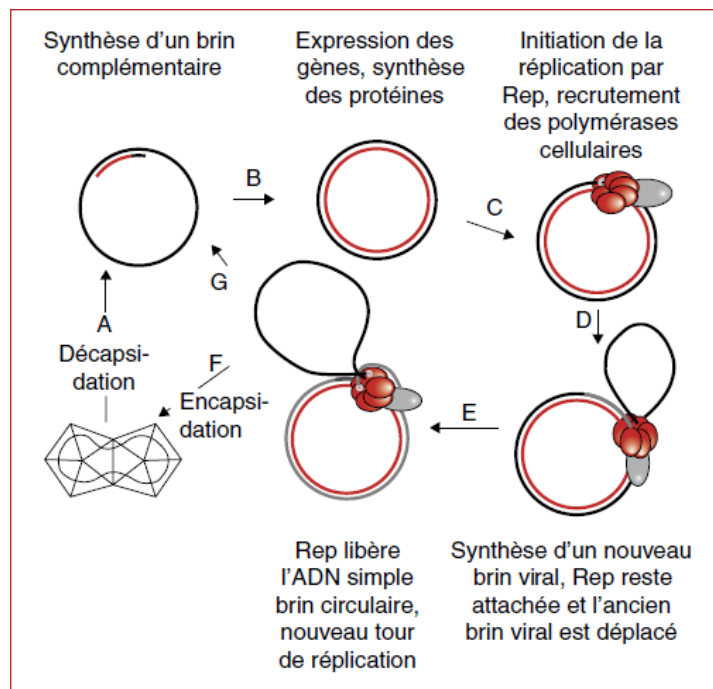


## 2.7. Réplication du FBNSV

Comme tous les autres membres de cette famille, le FBNSV est supposé être restreint au phloème (Timchenko and Bernardi, 2007) ; il est ainsi supposé se répliquer au sein des cellules compagnes et peut-être également au sein des cellules parenchymateuses du phloème.

Les nanovirus se répliquent dans le noyau des cellules infectées selon un mécanisme dit de « cercle roulant » initié par la protéine virale Master Rep (Figure 9). Les détails de ce mécanisme sont décrits par homologie avec ceux étudiés de manière plus approfondie chez les géminivirus (Gronenborn, 2004). Après décapsidation, l'ADN simple brin est converti en double brin par synthèse d'un brin complémentaire. La présence de petites molécules d'ADN encapsidées complémentaires de la région commune CR-M servirait d'amorce pour cette étape tout du moins pour le *Banana bunchy top virus* (Hafner et al., 1997). Cette forme double brin est une forme intermédiaire qui servirait également à la transcription. Suite à la formation de cette forme double brin, la Master Rep reconnaît la présence d'itérons situés au niveau de l'origine de réplication, séquences qui joueraient un rôle majeur dans la spécificité de la réplication des nanovirus (Timchenko and Bernardi, 2007). L'ADN nouvellement synthétisé est coupé par la Master Rep entre les nucléotides T et A dans la séquence conservée de la boucle AG/TTATTAC. Le 3'OH sert alors d'amorce pour la synthèse d'un brin d'ADN dans le sens viral par une ou plusieurs polymérases de l'hôte en utilisant le brin circulaire complémentaire comme matrice. Une fois le tour de réplication terminé, la Master Rep coupe entre les nucléotides T et A de ce nouveau brin viral. Cette coupure est suivie d'une réaction de transestérification entre l'extrémité 5' de l'adénine et l'extrémité 3'-OH de la thymine, libérant ainsi l'ADN simple brin nouvellement synthétisé. Cet ADN simple brin peut alors soit être encapsidé soit de nouveau transformé en double brin (Timchenko and Bernardi, 2007).

L'origine de réplication des 8 segments du FBNSV est supposée être au sein d'une région commune. Cette région est identique à 100% entre les segments R, S, M, N, U<sub>2</sub> et U<sub>4</sub> et entre les segments C et U<sub>1</sub>. Ces deux groupes partagent 83,6% d'identité au niveau de cette région (Grigoras et al., 2010b). Une régulation différentielle de la réplication pourrait donc résulter *a priori* en une abondance différente des segments entre ces deux groupes.



**Figure 9: Réplication des virus à ADN simple brin selon le mécanisme du cercle roulant**

Après décapsidation (étape A), l'ADN simple brin est converti en ADN double brin par extension d'amorce (étape B), le brin néosynthétisé est représenté en rouge. La transcription et la traduction des gènes viraux permettent l'induction de la réplication de l'ADN viral sous forme répllicative double brin. La protéine Rep (un oligomère formé de plusieurs molécules, en rouge) introduit une coupure au niveau de l'origine de réplication (sur un brin viral, en noir) et recrute des enzymes cellulaires de polymérisation de l'ADN (en gris) (étape C). Les enzymes cellulaires vont allonger le brin coupé donnant le brin néosynthétisé (en gris, étape D). La protéine Rep, accrochée à l'extrémité 5' de l'ADN par une tyrosine (le point blanc), fermera le cercle d'ADN déplacé (étape E). L'ADN simple brin pourra être encapsidé (icosaèdre simple pour les nanovirus, étape F) ou à nouveau transformé en double brin (étape G). Reproduit de (Timchenko and Bernardi, 2007)

## 2.8. Evolution du FBNSV

Différentes études sur les taux de mutation, de recombinaison et de réassortiment semblent suggérer que le FBNSV évoluerait rapidement. En effet, un taux de substitution de  $1,78 \times 10^{-3}$  par nucléotide par an a été déterminé par (Grigoras) et ses collègues en 2010, plaçant ainsi le FBNSV parmi les virus à ADN simple brin –que ce soit les virus de plantes ou d'animaux– évoluant le plus rapidement (Grigoras et al., 2010b). Ce taux de substitution diffère entre les segments. Ainsi les segments R et S du FBNSV semblent être les segments les moins variables tandis que le segment N présente le plus fort taux de substitutions.

En plus de ce taux élevé de substitutions, la recombinaison semble être une force majeure de l'évolution des *Nanoviridae*. En effet, des études récentes ont montré que des événements de recombinaison avaient lieu relativement fréquemment chez les nanovirus et les babuvirus. Ainsi, des recombinaisons à la fois inter et intra-segments ont été démontrées, révélant des hotspots de recombinaison au niveau des deux régions communes partagées par les différents segments (Grigoras et al., 2014; Stainton et al., 2012). Les événements répétés de recombinaison entre les différents segments pourraient être à l'origine d'une homogénéisation des régions non-codantes des divers segments (Hughes, 2004). Il y aurait ainsi un processus d'évolution concertée entre les régions non-codantes des différents segments des nanovirus, suggérant ainsi que certaines régions non-codantes seraient soumises à des contraintes fonctionnelles (Hughes, 2004). Par ailleurs, une étude menée sur le *Banana bunchy top virus* a révélé des taux de recombinaison différents entre les six segments composant son génome : le segment R étant le plus conservé et le segment U3 le plus sujet à la recombinaison (Stainton et al., 2012).

Enfin, des événements de réassortiment jalonnent également l'évolution des *Nanoviridae*. Certaines similarités ont pu être dressées concernant les événements de réassortiment chez les nanovirus et les babuvirus. Tout d'abord, un seul segment génomique est généralement échangé.

Cet échange semble être relativement fréquent pour certains segments comme le segment M qui montre le plus fort taux de réassortiment (respectivement 37,5% et 25% chez les babuvirus et les nanovirus) tandis qu'il semble être plus rare pour les segments C et S (Grigoras et al., 2014; Hu et al., 2007; Stainton et al., 2012). Ces différences pourraient être, au moins en partie, le résultat de contraintes biologiques concernant les capacités des différents segments et/ou des protéines qu'ils codent à fonctionner normalement dans le contexte d'un génome étranger (Grigoras et al., 2014).

Les différents segments composant le génome des *Nanoviridae* seraient ainsi susceptibles d'évoluer rapidement et à des rythmes différents.

**CHAPITRE I :**

**LE NOMBRE DE COPIES DE GENES EST REGULE  
DIFFERENTIELLEMENT CHEZ UN VIRUS  
MULTIPARTITE**

# Chapitre I : Le nombre de copies de gènes est régulé différemment chez un virus multipartite

---

## 1. Contexte et objectifs

Chez un virus monopartite ou segmenté, le succès de l'infection d'une cellule dépend du (des) variant(s) génétique(s) entrant. Si un seul variant entre par cellule et que ce variant est défectif pour au moins une fonction telle que la réplication, l'encapsidation, etc... l'infection sera avortée. En revanche, si ce variant défectif pénètre dans une cellule dans laquelle au moins un autre variant complémentaire (ou sauvage) est présent, les deux variants complémentaires pourront se répliquer, être encapsidés et infecter ainsi de nouvelles cellules. Le succès de l'infection dépend donc directement (i) de la nature du variant entrant et (ii) du nombre de variants entrant et se répliquant (ici dénommé multiplicité d'infection cellulaire ou MOI). La nature du variant entrant dépend à la fois de sa capacité intrinsèque à pénétrer dans la cellule et de sa disponibilité au sein de la population c'est-à-dire de sa fréquence au sein de la population virale.

Appliqué aux virus multipartites, ce raisonnement (bien que très simplifié) nous fait prendre conscience de l'importance de la MOI (abordée dans le chapitre 4) et du poids des fréquences des différents segments génomiques.

A l'instar des geminivirus, il est supposé que le mouvement systémique des nanovirus s'effectue sous forme encapsidée (Timchenko and Bernardi, 2007). Toutes les particules virales étant a priori identiques dans une plante infectée par un nanovirus, car composées de la même protéine de capsid et renfermant des segments de taille similaire, tous les segments ont a priori la même capacité à pénétrer dans une cellule donnée. Une question qui se pose alors est de savoir si chaque segment est représenté en quantité équivalente au sein de la population ou s'il existe une distorsion en faveur de certains segments ? Si tel est le cas, cette distorsion de fréquence est-elle répétable ou complètement aléatoire ? Ces questions paraissent importantes à aborder puisqu'une distorsion des quantités en faveur d'un ou plusieurs segments pourrait lui (leur) permettre d'avoir plus de chances de pénétrer au sein d'une cellule et donc d'être répliqué(s) et de pouvoir ainsi se propager. A contrario, des segments présents en fréquence

plus faible risqueraient d'être perdus au cours de l'infection. Si ces segments rares sont indispensables à l'infection, alors les nombreuses cellules où ils sont absents, et où d'autres segments sont entrés, avorteront l'infection, résultant en une perte (ou « gaspillage ») d'une portion importante de la descendance virale.

On peut présupposer que pour éviter la perte des trois segments « facultatifs » à l'infection de la fève dans des conditions de laboratoire, les segments C, N et U4, seront largement surreprésentés par rapport aux autres, augmentant ainsi leur chance de présence dans chaque cellule, avec les segments obligatoires plus rares.

Ce sont ces questions basées sur des concepts très simples qui ont initialement motivé la mise en place des expérimentations décrites dans ce chapitre, même si les résultats nous ont amenés à des conclusions plus complexes sur le rôle du contrôle du nombre de copies des gènes dans l'expression du génome des virus multipartites. Nous avons décidé d'étudier les fréquences relatives des huit segments du FBNSV au cours de l'infection de son hôte, la fève. Les résultats obtenus sont présentés dans l'article ci-après, publié en 2013.

## **ARTICLE 1**

# **GENE COPY NUMBER IS DIFFERENTIALLY REGULATED IN A MULTIPARTITE VIRUS**

Anne Sicard, Michel Yvon, Tatiana Timchenko, Bruno Gronenborn, Yannis Michalakis, Serafin Gutierrez & Stéphane Blanc

Publié dans *Nature communications* le 5 août 2013

ARTICLE

Received 13 Feb 2013 | Accepted 5 Jul 2013 | Published 5 Aug 2013

DOI: 10.1038/ncomms3248

# Gene copy number is differentially regulated in a multipartite virus

Anne Sicard<sup>1</sup>, Michel Yvon<sup>1</sup>, Tatiana Timchenko<sup>2</sup>, Bruno Gronenborn<sup>2</sup>, Yannis Michalakis<sup>3</sup>, Serafin Gutierrez<sup>1,\*</sup> & Stéphane Blanc<sup>1,\*</sup>

Multipartite viruses have a genome divided into several nucleic acid segments, each encapsidated separately. An evident cost for these viral systems, particularly if some segments are rare, is the difficulty of gathering one copy of each segment to ensure infection. Here, we investigate the segment frequency-related cost by monitoring the copy number of the eight single-gene segments composing the genome of a plant nanovirus. We show that some viral genes accumulate at low frequency, whereas others dominate. We further show that the relative frequency of viral genes impacts both viral accumulation and symptom expression, and changes specifically in different hosts. Earlier proposed benefits of viral genome segmentation do not depend on the segments' frequency and cannot explain our observations. We propose that the differential control of gene/segment copy number may represent an unforeseen benefit for multipartite viruses, which may compensate for the extra costs induced by the low-frequency segments.

<sup>1</sup>INRA, UMR BGPI INRA-CIRAD-SupAgro, Cirad TA-A54/K, Campus International de Baillarguet, 34398 Montpellier, France. <sup>2</sup>Institut des Sciences du Végétal, CNRS, 91198 Gif-sur-Yvette, France. <sup>3</sup>UMR MIVEGEC 5290, CNRS-IRD-UMI-UM2, IRD, 911 Avenue Agropolis, B.P. 64501, 34394 Montpellier, France. \* These authors contributed equally to the work. Correspondence and requests for materials should be addressed to S.B. (email: blanc@supagro.inra.fr).



Viruses exhibit a wide diversity of genome organizations, mechanisms of replication and gene expression strategies. Viral genomes can be composed of single- or double-stranded molecules of RNA or DNA, expressing genes from mono- or polycistronic mRNAs, which can be sub-genomic or full-genome-length mRNAs. Within this diversity, the so-called segmented viruses—DNA or RNA viruses infecting bacteria, animals or plants<sup>1</sup>—have a genome composed of more than one nucleic acid molecule, with 2 to 12 genome segments depending on the viral species. Among the segmented viruses, the most puzzling biological systems are found in the multipartite viruses<sup>2</sup> (described in plants and fungi), where the genome segments are not co-packaged in a single viral particle but are encapsidated individually, forming an ensemble of particles that must be transmitted together in order to infect new cells.

The functioning of these biological systems has long intrigued virologists and evolutionary biologists, who have striven to understand how such systems have evolved by modelling the parametric space in which the cost/benefit ratio is positive. Proposed advantages of genome compartmentalization are greater stability of smaller-sized segments<sup>3</sup>, a potential faster replication of small genomic segments<sup>4</sup>, or the increased genome shuffling that could result from genome segmentation and ‘multipartitism’<sup>5–10</sup>.

An obvious cost to genome compartmentalization is the necessity to either package together all segments, for segmented viruses, or to ensure the co-entry of an ensemble of virus particles containing at least one copy of each genomic segment<sup>3,6,11</sup>, for multipartite viruses. This cost increases dramatically with the number of segments constituting the viral genome, as amply discussed in the related literature, and recently reinvestigated in a theoretical study<sup>12</sup>.

One aspect of multipartite virus biology that has not been investigated, empirically or theoretically, is the potential regulation of the relative frequencies of different genome segments. Indeed, all else being equal, the probability of infecting a host cell successfully, that is, the probability that all genome segments are transmitted in at least one copy, would be maximized if all segments occurred at equal frequency; deviations from this situation would increase the cost of genome compartmentalization.

To address this knowledge gap, we tested whether the predicted situation of minimum cost is indeed observed in populations of the multipartite Faba bean necrotic stunt virus (FBNSV), or whether additional unknown constraints drive the frequency of different segments to different values. FBNSV is a member of the genus *Nanovirus* in the family *Nanoviridae*<sup>2</sup>. The FBNSV genome is composed of eight circular single-stranded DNA molecules of about 1 kb (segments C, M, N, R, S, U1, U2 and U4), each encoding a single gene and encapsidated separately<sup>13</sup>. The genome organization and function of each gene have been discussed previously in detail<sup>13,14</sup>. Briefly, C encodes a protein that interferes with the cell cycle and is a general enhancer of DNA replication; M produces the movement protein allowing viral cell-to-cell and long-distance transport within host plants; N encodes a nucleo-cytoplasmic shuttle protein of unknown biological function; R encodes the protein that governs replication of all viral genome segments by an unknown cellular DNA polymerase; S encodes the capsid protein encapsidating the different viral genome segments individually; Finally, U1, U2 and U4 encode proteins of unknown function. The infection of host plants by FBNSV as well as by other nanoviruses is restricted to phloem tissues and, because the coat protein appears mandatory for intra-plant movement<sup>14</sup>, it is generally assumed that individual segments move in the form of virus particles throughout the plant vasculature. In many host plants, member species of the genus *Nanovirus* induce severe stunting, often totally inhibiting further plant growth<sup>15</sup>.

Here we report monitoring of the relative frequency of the eight single-gene-encoding segments constituting the genome of FBNSV during infection of host plants. We show that the various genes accumulate with very different frequencies within individual host plants. Moreover, starting from distinct inoculation conditions and relative segment frequencies, the viral system reproducibly reaches a comparable state, designated the ‘setpoint genome formula’, where each genome segment (or gene) accumulates to its specific relative frequency. Our results further suggest that the setpoint genome formula corresponds to a state resulting in increased viral accumulation and enhanced symptoms for the FBNSV system. Finally, we show that the FBNSV genome formula reaches a different setpoint in distinct host plant species. Our results hence indicate that multipartite viruses can differentially control the copy number of their different genes (or segments). We propose that this is an unforeseen potential benefit for multipartite viruses.

## Results

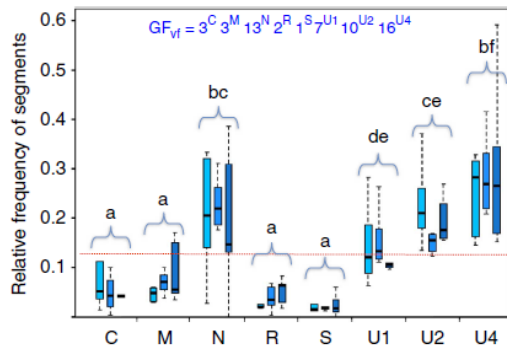
**Defining the genome formula of FBNSV.** A set of 50 *Vicia faba* plants was agro-inoculated with equal amounts of each of the eight genomic segments of the FBNSV genome<sup>13</sup>. The frequency of each segment, relative to total viral DNA, was estimated in the uppermost leaf of each plant by real-time quantitative PCR (Q-PCR) ~4 weeks post-infection, when the infection had fully developed and arrested further development of the plant. In three independent experimental replicates, despite a large variance between individual plants, each segment accumulated reproducibly at a specific median frequency (Fig. 1 and Supplementary Table S1), which ranged from ~2 to 30% (segments S and U4, respectively).

The reproducible accumulation of specific relative copy numbers for different viral genes poses the question of how best to describe the genome of a multipartite virus during infection of its host. To propose an easy and biologically sound description, and for practical use in this report, we define the term ‘genome formula’ by associating a relative copy number to each genomic segment. Thus, when applied to the FBNSV infecting *V. faba* plants under the experimental conditions shown in Fig. 1, the genome formula is  $3^C 3^M 13^N 2^R 1^S 7^{U1} 10^{U2} 16^{U4}$ , where numbers indicate the median copy number of each segment, rounded to the nearest integer, relative to that of the least abundant segment, and superscript letters indicate the name of the segment.

The use of mean frequency values (instead of median) gives a very similar genome formula. It must be noted, however, that some segments are dispensable for infection under laboratory conditions (see Methods for details) and that they can occasionally be lost upon artificial agroinoculation of plants<sup>13,14</sup>. Hence, we prefer the use of median values because they are less affected by these occasional segment losses than is the mean.

As all the results presented in this report were obtained with total viral DNA extracted from infected plant tissues, we assessed whether the estimated genome formula was similar for DNA encapsidated into virus particles. In independent plant sets, we compared the relative frequency of segments in total versus encapsidated DNA extracted from the same leaf samples (see Methods). The results shown in Supplementary Fig. S1 demonstrate that the genome formula in total DNA extractions closely reflects that in encapsidated DNA.

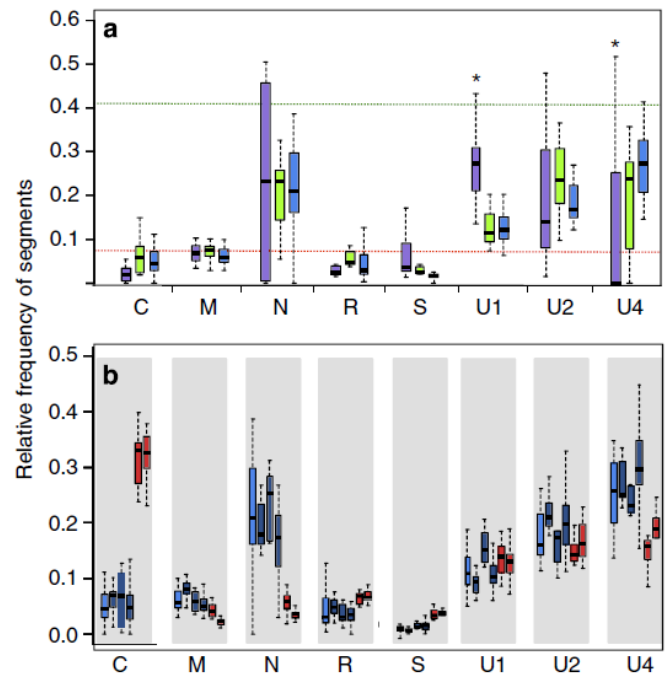
**The genome formula does not depend on initial conditions.** We first tested whether the genome formula in systemically infected plants depends on the initial frequency of the respective segments at inoculation. For this, we prepared two inocula with different segment proportions. Figure 2a shows the results of



**Figure 1 | The genome formula of FBNSV.** The relative frequency of each genome segment of FBNSV was evaluated by Q-PCR in systemically infected leaves of *V. faba* (Vf). The name of each segment is indicated below the graph. Results from three-independent replicates, named Vf-Agr1, Vf-Agr2 and Vf-Agr3, are plotted from left to right (with increasing shades of blue) for each segment. The number (*n*) of plants successfully infected and analysed 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 tests (all *P*-values are given in the Supplementary Table S1). The genome formula of FBNSV in *V. faba*, noted  $GF_{Vf}$  above the graph, was calculated with the data pooled from Vf-Agr1-3.

two-independent plant sets infected with either a fivefold excess of segment S (S + +) or of segment C (C + +). As previously, the uppermost leaf level was analysed for each plant set when the infection had developed fully and provoked the arrest of plant growth. The frequency of the different FBNSV segments in these two plant sets proved very similar to that in plants inoculated with equal quantities of all segments. The only significant differences detected were for segments U1 and U4 in set S + + (Fig. 2a). The very low median frequency observed for segment U4 in the plant set S + + is due to the 9 out of 12 plants where segment U4 was lost at inoculation—a situation sometimes observed when segments that are dispensable for infection under laboratory conditions are inoculated at too low a frequency. In the C + + set, only 2 out of 7 plants had lost U4.

To further assess whether inoculation mode could affect the genome formula of FBNSV establishing in *V. faba*, we switched from agro-inoculation to aphid transmission, that is, the natural transmission mode. We fed aphids for 3 days on four, randomly chosen, infected plants. Aphids were then pooled, and used to inoculate young test plants for another 3-day period (five aphids per test plant). Four weeks later, when plants showed symptoms of systemic infection and stopped developing new leaves, the uppermost leaf level was analysed by Q-PCR. The results of three-independent replicates of the experiment (noted as Vf-At1, Vf-At2 and Vf-At3 in Fig. 2b) show that the FBNSV genome formula obtained after aphid inoculation is close to that obtained after agro-inoculation (*cf.* genome formulae in *V. faba* in Figs 1 and 2b). Whatever the segment considered, no significant



**Figure 2 | Factors affecting the genome formula of FBNSV.** (a) The relative frequency of segment S (S + +, purple) or segment C (C + +, green) was increased fivefold in the inocula. The resulting frequencies in the uppermost leaf of infected plants are summarized in the box-plots (black bar is the median and vertical dotted line delineates 1.5 times the distance between the 25th and 75th percentile of the distribution). For comparison, data from Vf-Agr1, Vf-Agr2 and Vf-Agr3 (Fig. 1) were pooled and plotted in blue. The number (*n*) of infected plants analysed in each set is *n* = 12 and *n* = 7 for S + + and C + +, respectively. The upper green dotted line indicates the frequency of S or C in the inoculum, whereas the lower dotted red line indicates the frequency of all other segments. Between plant sets, significant differences (\*) were detected by an analysis of variance (ANOVA) test (plant set effect on the frequency of U1 and U4:  $F_{1, 27} = 33$  and 9.04, *P*-values =  $3.6 \times 10^{-6}$  and 0.0045, respectively). (b) The relative frequency of FBNSV segments in *V. faba* plants inoculated by aphids is plotted in dark blue and compared with that in agro-inoculated plants (light blue on the left, same Vf-Agr1-3 data as in a). Three-independent aphid inoculation experiments named Vf-At1 (*n* = 10), Vf-At2 (*n* = 6) and Vf-At3 (*n* = 32) are, respectively, plotted from left to right. The frequency of the segments was not affected by the mode of inoculation, as demonstrated by ANOVA tests detailed in Supplementary Table S2. The segments' frequency in *M. truncatula* plants inoculated by aphid was estimated in two plant sets, Mt-At1 (right, *n* = 12) and Mt-At2 (left, *n* = 13), and is plotted in red. The host species had a significant effect on all segments except U1, as shown by ANOVA tests detailed in Supplementary Table S4. The genome formulae of FBNSV in *V. faba* ( $GF_{Vf}$ ) and in *M. truncatula* ( $GF_{Mt}$ ) were calculated by pooling all Vf-At and Mt-At data sets, respectively, and are markedly distinct:  $GF_{Vf} = 3^C 3^M 9^N 2^R 1^S 6^{U1} 11^{U2} 15^{U4}$  and  $GF_{Mt} = 12^C 1^M 1^N 2^R 2^S 5^{U1} 6^{U2} 7^{U4}$ .

frequency difference could be detected when comparing the two modes of inoculation (Fig. 2b and Supplementary Table S2).

Altogether, these results demonstrate that the various segments of the FBNSV genome reproducibly adjust to a specific relative frequency during systemic invasion of the host plant *V. faba*, whatever the initial conditions at inoculation. We propose to refer to these adjusted frequencies of the different segments as the 'setpoint genome formula'.

**The genome formula converges on a setpoint.** An intriguing observation from Figs 1 and 2 is the large variance in segment

frequencies observed among individual plants within all plant sets but around one single-dominant situation.

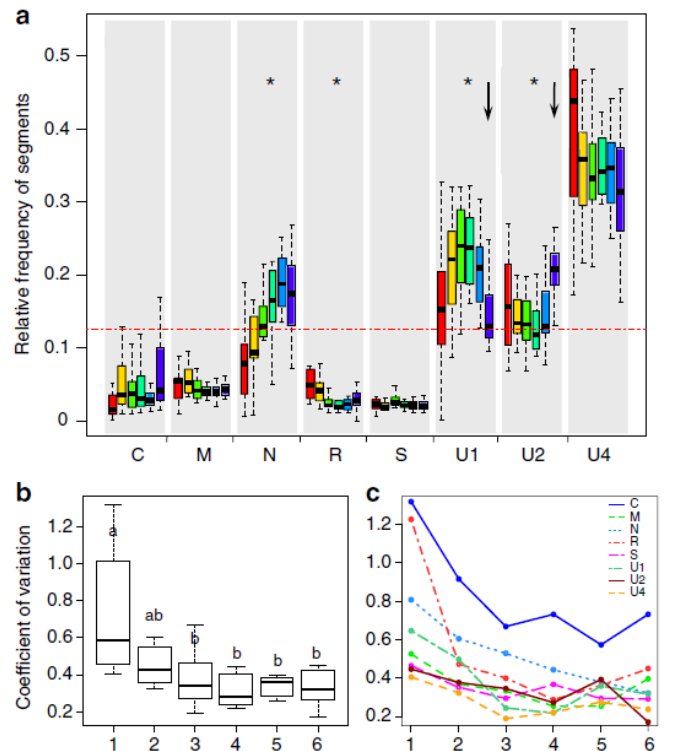
We further used the plant set with the highest number of infected plants, that is, the set Vf-At3 inoculated by aphids (Fig. 2b), to assess a possible temporal evolution of the genetic composition of within-host FBNSV populations. In this set, the plants were all inoculated at the second leaf stage but had their development arrested by virus infection at either the 5th ( $n = 2$ ), 6th ( $n = 16$ ), 7th ( $n = 10$ ) or 8th ( $n = 4$ ) leaf stage, suggesting potentially different kinetics of infection in different plants. To compare plants within a homogeneous sub-set, we focused on the 16 plants whose development was arrested at the six-leaf stage. As leaves appear consecutively during the development of viral disease until full appearance of symptoms provokes plant development arrest, lower leaves are likely to have been colonized by the virus earlier than upper leaves. A time series of the genetic structure of the virus population could thus be reconstituted by comparing the segment frequencies in the six successive leaf levels. Figure 3a shows that the median frequency did not change significantly across leaf levels for segments C, M, S and U4 (Fig. 3a and Supplementary Table S3), whereas N and R stabilized at leaf level 4, and only segments U1 and U2 were still significantly changing at later stages of infection (between leaf levels 4 and 6, Fig. 3a). This observation suggests that the median genome formula within a plant set stabilizes rapidly around the setpoint formula during disease progression.

An interesting follow-up question is whether the observed fluctuations of segment frequencies around the median values are maintained over time, or whether FBNSV populations within each individual plant tend to converge to the setpoint genome formula. In the same plant set as in Fig. 3a, we calculated the coefficient of variation of the frequency of all eight segments and found that it decreased significantly from lower to upper leaf levels (Fig. 3b). Noticeably, the same pattern could be observed for each individual segment (Fig. 3c), suggesting that all are submitted to frequency-dependent selection.

Frequency co-variation was detected in only 2 out of 28 possible pairs of segments (C-U1 and N-U4) in the data set presented in Fig. 3 (Spearman's rank correlation, for  $P$ -values for all pair wise combinations, see Supplementary Table S4), suggesting the absence of one major regulatory segment that would control the frequency of others.

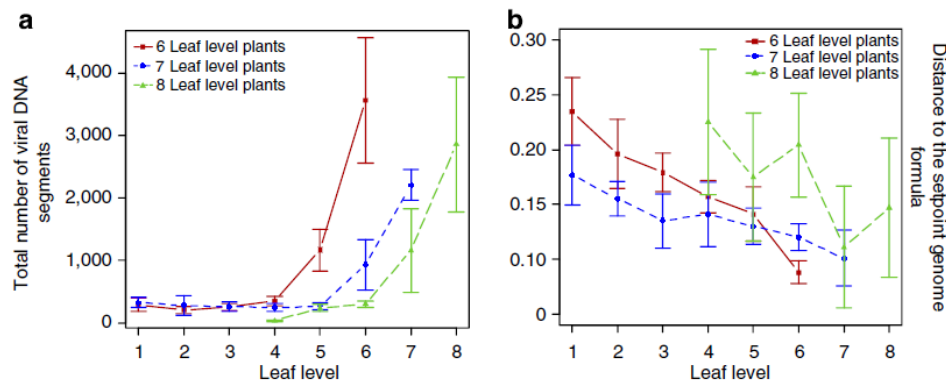
**The genome formula may affect virus accumulation and symptoms.** The previous observations indicate that, under these experimental conditions, the segment frequencies within each plant, and hence across plants, converge towards the setpoint genome formula. We thus asked whether the genome formula has any adaptive significance for FBNSV. To that end, we looked for a possible relationship between the proximity of the FBNSV population to the setpoint genome formula and viral accumulation *in planta*.

In the plant set Vf-At3, both virus accumulation and the distance of FBNSV populations from the setpoint formula were calculated in all leaves of plants with 6, 7 or 8 leaves (Fig. 4: blue, red and green lines, respectively). The viral load was calculated simply by summing the Q-PCR estimates of all eight segments (relative to a reference plant gene, see Methods). The distance to the setpoint genome formula was estimated by summing up the absolute difference between each segment frequency and its specific setpoint formula value<sup>16</sup> (see Methods, distance to the setpoint genome formula (dGF1)). The setpoint genome formula values used here to estimate the distance were calculated by pooling all independent *V. faba* plant sets shown in Figs 1, 2 and 3, except set Vf-At3.



**Figure 3 | Convergence of FBNSV within-host populations to the setpoint genome formula.** (a) The frequency of all FBNSV segments was analysed in 16 plants from Vf-At3 set (plant set used in Fig. 2b), which developed six leaf levels before the virus induced developmental arrest. For each box plot, the horizontal black bar represents the median of the distribution and the vertical dotted line delineates 1.5 times the distance between the 25th and 75th percentile of the distribution. For each segment, the results from leaf levels 1 to 6 are represented in different colours from left to right. The segments indicated with an asterisk accumulated at different frequency in different leaf levels (see analysis of variance (ANOVA) tests of the effect of leaf-levels for each of the eight segments in Supplementary Table S3). However, only the median frequency of U1 and U2 was still changing significantly at later stages of infection, between leaf 4 and leaf 6 (black arrows, Tukey HSD pair test,  $P$ -values = 0.014 and 0.00041, respectively). (b) Each box-plot represents the distribution of the coefficients of variation (CV) of the eight segments in a given leaf level among plants. The significant differences among leaf levels was confirmed by an ANOVA test:  $F_{5, 41} = 4.19$ ,  $P$ -values = 0.0036. Letters indicate significant differences between leaf levels, which were estimated by Tukey HSD pair tests ( $P$ -value between leaf levels 1-3, 1-4, 1-5 and 1-6 = 0.017, 0.0086, 0.049, 0.046, respectively). Black bars within box plots represent medians and vertical dotted lines delineate 1.5 times the distance between the 25th and 75th percentile of the distribution. (c) In successive leaf levels noted on the x axis, changes of the CV of each segment are represented by different coloured lines as indicated.

In the three sub-sets where plants developed 6, 7 or 8 leaves, respectively, before arresting development, several observations suggest a causal link between the distance of FBNSV populations to the setpoint genome formula, the rapid increase in viral load, and the inhibition of plant development: (i) in all plant subsets (all three curves: green, red and blue), the distance of the infecting FBNSV populations to the *V. faba* setpoint genome formula decreased monotonically as infection progressed (Fig. 4b), implying that, within each plant, the FBNSV population evolved continuously towards the setpoint formula; (ii) viral accumulation increased with time, as seen by leaf level (Fig. 4a); (iii) viral accumulation rose earlier in plant sets where the distance to the



**Figure 4 | Genome formula may impact on viral accumulation and plant development.** (a) Viral accumulation was measured in all leaves of plants that developed 6 (red square,  $n=16$ ), 7 (blue diamond,  $n=11$ ) or 8 (green triangle,  $n=4$ ) leaf levels. The scale on the left indicates the total number of DNA circles per leaf-sample accumulated at each leaf level. (b) The distance of FBNSV populations to the setpoint genome formula was calculated in all leaves of plants that developed 6 (red square), 7 (blue diamond) or 8 (green triangle) leaf levels. The scale on the left indicates the distance of FBNSV populations to the genome formula, calculated as described in Methods. Error bars are s.d.

genome formula reached its lowest value more rapidly; and (iv) viral accumulation rose earlier in plant sets where development was arrested earlier (Fig. 4a). In an alternative calculation, where the distance to the setpoint formula was weighted for all segments as the proportional deviation from the setpoint value (see dGF2 in Methods), the pattern observed in Fig. 4b was not qualitatively affected (Supplementary Fig. S2).

A direct correlation between viral accumulation and distance to the setpoint genome formula could not be detected in individual leaves of a comparable leaf level in the plant set Vf-At3 (Spearman's rank correlation coefficient at leaf level 4, 5 or 6:  $P$ -values = 0.087, 0.57, 0.97, respectively). Possible reasons for this are discussed below.

**The setpoint genome formula of FBNSV is host-dependent.** Finally, in order to test for a possible host effect on the setpoint genome formula, we used the aphid inoculation procedure described above to transmit FBNSV from infected *V. faba* plants to a set of plantlets of the related legume species *Medicago truncatula*. Symptoms of leaf yellowing and curling revealed the success of systemic infection within 10–25 days post-inoculation but, in contrast to the situation observed in *V. faba*, persistent FBNSV infection did not inhibit growth and new leaf development in *M. truncatula*.

Four weeks post-inoculation, whole leaflets were collected from the uppermost leaf level of each infected *M. truncatula* plant and processed for Q-PCR analysis of viral segment frequencies as above (for easy comparison, these results are shown in red in Fig. 2b). The experiment was repeated twice using different infected *V. faba* source plants, different aphid cohorts and different *M. truncatula* plantlets (Fig. 2b Mt-At1 and Mt-At2), and consistently demonstrated an effect of the host on the genetic structure of FBNSV populations. Except for U1, all genome segments accumulated at significantly different frequencies in the two host species (Fig. 2b and Supplementary Table S5). This resulted in markedly distinct setpoint genome formulae in *M. truncatula* ( $12^C_1M_1N_2R_2S_5U_1^6U_2^7U_4$ ) and in *V. faba* ( $3^C_3M_9N_2R_1S_6U_1^{11}U_2^{15}U_4$ ), where the largest differences pertained to segments C, N and U4.

## Discussion

In all organisms, gene expression is regulated through various distinct pathways, including transcriptional and post-transcriptional regulation of mRNAs, as well as translational control and

post-translational modification of proteins. Upstream of all these mechanisms, the regulatory role of the actual gene copy number (GCN) is becoming increasingly recognized. The GCN, and changes thereof (that is, copy number variation), of specific genes present within individual cells is thought to greatly affect gene expression in most (if not all) organisms<sup>17,18</sup>. For example, copy number variation has been associated with both control of development and genetic disease in organisms ranging from humans<sup>19</sup> to insects<sup>20</sup>, and with stress adaptation processed in protozoa<sup>21</sup> and bacteria<sup>18</sup>. The hijacking of host regulatory processes to fine-tune viral infection cycles is well known and has received enormous attention<sup>22–27</sup>. However, despite increasing evidence that copy number variation is a key regulatory process<sup>28</sup>, its potential exploitation by viruses has not been explored to date. This is particularly surprising given the ubiquitous impact of GCN on phenotypic expression<sup>28</sup>, on population genetics and evolution<sup>29</sup>, and when considering the enormous potential of such regulation in the diverse genome structures and expression strategies of viruses. We believe that regulation of GCN might explain the establishment of a setpoint genome formula of FBNSV during systemic host infection.

The mechanism actually driving adjustment of FBNSV populations to the setpoint genome formula in a given host remains enigmatic. A hypothesis (H1) that fits our results postulates both stochastic events at the segment (gene) level, and positive selection at the system (genome) level: at early steps of the plant infection, different cells could be infected by distinct ensembles of virus particles, among which the relative frequency of each segment would vary randomly (particularly if the number of particles infecting each individual cell is relatively small). Infecting ensembles with relative segment frequencies resulting in higher replication would produce more offspring. In such a scenario, if the eight FBNSV genome segments were replicated at the same rate and have similar chances to be transmitted to new cells, the pattern of relative segment frequencies established in newly infected tissues should reflect that from cells producing more viral offspring, and should thus progressively adjust to and stabilize at the genome formula maximizing replication in a given environment. It seems reasonable to assume that the different segments have the same replication efficiency, because they are similar in length (980–1,003 nucleotides) and all harbour a conserved origin of replication interacting with the same viral 'replication factor' (REP) encoded by segment R<sup>13</sup>. Likewise, once individually encapsidated (or associated) with the same coat protein (encoded by segment S), all segments can move similarly

through plasmodesmata, and can thus exit the cell, and enter sieve elements or adjacent uninfected cells. Hypothesis H1 is supported partially by the results described in Figs 3 and 4. Indeed, numerous FBNSV populations in individual *V. faba* plants converge to the same genome formula, suggesting selection of a specific relative frequency of genome segments in a given environment. Moreover, the viral load within leaves increased in parallel to adjustment of the population around the setpoint genome formula, possibly revealing a link between the two phenomena.

Not all our results, however, reflect this causal link between the distance of FBNSV populations to the setpoint genome formula (as calculated here) and viral load. Specifically, there are two such observations: (i) we detected no significant correlation between viral load and distance to the setpoint genome formula within comparable leaf levels of our plant sets and (ii) plants in which development was arrested at seven leaves were initially closer to the setpoint formula than those in which development was arrested at six leaves (Fig. 4b), yet convergence to the setpoint formula appeared slower in the former case and the increase in viral load was delayed accordingly. One explanation for this discrepancy may be the very nature of the distance measure we used, which ascribes the same weight to variation in all segments, and assumes a similar effect when a segment is either under- or over-represented. We felt that current knowledge of the FBNSV replication cycle is too fragmentary to justify any other weighting. It is highly unlikely, however, that quantitative variation in relative frequency above or below the setpoint value is strictly equivalent for all segments. It is much more likely that subtle co-variations in the frequency of several segments, yet to be identified, govern the efficiency of viral replication within host cells. We thus believe that our distance calculations only partly capture the effect of segment frequency variation when making comparisons at the stage at which plant development was arrested (Fig. 4), and that important aspects of the dynamics of how viral populations arrive at the point where they achieve high viral loads and eventually arrest plant development (in *V. faba*) are still obscure.

A simple alternative mechanism to explain adjustments to the genome formula in FBNSV populations could be differences in segment replication rates. Indeed, the sequences of the eight segments diverge considerably, even in non-coding regions<sup>13,30</sup>, and a distinct regulation during replication cannot be excluded. Under this hypothesis (H2), the replication rate of each segment would directly reflect the observed differences in accumulated GCNs. The segment-specific replication rate would have been selected to give a constant optimal genome formula and, given the same assumptions as in H1 (viral particles have similar chances to exit and enter cells whatever segment they contain), any initial segment frequency would return rapidly to the endpoint formula. A major caveat of H2 is that segments replicating slowly would be outcompeted rapidly by the others and lost. Our observation is that all segments detected in lower leaf-levels are generally maintained in upper leaf-levels, thus H2 is untenable without invoking additional unknown mechanisms ensuring maintenance of slowly replicating segments.

Considering that the genome formula described here applies similarly to encapsidated DNA (Supplementary Fig. S1), a most intriguing question for FBNSV, and multipartite viruses in general<sup>4,12</sup>, is that of the number of virus particles entering individual cells. Assuming that the probability for a genome segment to enter an individual cell depends solely on its relative frequency within the population, infection of 95% of susceptible cells by at least one copy of each segment would require the entry of approximately 160 virus particles per cell if the population is at the setpoint genome formula. The multiplicity of cell infection

(MOI), defined as the number of virus genomes initiating infection within individual cells, has been estimated for a small number of monopartite virus species infecting animals<sup>31–33</sup> and plants<sup>34–36</sup> to range from 1–13. In most cases, the virus is believed to infect new cells in the form of virus particles, thus connecting the MOI value to the number of particles penetrating each cell, and suggesting that the latter is of the order of a few units to a few tens of units. Genome multi-compartmentalization, augmented by the observed variation in frequency across segments of FBNSV, induces an extra cost at the step of cell infection because the number of virus particles required is one or two orders of magnitude higher. In fact, the required number of particles per cell for FBNSV appears so high that it is tempting to imagine a regulatory mechanism that would group one or more copies of each segment into a reasonably sized infectious unit. Further investigation on how multipartite viruses deal with the acute problem of MOI is clearly required.

The dependence of setpoint genome formula on host species opens another avenue for future investigation. On the one hand, one could argue that multipartite viruses are highly polymorphic biological systems, and that the distinct formulae observed in *V. faba* and in *M. truncatula* stem simply from a different optimal composition of FBNSV populations in these two hosts, as explained by scenario H1 described above. On the other hand, the different formulae observed in *M. truncatula* may be suboptimal, and may more trivially reflect mal-adaptation of this isolate of FBNSV to this particular host (the FBNSV isolate used here was isolated originally from *V. faba*<sup>13</sup>). As no consensus sequence change could be observed in FBNSV populations after two successive passages in *V. faba*<sup>30</sup>, longer experimental evolution both in *M. truncatula* and in *V. faba* may be required to assess whether the genome formula evolves, and to identify adaptive mutations responsible for the putative formula change.

A recent study described the phenomenon of ‘gene-accordion’ in the monopartite vaccinia poxvirus<sup>37</sup>. This latter study demonstrated that, under specific selection pressure, a vaccinia virus gene was amplified rapidly by successive duplications, which expanded the single-molecule genome and was immediately beneficial for the virus. The same authors further demonstrated that, once amplified, one of the copies of the corresponding gene eventually acquired a beneficial mutation decreasing the need for multiple extra copies, which were then progressively eliminated, compacting back the viral genome. Multipartite viruses appear particularly well adapted to the gene-accordion phenomenon, which can be implemented through the plasticity of the genome formula. The immediate change in GCN when FBNSV is transmitted from *V. faba* to *M. truncatula* appears as a spectacular illustration of the ease with which multipartite viruses can use gene-accordion, and could illustrate a major advantage of this peculiar genome architecture.

Gene or segment copy number regulation is similarly possible for all multipartite viruses, raising the question of the significance of our discovery for viral species other than FBNSV. Although no direct investigation has addressed this question, some data indicating unequal accumulation of different genome segments can be found, for example, in studies on cucumber mosaic virus<sup>38–41</sup>, tomato aspermy virus<sup>42,43</sup> and brome mosaic virus<sup>44</sup>—all positive-sense ssRNA viruses unrelated to FBNSV. These earlier studies were either purely descriptive, analysing the physical and/or chemical properties of the virus particle<sup>38,39,42,44</sup>, or focused on technical developments for RNA quantification<sup>40,41,44</sup>. Nevertheless, they suggest that the regulation of segment copy number may be a general feature of multipartite viruses. It will be of interest to extend the question of the control of GCN to those segmented viruses encapsidating one copy of each genome segment in a single virus particle<sup>45</sup>.

A series of studies on the switch from lytic to lysogenic cycle of the monopartite bacteriophage  $\lambda$  (refs 46,47) demonstrated that the number of genomes entering a host cell is instrumental in determining the phenotype expressed by the phage, that is, replicate fast and kill the host (lytic), or integrate into the host genome and protect it (lysogenic). So, whether infection is initiated with one or more copies of a monopartite genome can alone decide the fate of both virus and host cell. Focusing on the initial step of cell infection, segmented viruses co-encapsidating one copy of each segment resemble the monopartite situation with an equal copy number of each gene. In contrast, multipartite viruses can potentially adjust the copy number of each segment, even at cell entry. Whether this additional level of regulation at the very first step of infection benefits multipartite viruses is unclear, but it would be precluded by all other types of viral genome organization.

The potential benefit of differential control of GCN in segmented and multipartite viruses imposes direct constraints on the frequencies of individual genome segments. This clearly comes at additional cost and thus questions regarding the origin, and reasons for the evolution, of such biological systems deserve to be revisited in the light of the results described here on FBNSV.

## Methods

**Virus isolate and inoculation.** The isolate of FBNSV used in this study was originally isolated from *V. faba* in Ethiopia<sup>13</sup>. From this isolate, agrobacterium-compatible clones of each of the eight FBNSV segments (DNA C, M, N, R, S, U1, U2 and U4) were prepared, and shown to together constitute a fully infectious and transmissible clone of FBNSV<sup>13</sup>. Briefly, Agrobacteria containing tandem repeat clones of each segment were grown in NZY+ medium (0.1% NZ amine, 0.5% yeast extract, 0.5% NaCl, 12.5 mM MgCl<sub>2</sub>, 12.5 mM MgSO<sub>4</sub> and 0.4% glucose at pH 7.5) until an optical density (OD<sub>600</sub>) of 2–3 was reached. Mixtures of these bacterial suspensions were then prepared in *ad hoc* proportions, centrifuged and resuspended in 40 ml of MS buffer<sup>48</sup> containing 30  $\mu$ M of acetosyringone and 1 mM of MES (morpholinoethanesulfonic acid). These mixtures were then needle-inoculated into the stem of 10-day-old *V. faba* plants at the two-leaf stage, and resulted in infection of around 20% of the inoculated plants.

An earlier report demonstrated that the closely related Faba bean necrotic yellow virus (FBNYV) can replicate and progress systemically within its *V. faba* host plants, under laboratory conditions, when either one of C, N, U2 or U4 is missing<sup>14</sup>. However, how the absence of each of these segments induces differences in the infection kinetics and aphid transmission remains to be characterized in detail for FBNYV and for the FBNSV isolate used here. Despite these observations, and because in all described natural isolates of FBNYV and FBNSV the eight 'core' genome segments were always present, all eight are considered to be integral parts of the viral genome<sup>30</sup>.

**Plant growth conditions.** Broad bean (*V. faba*, var. 'Sevilla' from Vilmorin) and *M. truncatula* (Jemalong A17) plants were grown in soil treated with Trigrad 5W (2 g per 5 l) in a growth chamber with a 13/11 h day/night photoperiod, a temperature of 26/20 °C day/night and 70% hygrometry within an S2 restricted-access confinement facility.

**Aphid growth conditions.** The *Acyrtosiphon pisum* colony was maintained on *V. faba* plants in environmental growth chambers at a temperature of 23/21 °C and a photoperiod of 13/11 h (day/night), ensuring reproduction through parthenogenesis.

Cohorts of aphids were allowed a 3-day acquisition period on *V. faba* plants infected by FBNSV for 30 days (30 dpi). Then, aphids were transferred to 10-day-old *V. faba* plantlets, or on *M. truncatula* at the three-leaf stage, and allowed an inoculation period of three additional days. Aphids were finally killed by spraying the young test plants with Pirimor G (1 g l<sup>-1</sup> in water).

Whether agro-inoculation or aphid-transmission was used, symptoms of systemic infection appeared on newly developed leaves between 10 and 25 dpi, as inward leaf-curling and yellowing. The apical growth in *V. faba* plants was totally inhibited shortly after systemic symptoms appeared, but the plants survived for several additional weeks. Similar symptoms appeared on *M. truncatula* leaves between 10 and 25 dpi, but the viral infection did not stop apical development in this species and symptomatic plants continued growing indefinitely.

**DNA extraction and Q-PCR conditions.** One *V. faba* leaflet per leaf level of symptomatic plants was punched four times on the main vein to collect four leaf disks (6-mm  $\phi$  each), except for the youngest leaf level for which one entire leaflet was harvested. Concerning *M. truncatula*, one leaflet per leaf level was harvested

and analysed. Sampling was carried out between 30 and 35 dpi for both *V. faba* and *M. truncatula*. Total DNA from these samples was extracted according to a modified Edwards protocol<sup>49</sup> with an additional washing step with 70% ethanol. DNA was resuspended with 200  $\mu$ l of water, and 10-fold diluted samples were used for Q-PCR.

For extracting encapsidated DNA, the leaf disks were ground in 400  $\mu$ l of a mild extraction buffer (200 mM Tris HCl, 250 mM NaCl, 1 mM PVP-40, 0.05% Tween 20) described earlier<sup>50</sup>. Samples were vortexed and then clarified for 5 min at 10,000 g. To remove all unencapsidated DNA, samples were incubated with one volume of nucleic acid-digestion solution (200 mM Tris-HCl, 1 mg ml<sup>-1</sup> DNase I, 1 mg ml<sup>-1</sup> RNase A) for 30 min. at 37 °C. After addition of 5 mM EDTA, 0.2% SDS and 2 mg ml<sup>-1</sup> proteinase K (final concentrations) and further incubation for 30 min, the proteinase K was denatured by heating samples at 95 °C for 15 min. DNA was finally precipitated with isopropanol, resuspended in water and stored at -20 °C until use for Q-PCR. In these samples, no plant DNA (tested with *V. faba* legumin B gene *LeB4*, GenBank accession Nb X03677) could be Q-PCR-amplified, controlling for the efficient elimination of unprotected (unencapsidated) DNA.

All Q-PCR reactions (40 cycles of 95 °C for 10 s, 63 °C for 10 s and 72 °C for 10 s) were carried out using the LightCycler FastStart DNA Master Plus SYBR green I kit (Roche) in a LightCycler 480 thermocycler (Roche), following the manufacturer's instructions. The primers (Supplementary Table S6) were used at a final concentration of 0.3  $\mu$ M. In each PCR plate, fluorescence data were normalized with a calibrated reference sample for each segment (one of the plasmid dilutions initially used to establish standard curves), and analysed with the LinRegPCR program<sup>51</sup>. Relative frequencies were then calculated in each sample by dividing the estimated number of DNA copies of a given segment by the sum of the number of copies of all eight segments.

To determine overall viral accumulation, the sum of the estimated number of copies of each segment was divided by that of a host plant gene (*V. faba* legumin B gene *LeB4*, GenBank accession Nb X03677), normalizing the amount of plant material analysed in all samples.

**Statistical analysis.** All statistical tests were carried out with the R and JMP softwares (R Development Core Team, 2011, version 2.12.0; JMP10). The nature and results of each statistical test are indicated in the Results, figure legends and Supplementary Tables. For analysis of variance tests, we used fixed- or mixed-effect models for analysing differences in the frequency of the segments, or differences in their coefficient of variation, depending on the nature of the segment, the experimental replicate, the conditions at inoculation or the leaf level. Whenever several leaf-levels were analysed in the same plant set (that is, repeated measures in the same individuals), the plant effect was included in the model as a random effect.

The dGF in each sample was calculated as  $dGF1 = \sum_{i=1}^8 |p_i - pf_i|/2$  according to Manly (eq. 5.7 p. 68)<sup>16</sup>, or alternatively  $dGF2 = \sum_{i=1}^8 |p_i - pf_i|/pf_i$ , where  $i$  is the segment,  $p$  is the relative frequency of the segment in the sample and  $pf$  is the relative frequency of the segment in the setpoint genome formula.

## References

- Hull, R. *Matthews' Plant Virology* 4th edn, vol. 1 (Academic Press, 2001).
- van Regenmortel, M. H. V. *Virus Taxonomy: The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses* (Academic Press, San Diego, 2000).
- Ojosnegros, S. *et al.* Viral genome segmentation can result from a trade-off between genetic content and particle stability. *PLoS Genet.* 7, e1001344 (2012).
- Nee, S. The evolution of multicompartmental genomes in viruses. *J. Mol. Evol.* 25, 277–281 (1987).
- Pressing, J. & Reaney, D. C. Divided genomes and intrinsic noise. *J. Mol. Evol.* 20, 135–146 (1984).
- Chao, L. Evolution of sex in RNA viruses. *J. Theor. Biol.* 133, 99–112 (1988).
- Chao, L. Levels of selection, evolution of sex in RNA viruses, and the origin of life. *J. Theor. Biol.* 153, 229–246 (1991).
- Chao, L. Evolution of sex in RNA viruses. *Trends Ecol. Evol.* 7, 147–151 (1992).
- Chao, L. & Tran, T. T. The advantage of sex in the RNA virus Phi6. *Genetics* 147, 953–959 (1997).
- Michod, R. E., Bernstein, H. & Nedelcu, A. M. Adaptive value of sex in microbial pathogens. *Infect. Genet. Evol.* 8, 267–285 (2008).
- Escrú, F., Fraile, A. & Garcia-Arenal, F. Constraints to genetic exchange support gene coadaptation in a tripartite RNA virus. *PLoS Pathog.* 3, e8 (2007).
- Iranzo, J. & Manrubia, S. C. Evolutionary dynamics of genome segmentation in multipartite viruses. *Proc. Biol. Sci.* 279, 3812–3819 (2012).
- Grigoras, I. *et al.* Reconstitution of authentic nanovirus from multiple cloned DNAs. *J. Virol.* 83, 10778–10787 (2009).
- Timchenko, T. *et al.* Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus. *J. Gen. Virol.* 87(Pt 6): 1735–1743 (2006).
- Franz, A., Makkouk, K. M. & Vetten, H. J. Host range of faba bean necrotic yellows virus and potential yield loss in infected faba bean. *Phytopathol. Medit.* 36, 94–103 (1997).

16. Manly, B. F. J. *Multivariate Statistical Methods. A primer* 3rd edn (Chapman & Hall/CRC, Boca Raton, Fla., 2005).
17. Stranger, B. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
18. Hastings, P. J., Lupski, J. R., Rosenberg, S. M. & Ira, G. Mechanisms of change in gene copy number. *Nat. Rev. Genet.* **10**, 551–564 (2009).
19. Almal, S. H. & Padh, H. Implications of gene copy-number variation in health and diseases. *J. Hum. Genet.* **57**, 6–13 (2012).
20. Nordman, J. & Orr-Weaver, T. L. Regulation of DNA replication during development. *Development* **139**, 455–464 (2012).
21. Nair, S. *et al.* Adaptive copy number evolution in malaria parasites. *PLoS Genet.* **4**, e1000243 (2008).
22. White, J. P. & Lloyd, R. E. Regulation of stress granules in virus systems. *Trends Microbiol.* **20**, 175–183 (2012).
23. Ding, S. W. & Lu, R. Virus-derived siRNAs and piRNAs in immunity and pathogenesis. *Curr. Opin. Virol.* **1**, 533–544 (2012).
24. Brosnan, C. A. & Voinnet, O. The long and the short of noncoding RNAs. *Curr. Opin. Cell Biol.* **21**, 416–425 (2009).
25. Walsh, D. & Mohr, I. Viral subversion of the host protein synthesis machinery. *Nat. Rev. Microbiol.* **9**, 860–875 (2011).
26. Makinen, K. M. & Ivanov, K. I. Phosphorylation analysis of plant viral proteins. *Method Mol. Biol.* **451**, 339–359 (2008).
27. Choi, A. G., Wong, J., Marchant, D. & Luo, H. The ubiquitin-proteasome system in positive-strand RNA virus infection. *Rev. Med. Virol.* **23**, 85–96 (2011).
28. Mileyko, Y., Joh, R. I. & Weitz, J. S. Small-scale copy number variation and large-scale changes in gene expression. *Proc. Natl Acad. Sci. USA* **105**, 16659–16664 (2008).
29. Sjodin, P. & Jakobsson, M. Population genetic nature of copy number variation. *Method Mol. Biol.* **838**, 209–223 (2012).
30. Grigoras, I. *et al.* High variability and rapid evolution of a nanovirus. *J. Virol.* **84**, 9105–9117 (2010).
31. Bull, J. C., Godfray, H. C. & O'Reilly, D. R. Persistence of an occlusion-negative recombinant nucleopolyhedrovirus in *Trichoplusia ni* indicates high multiplicity of cellular infection. *Appl. Env. Microbiol.* **67**, 5204–5209 (2001).
32. Jung, A. *et al.* Multiply infected spleen cells in HIV patients. *Nature* **418**, 144 (2002).
33. Josefsson, L. *et al.* Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proc. Natl Acad. Sci. USA* **108**, 11199–11204 (2011).
34. Gonzalez-Jara, P., Fraile, A., Canto, T. & Garcia-Arenal, F. The multiplicity of infection of a plant virus varies during colonization of its eukaryotic host. *J. Virol.* **83**, 7487–7494 (2009).
35. Miyashita, S. & Kishino, H. 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. *J. Virol.* **84**, 1828–1837 (2010).
36. Gutierrez, S. *et al.* Dynamics of the multiplicity of infection in a plant virus. *PLoS Pathog.* **6**, e1001113 (2010).
37. Elde, N. C. *et al.* Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* **150**, 831–841 (2012).
38. Lot, H. *et al.* Evidence for three functional RNA species in several strains of cucumber mosaic virus. *J. Gen. Virol.* **22**, 81–93 (1974).
39. Lot, H. & Kaper, J. M. Further studies on the RNA component distribution among the nucleoproteins of cucumber mosaic virus. *Virology* **74**, 223–226 (1976).
40. Feng, J. L. *et al.* Quantitative determination of cucumber mosaic virus genome RNAs in virions by real-time reverse transcription-polymerase chain reaction. *Acta Bioch. Bioph. Sin.* **38**, 669–676 (2006).
41. Feng, J. L., Zeng, R. & Chen, J. S. Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. *Biotechniques* **44**, 901–912 (2008).
42. Stace-Smith, R. & Tremaine, J. H. Biophysical and biochemical properties of tomato aspermy virus. *Virology* **51**, 401–408 (1973).
43. Habili, N. & Francki, R. I. Comparative studies on tomato aspermy and cucumber mosaic viruses. III. Further studies on relationship and construction of a virus from parts of the two viral genomes. *Virology* **61**, 443–449 (1974).
44. Kwon, S. J., Seo, J. K. & Rao, A. L. Detection and quantification of viral and satellite RNAs in plant hosts. *Method Mol. Biol.* **883**, 131–141 (2012).
45. Chou, Y. Y. *et al.* One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl Acad. Sci. USA* **109**, 9101–9106 (2012).
46. Weitz, J. S., Mileyko, Y., Joh, R. I. & Voit, E. O. Collective decision making in bacterial viruses. *Biophys. J.* **95**, 2673–2680 (2008).
47. Zeng, L. *et al.* Decision making at a subcellular level determines the outcome of bacteriophage infection. *Cell* **141**, 682–691 (2010).
48. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio-assays with tobacco tissue cultures. *Physiol. Plantarum.* **15**, 473–497 (1962).
49. Edwards, K., Johnstone, C. & Thompson, C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucl. Acids Res.* **19**, 1349 (1991).
50. Rampersad, S. N. & Umaharan, P. Detection of begomoviruses in clarified plant extracts: a comparison of standard, direct-binding, and immunocapture polymerase chain reaction techniques. *Phytopathology* **93**, 1153–1157 (2003).
51. Ruijter, J. M. *et al.* Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **37**, e45 (2009).

#### Acknowledgements

*M. truncatula* seeds were kindly provided by Jean-Marie Prosseri and Pascal Joly, UMR DiA-PC, Domaine de Melgueil, 34130 Mauguio, France. We are grateful to Dr. D.K. Willis and Dr. T.L. German for fruitful discussions on Q-PCR set up. A.S., M.Y., S.G. and S.B. acknowledge support from INRA Department SPE; T.T. and B.G. from CNRS and Y.M. from CNRS and IRD. This work was specifically funded by an INRA-SPE 'MOI' starting Grant and by ANR Grant N° 2010-BLANC-1704 01. We are very grateful to the group APRO-Lacombe for repeated inspiring think-tank sessions.

#### Author contributions

S.B., S.G., Y.M., M.Y. and A.S. conceived the work and designed experiments. A.S., M.Y., T.T. and S.G. performed experiments. A.S., S.G., M.Y., Y.M. and S.B. analysed the data. S.B., Y.M., T.T., B.G. and S.G. wrote the paper.

#### Additional information

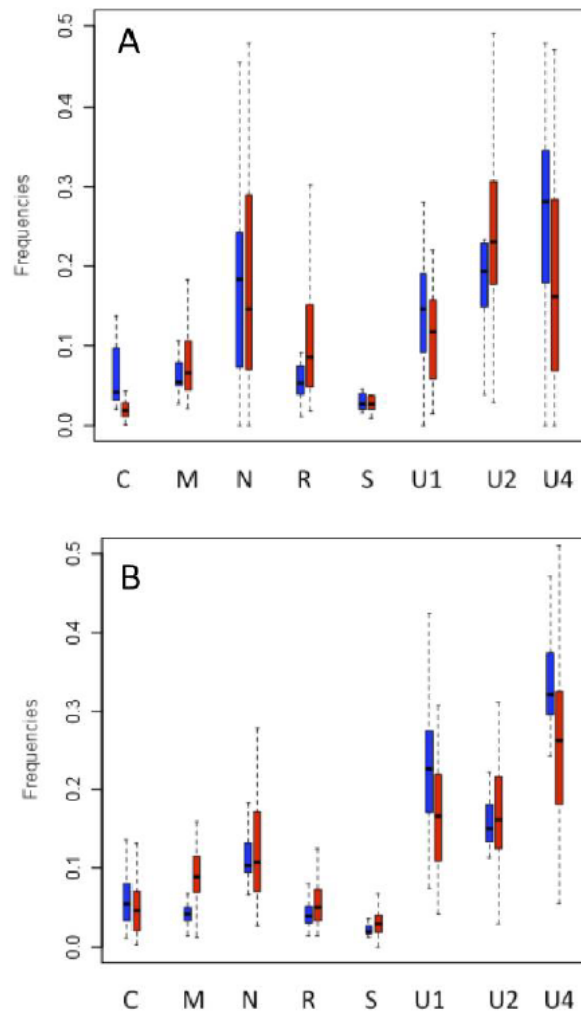
**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Sicard, A. *et al.* Gene copy number is differentially regulated in a multipartite virus. *Nat. Commun.* **4**:2248 doi: 10.1038/ncomms3248 (2013).

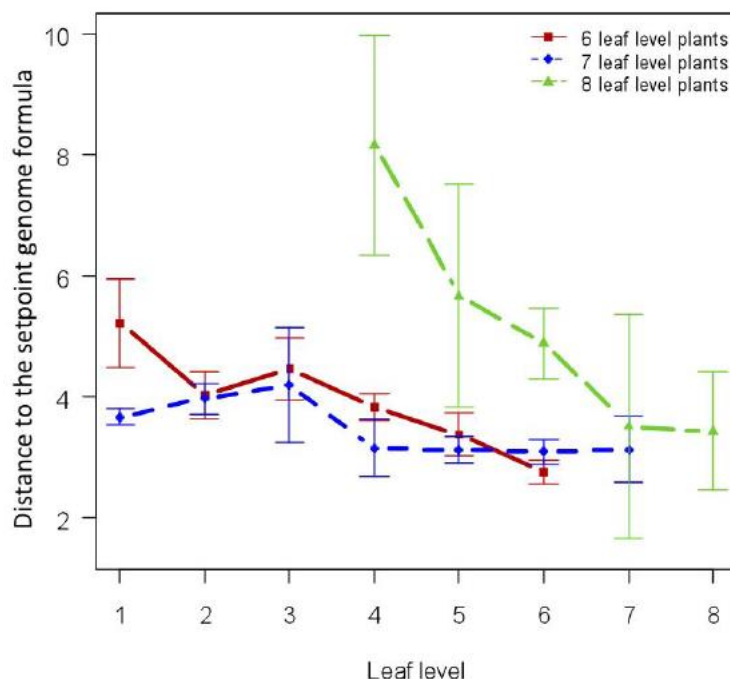
## Supplementary figures



### Supplementary Figure S1. Comparison of the relative frequency of various FBNSV genomic segments in total viral DNA and encapsidated DNA

In complement to Figure 1, the relative frequency of each genome segment of FBNSV has been compared by Q-PCR in both total (blue) and encapsidated viral DNA (red, extracted as described in the Materials and Methods) of systemically infected leaves of *Vicia faba*. The name of each segment is indicated below the graph. The horizontal black bar represents the median of the distribution and the vertical dotted line delineates 1.5 times the distance between the 25<sup>th</sup> and 75<sup>th</sup> percentile of the distribution. Results from independent experiments are plotted in A and B. The results in A were obtained from the uppermost leaf level of 20 plants, whereas those in B were from 29 samples collected from the three upper leaf levels of 11 plants. In both A and B, the relative frequency of the eight FBNSV genome segments proved very similar in total viral DNA and in encapsidated DNA. To control for the absence of unprotected (unencapsidated) DNA in the samples plotted in red, we verified that the Q-PCR amplification of plant DNA (amplification of *Vicia faba* legumin B gene LeB4, GenBank accession Nb X03677) was not detectable.





**Supplementary Figure S2. Distance to the setpoint genome formula in FBNSV populations across leaf levels**

In complement to Figure 4B, the distance of FBNSV populations to the setpoint genome formula was alternatively calculated as indicated for dGF2 in the Materials and Methods. The distance dGF2 of FBNSV populations to the setpoint genome formula was calculated in all leaves of plants that developed 6 (red square), 7 (blue diamond) or 8 (green triangle) leaf levels. Error bars represent the standard deviations.

**Supplementary Table S1. The different genome segments of FBNSV accumulate at different relative frequency within infected host plants.**

Segments	C	M	N	R	S	U1	U2
C							
M	0.9999949						
N	0.0000015	0.0000004					
R	0.9578789	0.9915555	0.0000000				
S	0.4485940	0.6285719	0.0000000	0.9782571			
U1	0.2009629	0.1100894	0.0287565	0.0104511	0.0003021		
U2	0.0000633	0.0000199	0.9903335	0.0000004	0.0000000	0.2343362	
U4	0.0000000	0.0000000	0.0409977	0.0000000	0.0000000	0.0000001	0.0026887

This table provides the p-values from Tukey HSD pair tests comparing the frequency of each FBNSV genome segment to that of the others, in the uppermost systemically infected leaves of *Vicia faba* host plants (in complement to Figure 1)

All values indicating a significant difference are in red.

**Supplementary Table S2. The genome formula of FBNSV is not affected by the mode of inoculation.**

Segments	C	M	N	R	S	U1	U2	U4
$F_{1,64}$	0.27	0.057	1.97	0.22	0.19	0.0046	1.04	0.6
p-value	0.6	0.81	0.16	0.64	0.67	0.95	0.31	0.44

This table shows the results of ANOVA tests for the effect of the mode of FBNSV inoculation (Agro-inoculation versus aphid-transmission) on the relative frequency of each of the eight segments (in complement to Figure 2B).

F indicates the results of the F statistics with subscripts indicating the degrees of freedom of the numerator and the denominator, respectively.

No significant effect could be detected, on any segment.

**Supplementary Table S3. The setpoint genome formula is reached rapidly during FBNSV progression into *Vicia faba* host plants**

Segments	C	M	N	R	S	U1	U2	U4
$F_{1,72}$	2.13	1.35	3.74	3.22	2.03	4.35	5.48	1.47
p-value	0.07	0.25	0.0044	0.011	0.084	0.0015	0.00023	0.21

This table shows the results of ANOVA tests for the effect of the leaf level on the frequency of FBNSV segments (in complement to Figure 3A).

F indicates the results of the F statistics with subscripts indicating the degrees of freedom of the numerator and the denominator respectively.

All p-values indicating a significant effect of the leaf level are in red.

**Supplementary Table S4. Co-variance of the relative frequency of different FBNSV segments**

	Segments	C	M	N	R	S	U1	U2
<i>p</i> de Spearman	<b>M</b>	0.25						
Prob.>  <i>p</i>		0.35						
<i>p</i> de Spearman	<b>N</b>	-0.3235	0.3941					
Prob.>  <i>p</i>		0.2216	0.1309					
<i>p</i> de Spearman	<b>R</b>	0.2647	-0.3353	-0.25				
Prob.>  <i>p</i>		0.3218	0.2043	0.3504				
<i>p</i> de Spearman	<b>S</b>	-0.0824	-0.0824	-0.4147	0.0588			
Prob.>  <i>p</i>		0.7617	0.7617	0.1102	0.8287			
<i>p</i> de Spearman	<b>U1</b>	-0.5655	-0.2784	-0.1708	-0.1325	0.3255		
Prob.>  <i>p</i>		0.022	0.2965	0.527	0.6246	0.2186		
<i>p</i> de Spearman	<b>U2</b>	-0.1972	-0.3311	-0.1251	0.1634	-0.2752	0.3589	
Prob.>  <i>p</i>		0.4641	0.2103	0.6444	0.5455	0.3023	0.1722	
<i>p</i> de Spearman	<b>U4</b>	-0.0559	-0.2853	-0.5059	-0.0088	0.0853	-0.1885	-0.39
Prob.>  <i>p</i>		0.8371	0.2841	0.045	0.9741	0.7535	0.4844	0.1354

This table shows the Spearman's rank correlation coefficients of the relative frequencies of all possible pairs of FBNSV segments.

All p-values indicating a coefficient significantly different from zero are in red.

**Supplementary Table S5. The genome formula of FBNSV is dependent on the host plant species**

Segments	C	M	N	R	S	U1	U2	U4
$F_{1, 72}$	386.1	29.3	102.7	38.1	137.8	2.9	10.5	36.7
p-value	<2.2e-16	7.7e-7	1.7e-15	3.6e-8	<2.2e-16	0.093	0.0018	5.67e-8

This table shows the results of ANOVA tests for the effect of the host plant species on the relative frequency of each of the eight FBNSV segments (in complement to Figure 2B).

F indicates the results of the F statistics with subscripts indicating the degrees of freedom of the numerator and the denominator respectively.

All p-values indicating a significant effect are in red.

The relative frequency of all segments but segment U1 is host-dependent.

**Supplementary Table S6. Sequence of all primers used for Q-PCR**

Targeted Segments	Direction	Sequence 5'-3'
C	Forward	GACCCATCAACTGAAGACCTGTTGTCTC
	Reverse	AGCAACTCTCTACATGCAAGGTCTTCCA
M	Forward	ACGCTGCGTATCAAGACGACGG
	Reverse	TCCCGGAAGATAGCATGCAAGCA
N	Forward	GCATGAAAGACAAGCTCAACG
	Reverse	GTTCTGTTTCCACCATAGAACTAC
R	Forward	GGAGAACTGAAGAAGAAGAAAGCTAT
	Reverse	AAGCAAAGCATCAACCTCG
S	Forward	AGAAAGATGACGTGTCTGGTAG
	Reverse	TCAGATAATTCACCAACTCTCCCG
U1	Forward	AGCTATTGATTTGGAAGACAGAGT
	Reverse	CGACAGCCAAGAAGCGT
U2	Forward	ATTAGAGTAGATGAATGCAAACTTGTAT
	Reverse	TAAACCTTCTAGACTGCCATCGT
U4	Forward	CCAGATTAAGAAGTTACTGTCAGGTAATA
	Reverse	ATTCAGCATCCTCGAACGG

### 3. Complément à l'article 1

Il a été démontré que les huit segments du FBNSV présentait des fréquences relatives spécifiques et reproductibles au sein d'un hôte donné. Contrairement à ce qui était supposé en introduction, aucun lien entre la fréquence relative d'un segment et son indispensabilité à l'infection de la fève n'a été apparue. Nous ne pouvons apporter aucune explication à cette observation, si ce n'est le fait que la réflexion initiale qui nous avait amenés à prédire une fréquence élevée pour des segments non indispensables est peut être tout simplement erronée. En effet, elle était basée sur deux postulats : (i) tous les segments se répliquent ensemble au sein de cellules individuelles, et (ii) leur chance d'entrée au sein de ces cellules avec les autres et donc de participer à ce « pool répliatif » dépend de leur fréquence dans la population. Ces postulats de base, qui sont généralement issus de la littérature disponible dans ce domaine (cf Revue bibliographique en Introduction) seront remis en question dans le Chapitre III.

#### *3.1. Essai de mise en place de l'évolution expérimentale du FBNSV*

Le concept de « formule génomique » correspond à la série de chiffres qui associe chaque segment génomique à son nombre relatif de copies au sein de la population virale. Comme discuté dans l'article, la spécificité de cette formule en fonction de l'hôte dans lequel le virus se trouve pourrait résulter d'une possibilité du virus à ajuster rapidement et de façon optimale le nombre de copies de ses gènes en fonction de son environnement. L'ajustement immédiat du virus à son hôte s'opérerait ainsi grâce à une plasticité phénotypique sans modification génétique, et c'est pour cela que nous n'utilisons pas ici le terme « adaptation ». Dans ce cas, le virus serait une sorte de généraliste, pré-adapté à des changements d'hôte fréquent.

*A contrario*, la formule génomique différente observée au sein de *Medicago truncatula* pourrait être le reflet d'une maladaptation du clone de FBNSV à cet hôte. En effet, ce clone ayant été isolé puis maintenu pendant trois ans sur fève, l'inoculation par pucerons de *Medicago truncatula* pourrait constituer une première rencontre avec un environnement auquel le virus n'est pas adapté. Dans ces conditions, on peut penser que le FBNSV développe une formule génomique non-optimale chez *Medicago truncatula* (d'où les différences avec celle observée dans la fève) et qu'il requiert davantage de temps pour adapter des fréquences optimales à son fonctionnement chez cet hôte. Si tel est le cas, on s'attend à voir les fréquences relatives changer au cours du temps si le FBNSV est constamment maintenu dans *Medicago truncatula*, et à ce

que ces changements de formule génomique s'accompagnent de modifications de séquences (de mutations adaptatives). En parallèle, la valeur sélective du virus devrait augmenter, ce qui peut être déterminée par la mesure de la charge virale *in planta* et celle du taux de transmission.

Afin de déterminer si la formule génomique d'un virus multipartite est effectivement un paramètre susceptible d'évoluer et, le cas échéant, si cette évolution est adaptative, nous avons décidé de suivre la formule génomique de douze lignées indépendantes du FBNSV, maintenue au sein de l'hôte *Medicago truncatula* au cours de passages successifs par pucerons vecteurs. Cette expérience constituait une des perspectives immédiates de l'article présenté. Le protocole de cette expérience est très brièvement exposé ci-après, puisqu'un premier essai de ce protocole a déjà été effectué, et a totalement échoué. Les raisons possibles de cet échec sont discutées à la fin de ce chapitre I.

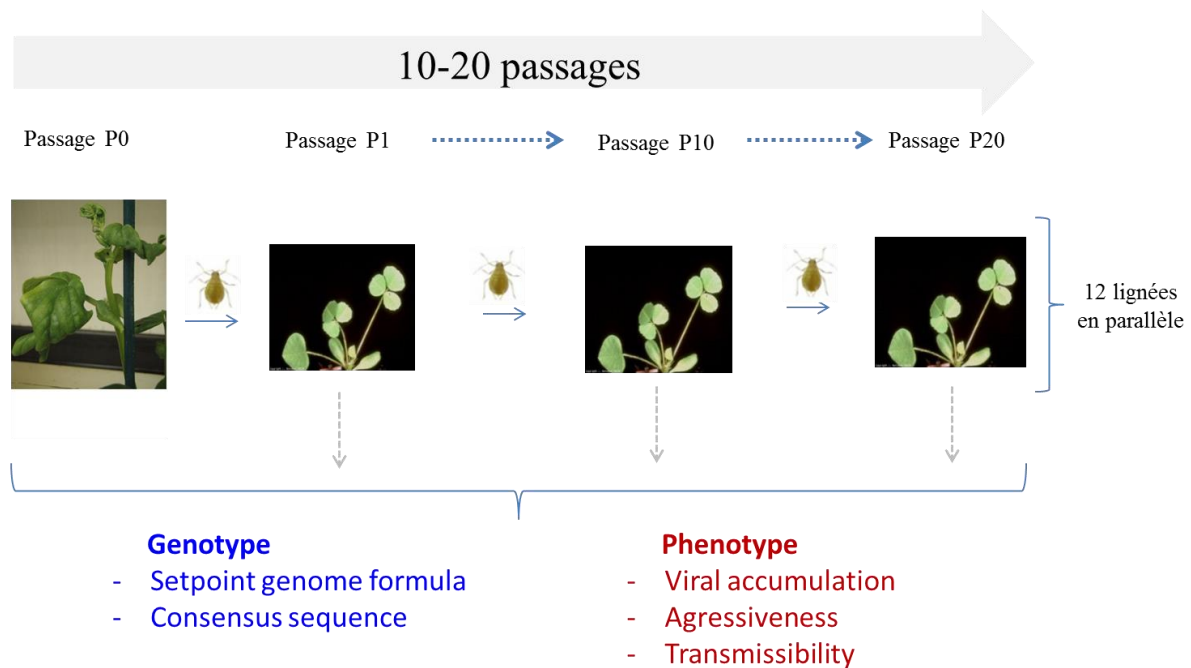
### 3.2. Protocole

#### **Choix de l'espèce vectrice**

L'espèce *Myzus persicae* a été choisie comme espèce vectrice du fait de sa facilité de manipulation, et de sa moindre sensibilité au problème de surpopulation (notamment rencontré pour le démarrage des lignées), contrairement à l'espèce *Acyrtosiphon pisum* pour laquelle une surpopulation entraîne la mort de tous les pucerons.

#### **Démarrage des lignées**

Nos échecs répétés d'agroinoculation ou d'agroinfiltration des plantes de *Medicago truncatula* nous ont conduits à utiliser une fève agro-inoculée comme plante source initiale, et la population de FBNSV qu'elle contient comme population ancestrale. Des centaines de pucerons ont ainsi été mis en acquisition sur l'apex d'une même fève infectée source, préalablement contrôlée par qPCR, pour la présence des huit segments. Après deux jours d'acquisition du virus sur cet apex, les pucerons ont été transférés à raison de dix pucerons par plante sur 48 plantules de *Medicago truncatula* pour trois jours d'inoculation. Douze plantes parmi ces quarante-huit inoculées ont alors été choisies comme plantes sources pour le démarrage des 12 lignées qui sont maintenues en parallèle (Figure 10).



**Figure 10 : Evolution expérimentale de la formule génomique du FBNSV**

Les clones infectieux du FBNSV sont agroinoculés à la fève (P0). Tous les trente jours, les lignées virales sont transférées à dix plantes de *Medicago truncatula* receveuses en utilisant dix pucerons par plante. Douze lignées en parallèle sont suivies. A P1, la formule génomique et la séquence consensus sont déterminées. De même, l'accumulation virale, l'agressivité et la transmissibilité sont établies. Des analyses similaires du phénotype et du génotype des lignées virales seront menées à des passages intermédiaires et à la fin du protocole. L'idée est de laisser les lignées virales évoluées pendant 10 à 20 passages, la fin de l'expérience dépendant de la rapidité de l'évolution du FBNSV dans *Medicago truncatula*.

### Maintien des lignées

Tout au long de l'expérience, le maintien des 12 lignées de FBNSV s'effectue par des passages successifs (toujours sur *Medicago truncatula*) qui ont lieu tous les trente jours, date à laquelle les plantes ont pu pleinement développer des symptômes et atteindre une taille suffisante pour recevoir une centaine de pucerons. Dix pucerons par plante receveuse sont utilisés tout au long de ces passages successifs. Ce nombre a été déterminé comme un compromis entre : (i) un nombre suffisant de pucerons pour éviter un taux de transmission faible qui aurait pu engendrer la perte de lignées, (ii) un nombre de pucerons aussi élevé que possible pour le passage des lignées du FBNSV afin de limiter les goulots d'étranglement génétique et populationnel lors de la transmission, et éviter une trop forte dérive génétique qui ralentirait l'adaptation, (iii) un nombre limité de pucerons afin d'éviter la surpopulation des pucerons sur les plantes. Au delà

de cette attention portée au nombre de pucerons utilisés pour les passages de nos lignées de plante à plante, nous avons aussi essayé d'optimiser le nombre de plantes utilisées pour chaque lignée. Ce nombre est aussi le résultat d'un compromis entre l'utilisation d'un nombre suffisant de plantes pour éviter la perte des lignées même lors de passages où le taux de transmission est faible et la place disponible et nécessaire pour cette expérience de longue haleine. Pour chaque lignée et à chaque passage, dix plantes receveuses sont inoculées chacune par 10 pucerons ayant acquis le virus sur les plantes infectées du passage précédent.

### **Suivi de la formule génomique au cours du temps**

Des prélèvements de toutes les plantes –symptomatiques ou non- de chaque lignée sont effectués à chaque passage et conservés. Le prélèvement des plantes asymptomatiques vise à s'assurer que l'absence de symptôme est liée à une absence de virus et non pas à une diminution de la virulence au cours du temps. Seules les plantes effectivement infectées (vérification par simple PCR sur le segment R) sont alors utilisées comme plantes sources pour le passage suivant. La prévision initiale est de laisser ce clone évoluer durant environ un an, ce qui correspond à une dizaine de passages de plante à plante. Après cinq et dix passages, une analyse approfondie de toutes les plantes infectées est prévue afin de voir si la formule génomique a évolué dans chaque lignée. Si un changement par rapport à la formule initialement établie (sur les 48 premières plantes infectées pour la mise en place des lignées, cf ci-dessus) était observé, nous analyserions alors la formule génomique de la ou des lignée(s) dans laquelle (lesquelles) un changement s'est produit lors des passages précédents afin de déterminer quand ce changement a eu lieu. Il est aussi prévu d'établir la séquence consensus du FBNSV et de la comparer dans chaque lignée au 5<sup>ième</sup> et 10<sup>ième</sup> passage afin de déterminer si des mutations sont apparues, et si ces mutations peuvent être considérées comme adaptatives : si ce sont des mutations parallèles qui sont apparues de manière identique dans plusieurs lignées indépendantes.

Le but premier est ici de voir si une même évolution de la formule génomique se produit dans plusieurs lignées en parallèle, voire toutes, ce qui serait un signe d'un changement adaptatif, indiquant un lien entre formule génomique dans un hôte donnée et valeur sélective du FBNSV dans ce même hôte. Si tel était le cas, le FBNSV évolué chez *M.truncatula* pourrait alors être testé chez *V. faba* pour tester la possibilité que l'adaptation à un hôte s'accompagne d'une perte d'adaptation à un autre (spécialisation). De même, nous espérons que le séquençage des lignées évoluées permettra d'identifier les mutations adaptatives responsables du changement de



formule (testable par génétique inverse sur le clone infectieux), ce qui informerait sur les segments et les séquences qui la contrôlent.

### **Suivi de la valeur sélective au cours du temps**

- Charge virale *in planta*

Les prélèvements des folioles infectées, effectués pour toutes les lignées au fil des passages successifs, permettront également de déterminer la charge virale *in planta*. En fait, les suivis de formule décrits ci-dessus permettent de suivre la charge virale totale sur les mêmes échantillons, en calculant la somme du nombre de copies de chaque segment.

- Taux de transmission

Afin de déterminer les changements possibles du taux de transmission au cours de l'évolution du FBNSV dans *Medicago truncatula*, vingt-quatre plantes supplémentaires seront inoculées à raison de deux pucerons par plante, tous les 2 passages et pour chaque lignée. Nous pourrions ainsi avoir un suivi du taux de transmission du FBNSV non seulement par lignée mais aussi globalement sur l'ensemble des 12 lignées.

### *3.3. Résultats et discussion*

Malheureusement, le protocole décrit ci-dessus ne s'est pas déroulé comme prévu. Les 12 lignées ont pu être établies comme prévu, à partir d'une fève unique agro-inoculée avec le clone infectieux. Après seulement un passage de *M.truncatula* à *M.truncatula*, 9 lignées sur 12 ont été perdues (Passage P2) ; elles n'ont tout simplement pas été transmises malgré l'utilisation de 10 pucerons par plante. L'analyse de la formule génomique et de la charge virale (ce dernier paramètre étant souvent corrélé à la transmission) dans les plantes sources qui ont conduit à la perte de ces 9 lignées ne nous permet pas d'expliquer cette perte. Par contre, le faible nombre de pucerons retrouvé après trois jours de période d'inoculation sur les différentes plantes receveuses nous amène à penser que les pucerons ont pu souffrir d'un problème de confinement. Les plantes ont été recouvertes par des tubes Falcon de 50 mL pour éviter la fuite des pucerons durant cette période. Bien que testée auparavant et ne semblant pas poser de problème particulier, l'utilisation de tubes Falcon pourrait être la source principale du problème, la nocivité du confinement pouvant dépendre de conditions d'humidité ou d'autres facteurs non identifiés et donc non-contrôlés. Nous pensons maintenant que ce confinement n'est pas nécessaire dans la mesure où les pucerons de l'espèce *Myzus persicae* placés sur une plantule

de *Medicago* y restent (ce qui n'est pas nécessairement le cas pour d'autres espèces de pucerons vecteurs).

Le faible taux de transmission qui a conduit à la perte de nos lignées pourrait être aussi lié à une transmission erratique du virus, un constat vérifié lors nos différents essais de transmission et signalé depuis plus de 15 années (Franz et al., 1998). Cependant, l'utilisation de dix pucerons par plante était une des précautions qui devait limiter ce problème.

Quelques pistes d'amélioration pour les expériences futures sont présentées ci-dessous :

Tout d'abord, une période d'acquisition du virus par les pucerons de 48 heures a été utilisée, alors que nos conditions classiques sont plutôt de 72h. Nous avons mené une étude comparative préalable pour regarder l'effet de la période d'acquisition sur le taux de transmission du FBNSV dans nos conditions. Cette étude avait été menée sur fève avec cinq pucerons par plantes ayant acquis sur la même plante source. Une faible différence au niveau des taux de transmission avait été constatée en comparant des acquisitions de 48 et 72 heures. Une étude évaluant l'effet de la période d'acquisition sur le taux de transmission du *Banana bunchy top virus* semble également montrer que 48h soit une durée suffisante pour que le virus soit transmis de manière optimale (Anhalt and Almeida, 2008). Il est donc peu probable qu'une période plus longue puisse augmenter le taux d'infection, mais ceci devra être testé au préalable, spécifiquement de *Medicago* à *Medicago*.

Un paramètre qui pourrait avoir un effet significatif sur le taux de transmission est le stade de développement du puceron au moment de la transmission. Par exemple, il a été montré que pour être transmis, le *Tomato spotted wilt virus* devait être acquis par des thrips au cours du premier stade larvaire (Wetering, 1996). Il a également été montré pour un nanovirus et un babuvirus que le stade de développement du puceron vecteur jouait un rôle au niveau du succès de transmission. Anhalt et Almeida (2008) ont ainsi démontré que le *Banana bunchy top virus* était transmis plus efficacement par les adultes que par les stades larvaires tandis que d'après des travaux de Franz et al. (1998), la tendance est inverse pour la transmission du *Faba bean necrotic yellows virus*. Ce phénomène devrait donc être déterminé au cas par cas, et il sera obligatoire de clarifier quel stade de développement de puceron nous devons utiliser pour optimiser la transmission de nos lignées.

De même, si *A.pisum* semble être inapproprié dans cette expérience pour la raison évoquée plus haut (trop forte sensibilité à la surpopulation), l'espèce *Aphis craccivora* pourrait également être utilisée ici. Une comparaison du taux de transmission des deux espèces de pucerons *M.persicae* et *A.craccivora* devra donc également être menée afin de sélectionner le meilleur vecteur.

Un autre paramètre jouant sur le taux de transmission des nanovirus est la température (Anhalt and Almeida, 2008). Ainsi, d'après les travaux menés par ces auteurs, une température constante de 25°C au cours de l'acquisition/inoculation du *Banana bunchy top virus* augmente la transmission de 300%, par rapport à une température constamment réglée à 20°C. Les températures utilisées au sein de notre expérience étaient de 26°C le jour et 20°C la nuit. Il pourrait être important de vérifier si une température constante réglée à 25 ou 26°C pendant les périodes d'acquisition et d'inoculation augmenterait les taux de transmission du FBNSV.

Outre ces mises au point pour sélectionner les meilleures conditions de transmission par les pucerons, il semble important de tenir compte de l'état des plantes sources. En effet, il est possible que trente jours après infection soit un temps trop tardif pour l'acquisition, les tissus étant déjà très abîmés par l'infection et le durcissement des feuilles pouvant rendre l'acquisition du virus plus difficile. Un test de transmission à partir de plantes sources sélectionnées à différents stades de l'infection pourrait être utile.

Enfin, des espèces végétales d'une même famille pouvant présenter des susceptibilités différentes à l'égard d'un virus donné, il pourrait être intéressant de tester d'autres espèces afin de voir s'il n'est pas possible de trouver un hôte plus susceptible que *Medicago truncatula*. L'objectif de cette expérience étant de regarder l'évolution de la formule génomique au sein d'un hôte autre que la fève et, la formule génomique étant hôte spécifique, il n'est donc pas exclu de tester un autre hôte.

Enfin, et ceci fait le lien avec le chapitre suivant, une question sous-jacente à cette étude est de savoir si la formule génomique observée au sein d'une plante est maintenue au cours des cycles successifs de transmission. L'espèce de pucerons vecteurs choisie pourrait éventuellement avoir aussi un impact répété sur l'évolution de la formule, et nous avons jugé intéressant de tester la possibilité d'un ajustement de la formule durant le passage du FBNSV dans le puceron vecteur.

Si tel était le cas, le virus devrait constamment modifier la composition génétique de ses populations, à chaque passage de la plante vers le puceron et vice versa. Cette question est abordée en détail dans le chapitre II.

**CHAPITRE II :**

**LA TRANSMISSION CIRULANTE NON-  
PROPAGATIVE DES NANOVIRUS : UNE VISION  
SIMPLIFICATRICE**

## Chapitre II : La transmission circulante non-propagative des nanovirus : une vision simplificatrice

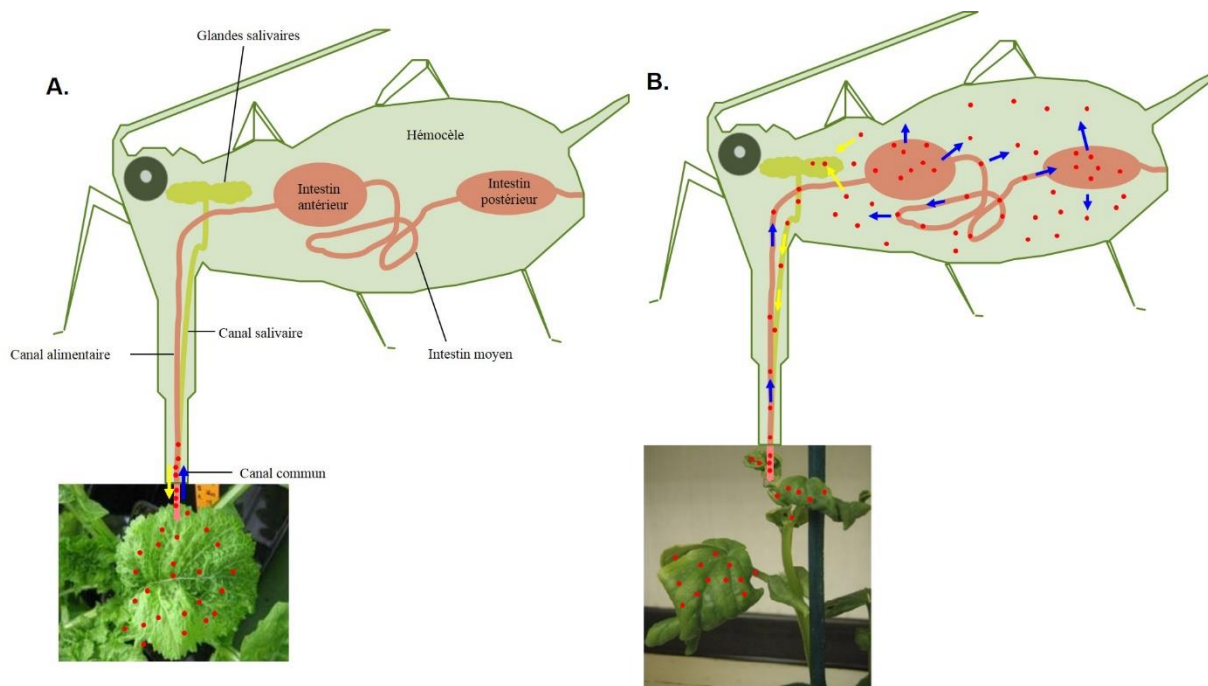
---

### 1. Contexte et objectif

Les phytovirus doivent faire face à deux problèmes majeurs lors de leur transmission : l'immobilité de leurs hôtes, et la paroi pecto-cellulosique entourant les cellules végétales, une barrière que les virus doivent franchir à la fois pour sortir et pour pénétrer dans une plante. Pour s'affranchir de ces problèmes, la majorité des virus de plante utilisent des vecteurs pour leur transmission, qui sont définis comme des organismes parasites ou prédateurs des plantes, capables de créer une blessure pour se nourrir et de se déplacer d'une plante à une autre en transportant des virus. Des vecteurs ont été décrits chez des champignons phytopathogènes, des nématodes des racines, des acariens et des insectes, dont les plus importants sont les insectes phloémophages de l'ordre des Hémiptères, et en particulier les pucerons (pour une revue récente voir Blanc et al., 2014). Les interactions virus-vecteurs sont historiquement regroupées sous trois grandes catégories ou modes de transmission qui ont été définis sur la base de l'étude des insectes phloémophages (Figure 11), mais qui peuvent être adaptées aux autres types de vecteurs :

- i) Le mode « non-circulant » dans lequel le virus s'attache à des récepteurs externes situés sur la cuticule, au niveau des stylets ou de l'œsophage de l'insecte, sans jamais pénétrer son milieu intérieur. L'association est éphémère et le virus n'est spécifiquement retenu que quelques minutes à quelques heures suivant les cas, puis détaché par un mécanisme inconnu et inoculé à une nouvelle plante hôte. Dans le cas des vecteurs de type hémiptères, les virus sont inoculés durant la phase de salivation dans les tissus végétaux, mais les mécanismes moléculaires en jeu restent obscurs.
- ii) Le mode « circulant-propagatif » dans lequel le virus effectue un cycle à l'intérieur du corps de l'insecte vecteur, depuis l'intestin où il est internalisé sous l'action de récepteurs spécifiques, jusqu'aux glandes salivaires d'où il est inoculé. Le virus se réplique durant ce cycle dans l'insecte et on peut donc considérer le vecteur comme un hôte alternatif.

iii) Le mode « circulant non-propagatif » dans lequel le virus effectue un cycle similaire à l'intérieur du vecteur, depuis l'intestin jusqu'aux glandes salivaires, mais durant lequel il ne se réplique jamais. Dans ce mode de transmission, pour autant que les couples virus/vecteur soient pleinement compatibles, les virus sont perçus comme traversant simplement les barrières cellulaires de leurs vecteurs, qui sont alors des sortes de passoires spécifiques dans lesquelles les virus sont temporairement stockés et protégés.



**Figure 11: Comparaison des modes de transmission non-circulant (A) et circulant (B) chez le puceron**

Les légendes correspondant à l'anatomie des pucerons sont identiques sur les deux schémas, mais par souci de clarté elles ne sont représentées que dans le cas de la transmission non-circulante (A). Les ronds rouges illustrent la localisation des particules virales à l'intérieur de la plante et du puceron, et les flèches bleues et blanches représentent respectivement leur trajet entrant et sortant. Ce trajet est limité au canal commun des stylets pour les virus non-circulants, et concerne l'ensemble du corps de l'insecte pour les virus circulants. La distinction entre propagatif et non-propagatif dans le cas des virus circulants n'est pas représentée ici.

Les espèces virales de la famille *Nanoviridae*, à l'instar de celle des *Geminiviridae* et des *Luteoviridae*, ont été décrites comme étant transmises de manière circulante non-propagative. Ces trois familles semblent effectivement partager les caractéristiques propres à ce mode de transmission : les temps d'acquisition et d'inoculation de ces virus sont de quelques heures voire quelques jours, leur temps de rétention au sein de l'insecte dure plusieurs jours voire la vie entière de l'insecte, une période de latence de plusieurs heures à quelques jours empêche leur transmission directement après leur acquisition, et enfin ni réplication ni transcription au sein de l'insecte ne semblent pouvoir être détectées (Bressan and Watanabe, 2011).

De façon surprenante et intéressante, les données accumulées au cours de ma thèse indiquent que les interactions entre le FBNSV (membre des *Nanoviridae*) et son ou ses pucerons vecteurs pourraient être plus complexes que ce qui a été décrit jusqu'à aujourd'hui. Dans ce chapitre, je propose donc de reconsidérer la vision que l'on a de la transmission des *Nanoviridae* sur la base de l'étude des fréquences relatives du *Faba bean necrotic stunt virus* au sein de ses pucerons vecteurs. Ce chapitre est écrit sous la forme d'un article qui est actuellement soumis pour publication.



## **ARTICLE 2**

# **CIRCULATIVE NON-PROPAGATIVE APHID-TRANSMISSION OF NANOVIRUSES : AN OVERSIMPLIFIED VIEW**

Anne Sicard, Jean-Louis Zeddou, Michel Yvon, Yannis Michalakis, Serafin Gutierrez and  
Stéphane Blanc.

Soumis pour publication

## 2. Circulative non-propagative aphid-transmission of nanoviruses : an oversimplified view (Article 2)

Anne Sicard<sup>1</sup>, Jean-Louis Zeddam<sup>2</sup>, Michel Yvon<sup>1</sup>, Yannis Michalakis<sup>3</sup>, Serafin Gutierrez<sup>4</sup> and Stéphane Blanc<sup>1</sup>.

*1. INRA, UMR BGPI, Cirad TA-A54/K, Campus International de Baillarguet, 34398 Montpellier, France.*

*2. IRD, UMR RPB, IRD-CIRAD-UM2, Centre IRD, 911, Avenue Agropolis, BP 64501, Montpellier Cedex 5, France*

*3. CNRS, UMR MIVEGEC 5290, CNRS-IRD-UM1-UM2, 911 Avenue Agropolis, B.P. 64501, 34394 Montpellier, France.*

*4. CIRAD, CMAEE Unit, CIRAD, TA-A54/K, Campus International de Baillarguet, 34398 Montpellier, France*

Corresponding author : [blanc@supagro.inra.fr](mailto:blanc@supagro.inra.fr)

### Key words:

Nanovirus, Multipartite viruses, Faba bean necrotic stunt virus, aphid transmission, gene copy number, arbovirus, genome formula

## ABSTRACT

Plant virus species of the family *Nanoviridae* have segmented genomes with the highest known number of segments encapsidated individually. They thus likely represent the most extreme cases of the so-called multipartite or multicomponent viruses. All species of this family are believed to be transmitted in a circulative non-propagative manner by aphid vectors, meaning that the virus simply crosses cellular barriers within the aphid body, from the gut to the salivary glands, without replicating or even expressing any of its genes. However, this assumption is largely based on analogy with the transmission of other plant viruses such as geminiviruses or luteoviruses, and the details of the molecular and cellular interactions between aphids and nanoviruses is actually poorly investigated. When comparing the relative frequencies of the eight genome segments in populations of the species *Faba bean necrotic stunt virus* (FBNSV, genus *Nanovirus*) within host plants and within aphid vectors fed on these plants, we evidenced reproducible changes of the frequency of some specific segments. We further show that these changes occur at early stages of the virus cycle in the aphid, within the gut, and not later when the virus is translocated into the salivary glands. This peculiar observation, which was confirmed in the three aphid vector species *Acyrtosiphon pisum*, *Aphis craccivora* and *Myzus persicae*, calls for a revisiting of the mechanisms of nanovirus transmission. It reveals an unexpected intimate interaction that may not fit the canonical circulative non-propagative transmission.

## INTRODUCTION

Plant viruses of the families *Luteoviridae*, *Geminiviridae* and *Nanoviridae* are transmitted in a circulative (persistent) non-propagative manner by their insect vectors (Blanc et al., 2014; Hogenhout et al., 2008). This transmission mode is characterized by the internalization of the virus within the insect without replication or even transcription of the viral genome. After feeding on an infected plant, the virus is ingested by the vector along with the sap. The virus is then transported across the insect gut and released into the haemocoel, before penetrating into the salivary glands and being injected with saliva during feeding on a new plant. Though not replicating, the virus can remain infectious in the vector for several days, and even throughout the vector's lifespan in some cases. The transcytosis mechanisms by which luteoviruses cross the gut and salivary gland epithelia in aphid vectors have been relatively well described by electron microscopy. The transmitted virus particles are transported across cells into membrane vesicles preventing any contact between the virus and the insect cell cytoplasm (Brault et al., 2007). However, no such evidence is available for either geminiviruses or nanoviruses, while putative deviations from the canonical circulative non-propagative transmission model can be suspected from the related literature. For example, in several studies, transient replication of the geminivirus *Tomato yellow leaf curl virus* (TYLCV) within its whitefly vector could not be totally excluded (Czosnek et al., 2001; Mehta et al., 1994; Sinisterra, 2005; Díaz-Pendón et al., 2010). In another example, an unknown virus-encoded factor has been demonstrated to be mandatory for successful access and/or passage of the nanovirus *Faba bean necrotic yellows virus* (FBNYV) through the salivary glands of the aphid vectors (Franz et al., 1999). The involvement of this additional factor indicates in the least that the nanovirus-aphid relationship might be more complex than (or different from) that described for luteoviruses.

Virus species of the family *Nanoviridae* have their genome segmented into 6-8 segments of ~1 kb each. A single copy of each segment is encapsidated into individual icosahedral virus particles of 18-20 nm diameter (Grigoras et al., 2014). Five segments are common to all nanovirus species: the segments C, M, N, R and S. The segment C encodes a protein interacting with the cell cycle, the segment M encodes the movement protein, the segment N encodes a nuclear-shuttle protein, the segment R encodes a protein initiating replication of all segments, and the segment S encodes the capsid protein packaging each segment individually. The other segments, named U1, U2 and U4 for the genus *Nanovirus* and U3 for the genus *Babuvirus*, encode proteins of unknown functions (Timchenko and Bernardi, 2007). All member species

of the family *Nanoviridae* are aphid-transmitted, and the viral form crossing gut and salivary gland barriers has been proposed to be the viral particle (Chu and Helms, 1988; Harding et al., 1991; Katul et al., 1993). Though nanovirus particles have never been directly visualized by electron microscopy inside the insect vector, studies using immuno-fluorescence against the coat protein of *Banana bunchy top virus* have consistently demonstrated that this virus accumulates within the aphid anterior midgut, before being released in the haemolymph and translocated in the principal salivary glands (Watanabe and Bressan, 2013).

We have recently shown that each of the eight segments composing the genome of the Faba bean necrotic stunt virus (FBNSV) reproducibly accumulates at a specific relative frequency within infected host plants, the corresponding segment-frequency pattern defining the “genome formula” (Sicard et al., 2013). Because in the same study the genome formula appeared to be specific of the host plant species, we questioned whether it could also be affected in a specific way during the transit of the virus within the body of aphid vectors. We here show that the FBNSV genome formula within aphids is significantly different from that observed within source plants. The observed frequency changes affect primarily the segments N and U2. They are similar in three distinct aphid vector species, *Acyrtosiphon pisum*, *Aphis craccivora* and *Mysus persicae*, and occur very early during the internalization of the virus within midgut cells, with no further changes at later stages of the virus transfer across the aphid body. Although we cannot explain such changes in the viral genome formula at this stage, they suggest the existence of unforeseen intimate interactions between the FBNSV and its vectors.

## **MATERIALS AND METHODS**

### ***Host Plants***

*Vicia faba* (var. “Sevilla”, Vilmorin) and *Medicago truncatula* (Jemalong A17) plants were grown within an S2 restricted-access confinement facility, in a growth chamber with a 13/11h day/night photoperiod, a temperature of 26/20°C (day/night), and 70% hygrometry.

### ***Virus Isolate***

The FBNSV infectious clone used in this study is derived from an Ethiopian isolate maintained by serial passages on *Vicia faba* during three years, and corresponds to the infectious clone used in previous studies (Grigoras et al., 2009, 2010; Sicard et al., 2013) Briefly, each of the eight FBNSV genome segments is cloned as a tandem-repeat in pBin19 plasmid and transformed into

the *Agrobacterium tumefaciens* COR308 strain. To inoculate faba bean plants, eight agrobacteria cultures (each containing one of the eight cloned segments) were grown separately in NZY+ medium before being mixed in equal proportions (equal OD<sub>600</sub>), centrifuged and resuspended in a water solution containing 10 mM MgCl<sub>2</sub> and 150 µM Acetosyringone. Young *V. faba* plants were then needle inoculated in the stem as previously described (Grigoras et al., 2009; Sicard et al., 2013; Timchenko, 2006).

Young *Medicago truncatula* plants were inoculated by *Acyrtosiphon pisum* as described (Sicard et al., 2013), and characteristic FBNSV symptoms appeared on both plant species between 10 and 25 dpi.

### ***Aphid rearing and transmission of FBNSV***

Colonies of the aphid species *Acyrtosiphon pisum* and *Aphis craccivora* were reared on *Vicia faba* plants, whereas those of *Myzus persicae* and *Aphis gossypii* were respectively maintained on eggplants and zucchini. All reared aphids were maintained under similar controlled conditions (25/19°C day/night and a photoperiod of 13/11 h) ensuring reproduction through parthenogenesis.

For transmission of FBNSV, aphids were given an acquisition access period (AAP) of three days on infected *V. faba* or *M. truncatula* plants (at 30 days post inoculation (dpi)). Aphids were then transferred to two-leaf stage healthy *V. faba* plants or to three-leaf stage *M. truncatula* for an inoculation access period (IAP) of three additional days, except otherwise indicated. The aphids were then collected individually or by groups of five for further analysis of their viral contents, and the plants were finally treated with the insecticide Pirimor G (1 g/l in water). It should be noted that, because some aphids run off the plants or eventually die, we were not always able to find them all after three days of inoculation access period. This explains most differences observed between the number of aphids used for inoculation and the final number of aphids used in various analyses.

### ***Aphid dissection***

Guts of *A. pisum* individuals were pulled out of the bodies under a stereomicroscope, and the corresponding heads were severed and set aside because they contain the salivary glands. DNA from the gut and the head of each individual was extracted with PureLink Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, USA) and analyzed by QPCR.

### ***EDTA-facilitated exudation***

Phloem sap was collected as earlier described (King and Zeevaart, 1974). Apices of infected *V.faba* plants were severed with a razor blade and immediately rinsed in a solution of 20 mM EDTA (pH=7,0) to avoid contamination from wounded cells. Apices were then individually immersed in a tube containing 300 µl of the same solution for one hour in the dark at 26°C, and finally transferred in a second tube containing 300 µl of water for seven hours. Total DNA content from the sap exudates was extracted from both EDTA and water solutions according to (Edwards et al., 1991), with a longer centrifugation time after isopropanol addition (25 minutes). DNA from each plant apex initially collected was finally resuspended in 25 µL of water.

### ***DNA extraction from plants and whole aphids***

Total DNA from single aphids, or from groups of five, was extracted with a protocol described for *Bemisia tabaci* (Delatte et al., 2005). Briefly, single aphids were ground in 10 µL of extraction buffer containing 50mM KCl, 10mM Tris-base, pH 8, 0.45% Nonidet P-40, 0.45% Tween 20, and 500 µg/mL proteinase K. Fifteen µL of extraction buffer were then added to each of the extracts, which were then incubated at 65°C for 1h, followed by an incubation at 95°C for 15 min. After final addition of 35 µL of water per extract, they were stored at -20°C until use. The volume of extraction buffer and water was proportionally adjusted when extracting DNA from groups of five aphids.

One leaflet of the last leaf level of each plant of *M.truncatula* or *V.faba* was removed at 30 dpi and its DNA extracted as described by (Edwards et al., 1991) with an additional washing step with ethanol 70%. The DNA of leaflets of *M.truncatula* was extracted in 300 µL of extraction buffer whereas the DNA of leaflets of *V.faba* was extracted in 400 µL of extraction buffer.

For DNA extracted from both aphids and plants, Q-PCR analyses were carried out on ten-fold diluted samples.

### ***Q-PCR conditions***

All Q-PCR reactions were carried out using the LightCycler FastStart DNA Master Plus SYBR green I kit (Roche) on the LightCycler 480 thermocycler (Roche), exactly as previously described (Sicard et al., 2013). The primers used at a final concentration of 0.3 µM are those

described in the supplementary table S6 in (Sicard et al., 2013). For DNA extracted from both aphids and plants, Q-PCR analyses were carried out on ten-fold diluted samples. Fluorescence data were normalized through standard curves and analyzed with the LinRegPCR program (Ruijter et al., 2009)

The relative frequency of each FBNSV genome segment was calculated as the estimated copy number of a given segment divided by the total copy number of the eight segments. The total viral accumulation was obtained by summing the estimated number of copies of each of the eight segments.

### ***Statistical analysis***

All statistical tests have been carried out with the R software (R Development Core Team, 2011, version 2.12.0) - except for the correlation test for which JMP10 software was used. The nature and the results of the test are indicated throughout the text.

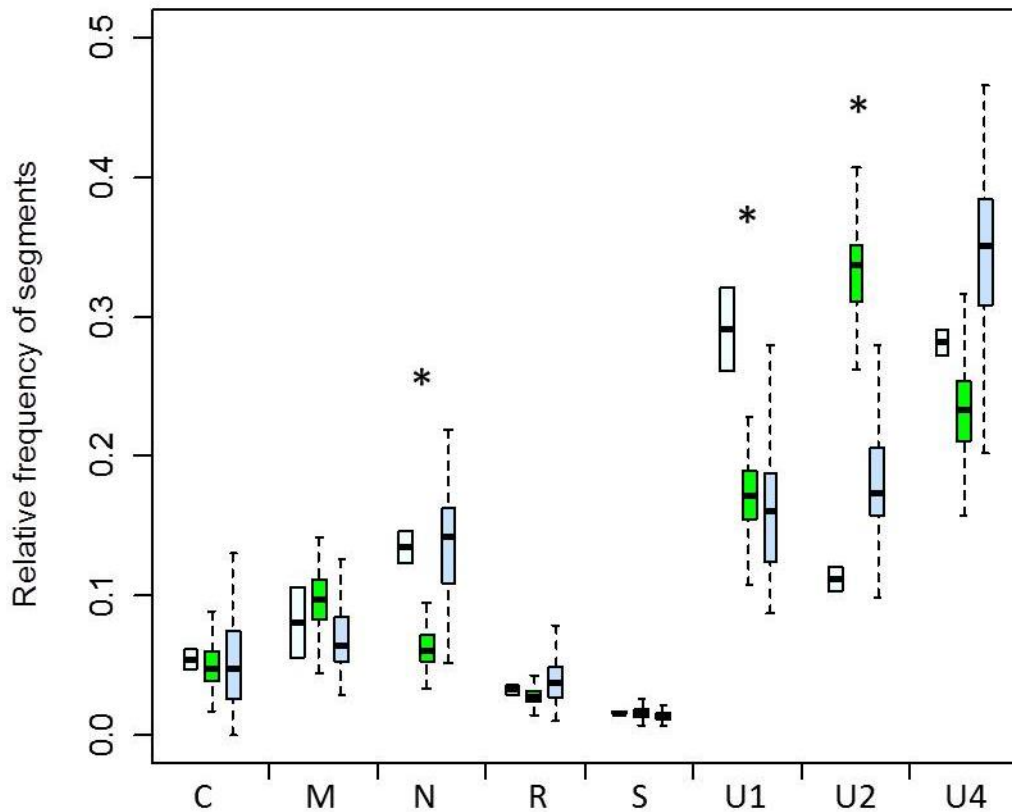
## **RESULTS**

### ***The relative frequency of the FBNSV genome segments changes within aphids.***

One hundred aphids of the species *Acyrtosiphon pisum*, an efficient vector of FBNSV (Grigoras et al., 2009), were allowed a 3-days acquisition access period (AAP) on the apex of an agroinfiltrated symptomatic faba bean “source” plant at 30 dpi. At the end of the AAP, aphids were transferred individually on one hundred healthy faba bean “test” plants for an additional 3-days inoculation access period (IAP). Within 10 to 20 days, nearly half of the test plants proved successfully infected in two repeats of the experiments (respectively 38% and 51% of test plants infected). The relative frequency of each of the eight FBNSV genome segments was then estimated by QPCR in the last leaf level of both the source and the symptomatic test plants, and in individual aphids. Over the two experiment repeats we could successfully quantify all eight FBNSV genome segments from the two source plants, from 54 infected test plants and from 117 individual aphids (Figure 1). Consistent with previously reported results, the relative frequencies of all segments were similar in both source and test plants, as the FBNSV genome formula is highly reproducible in a given host species (Sicard et al., 2013), except for segment U1. However, we found that the frequency of segments N and U1 decreased significantly within aphids, as compared to the source plants, whereas the frequency of segment U2 sharply



increased (Figure 1). No significant difference between source plants and aphids was found for the five other segments.



**Figure 1: Relative frequency of FBNSV segments in host plants and aphid vectors**

One hundred aphids of the species *Acyrtosiphon pisum* were fed on the same FBNSV-infected faba bean plant during three days before individual transfer onto test plants for three additional days. This experiment was repeated twice and the results have been pooled in this figure. The relative frequency of each segment was evaluated within source plants (light blue boxplot, n=2), individual aphids (green boxplot, n=117) and test plants (blue boxplot, n=54). The virus was detectable in almost all aphids, whereas only the 46% of the test plants developing symptoms have been analyzed. However, some aphids and plants were lost due to death, extraction or analysis problem. The name of each segment is indicated under the graph. Significant differences between source plants and aphids were found for segments N, U1 and U2 indicated by an asterisk on the graph (p-value= 4.8e-09, 1.2e-4 and 4.7e-09 respectively by unpaired Student's t test, p-values>0.29 for all other segments). All segments but U1 (p-value= 2.9e-4 by unpaired Student's t test) show similar frequencies in source and test plants.

The results shown in Figure 1 suggest a change of the FBNSV genome formula within aphids. However, the genome formula in source and test plants is deduced from whole leaflets containing virus DNA not only from the sieve tubes, but also from companion and phloem parenchyma cells. Because during the 3-days AAP the aphids are actually operating long sustained feeding in sieve tubes, one cannot exclude that the observed differences in aphids may reflect a distinct genome formula in the phloem sap, specifically ingested by aphids in this experiment. To evaluate this possibility, we thus decided to compare the FBNSV genome formula in whole leaf extracts and in phloem sap exudates. The apices of seventeen infected faba bean plants were cut, rinsed, and used individually to collect phloem sap (see Materials and Methods), prior to grinding for total DNA extraction from whole leaves from the apex. The segments' frequencies in the sap exudates were calculated and compared to those in whole corresponding leaves (supplementary Figure S1). The FBNSV genome formula in the exudates appears similar to that in whole leaves, confirming that the differences observed in aphids do not stem from differences in distinct plant compartments.

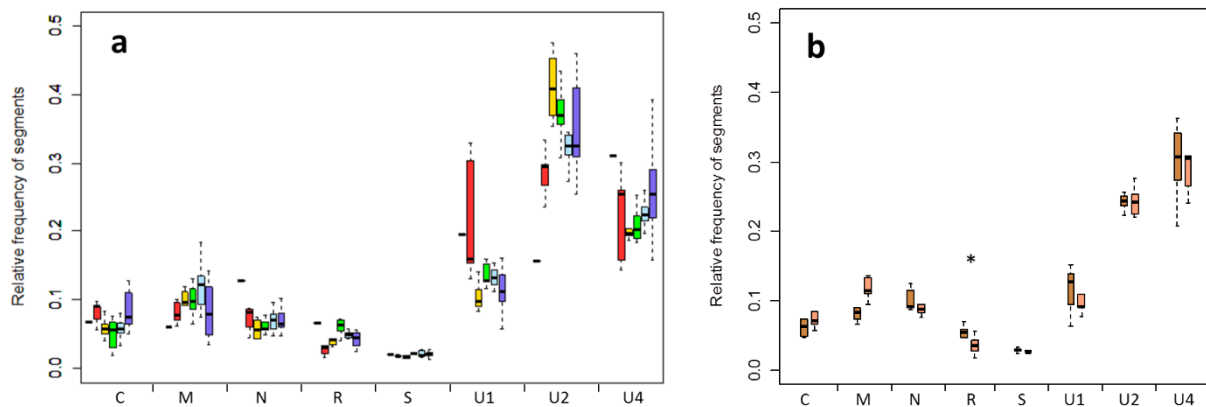
#### ***Changes of FBNSV genome formula occur at early stage within the aphid gut.***

To determine whether these frequency changes in aphids are a continuous process or only transient, we allowed aphids to feed on the apex of a faba bean infected plant and collected them at different time points. Collected aphids were then “purged” for at least four hours on healthy plants before DNA extraction from groups of five individuals and QPCR analysis. The different FBNSV genome segments did not reach sufficient amounts and could not be reliably quantified in aphids before a minimum AAP of 6 hours, a time at which the frequency changes were already completed (Figure 2a), with a remarkable drop of the frequency of N and rise of that of U2. No other significant changes of the segments' frequency were observed at later times, with the exception of segment U1 where a decrease in frequency was only significant from the second time point (figure 2a).

To support the conclusion that the change in genome formula is rapid and occurs at early stages of the aphid-FBNSV interaction, we decided to compare it within dissected aphid guts and within corresponding severed heads, containing the salivary glands. Guts and heads of six *A. pisum* individuals were separated from the rest of the aphid body and individually analyzed by

QPCR. As illustrated in Figure 2b, the frequency of the FBNSV segments in aphid heads -most likely in the salivary glands- proved very similar to that found within guts. We thus conclude

that the FBNSV genome formula is altered either directly in the gut lumen, during internalization into gut cells, or soon after, and that it subsequently remains stable during release into the haemolymph and translocation into the salivary glands. This phenomenon and the possible underlying mechanisms are further discussed later.



**Figure 2: Changes in the relative frequency of FBNSV segments occur at early stages within aphids**

**(a)** Hundreds of aphids of the species *Acyrtosiphon pisum* were fed on the same faba bean source plant. Seven groups of five aphids were removed after an AAP of 6h (red), 10h (yellow), 25h (green), 49h (light blue) and 74h (dark blue). For all time points, aphids were then purged on healthy plants for at least four hours before analyzing their viral content. The relative frequency of each segment in the source plant is indicated by a horizontal black bar. Though sometimes small, significant frequency differences between source plant and aphids are evidenced from the first time point for all segments except C and U1 (p-value < 0.041 for segments M, N, R, S, U2 and U4 by unpaired Student's t test), the most striking changes being the drop of N and the rise of U2. No drastic further changes were observed at later time points, apart from U1 for which a significant difference was observed along time (Anova test,  $F_{1,36}=4.56$ , p-value = 0.039 for segment U1).

**(b)** Dissected guts (left boxplots for each segment) and heads (right boxplots) of six aphids of the species *A. pisum* were analyzed for their viral contents after an AAP on an infected plant and an IAP on a healthy plant of three days each. No significant frequency change between guts and heads could be detected for the different segments, except for a slight drop of segment R frequency in the head indicated by an asterisk on the graph (p-value =  $4.5e-4$  for segment R, pvalue > 0.11 for other segments by paired Student's t test).

### ***Genome formula in aphids depends only partly on the initial formula in source plants***

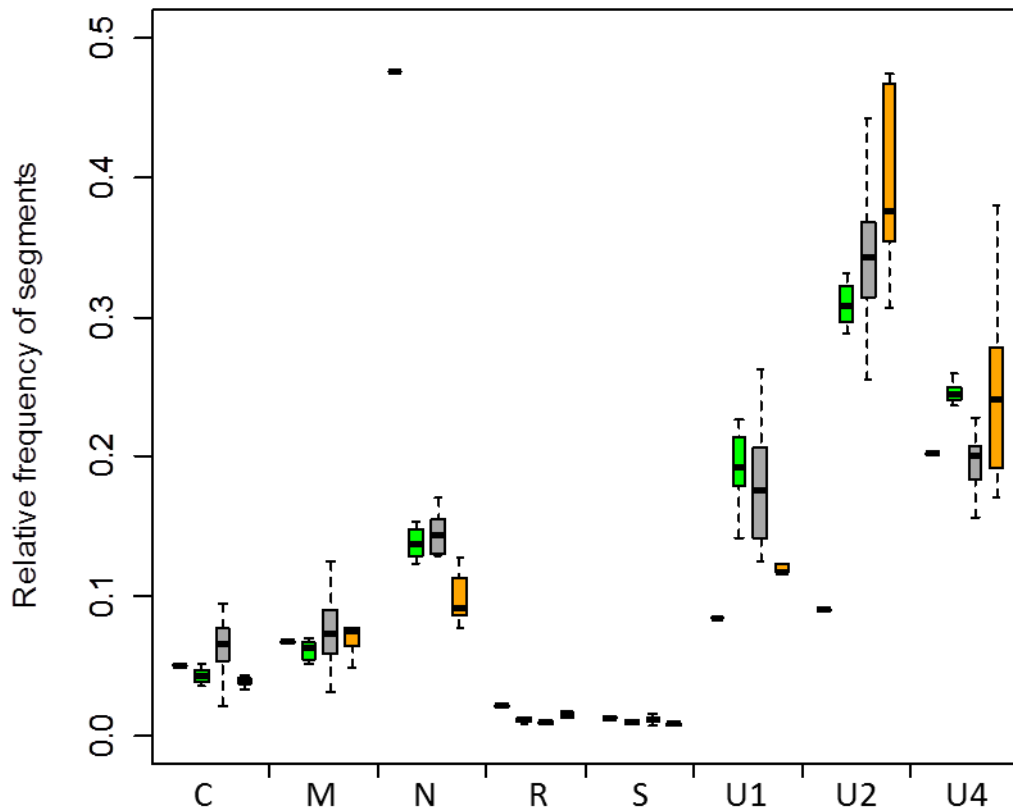
An important follow up question is whether the FBNSV genome formula, and particularly the frequency of segments N and U2 (those primarily changing in aphids, see Figures 1 and 2), tends to converge toward a specific end point in aphids whatever the initial situation in the ingested plant sap. For this, we tested whether the distribution of the frequencies of these two segments appears more constrained within aphids than it is in the corresponding source plants. Six independent source plants have been used for this analysis. The source plants already used in Figure 1 (2 source plants), those used in Figures 3 (1 source plant) and 4b (1 source plant), plus 2 source plants from two additional experiments which details are provided in supplementary Figure S2. By comparing the coefficient of variation of the most affected segments, N and U2, within these source plants and within the corresponding aphids, we found that the frequency distribution of segments N and U2 was less dispersed among aphids (respectively 38,8% and 18,1%) than among source plants (respectively 73,1% and 69,5%). This result suggests the existence of a yet unknown aphid-related constraint that tends to decrease the relative frequency of N and increase that of U2, whatever the initial situation within source plants.

We also tested a possible correlation between the decrease of the relative frequency of N and the increase of that of U2, because we thought that the drop of one could actually “compensate” the rise of the other, or vice versa. No significant correlation could be detected (Pearson’s correlation test,  $\rho = 0.083$ ,  $p\text{value} = 0.29$ ), showing that the changes in the relative frequency of these two segments are independent.

### ***Aphid-related FBNSV genome formula is conserved in distinct aphid vector species***

In addition to *A. pisum* used in this study, the FBNSV has been reported to be transmitted by other aphid species, including *Aphis craccivora* (Grigoras et al., 2009), which we decided to test for a similar effect on the FBNSV genome formula. In parallel, we tested a third aphid species, *Myzus persicae*, because it has earlier been described as a non-vector of the related nanovirus species *Faba bean necrotic yellows virus* (FBNYV) (Katul et al., 1993). Somewhat to our surprise, *Myzus persicae* proved to be as efficient a vector of FBNSV as are *A. pisum* and *A. craccivora*. Depending on the experiment, and using 5 aphids per test plant, *A. pisum* transmitted the FBNSV to 18% to 100% of the test plants, *A. craccivora* to 81% to 90% and *M. persicae* to 31% to 100%.

Over 100 individuals of each of these three aphid species were caged together on the same infected faba bean source plant for a 3-day AAP. Then, groups of five individuals of the same species were each transferred to 16 or 17 test plants per species for a 3-days IAP. At the end of the IAP, total DNA was extracted from each group of five aphids to determine the FBNSV genome formula in *A.pisum*, *A.craccivora* and *M.persicae*, and compare it to that in the unique source plant (Figure 3). Although some viral genome segments proved to reach a slightly different frequency, the overall genome formula was very similar across the three aphid species and quite distinct from that in the initial source plant. In particular, this experiment confirms that the more drastic and reproducible changes, whatever the vector species, are a sharp drop and increase of the frequency of the segments N and U2, respectively.



**Figure 3: FBNSV genome formula is similar in distinct aphid vector species**

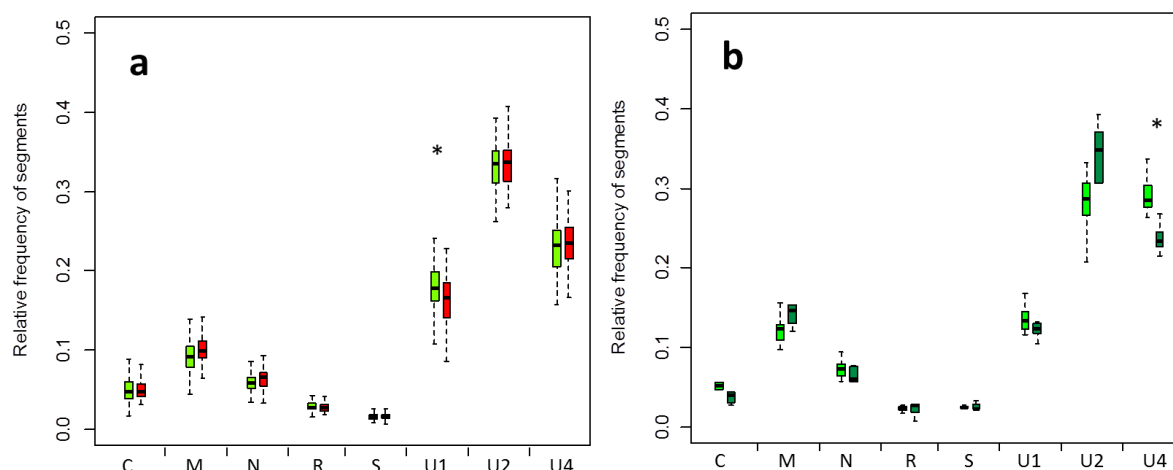
The relative frequency of the FBNSV genome segments was determined in three distinct aphid vector species, *Acyrthosiphon pisum* (left green boxplots,  $n=11*5$ ), *Aphis craccivora* (center grey boxplots,  $n=13*5$ ) and *Myzus persicae* (right orange boxplots,  $n=8*5$ ). All aphids were fed during three days on the apex on the same faba bean source plant. They were then transferred as groups of five of the same species onto healthy plants for an additional three days inoculation period. Finally, they were collected for DNA extraction and QPCR. The unique source plant has also been analyzed, and the horizontal black bars indicate the corresponding frequency values for each segment. Although the three aphid species show similar relative frequencies for each segment, slight but significant differences were found between aphid species for four out of the eight segments (ANOVA test,  $F_{2,26} = 9.97, 4.32, 6.60$  and  $6.50$ ;  $p$ -value=  $6.1e-04, 2.4e-2, 4.8e-3$  and  $5.2e-3$  for the segments C, U1, U2 and U4 respectively). The relative frequencies of the segment C, N, R, U1 and U2 proved to be significantly different between the source plant and the three aphid species (unpaired Student's  $t$ -test,  $p$ -values  $<0.043$  for all those comparisons), with segments N and U2 showing the most striking differences.

### ***Segments' frequency within aphids and transmission efficiency***

Finally, we questioned whether the changes of the FBNSV genome formula within aphids play a role in the success of transmission. We first further analyzed the data set presented in Figure 1. There, *A. pisum* aphids had been tested individually for transmission (a single aphid per test plant). It is thus possible to compare the FBNSV genome formula within those aphids that actually successfully transmitted the virus ( $n=54$ ) to that within those that failed to transmit ( $n=63$ ). Figure 4a shows that the genome formula is similar in both aphid sets, indicating that,

at least in this experiment, the success or failure in transmission by individual aphids was not related to an altered genome formula of the acquired FBNSV. We also compared the total viral load within aphids by summing up the copy number of all eight segments. This viral load proved similar in the two aphid sets (not shown).

We then tested whether the genome formula observed in aphid vectors differ in non-vector species. Aphids of the species *Aphis gossypii* (considered a non-vector species) were caged together with aphids of the species *A. pisum* on the apex of an infected faba bean source plant for a 3-day AAP. Groups of five aphids of the same species were then each transferred onto a test plants for a 3-day IAP. In this experimental setting, *A. pisum* transmitted successfully the virus to 7 out of 23 test plants, whereas *A. gossypii* failed to transmit to any of fifteen test plants. Despite the fact that *A. gossypii* proved to be a non- or poor-vector species in our experimental conditions, similar amounts of virus accumulated in both *A. pisum* and *A. gossypii* (data not shown). Moreover, a similar genome formula was estimated for both species (Figure 4b, and discussed below).



**Figure 4: FBNSV genome formula within aphids and transmission success**

**(a)** Comparison of the relative frequency of FBNSV segments in individual *A. pisum* that have efficiently transmitted the virus (in green,  $n=54$ ) in the experiment described in Figure 1, and in those that have not (in red,  $n=63$ ). No significant difference can be detected between these two groups except for the segment U1 (unpaired Student's t-test,  $p$ -value = 0.014 for segment U1 ;  $p$ -value  $> 0.17$  for the seven other segments).

**(b)** Comparison of the relative frequency of FBNSV segments within the vector species *Acyrthosiphon pisum* (light green,  $n=10*5$ ), and the non-vector, *Aphis gossypii* (dark green,  $n=5*5$ ). Both aphid species were given an AAP of three days on the same faba bean individual source plant and were then transferred by groups of five aphids on healthy plants for an additional IAP of three days. Apart from segment U4, no significant difference was found between the two aphid species (ANOVA test,  $F_{1,13}=20.7$  and  $p$ -value =  $5.45e-4$  for segment U4 ;  $p$ -value  $> 0.065$  for the seven other segments).



## Discussion

The process underlying the changes of the genome formula of FBNSV within aphids are thus far unknown. One obvious possibility, although not consistent with the circulative non-propagative mode of transmission generally assumed for nanoviruses, is that some -or all- FBNSV genome segments could be decapsidated within the aphids so that replication or transcription could occur. Mechanisms similar to those we earlier proposed (Sicard et al., 2013) for the establishment of the setpoint genome formula within host plants could then also operate within aphids. That the relative frequency of the most affected segments N and U2 appears more constrained among the various aphid sets analyzed than among the corresponding source plants may support this possibility. Because the replication cycle of some dsDNA virus can be completed within a few minutes (De Paepe and Taddei, 2006), it is impossible with the currently available data to exclude a very short and transient phase of FBNSV replication within aphids' gut cells. The absence of further changes of the genome formula at later time points (e.g. in the salivary glands, Figure 2), would be consistent with this putative replication being early and short-lived, and overlooked in the previous investigation on nanovirus transmission (Hafner et al., 1995).

Alternatively, the observed changes in aphids could be attributed to a differential decapsidation and degradation of some segments. But this would involve unknown distinct physical properties (stability) of virus particles depending on the contained genome segment, N- and U2-containing particles being highly labile and stable, respectively. While this particle stability-related hypothesis fits the case of segments N and U2, respectively decreasing and increasing in relative frequency in all experimental repeats, it cannot explain why for example U1 and U4 behave in a more inconstant way, sometimes decreasing and sometimes increasing relative to the source plant, depending on the experiment (see detailed data in supplementary Figure S2).

As a third possibility, physicochemical differences of virus particles depending on the contained segments could also account for a differential interaction with the aphid midgut putative receptor(s). It is widely recognized that the first cellular barrier in insect vectors encountered by circulative viruses is the gut barrier. The time at which a nanovirus passes this barrier has not yet been reported, but geminiviruses and luteoviruses cross it within a few hours: respectively 1 hour and 12-16 hours (Garret et al., 1996; Ghanim et al., 2001). FBNSV could thus enter the aphid gut cells very rapidly and a differential uptake for different segment-containing particles could explain the observed changes in genome formula. Nevertheless, this

hypothesis suffers the same drawbacks as that mentioned above for the differential degradation: while it could fit the observation for N and U2, it cannot explain the inconstant behavior of U1 and U4.

At this stage, we do not know which of the above hypotheses explains our results. It is worth noting, however, that segments N and U2, whose frequency is mostly affected in aphids, have recently been shown to be those controlling the efficiency of aphid-transmission of FBNSV (Grigoras et al 2015, in press). In this study, the authors demonstrate that the absence of N has little effect on virus accumulation and symptom expression *in planta* but totally abolishes vector-transmission. They conclude that N (either the segment or the protein) is the helper factor of nanoviruses for aphid transmission, which requirement was first suspected by Franz and coworkers (Franz et al., 1999). In the same study they also showed a strong negative effect of the absence of U2 in source plants on the efficiency of aphid-transmission. With all possible cautions, we here propose that this study by Grigoras et al (Grigoras et al 2015, in press), and the uncorrelated changes of the frequency of N and U2 in aphids shown here, together suggest that these two genome segments have a distinct function in the regulation of FBNSV-aphid interaction.

The ultimate question is whether such changes of the FBNSV genome formula within aphids are functional and actually play a role during transmission. The facts that the presence/absence of N and U2 in infected source plants greatly affect transmission (Grigoras et al 2015, in press) and that we show here that these same segments are those most affected in the aphid-specific FBNSV genome formula support this hypothesis. Moreover, all aphid vector species tested induce similar frequency changes of the FBNSV genome formula. The only non-vector species (*Aphis gossypii*) tested here is internalizing and accumulating the FBNSV in amounts similar to those detected in vector species. After a 3-days AAP on infected plants and an additional 3-days IAP on healthy plants, *A. pisum* and *A. gossypii* accumulated  $1.01 \times 10^6$  ( $\pm 1.24 \times 10^6$ ) and  $4.01 \times 10^5$  ( $\pm 3.56 \times 10^5$ ) total copies of FBNSV DNA, respectively. The similar changes of the FBNSV genome formula observed in both vector and non-vector aphids could thus be due to compatible early interactions in all cases followed by a transmission blocking at a later step for non-vectors only. Consistently, it has been shown that the passage across the salivary gland barrier can specifically block both luteoviruses and geminiviruses (Cohen, 1989; Gildow, 1993; Gray and Gildow, 2003). Further investigation is required to determine where the interaction is effectively blocked during the transit of the FBNSV within the *A. gossypii*.

At this point, we can only propose the speculations above to explain our observations. What is clear, however, is that we reveal intimate unknown interactions, which reproducibly change the relative frequency of some genome segments, and putatively regulate the virus-vector relationship. Investigation of the localization of segments N and U2, and their RNA and protein products, within the body of aphid vectors will probably help unraveling an unforeseen nanovirus-aphid interaction that may not fit into any of the pre-existing and long-standing category of plant virus vector-transmission.

### **Acknowledgments**

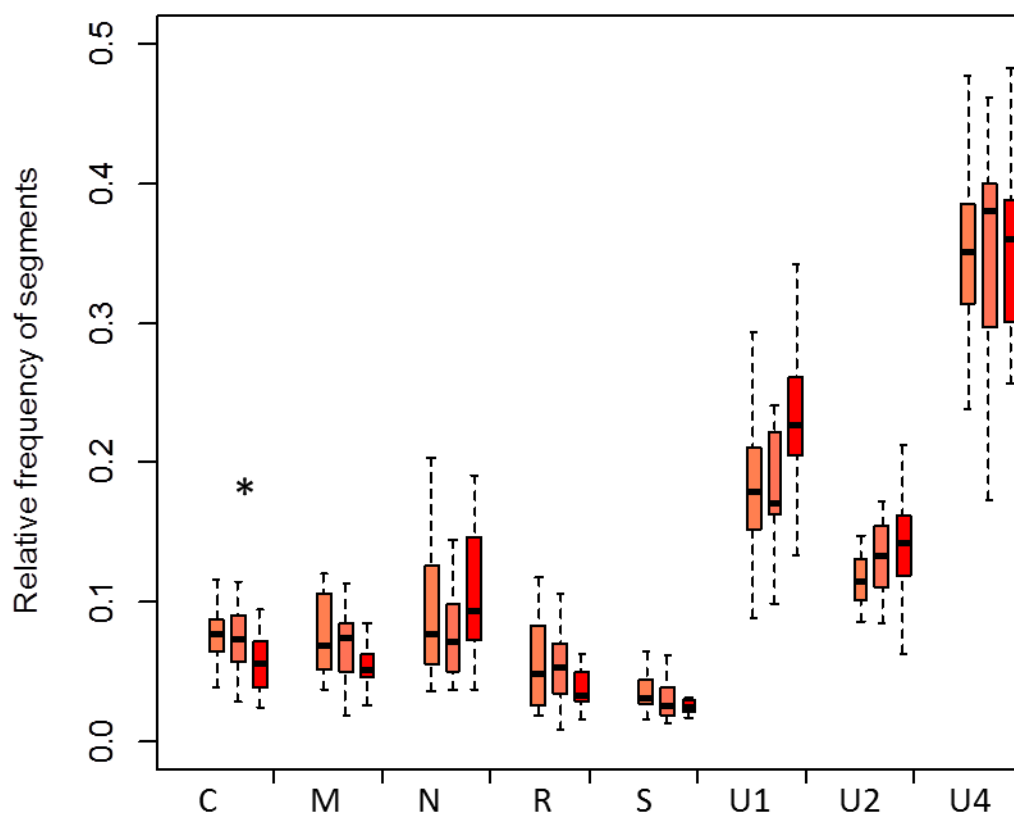
We are grateful to B. Gronenborn and T. Timchenko for technical help and advices. AS, SB MY and SG acknowledge support from INRA SPE division (project NanoVigs-SPE) and from the ANR-MOI grant (N°2010BLAN170401), YM from CNRS, and JLZ from IRD.

## References

- Blanc, S., Drucker, M., and Uzest, M. (2014). Localizing Viruses in Their Insect Vectors. *Annu. Rev. Phytopathol.* 52, 403–425.
- Brault, V., Herrbach, É., and Reinbold, C. (2007). Electron microscopy studies on luteovirid transmission by aphids. *Micron* 38, 302–312.
- Chu, P.W.G., and Helms, K. (1988). Novel virus-like particles containing circular single-stranded DNAs associated with subterranean clover stunt disease. *Virology* 167, 38–49.
- Cohen, S. (1989). Acquisition, Interference, and Retention of Cucurbit Leaf Curl Viruses in Whiteflies. *Phytopathology* 79, 109.
- Czosnek, H., Ghanim, M., Morin, S., Rubinstein, G., Fridman, V., and Zeidan, M. (2001). Whiteflies: Vectors, and victims (?), of geminiviruses. In *Advances in Virus Research*, (Elsevier), pp. 291–322.
- Delatte, H., Reynaud, B., Granier, M., Thornary, L., Lett, J.M., Goldbach, R., and Peterschmitt, M. (2005). A new silverleaf-inducing biotype Ms of *Bemisia tabaci* (Hemiptera: Aleyrodidae) indigenous to the islands of the south-west Indian Ocean. *Bull. Entomol. Res.* 95.
- Díaz-Pendón, J.A., Cañizares, M.C., Moriones, E., Bejarano, E.R., Czosnek, H., and Navas-Castillo, J. (2010). Tomato yellow leaf curl viruses: ménage à trois between the virus complex, the plant and the whitefly vector: Tomato yellow leaf curl viruses. *Mol. Plant Pathol.* 11, 441–450.
- Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19, 1349.
- Franz, A.W., van der Wilk, F., Verbeek, M., Dullemans, A.M., and van den Heuvel, J.F. (1999). Faba bean necrotic yellows virus (genus Nanovirus) requires a helper factor for its aphid transmission. *Virology* 262, 210–219.
- Garret, A., Kerlan, C., and Thomas, D. (1996). Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, *Myzus persicae* Sulz. *Arch. Virol.* 141, 1279–1292.
- Ghanim, M., Morin, S., and Czosnek, H. (2001). Rate of Tomato yellow leaf curl virus translocation in the circulative transmission pathway of its vector, the whitefly *Bemisia tabaci*. *Phytopathology* 91, 188–196.
- Gildow, F.E. (1993). Evidence for Receptor-Mediated Endocytosis Regulating Luteovirus Acquisition by Aphids. *Phytopathology* 83, 270.
- Gray, S., and Gildow, F.E. (2003). Luteovirus-aphid interactions. *Annu. Rev. Phytopathol.* 41, 539–566.
- Grigoras, I., Timchenko, T., Katul, L., Grande-Perez, A., Vetten, H.-J., and Gronenborn, B. (2009). Reconstitution of Authentic Nanovirus from Multiple Cloned DNAs. *J. Virol.* 83, 10778–10787.

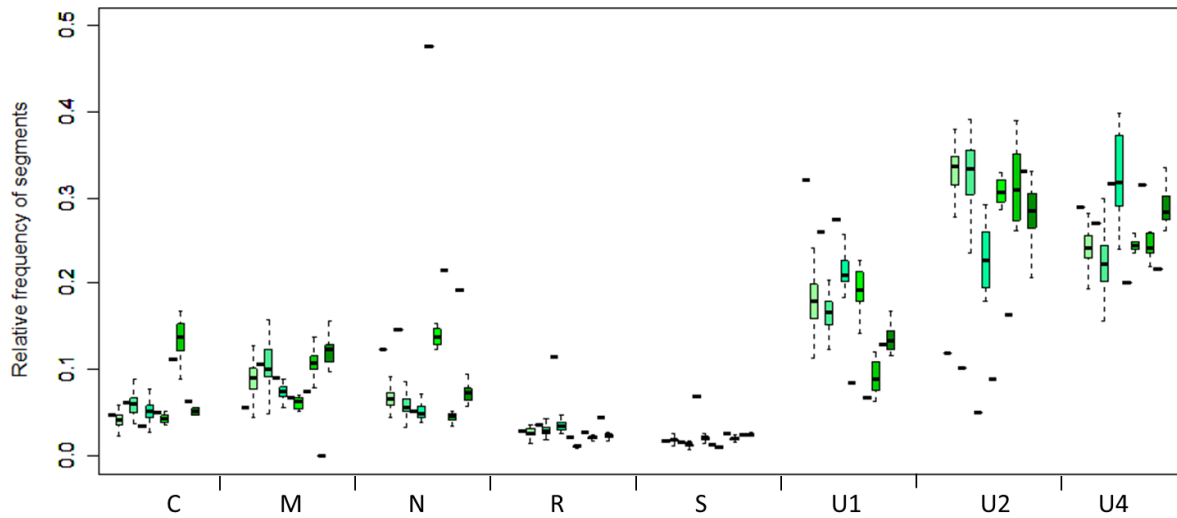
- Grigoras, I., Timchenko, T., Grande-Perez, A., Katul, L., Vetten, H.-J., and Gronenborn, B. (2010). High Variability and Rapid Evolution of a Nanovirus. *J. Virol.* *84*, 9105–9117.
- Grigoras, I., Ginzo, A.I. d. C., Martin, D.P., Varsani, A., Romero, J., Mammadov, A.C., Huseynova, I.M., Aliyev, J.A., Kheyr-Pour, A., Huss, H., et al. (2014). Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *J. Gen. Virol.* *95*, 1178–1191.
- Hafner, G.J., Harding, R.M., and Dale, J.L. (1995). Movement and transmission of banana bunchy top virus DNA component one in bananas. *J. Gen. Virol.* *76*, 2279–2285.
- Harding, R.M., Burns, T.M., and Dale, J.L. (1991). Virus-like particles associated with banana bunchy top disease contain small single-stranded DNA. *J. Gen. Virol.* *72 ( Pt 2)*, 225–230.
- Hogenhout, S.A., Ammar, E.-D., Whitfield, A.E., and Redinbaugh, M.G. (2008). Insect Vector Interactions with Persistently Transmitted Viruses\*. *Annu. Rev. Phytopathol.* *46*, 327–359.
- KATUL, U., Vetten, H.J., Maiss, E., Makkouk, K.M., LESEMANN, D.-E., and Casper, R. (1993). Characterisation and serology of virus-like particles associated with faba bean necrotic yellows. *Ann. Appl. Biol.* *123*, 629–647.
- King, R.W., and Zeevaart, J.A.D. (1974). Enhancement of phloem exudation from cut petioles by chelating agents. *Plant Physiol.* *53*, 96–103.
- Mehta, P., Wyman, J.A., Nakhla, M.K., and Maxwell, D.P. (1994). Transmission of Tomato Yellow Leaf Curl Geminiviruses by *Bemisia tabaci* (Homoptera: Aleyrodidae). *J. Econ. Entomol.* *87*, 1291–1297.
- De Paepe, M., and Taddei, F. (2006). Viruses' Life History: Towards a Mechanistic Basis of a Trade-Off between Survival and Reproduction among Phages. *PLoS Biol.* *4*, e193.
- Ruijter, J.M., Ramakers, C., Hoogaars, W.M.H., Karlen, Y., Bakker, O., van den Hoff, M.J.B., and Moorman, A.F.M. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* *37*, e45–e45.
- Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakos, Y., Gutierrez, S., and Blanc, S. (2013). Gene copy number is differentially regulated in a multipartite virus. *Nat. Commun.* *4*.
- Sinisterra, X.H. (2005). Differential transcriptional activity of plant-pathogenic begomoviruses in their whitefly vector (*Bemisia tabaci*, Gennadius: Hemiptera Aleyrodidae). *J. Gen. Virol.* *86*, 1525–1532.
- Timchenko, T. (2006). Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus. *J. Gen. Virol.* *87*, 1735–1743.
- Timchenko, T., and Bernardi, F. (2007). Les nanovirus, petits virus de plantes : similitudes et différences avec les géminivirus.
- Watanabe, S., and Bressan, A. (2013). Tropism, compartmentalization and retention of banana bunchy top virus (Nanoviridae) in the aphid vector *Pentalonia nigronervosa*. *J. Gen. Virol.* *94*, 209–219.

## Supplementary figures



**Supplementary Figure S1: Relative frequency of the FBNSV genome segments within whole leaf extracts and within sap exudates**

The relative frequency of each FBNSV genome segment, specifically within phloem sap, was analyzed by EDTA-facilitated exudation. After rinsing and immersion in a first solution of EDTA for one hour, the apices of seventeen faba bean plants were immersed in water for seven additional hours. The frequency of each segment was estimated both in exudates in the first EDTA solution, in exudates in the subsequent water solution (respectively the first and second boxplots) and within the corresponding whole leaves (third boxplots). The relative frequency of all segments was not significantly different in exudates (plant sap) and in whole leaves (ANOVA test,  $p$ -values  $> 0.058$  except for segment C,  $p$ -value = 0.029).



**Supplementary Figure S2: The aphid genome formula in aphids and in corresponding source plants: detail of six experimental repeats.**

Aphids of the species *A.pisum* were fed on *Vicia faba* plants for three days before being transferred to healthy plants. Six different aphid batches (in green) and their corresponding source plants (horizontal black bar immediately preceding each boxplot) are here represented. For aphid batches, from left to right : first and second boxplots correspond to the two aphid sets pooled in figure 1 (n=62\*1 and n=55\*1 respectively), the third boxplot is derived from the *A.pisum* set of Figure 3 (n=11\*5), the fourth boxplot corresponds to the *A.pisum* set of Figure 4b (n=10\*5), the fifth and sixth boxplots are additional experiments not used in other Figures (n=10\*5 and 19\*5 respectively).

### 3. Conclusion

Dans ce chapitre, nous avons démontré que les pucerons n'étaient pas de simples « passoires » traversées par le *Faba bean necrotic stunt virus* avant qu'il ne soit relargué au sein d'une autre plante. Une modification de la composition génétique du pool de particules virales ingérées par le puceron semble effectivement indiquer que le virus interagit de manière plus intime avec son vecteur que ce qui a été précédemment décrit. Le(s) mécanisme(s) à l'origine de cette modification reste(nt) pour le moment totalement inconnu(s) et ouvre ainsi de nouvelles pistes de recherche pour le laboratoire. Outre le(s) rôle(s) que peuvent jouer les segments les plus « affectés » au sein des vecteurs, une des questions principales reste donc de déterminer si le FBNSV se réplique au sein de son vecteur, classant ainsi les *Nanovirus* parmi les virus transmis de manière circulante propagative, ou si ces changements de formule génomique doivent être attribués à d'autres mécanismes qui seraient à ce jour inclassables.

Quoi qu'il en soit, la reproductibilité de ces changements de fréquences relatives au sein des différentes espèces de pucerons vecteurs invoque un mécanisme déterministe à l'origine de ces changements. Les modifications de formule génomique au sein des vecteurs du FBNSV pourraient être encore le résultat d'un ajustement rapide de ce virus au sein d'un nouvel environnement, mais cette interprétation n'est possible que s'il y a effectivement répllication du virus, même de manière transitoire, ce qui n'est pas démontré.

Les différences de fréquences relatives entre les segments, et notamment la présence de segments rares observées à la fois au sein des plantes (chapitre I) et au sein des pucerons (chapitre II), ramènent toujours à la même contrainte et au même questionnement : comment le virus « s'assure » que tous les segments co-infectent suffisamment de nouvelles cellules lors de l'invasion de la plante, et comment l'ensemble de ces segments peut être co-transmis à de nouveaux hôtes dans le paysage avec suffisamment d'efficacité. Ceci a suscité une réflexion importante au cours de ma thèse, pour laquelle nous voyons deux possibilités totalement opposées. La première est assez classique et suggère que le virus doit avoir mis en place des mécanismes spécifiques pour éviter la perte de segments à la fois au cours de la colonisation de la plante et lors de la transmission, mais de tels mécanismes ne sont pour l'instant pas « visibles » dans la littérature (cf la revue bibliographique qui traite spécifiquement de ce point).



La deuxième est plus « provocatrice », mais très simple, et propose que cette fameuse contrainte n'existe pas, le système pouvant fonctionner même si les segments sont temporairement séparés. Les deux chapitres suivants abordent directement ou indirectement ces deux possibilités.

**CHAPITRE III :**

**LES VIRUS MULTIPARTITES : UN  
FONCTIONNEMENT DECENTRALISE**

# Chapitre III : Les virus multipartites : un fonctionnement décentralisé

---

## 1. Introduction

Une, pour ne pas dire la, vision classique en virologie est que la réplication virale s'effectue au sein de cellules permissives individuelles dans lesquelles un génome viral est exprimé et répliqué, puis transmis aux cellules voisines (ou plus distantes). Ensuite, ce phénomène se réitère de manière identique et entraîne ainsi la colonisation de l'hôte. Suivant cette vision, le cycle de réplication d'un virus est considéré « cellule-autonome », c'est à dire que l'unité spatiale de la réplication virale est la cellule individuelle. Il paraît évident que ce point de vue est valide (en tout cas dans l'immense majorité des cas) pour les virus monopartites, et c'est d'ailleurs clairement d'un modèle générique/idéal de virus monopartites que découle la conceptualisation la plus commune de ce qu'est un virus, dans la communauté scientifique et au-delà.

En accord avec ce cadre conceptuel, l'infection d'une cellule pour un virus multipartite nécessiterait donc l'entrée d'au moins une copie de chaque segment génomique pour pouvoir y exprimer/répliquer la totalité de son information génétique. Si l'entrée des particules virales dans les cellules s'effectue de manière aléatoire, quel que soit le segment génomique qu'elles contiennent, le nombre minimum de particules requises pour assurer la présence de tous les segments dépend donc à l'évidence du nombre de segments constituant le génome viral. De nombreuses études théoriques ont analysé ce phénomène et le coût potentiellement associé (cf partie 1 de l'Introduction). Très récemment, une de ces études a conclu que, dès lors que le nombre de segments génomiques excède 3 ou 4, le nombre de particules requis à l'entrée des cellules individuelles pour éviter la perte d'information génétique est si élevé et improbable (selon les auteurs) qu'il est difficile de concevoir comment de tels systèmes viraux ont pu évoluer (Iranzo and Manrubia, 2012).

Le problème lié à ce coût est plus surprenant encore si l'on considère la fréquence relative de chacun des segments génomiques. En effet, il a été noté très tôt que si les fréquences des différents segments sont inégales, avec des segments rares et d'autres fréquents, alors le nombre

de particules requises pour assurer l'entrée d'un génome entier dans chaque cellule s'accroît encore (Pressing and Reaney, 1984). Ceci explique certainement pourquoi cette étude théorique et celle qui suivirent ont toujours calculé ce coût en considérant que les différents segments s'accumulaient à fréquence égale. Cependant, il est possible de trouver des indices d'une accumulation différentielle des différents segments chez plusieurs espèces de virus multipartites dans plusieurs études expérimentales qui ne s'adressaient pas directement à cette question (Feng et al., 2008, 2006; Habili and Francki, 1974; Lot and Kaper, 1976; Lot et al., 1974; Stace-Smith and Tremaine, 1973), un phénomène qui a été directement confirmé récemment (Sanchez-Navarro et al., 2013; Sicard et al., 2013).

Donc, si l'on résume : plus le nombre de segments est élevé et plus les différences de fréquence sont importantes, plus le nombre de particules requises pour assurer l'infection est grand et constitue un coût majeur pour les virus multipartites. Ce coût semble donc être maximisé pour notre modèle d'étude, le *Faba bean necrotic stunt virus* (FBNSV). En effet, non seulement ce virus est un des virus multipartites constitué du plus grand nombre de segments découvert à ce jour, mais aussi chacun de ces segments présente des fréquences relatives différentes variant entre 2% et 30% au sein de la plante hôte *V. faba* rendant ainsi certains segments relativement fréquents et d'autres très rares (cf chapitre I). Il est important de rappeler que la majorité des estimations empiriques du nombre de particules virales infectant les cellules des hôtes pluricellulaires (MOI) ont été obtenues chez des virus monopartites d'insectes (Bull et al., 2001) de vertébrés (Josefsson et al., 2011; Jung et al., 2002) et de plantes (Gonzalez-Jara et al., 2009; Gutiérrez et al., 2010) et donnent des valeurs comprises entre 1 et 13. Pour le FBNSV et sa formule génomique dans la fève, si (i) le nombre de particules entrant en moyenne par cellule est de 13, valeur maximale trouvée chez les virus monopartites, comme indiqué ci-dessus (Gutiérrez et al., 2012) et (ii) si les segments rentrent de manière aléatoire et indépendante sous forme de particules indistinguables quel que soit le segment qu'elles contiennent, alors dans près de 99% des cellules infectées, au moins un segment sera manquant (calcul en Annexe II). Ainsi, trois hypothèses sur le fonctionnement du système peuvent être formulées ; soit le nombre moyen de particules entrant effectivement par cellules est extrêmement élevé pour les virus multipartites - ici pour le FBNSV (hypothèse H1), soit ces virus font appel à un mécanisme encore inconnu pour contrôler l'entrée au sein de la cellule, qui permet de « trier » un petit nombre de particules et/ou de segments avec au minimum une copie de chaque (H2), soit enfin ils fonctionnent différemment et ce coût théoriquement insurmontable n'existe pas dans la pratique (H3). Comme déjà indiqué précédemment, le nombre de particules requises

pour que le FBNSV développe une infection non-avortée apparaît totalement démesuré et dépasse plusieurs centaines par cellules, rendant l'hypothèse H1 improbable. L'hypothèse H2 repose sur le fait que les particules virales contenant les différents segments diffèrent par leurs propriétés électrophysicochimiques, et peuvent être triées à l'entrée des cellules, où encore que le mouvement se fait sous forme d'un complexe de plusieurs molécules d'acide nucléique (une de chaque). Comme cela est discuté dans l'introduction (partie 1), aucune donnée ne permet pour le moment de favoriser ce scénario, en particulier dans le cas spécifique du FBNSV. Cette hypothèse H2 suggérerait par ailleurs que seule une copie de chaque segment pénètre dans chaque cellule, tout comme pour les virus segmentés, ce qui n'expliquerait donc pas l'intérêt de la multienapsulation. La distinction entre H1 (entrée massive et aléatoire) et H2 (entrée spécifiquement triée) peut être basée sur les expérimentations de calcul de la MOI de chaque segment du génome, dont le principe sera décrit dans le Chapitre IV. Ce qui nous intéresse ici, c'est d'imaginer autre chose que H1 et H2 qui sont des hypothèses difficilement soutenues par les données expérimentales disponibles dans la littérature.

L'hypothèse H3 est inspirée d'une remarque de Iranzo et Manrubia (Iranzo and Manrubia, 2012) qui, suite au constat de la difficulté d'expliquer l'existence des virus multipartites au-delà de 3 ou 4 segments génomiques, nous invite à revisiter notre conception de l'infection chez les virus multipartites. Cette conception, largement basée sur le fonctionnement des virus monopartites, pourrait effectivement être inappropriée pour expliquer le fonctionnement des virus multipartites. Nous avons simplement imaginé que la complémentation permettant de coordonner les différentes fonctions nécessaires au cycle de l'infection pourrait avoir lieu sans que tous les segments n'aient besoin d'infecter la même cellule. Les segments pourraient effectivement entrer de manière aléatoire au sein des cellules, une information génétique partielle étant ainsi répliquée et exprimée dans des cellules différentes. Ceci diminuerait considérablement le coût à l'entrée des cellules. Cependant, ce fonctionnement hypothétique impliquerait de fait une sorte de communication à distance, entre les cellules infectées par différents segments, ou sous-groupes de segments, pour que l'intégrité du génome et sa fonctionnalité soient maintenues. Cette hypothèse impliquerait donc que (i) les segments ne sont pas toujours ensemble au sein des cellules, que (ii) les cellules infectées par des sous-groupes de segments ne sont pas des cellules « impasses » pour le virus mais bien des cellules au sein desquelles les segments présents sont exprimés et répliqués et (iii) que des produits viraux voyagent entre cellules assurant la disponibilité des facteurs requis pour l'expression et la réplication. Cette hypothèse peut paraître insensée d'autant que les mécanismes de

communication entre cellules infectées qu'elle implique sont totalement inconnus. Cependant, il s'agit d'une hypothèse testable expérimentalement. Les résultats attendus sont importants en termes de compréhension du fonctionnement des virus multipartites et questionnent certains concepts de base en virologie. Nous avons donc décidé de la tester.

Nous avons, dans un premier temps, étudié la distribution des segments au sein des cellules par hybridation *in situ* et, ce pour les paires de segments S et U4 d'une part et S et R d'autre part. Dans les deux cas, des cellules dans lesquelles seul un des deux segments était détectable ont été observées. Cette découverte nous a amenés à nous demander si les cellules qui paraissaient être mono-infectées -notamment les cellules dans lesquelles R qui code pour la Master Rep n'était pas détectable- étaient des cellules « impasses » c'est-à-dire dans lesquelles le système viral n'est pas complet et fonctionnel, et ne participe donc plus à l'infection. Afin de vérifier cela, nous tentons de démontrer que la fonction d'un gène peut être présente dans une cellule où le gène lui-même est absent. Nos premières approches visent à obtenir des indications qu'un segment du FBNSV peut se répliquer au sein d'une cellule où le segment R est absent. D'après les premiers résultats obtenus, il semblerait que les segments analysés (U4 et S) puissent être répliqués en l'absence du gène R, et que la protéine Master-Rep produite par le gène R puisse se déplacer, au-delà des cellules où le gène R est présent. Toutes les données présentées ci-dessous doivent encore être considérées comme préliminaires et restent donc à confirmer. Les différents résultats obtenus et les expériences complémentaires à mener sont présentés et discutés ci-après.

## 2. Matériel et méthodes

### **Matériel végétal**

Selon les expériences, des pois ou des fèves inoculées par pucerons ou agroinoculées avec les clones infectieux ont été utilisés comme matériel végétal. Les conditions de maintien et d'inoculation de ces plantes sont strictement identiques à celles décrites dans les chapitres précédents.

## Préparation de protoplastes

Des coupes transversales de 300 µm ont été réalisées à l'aide d'un microtome à vibrations (HM 650V Vibrating-Blade Microtome) au niveau de tiges apicales de fève. Après une pré-incubation des coupes de 15 min dans un milieu contenant 600mM de mannitol, 1 mM DL-dithiothreitol (DTT) et 25 mM MES/NaOH pH 5.7, les tissus ont été transférés dans un mélange enzymatique (400 mM de mannitol, 100 mM de KCl, 5 mM de MgCl<sub>2</sub>, 1 mM de DTT, 0.2% (p/v) polyvinylpyrrolidone-25, 0.5% bovine serum albumine, 0.5% (p/v) cellulase Onozuka RS (Yakult Honsha), 0.03% (p/v) pectolyase Y-23 (Seishin) et 25 mM de MES/NaOH pH 5.7 ; d'après Hafke et al., 2007. Des prélèvements ont été réalisés au bout de 7 et 9 heures et observées à l'aide d'un microscope inversé (Olympus CK2) afin de vérifier la présence et l'état des protoplastes obtenus. Les protoplastes ont ensuite été filtrés sur miracloth (Calbiogem) puis culottés par centrifugation pendant 3 min à 80 g. Après un lavage dans une solution contenant 400 mM de mannitol, 100 mM de KCl, 5 mM de MgCl<sub>2</sub>, 1 mM de DTT, 0.2% (p/v) polyvinylpyrrolidone-25, 0.5% bovine serum albumine et 25 mM de MES/NaOH pH 5.7, les protoplastes ont été fixés dans une solution à 2% de para-formaldéhyde (PFA) sous agitation pendant 20 minutes à température ambiante. Enfin, les protoplastes ont été lavés deux fois dans du tampon d'hybridation (20mM Tris-HCl pH 8.0, 0.9M NaCl, 0.01% (p/v) sodium dodecyl sulfate (SDS), 30% (v/v) formamide) avant d'être utilisés en hybridation *in situ*.

Dans un deuxième temps, des protoplastes de différents organes de pois –tige, racine et feuille- ont été obtenus suivant le protocole ci-après. Les feuilles entières de pois ont été infiltrées par seringue avec une solution enzymatique contenant 500 mM de mannitol, 25 mM de MES/NaOH pH 5.7, 5 mM de CaCl<sub>2</sub>, 2% (p/v) cellulase Onozuka RS (Yakult Honsha), 0.3% (p/v) macerozyme et 0.06% (p/v) de pectolyase Y-23 (Seishin). Pour les tiges et racines, des coupes transversales de 300 µm ont été mises à incuber directement dans ce milieu. Les protoplastes obtenus après six heures de digestion ont ensuite été filtrés sur miracloth puis culottés par centrifugation pendant 3 min à 80g. Après un lavage dans une solution contenant uniquement 500 mM de mannitol, les protoplastes ont été fixés dans une solution à 2% de PFA sous agitation pendant 20 minutes à température ambiante. Enfin, les protoplastes ont été lavés deux fois dans du tampon d'hybridation avant d'être utilisés en hybridation *in situ*.

## **Préparation des tissus cibles**

Les divers organes utilisés pour l'hybridation *in situ* (tige, racine, feuille et pétiole) ont été, dans un premier temps, fixés pendant 24h dans une solution de PFA à 4% sous agitation à 4°C. Après inhibition du PFA par une solution de PBS 1X –glycine à 0.1 M, les échantillons ont été placés dans de l'éthanol à 70% à 4°C jusqu'à utilisation.

Des coupes transversales de 300 µm ont été ensuite réalisées au microtome à vibrations sur les différents organes préalablement inclus dans de l'agarose NuSieve GTG à 8%. Les coupes obtenues ont ensuite été décolorées dans une solution de Carnoy (60% d'éthanol, 30% de chloroforme et 10% d'acide acétique glacial) pendant 2h à température ambiante sous agitation. Après 3 lavages dans du tampon d'hybridation puis un lavage dans une solution de PBS 1X, les coupes ont été utilisées en hybridation *in situ* (FISH) pour localiser les acides nucléiques, ou en immunofluorescence pour localiser les protéines, ou encore une combinaison des deux.

Comme spécifié dans les résultats, un traitement RNase (100µg/mL) pendant 1h à 37°C suivi de deux rinçages de 10 min dans du PBS 1X et de trois rinçages dans du tampon d'hybridation à température ambiante a été réalisé sur certaines coupes avant hybridation. Sauf indication contraire, les résultats présentés sont issus de coupes transversales non traitées par RNase.

## **Définition, préparation et choix de sondes**

L'ensemble des huit segments du génome du FBNSV ayant des parties plus ou moins conservées dans les régions non-codantes, toutes nos sondes pour l'hybridation *in situ* ont été ciblées dans les régions codantes qui elles ne présentent aucune homologie.

Deux approches différentes ont été mises en œuvre pour préparer des sondes spécifiques des segments génomiques ciblés. Le premier type de sonde a été construit sur la base d'oligonucléotides de 19-24 nt - correspondant aux amorces anti-sens utilisées en qPCR (Tableau S6, Article 1) spécifiques aux segments S, R et U4 - couplés au fluorochrome Alexa488 ou ATTO565. Un second lot de trois sondes spécifiques des régions codantes des segments R, S et U4 (de tailles respectives 810, 350 et 313 nucléotides) a été produit par amorçage aléatoire (*random priming*). Pour ce faire, les régions codantes de chaque segment ont tout d'abord été amplifiées par PCR (40 cycles de 95 °C pendant 45 s, 59 °C pendant 45 s et 72 °C pendant 45 s) grâce aux amorces répertoriées dans le tableau 1.



Séquence cible	Direction	Séquence 5'-3'
Segment R	Sens	ACATTAAATAATCCTCTCTCTCCTA
	Antisens	CCTATCATCACTAAACATGCC
Segment R1	Sens	ACATTAAATAATCCTCTCTCTCCTA
	Antisens	CTCTAAGCGTACTTGCAG
Segment R2	Sens	GAAGAAAGCTATTGCTACTTGG
	Antisens	CCTATCATCACTAAACATGCC
Segment S	Sens	AAATGGTGAGCAATTGGAA
	Antisens	GCCTATGATAGTAATCATATCTTGACA
Segment U4	Sens	AGCAGGTTATGGAATGTAGG
	Antisens	ATAGATTCCCACAATCGCT

<sup>a</sup> Ces amorces amplifient un fragment de 810nt situé dans la région codante du segment R

<sup>b</sup> Ces amorces amplifient un fragment de 382nt situé dans la première moitié de la région codante du segment R

<sup>c</sup> Ces amorces amplifient un fragment de 385nt situé dans la deuxième moitié de la région codante du segment R

<sup>d</sup> Ces amorces amplifient un fragment de 350nt situé dans la région codante du segment S

<sup>e</sup> Ces amorces amplifient un fragment de 313nt situé dans la région codante du segment U4

**Tableau 1 : Amorces utilisées en PCR pour amplifier les régions codantes de chaque segment**

Après purification sur gel avec le kit Wizard SV Gel and PCR Clean-Up System (Promega), ces produits de PCR ont servi de matrices pour la production des sondes par amorçage aléatoire et incorporation de dUTP marqués suivant les instructions du kit BioPrime DNA labeling (Invitrogen).

Brièvement, un mélange constitué de 400 ng du produit de PCR correspondant à la région ciblée d'un segment donné, et de 20 µL de la solution d'amorces aléatoires du kit, a été dénaturé

pendant 10 min à 92°C puis placé pendant 15 min dans de la glace. 10 µL de fluomix (composant du kit contenant un mélange des quatre bases azotées et de dUTP-Alexa488 ou dUTP-Alexa568), 1 µL de la solution du kit contenant le fragment de Klenow, et de l'eau ont été ensuite ajoutés au précédent mélange avant qu'il ne soit incubé une nuit à 37°C. Le lendemain, 5 µL de tampon stop du kit (contenant de l'EDTA) ont finalement été ajoutés au mélange réactionnel.

Les sondes spécifiques dénommées R1 et R2 ont été produites suivant le même processus, en utilisant comme matrices de l'amorçage aléatoire deux produits de PCR non chevauchant, tous deux issus de la séquence codante du gène R (composées respectivement de 382 et 385 nt). Ces sondes ciblent respectivement et spécifiquement la première (R1) et la seconde (R2) partie de la séquence codante du segment R ; elles servent ainsi de contrôle de co-localisation au sein des expériences.

Enfin, des sondes dirigées contre l'ensemble des huit segments (FBNSV-total) ont également été produites par incorporation de dUTP-Alexa488 par amorçage aléatoire, en utilisant un mélange équimolaire des produits de PCR issus des huit segments comme matrice.

### **Spécificité des sondes**

Afin de tester la spécificité des sondes, 1 µg de chaque plasmide contenant un des huit segments -préalablement dénaturé pendant 10 min à 95°C- a été déposé directement sur une membrane de nylon ainsi que 1 µL d'extraits de plante saine et de plante infectée obtenus selon le protocole d'(Edwards et al., 1991). L'ADN a ensuite été fixé aux ultraviolets par lampes ultraviolets UVllite (UVItec). Les membranes ainsi obtenues ont alors été pré-hybridées dans du tampon d'hybridation (20mM Tris-HCl pH 8.0, 0.9M NaCl, 0.01% (p/v) sodium dodecyl sulfate (SDS), 30% (v/v) formamide) contenant 1% (p/v) de BSA et 2 mg de sperme de hareng dénaturé. S'en est suivi une nuit d'incubation à 37°C dans 10 mL de tampon d'hybridation contenant 10 µL de sonde obtenue par amorçage aléatoire ou 10 pmol/mL de sonde oligonucléotidique. Les membranes ont alors subi deux rinçages dans 10 mL de tampon d'hybridation puis un dans 10 mL de TS1X. Après séchage des membranes dans du papier Whatman, la lecture des résultats s'est effectuée sur le phosphorimager Typhoon FLA 9000.

### **Hybridation *in situ***

Les échantillons fixés (coupes ou protoplastes) ont été mis à incuber durant une nuit à 37°C dans une solution de 300 µL de tampon d'hybridation (cf composition ci-dessus) contenant 10 µL de sonde obtenue par amorçage aléatoire ou 50 pmol d'oligonucléotides marqués. Trois rinçages d'une minute dans du tampon d'hybridation suivi d'un rinçage dans du PBS 1X ont été effectués avant observation.

### **Immunofluorescence**

Les mêmes coupes transversales de tissu végétal (fixées au PFA et décolorées au Carnoy), ont d'abord subi une pré-hybridation d'1h30 dans de la BSA 5% (p/v) dilué dans du PBS 1X avant d'être mises à incuber toute une nuit à 4°C avec des anticorps primaires, dilués au 1/500<sup>e</sup> dans du PBS 1X, dirigés contre la Master Rep du *Faba bean necrotic yellows virus* couplés à une histidine (Timchenko et al., 1999). Trois rinçages de 10 minutes dans une solution de PBS 1X et Tween 20 sous agitation à température ambiante ont suivi cette incubation. Les coupes ont ensuite été mises à incuber pendant 1h à 37°C dans une solution de PBS 1X contenant les anticorps secondaires Alexa Fluor 488 Goat Anti-Rabbit IgG (Invitrogen) dilués au 1/200<sup>e</sup> dans du PBS 1X. Afin de retirer l'excès d'anticorps secondaires, 3 rinçages de 10 minutes dans une solution de PBS1X ont été effectués sous agitation à température ambiante.

Dans les cas où nous souhaitons observer à la fois les molécules d'ADN et la protéine du segment R, les coupes ont finalement été rincées trois fois dans du tampon d'hybridation avant d'être mises à hybrider avec les sondes ADN correspondantes, comme décrit ci-dessus.

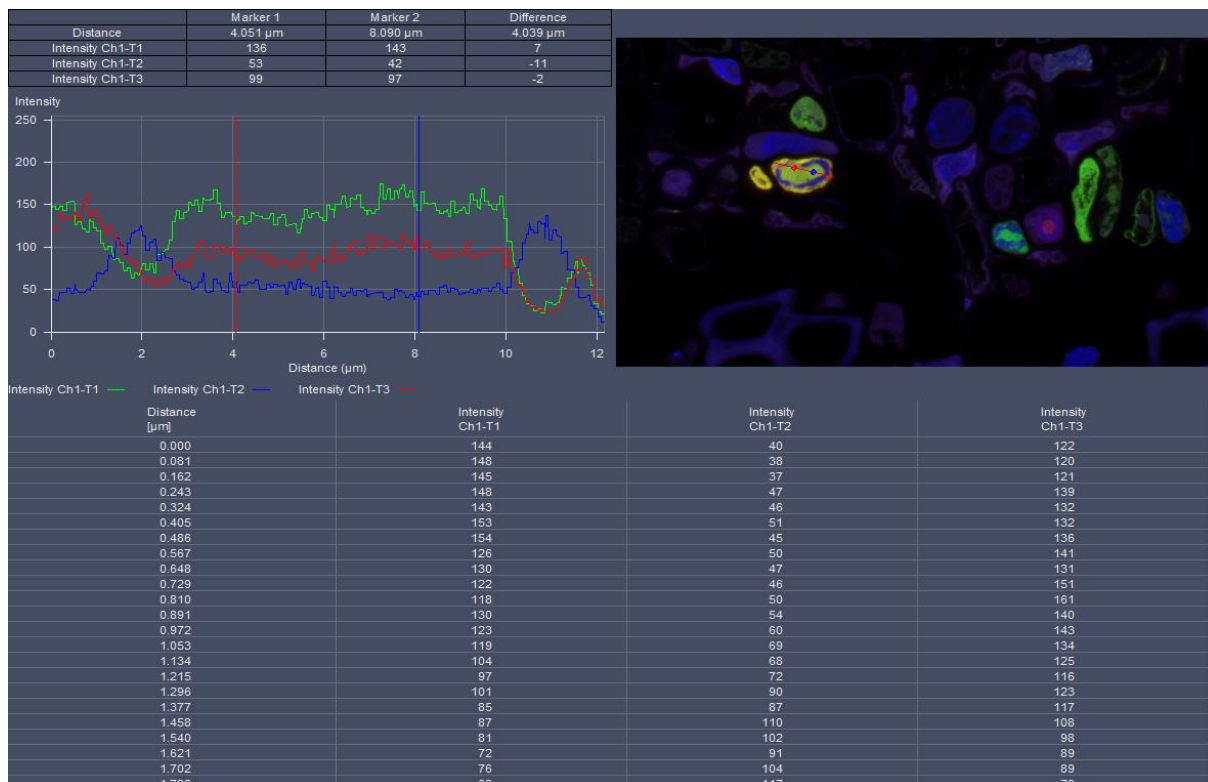
### **Détection et visualisation**

Les échantillons ont été montés entre lame et lamelle dans un milieu de montage pour fluorescence avec DAPI (Vectashield, Vector Laboratories Inc.) destiné à marquer le matériel génétique de manière non-séquence-spécifique, et donc principalement les noyaux des cellules. Avant montage, une incubation de quelques minutes dans une solution de perfluorodecaline

95% (Sigma) a été réalisée pour l'observation des coupes de feuilles afin d'augmenter la résolution. Les lames ont alors été observées au microscope confocal LSM 700 (ZEISS). Les images ont été prises avec le logiciel Zen 2009 (ZEISS). Les lasers utilisés excitent à 405 nm, 488 nm et 555 nm. Une configuration spécifique à chaque fluorochrome a été mise en place.

### **Mesures d'intensité de fluorescence**

Les mesures d'intensité de fluorescence ont été réalisées avec le logiciel Zen 2009 (ZEISS). Ainsi l'intensité de fluorescence des deux fluorochromes a été mesurée dans chaque cellule nucléée et étant infectée par au moins un segment. Une moyenne de l'intensité de fluorescence de chaque segment par cellule a été calculée à partir de différents points de mesures prises au sein de chaque cellule (Figure 12).



**Figure 12 : Mesure de l'intensité de fluorescence de chaque segment dans une cellule coinfectede par les segments S (en rouge) et U4 (en vert), 29 jours post-infection avec le logiciel ZEN 2009 (ZEISS)**

La flèche rouge traversant une cellule coinfectede par les deux segments (figure en haut à droite) indique la ligne de balayage et la direction de balayage suivie pour créer les profils d'intensité de fluorescence (en unités arbitraires, UA) des segments S (rouge) et U4 (vert) ainsi que du DAPI (bleu) visibles en haut à gauche. Le tableau en-dessous correspond aux premières mesures d'intensité de fluorescence pour les trois lasers utilisés à partir du premier point de mesure (distance 0  $\mu\text{m}$ ).

### Analyse statistique

Les analyses statistiques ont été réalisées avec le logiciel JMP10. La nature et les résultats des analyses sont indiqués dans le texte.

### Microscopie électronique à transmission

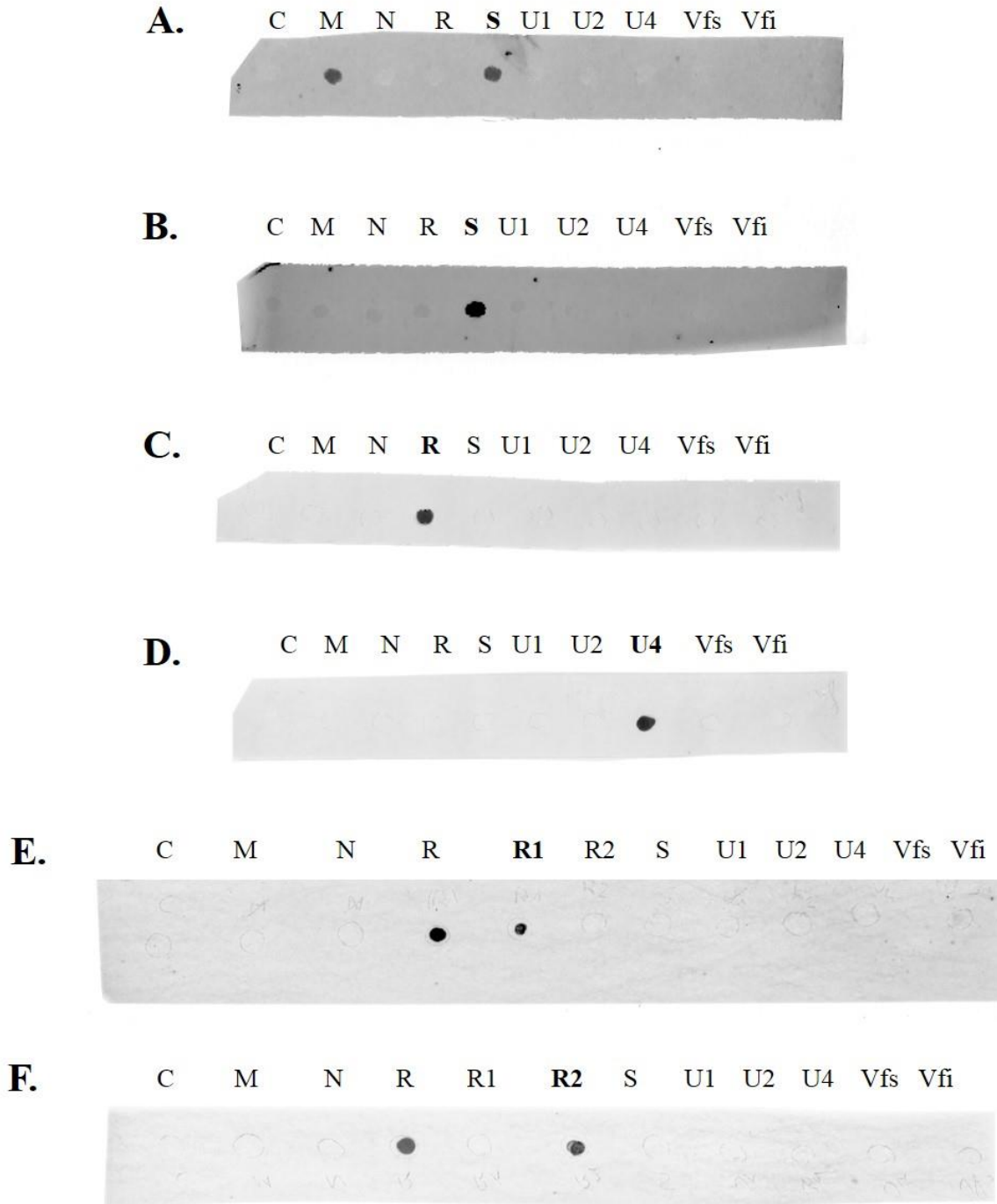
Des portions de pétioles de feuilles de fève saines et virosées, prélevées à 29 jours post-inoculation, ont été fixées dans du glutaraldéhyde à 4% dans du tampon cacodylate sodium 0.1M pH 7.2 pendant 4h à 4°C. Après trois rinçages dans le tampon de fixation, les échantillons ont été post-fixés dans du tétroxyde d'osmium à 1% pendant 1h à 4°C. Après trois rinçages dans de l'eau distillée, les échantillons ont ensuite été déshydratés progressivement dans des bains d'acétone à 50%, 70%, 90% puis 100%. Après imprégnation dans un gradient de résine

epoxy (TAAB812) dans l'acétone, les échantillons ont été inclus dans cette résine par polymérisation à 60°C pendant une nuit. Des coupes ultrafines de 60nm et semi-fines de 0,5 µm ont été réalisées à l'ultramicrotome. Les coupes semi-fines ont été observées en microscopie photonique après coloration au bleu de toluidine-AzurII. Les coupes ultrafines ont été observées au microscope électronique à transmission (Jeol JEM 100CX II) à 80 kV après avoir été contrastées à l'acétate d'uranyle à 5% et au citrate de plomb.

### 3. Résultats

#### **Choix des sondes**

Tout d'abord, nous avons testé la spécificité et l'intensité du signal fluorescent pour les deux types de sondes générées : les sondes obtenues par random-priming (sondes RP), et les sondes oligonucléotidiques (sondes ON). La spécificité des sondes a été contrôlée sur des plasmides contenant un des huit segments du FBNSV ainsi que sur des extraits de plantes saines et virosées. Quelle que soit la sonde utilisée, aucun signal n'a été détecté sur les extraits de plantes saines et virosées, les concentrations virales dans la plante infectée étant certainement trop faibles pour être détectables. La sonde ON conçue pour reconnaître le segment S s'est avérée aspécifique (hybridation sur les segments M et S) tandis que toutes les sondes RP étaient spécifiques (Figure 13). Par ailleurs, une comparaison *in situ* des sondes ON et RP a révélé que l'intensité de la sonde contre le segment R, RP-R était plus forte que celle de ON-R ; ce qui se traduisait à la fois par une observation plus facile des zones infectées et une puissance des lasers deux à trois fois moins élevée pour prendre des photographies des cellules infectées avec la sonde RP-R. La suite des expériences a donc été réalisée avec les sondes RP.



**Figure 13 : Spécificité des sondes S-oligonucléotidique (ON-S, A) , S-random priming (RP-S, B), R-random priming (RP-R, C), U4-random priming (RP-U4, D), R1-random priming (RP-R1, E) et R2-random priming (RP-R2, F)**

Chaque sonde a été testée sur les huit segments du FBNSV ainsi que sur des extraits de fève saine (Vfs) et infectée (Vfi). Les sondes R1 et R2 ont également été vérifiées sur des produits de PCR visant à amplifier leurs séquences respectives. Tandis que la sonde ON-S s'hybride sur les segments M et S, toutes les sonde RP sont spécifiques. La lecture des résultats s'est effectuée sur le phosphorimager Typhoon FLA 9000.

## Choix du matériel végétal

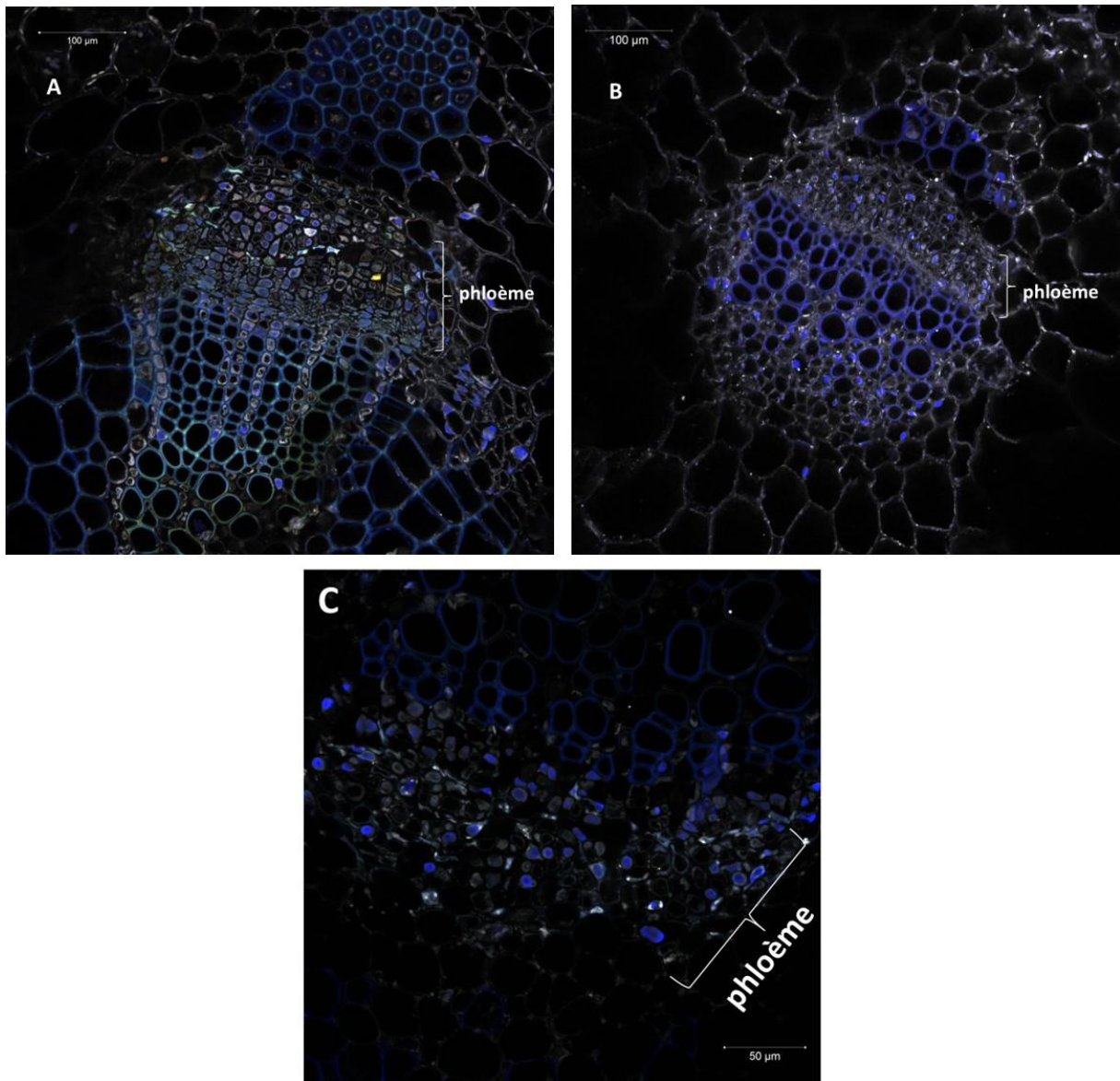
Deux approches différentes ont été menées dans le but de localiser ou co-localiser les segments génomiques du FBNSV dans des cellules individuelles. La première approche consistait à digérer les parois végétales afin d'obtenir des protoplastes et pouvoir ainsi marquer les protoplastes isolés. L'avantage principal de cette approche réside dans l'élimination de la fluorescence des parois et permet ainsi d'augmenter le ratio « signal/bruit ». Cette approche a été testée sur différents organes de plantes de pois (feuille, racine et tige) et sur tiges de fève. Très peu de protoplastes ont été obtenus à partir des racines de pois. En revanche, de nombreux protoplastes de toutes tailles ont été obtenus par digestion de tiges (de pois et de fève) et de feuilles de pois. Les essais de marquage fluorescent sur ces différents protoplastes ont révélé un pourcentage extrêmement faible de protoplastes marqués, tout au plus 1/1000, un nombre trop faible pour permettre de poursuivre nos études en routine avec cette technique.

Nous avons alors opté pour une seconde approche : le marquage *in situ*, directement sur coupes de tissus. Contrairement à l'approche sur protoplastes isolés, l'approche *in situ* permet d'avoir en plus une information spatiale quant à la répartition des cellules marquées. Cette approche a été conduite sur différents organes –tige, racine, feuille et pétiole- et sur différentes espèces végétales –pois, fève, luzerne, soja et lentille. Pour étudier la répartition du virus au sein de la plante, la sonde RP-totale (dirigée contre l'ensemble des huit segments) a été utilisée. Des marquages ont été observés exclusivement au niveau de la zone phloémienne, mais ceci dans tous les organes analysés, révélant ainsi que l'infection systémique touche tous les organes de la plante hôte. Nous avons préféré dans un premier temps étudier la co-localisation des segments au niveau du système foliaire. Effectivement, toutes nos données sur les fréquences relatives des segments –données que nous aimerions corrélérer avec les données obtenues par FISH (cf discussion)- ont été obtenues à partir de feuilles. Cependant, la forte autofluorescence présente dans les feuilles couplée à une plus grande difficulté à délimiter les cellules et à distinguer les cellules phloémiennes nous ont poussés à faire nos observations sur un autre organe. Le pétiole de feuille a été finalement choisi car c'est une pièce foliaire qui présente moins d'autofluorescence, de nombreuses cellules infectées et des zones phloémiennes facilement distinguables.

Toutes les observations ont été réalisées sur des coupes transversales de pétioles apicaux de fèves 20 ou 29 jours après inoculation (Figure 14). Aucun signal fluorescent n'apparaît sur tissus sains (Figure 14B) ou infectés en l'absence de sonde (Figure 14C). De plus, sur les coupes marquées avec les sondes RP-R et RP-S, aucun marquage n'a été observé hors de la zone



phloémienne sur tissus infectés (Figure 14A). Au-delà de leur valeur de contrôle, ces résultats confirment les travaux de (Katul et al., 1993) démontrant que les nanovirus sont restreints au phloème et se répliquent donc uniquement dans les cellules compagnes, et peut être dans des cellules du parenchyme phloémien.

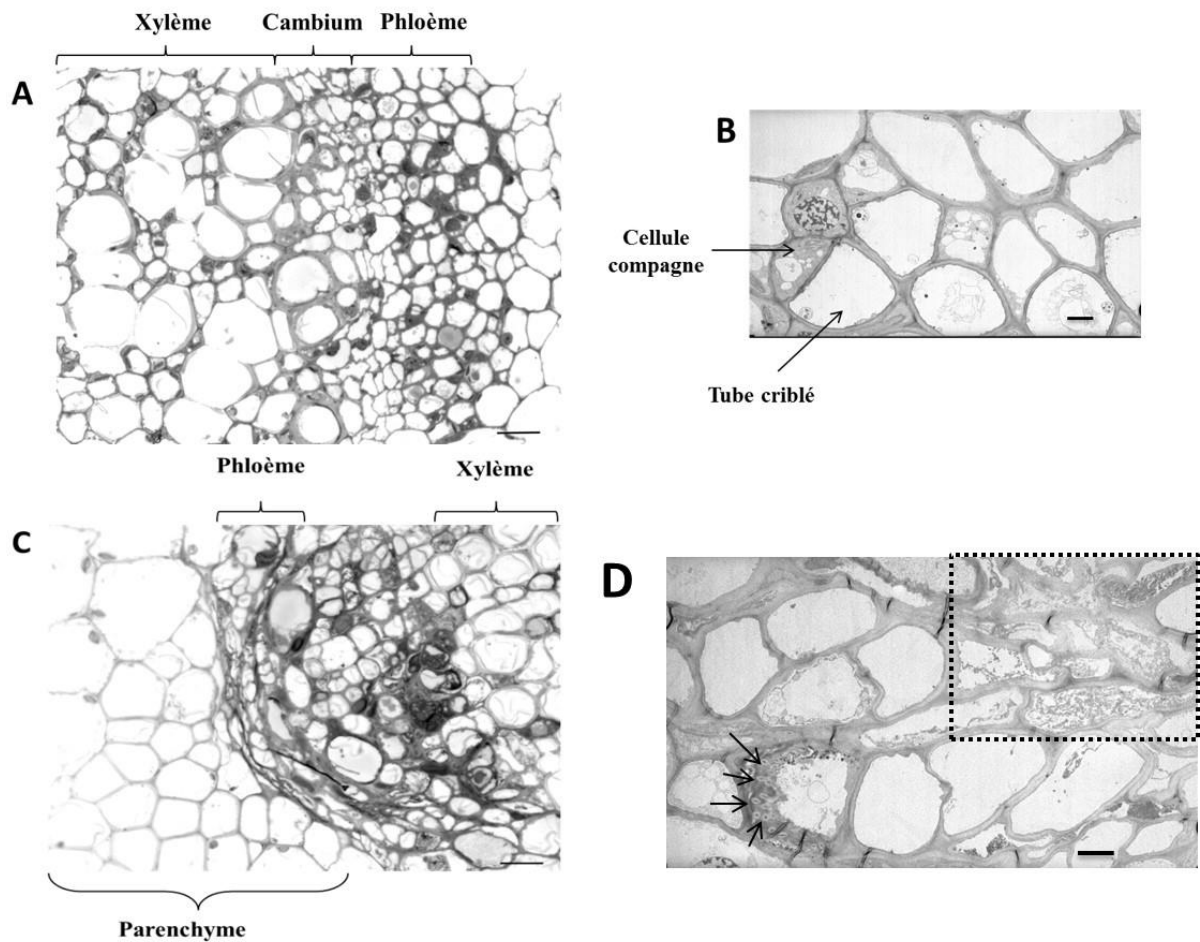


**Figure 14: Hybridation fluorescente in situ (FISH) sur faisceaux conducteurs de pétioles de *Vicia faba***

Des coupes de pétioles issus de plantes infectées depuis 29 jours avec (A) ou sans sonde (C), ou saines (B), sont observées après FISH dirigé contre les segments R (vert) et S (rouge). Les noyaux des cellules ont été marqués au DAPI (en bleu) et l'autofluorescence des tissus apparaît ici en blanc. Les cellules contenant majoritairement le segment R, le segment S ou les deux segments apparaissent respectivement en vert, rouge ou jaune. Aucun marquage n'a été détecté en dehors de la zone phloémienne, dans les tissus sains ou dans les tissus infectés en absence de sonde.

Sans coloration spécifique, nous avons eu beaucoup de difficulté à distinguer avec certitude les tubes criblés, les cellules compagnes et les cellules du parenchyme phloémien dans les coupes observées au microscope confocal. Nous avons tenté d'explorer l'anatomie de nos pétioles de manière plus résolutive par microscopie électronique à transmission. La Figure 15 montre que l'infection par le FBNSV désorganise considérablement les tissus du phloème et, que cette difficulté sera donc certainement très difficile à surmonter sans des colorations permettant l'identification de types cellulaires spécifiques. Nous avons donc coloré de façon systématique nos coupes à l'aide de DAPI pour marquer les noyaux. Ce marquage permet de différencier entre les tubes criblés (qui ne contiennent pas de noyau) et les autres types cellulaires.

Le point qui nous intéresse en premier lieu dans ce chapitre est de savoir si le virus se réplique avec les huit segments dans la même cellule ou pas. Puisque la réplication virale s'opère dans le noyau des cellules (Timchenko and Bernardi, 2007) donc hors des tubes criblés qui en sont dépourvus, nous analyserons en premier lieu les cellules où le noyau est clairement identifiable (Pour plus de précisions sur l'anatomie du phloème, voir l'annexe III). La situation dans les tubes criblés n'est donc pas une priorité ici, et ce point est abordé plus en détails dans la discussion.



**Figure 15: Effets de l'infection du FBNSV sur les tissus phloémiens de pétiole de fève 29 jours post-inoculation**

Photographies de microscopie électronique à transmission prises au niveau du deuxième pétiole de feuille en partant de l'apex de la plante. S'il est relativement aisé de distinguer les différents types cellulaires sur les tissus sains (A) et (B) qui présentent une organisation bien structurée et des contours bien délimités, il est en revanche plus difficile de pouvoir distinguer les types cellulaires au niveau de la zone phloémienne des pétioles infectés par le FBNSV 29 jours post-inoculation (C) et (D). La présence de cribles (indiqués par des flèches sur la photographie D) permet parfois de distinguer un tube criblé. Dans la zone délimitée par des pointillés (D), il est en revanche difficile de définir les types cellulaires. Pour les échelles, (A–C): 10  $\mu\text{m}$ ; (B–D): 3  $\mu\text{m}$ .

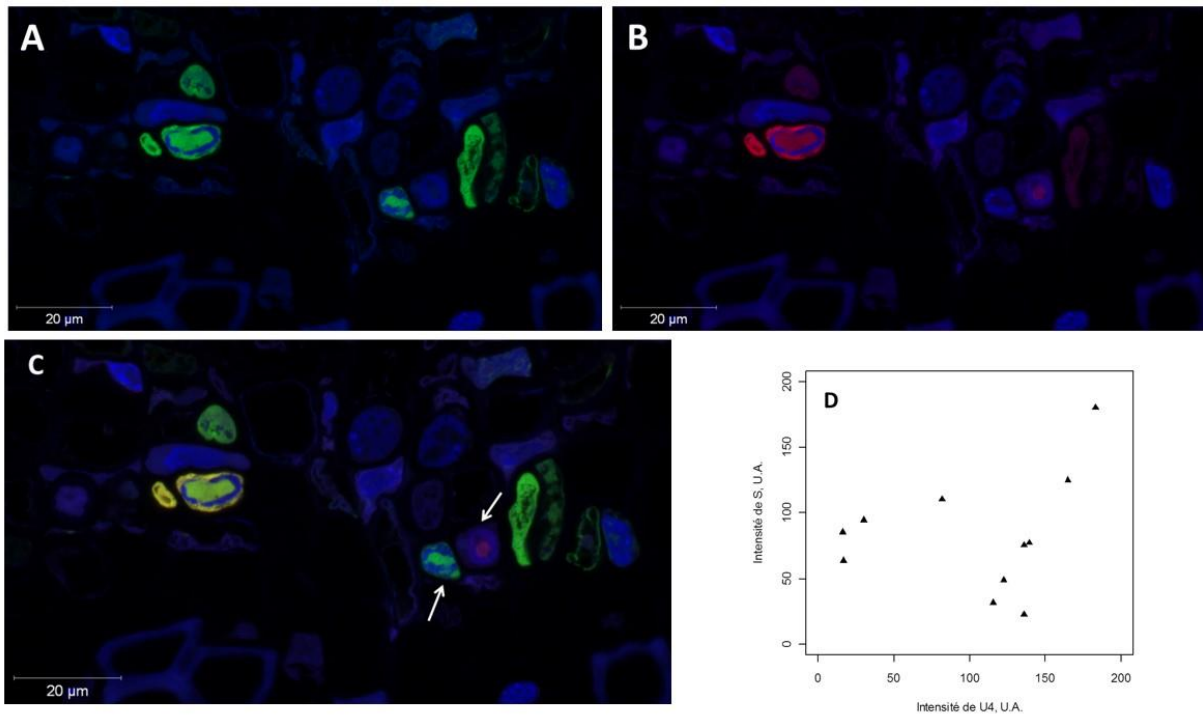
## Localisation des segments

Trois lasers étant à notre disposition, la co-localisation possible des segments n'a pu être testée qu'en prenant les segments deux à deux, le troisième laser étant utilisé pour visualiser les noyaux grâce à la coloration au DAPI. Le nombre de paires possibles à partir de huit segments étant trop élevé pour être toutes analysées (28 paires possibles), nous avons fait des choix qui sont justifiés dans les résultats et la discussion.

- Localisation des segments S et U4

Dans un premier temps, nous avons choisi des sondes s'hybridant au segment le plus abondant au sein de la fève, le segment U4, et au segment le moins fréquent, le segment S. Au-delà des cellules non marquées, nous avons observé des cellules contenant les deux segments en proportion variable, des cellules dans lesquelles seul le segment U4 semble présent et, moins fréquemment, des cellules où seul le segment S est présent (Figure 16 A-C). Il semble donc que ces segments ne soient pas toujours nécessairement retrouvés ensemble au sein de cellules individuelles.

Chaque technique ayant ses limites, il pourrait être opposé à cette conclusion, que le segment qui n'est pas observé dans une cellule infectée est simplement en-dessous de la limite de détection. Afin de pouvoir argumenter en faveur d'une dé-corrélation de la présence des différents segments au sein de cellules individuelles, qui ne soit pas liée aux limites de détection du microscope confocal, nous avons par ailleurs quantifié les différences d'intensité du signal fluorescent rouge et vert dans chaque cellule où l'un des deux pouvait être détecté de manière non équivoque. Un simple graphe montrant l'intensité de vert en fonction de l'intensité de rouge, en d'autres termes, montrant la quantité du segment U4 en fonction de celle du segment S a révélé une absence totale de corrélation (Figure 16 D, test de corrélation de Pearson,  $\rho=0.25$ ,  $pvalue=0.45$ ).

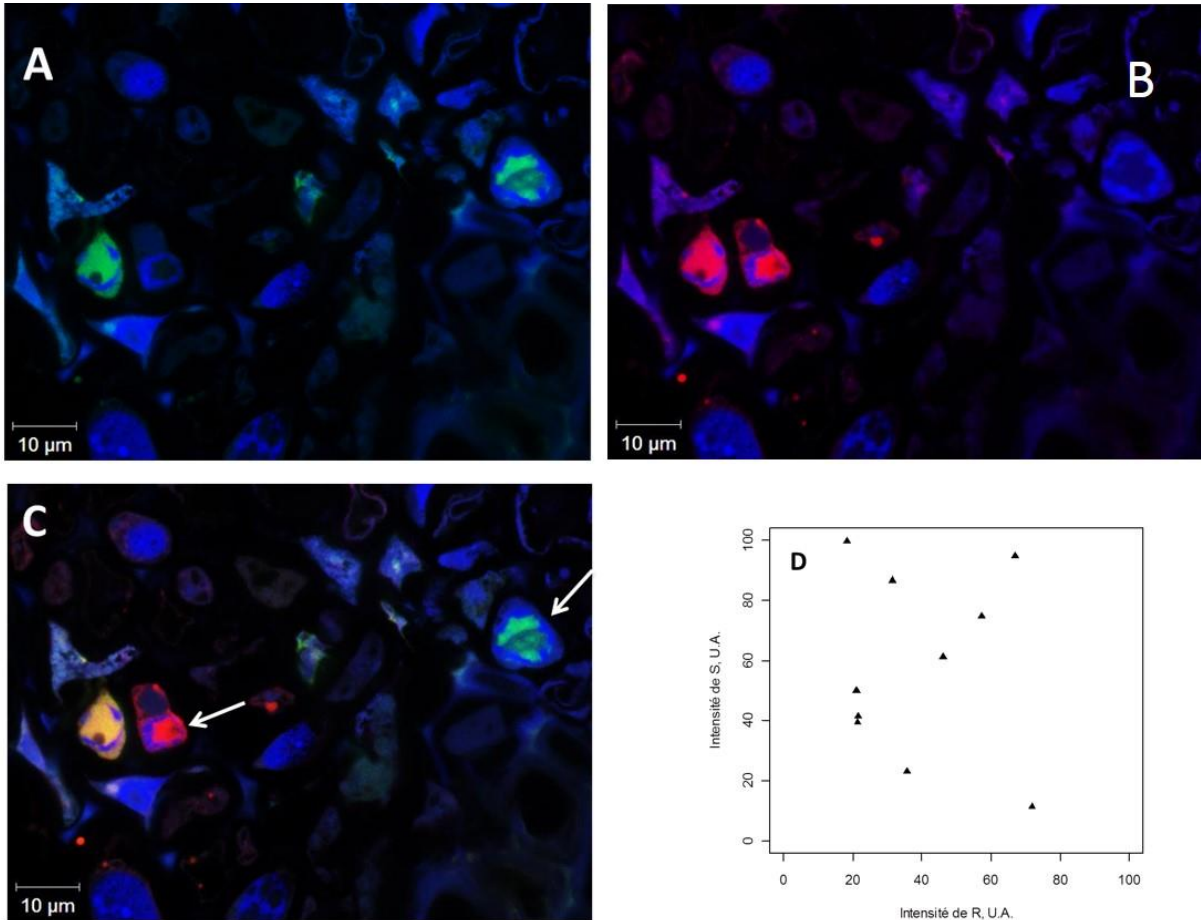


**Figure 16 : Localisation des segments S et U4 du FBNSV dans les cellules de pétioles de fève.**

Les coupes de pétioles de feuilles de fève infectés à 29dpi ont été hybridées avec un mélange des sondes RP-S et RP-U4. (A) Localisation du segment U4 par FISH (Alexa 488, en vert) (B) Localisation du segment S par FISH (Alexa 568, en rouge). (C) Fusion des deux photographies précédentes. Les noyaux des cellules ont été marqués au DAPI (en bleu). Deux cellules dans lesquelles seul S ou U4 est détectable sont indiquées par des flèches blanches en C. (D) Relation entre l'intensité de fluorescence des segments S et U4 dans onze cellules. Chaque triangle représente l'intensité moyenne de fluorescence mesurée au sein d'une de ces 11 cellules. Les intensités sont indiquées en unités arbitraires de fluorescence (U.A.).

- Localisation des segments R et S

Souhaitant savoir si les résultats obtenus avec les segments U4 et S étaient généralisables aux autres segments du FBNSV, la paire des deux segments les moins fréquents au sein de la fève et indispensables à son infection ont été choisis : la paire R et S. Comme précédemment, des pétioles de fève ont été hybridés avec des sondes complémentaires de ces deux segments et examinés. De nouveau, des cellules ne présentant aucun des deux segments, présentant les deux segments, ou un seul des deux, ont été observées (Figure 17). De manière similaire à la paire S/U4, l'intensité de fluorescence quantifiée pour R et S est décorrélée (Figure 17D, test de corrélation de Pearson,  $\rho = 0.31$ ,  $p\text{-value} = 0.34$ ), suggérant que la présence de l'un ne dépend pas de celle de l'autre.

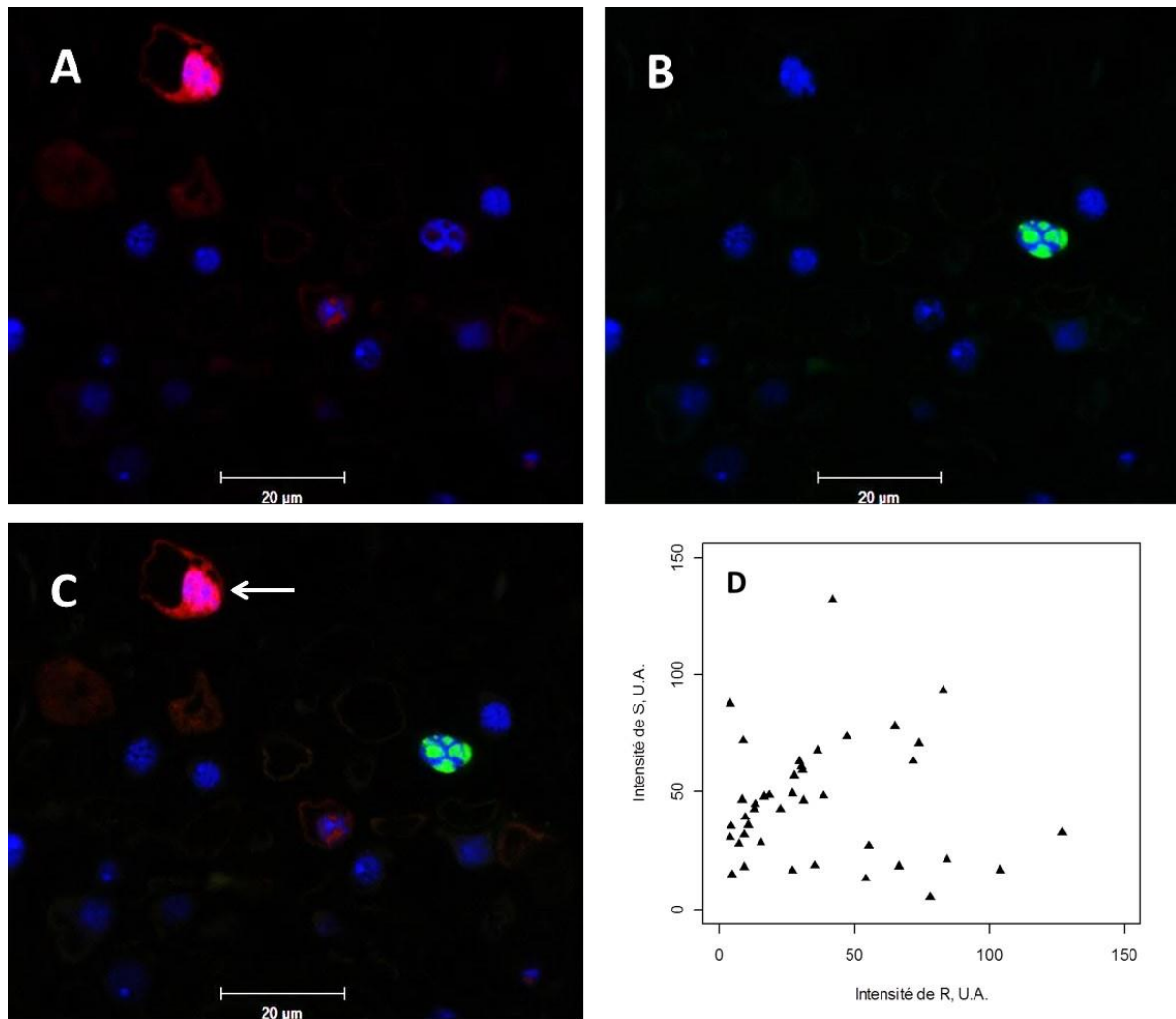


**Figure 17: Localisation des segments R et S du FBNSV dans les cellules de pétioles de fève**

Les coupes de pétioles de feuilles de fève infectés à 29dpi ont été hybridées avec un mélange des sondes RP-R et RP-S. Localisation du segment R par FISH (Alexa 488, en vert). (B) Localisation du segment S par FISH (Alexa 568, en rouge). (C) Fusion des deux photographies précédentes. Les noyaux des cellules ont été marqués au DAPI (en bleu). Deux cellules dans lesquelles seul R ou S est détectable sont indiquées par des flèches blanches en C. (D) Relation entre l'intensité de fluorescence des segments R et S dans onze cellules. Chaque triangle représente l'intensité moyenne de fluorescence mesurée au sein d'une de ces 11 cellules. Les intensités sont indiquées en unités arbitraires de fluorescence (U.A.)



Les sondes utilisées pouvant potentiellement s'hybrider à la fois à l'ADN et à l'ARN viral, nous avons ensuite voulu vérifier que les résultats obtenus n'étaient pas en partie biaisés par la présence d'ARN. Un traitement à la RNase a été effectué sur des coupes de pétiole avant hybridation avec les mêmes sondes RP-R et RP-S. Les différences d'intensité observées entre segments dans différentes cellules ont de nouveau été observées (Figure 18). Comme attendu, une décorrélation très nette de la quantité de R par rapport à celle de S a encore une fois pu être démontrée (Figure 18D, test de corrélation de Pearson,  $\rho = -0.0062$ ,  $p\text{value} = 0.97$ ).

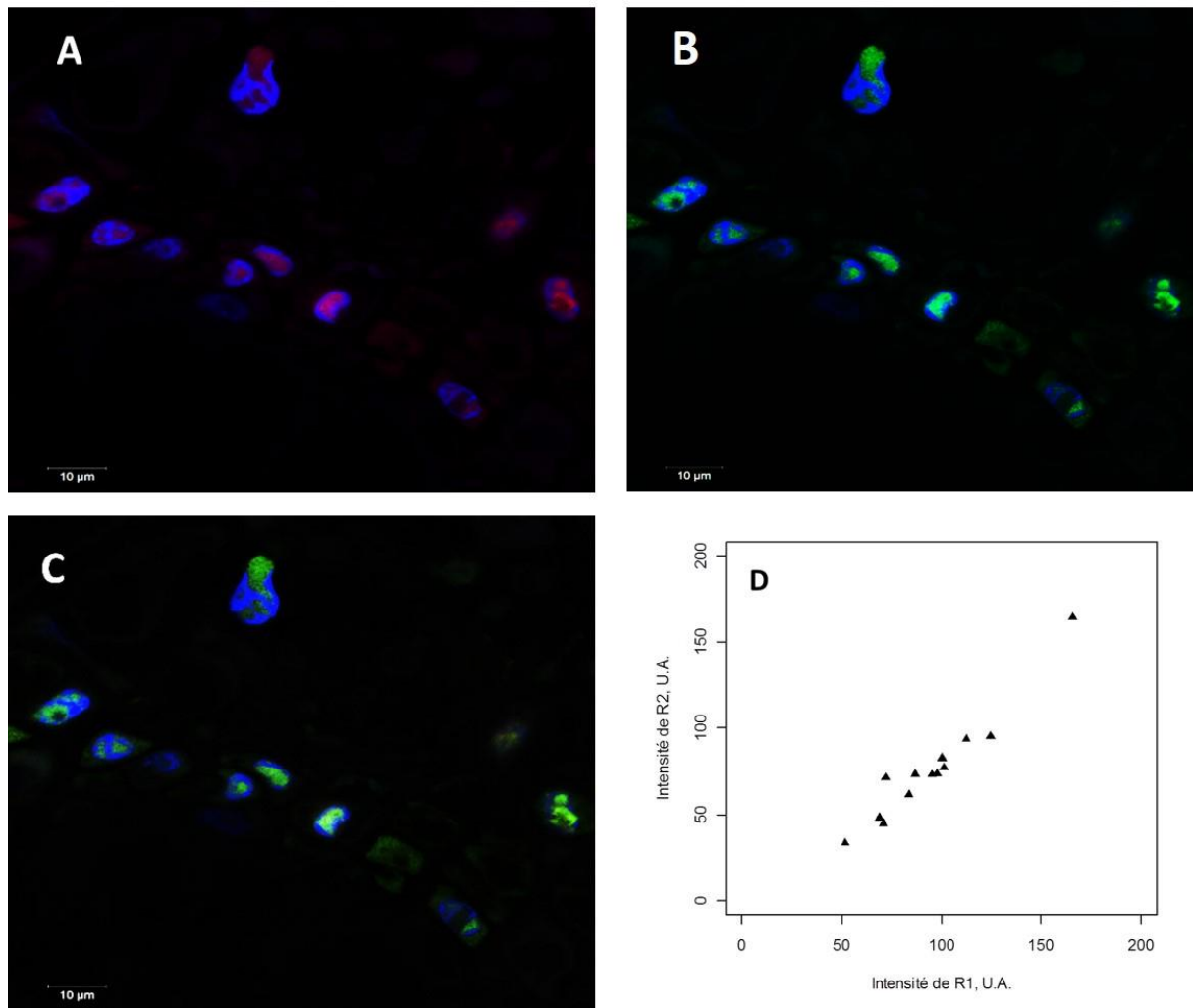


**Figure 18: Localisation des segments R et S dans les cellules de pétioles de fève après traitement RNase.**

Les coupes de pétioles de feuilles de fève infectés à 20dpi ont été hybridées avec un mélange des sondes RP-R et RP-S, après traitement à la RNase comme indiqué dans le matériel et méthodes. Localisation du segment R par FISH (Alexa 488, en vert). (B) Localisation du segment S par FISH (Alexa 568, en rouge). (C) Fusion des deux photographies précédentes. Les noyaux des cellules ont été marqués au DAPI (en bleu). Une cellule où seul S est détectable est indiquée par une flèche blanche en C. (D) Relation entre l'intensité de fluorescence des segments R et S dans quarante cellules. Chaque triangle représente l'intensité moyenne de fluorescence mesurée au sein d'une de ces 40 cellules. Les intensités sont indiquées en unités arbitraires de fluorescence (U.A.)

- Contrôle des résultats obtenus

Afin de consolider nos résultats, un contrôle important a été mis en œuvre. Ce contrôle était destiné à démontrer que les approches développées ici sont capables de détecter une colocalisation des séquences virales ciblées au sein des cellules, lorsque nous savons *a priori* que cette colocalisation est obligatoire. Pour ce faire, deux sondes ciblant chacune une des deux moitiés (non chevauchantes) de la région codante du segment R ont été produites et hybridées sur des pétioles de fèves infectées. Il a ainsi pu être vérifié que dans toutes les cellules positivement marquées pour une des moitiés du segment R, la moitié complémentaire était aussi détectable (Figure 19). De plus, une corrélation positive très nette a pu être statistiquement démontrée entre la quantité de chacune des deux moitiés (Figure 19D, test de corrélation de Pearson,  $\rho = 0.96$ ,  $p\text{value} < 0.0001$ ).

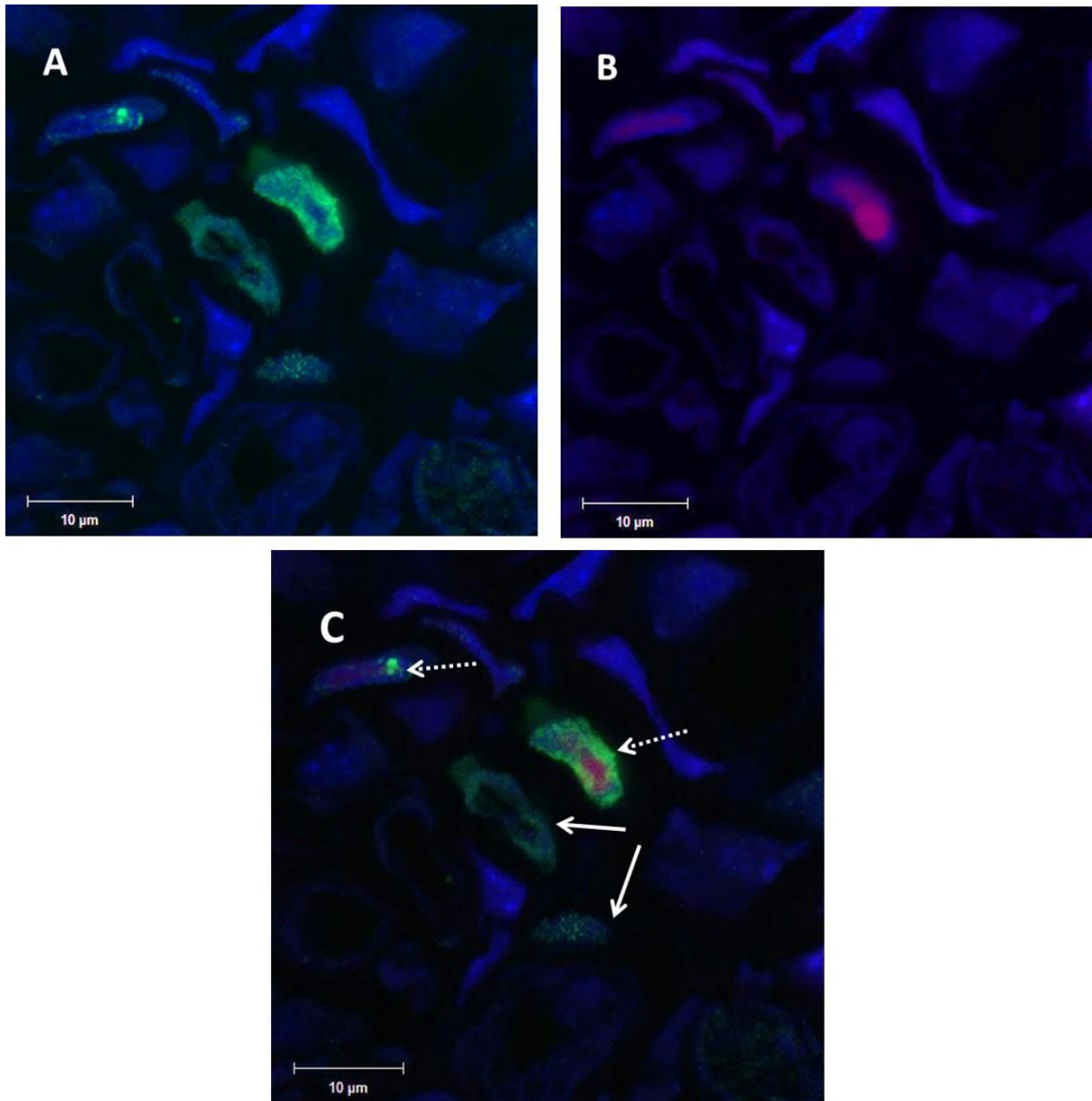


**Figure 19 : Colocalisation des deux moitiés du segment R du FBNSV au sein du phloème des pétioles de fèves, 29 jours post-inoculation**

Les coupes de pétioles de feuilles de fève infectés ont été hybridées avec un mélange des sondes RP-R1 et RP-R2. (A) Localisation de la moitié R2 par FISH (Alexa 488, en vert). (B) Localisation de la moitié R1 par FISH (Alexa 568, en rouge). (C) Fusion des deux photographies précédentes. Les noyaux des cellules ont été marqués au DAPI (en bleu). (D) Relation entre l'intensité de fluorescence de deux moitiés de segment R1 et R2 dans quatorze cellules. Chaque triangle représente l'intensité moyenne de fluorescence mesurée au sein d'une de ces 14 cellules. Les intensités sont indiquées en unités arbitraires de fluorescence (U.A.).

### **Co-localisation d'un segment génomique et de la protéine pour laquelle il code**

Nous avons enfin cherché à vérifier s'il existait une dissociation entre la localisation des segments et celle des protéines pour lesquelles ils codent au sein des cellules colonisées par le FBNSV. Ceci suggérerait que le produit d'expression d'un gène viral peut circuler entre les cellules et exporter sa fonction, et indiquerait donc une possibilité de complémentation fonctionnelle pour le système viral, au-delà des cellules où ce gène est présent. Pour cela, nous avons utilisé un anticorps dirigé contre la Master Rep du *Faba bean necrotic yellows virus*, protéine présentant 89% d'homologie avec la Master Rep de notre clone de FBNSV (anticorps fourni par T. Timchenko, IBP, CNRS, Paris). Après avoir immuno-marqué les coupes de pétioles de fève avec cet anticorps, les mêmes coupes ont été hybridées avec la sonde spécifique du segment R, la sonde RP-R. Au-delà des cas attendus où le gène R et sa protéine co-existent au niveau intracellulaire, nous avons aussi constaté que dans plusieurs cellules la protéine Master-Rep peut être détectable sans ambiguïté en l'absence de détection du segment R (Figure 20).



**Figure 20 : Localisation du segment R et de son produit d'expression dans des pétioles de fève infectée**

Des coupes de pétioles de fève infectée par le FBNSV à 29 dpi ont été analysées simultanément par FISH dirigé contre le segment R et par immunofluorescence dirigée contre la protéine correspondante Master-Rep. (A) Localisation de la protéine M-Rep par immunofluorescence (Alexa 488, en vert) (B) Localisation du segment R par FISH (Alexa 568, en rouge) (C) Fusion des deux photographies précédentes. Les noyaux des cellules ont été marqués au DAPI (en bleu). Les cellules où le segment R et son produit d'expression sont tous deux détectés sont indiquées avec une flèche en pointillés. Celles où la protéine Master-Rep est détectée en l'absence de détection du segment R sont indiquées par des flèches pleines.

## 4. Discussion

Cette étude constitue la première investigation visant à vérifier si tous les segments génomiques d'un virus multipartite doivent effectivement entrer dans la même cellule pour qu'il y ait réplication et expression, et donc pour que le système viral soit fonctionnel. S'il s'agissait initialement de vérifier expérimentalement l'existence d'un coût majeur à chaque infection de cellule, largement étudié et discuté (mais jamais remis en cause) dans la littérature théorique, la portée de nos résultats nous semble aller au-delà. En effet, outre le fait qu'ils réfutent au moins partiellement l'existence de ce coût, ils représenteraient la première indication que le cycle de réplication d'un virus peut ne pas être « cellule autonome ». Autrement dit, nous démontrerions que pour certaines espèces virales, l'unité spatiale du cycle de réplication n'est pas une cellule individuelle, mais un ensemble de cellules interconnectées où l'information génétique virale communique et se réplique de façon fragmentaire dans différentes cellules.

Nos données actuelles doivent être considérées comme préliminaires et seront consolidées. Néanmoins, ces données semblent démontrer que le maintien de l'intégrité du génome du FBNSV et la fonctionnalité de ce système viral ne repose pas sur la coinfection systématique des cellules par tous les segments.

Qu'il s'agisse des couples de segments S et U4 ou R et S, des cellules dans lesquelles seul un des deux segments est présent ont souvent été observées. La fréquence des cellules où les segments ciblés co-existent, comparée à celles où les paires de segments sont séparées, serait certainement une information importante et, à ce stade, il est très probable que trop peu de cellules aient été observées. Des approches à plus « haut débit » ont été développées par cytométrie de flux sur protoplastes, afin de trier les cellules infectées par un variant vert, un variant rouge ou les deux variants d'un virus monopartite (Lafforgue et al., 2012; Tromas et al., 2014). En imaginant une adaptation de cette approche à notre système FBNSV nous pourrions peut être compléter les résultats obtenus en FISH par l'analyse rapide du statut de l'infection d'un grand nombre de cellules (un segment, l'autre, ou les deux de la paire). Ceci nous permettrait notamment de tester si les segments les plus fréquents « infectent » plus de cellules que les segments rares, et ainsi de pouvoir confirmer expérimentalement l'hypothèse de l'entrée aléatoire des segments dans les cellules. Cependant, le faible nombre de protoplastes infectés obtenus lors de nos essais, sur différents tissus et différentes plantes hôtes, ne nous pousse pas à investir de gros efforts dans ce sens. Il est possible que le faible pourcentage de cellules

susceptibles, lié à la restriction du FBNSV au phloème (les cellules compagnes représentent environ 1% des cellules totales, d'après S. Dinant, communication personnelle), et une plus forte dégradation des cellules infectées, soient insurmontables pour conduire une étude sur protoplastes par FACS. Nous pensons somme toute qu'il sera plus facile de compter un nombre plus important de cellules par la technique de FISH sur coupe de tissus développée ici.

Lorsque nous déclarons qu'un segment est absent dans une cellule donnée, une critique majeure repose sur la sensibilité de la technique qui pourrait ne pas détecter les trop faibles quantités. Ainsi, afin d'apporter un argument supplémentaire qui ne dépende pas de la sensibilité de la technique du FISH, nous avons décidé de développer un test statistique de l'interdépendance des segments d'une paire donnée. L'idée étant que si la présence d'un segment au sein d'une cellule dépend de celle de l'autre, il est fort probable qu'il existe une corrélation positive entre leurs quantités respectives. Nos résultats ne montrent aucune corrélation entre les quantités des segments S et U4 d'une part, et celles des segments R et S d'autre part dans différentes cellules, apportant une indication supplémentaire de leur indépendance en termes de localisation et d'accumulation cellulaire. Cette décorrélation ne semble pas être affectée par un marquage des transcrits ARN puisqu'après traitement par RNase, des résultats similaires ont été observés.

Une autre critique possible serait de considérer que nous n'avons observé la situation qu'à un seul point de temps et qu'en fin d'infection tous les segments seront bel et bien ensemble dans les cellules. Il est alors important de souligner que les symptômes du FBNSV sur fève -dont l'arrêt de croissance de la plante- apparaissent entre 10 et 20 jours post-inoculation et que toutes nos observations ont été effectuées à 20 ou 29 dpi (la plupart à 29 dpi), peu de temps avant que la plante infectée n'entre en sénescence. A un temps aussi tardif de l'infection, les segments ségrégent toujours spatialement, nous poussant à dire qu'il est assez peu probable que les cellules soient *in fine* infectées par tous les segments. Une étude à des stades précoces de l'infection pour confirmer cette décorrélation à des temps précoces de l'infection est actuellement en cours.

A l'encontre de nos conclusions, l'accumulation indépendante des segments au sein des cellules peut amener à se demander si les cellules partiellement infectées ne constituent pas tout simplement des impasses pour les segments qui s'y trouvent « isolés », donc pour le système viral, et serait ainsi une matérialisation du fameux coût, conséquence directe de la multien capsidation. Une forte intensité du marquage du segment S (segment le plus rare chez la fève) dans certaines cellules où le segment R n'est pas détecté semble cependant contrer cette



critique. En effet, la présence de très nombreuses copies de S au sein d'une cellule pourrait être le résultat soit d'une entrée massive de ce segment exclusivement dans cette cellule, soit de sa réplication. La première hypothèse semble assez peu vraisemblable étant donné que ce segment est parfois indétectable dans les cellules phloémiennes voisines, pourtant permissives au virus. La deuxième hypothèse impliquerait que le segment S peut être répliqué dans des cellules sans que le segment R ne soit présent. Cette hypothèse suppose donc que la fonction d'un segment (i.e. R) peut être présente sans que le segment lui-même ne soit là, ce qui impliquerait que les produits d'expression –l'ARN ou les protéines- d'un segment donné peuvent transiter entre cellules. Si tel était le cas, on devrait pouvoir trouver des cellules dans lesquelles la protéine est présente sans que le segment correspondant n'y soit. Nos premières observations visant à regarder la colocalisation du segment R et de sa protéine Master Rep semblent aller dans ce sens, mais ces résultats devront être largement vérifiés pour les raisons suivantes:

i) Nous n'avons pas encore pu valider la spécificité de l'anticorps utilisé ici, par Western Blot ou par toute autre technique. Si, au cours des différentes observations réalisées, nous n'avons pas constaté de fixation de l'anticorps à d'autres endroits que dans les cellules du phloème de plantes infectées, il n'en reste pas moins que cet anticorps est dirigé contre la Master-Rep du FBNYV qui ne partage que 89% d'identité avec la Master Rep du FBNSV. Un autre anticorps dirigé contre la partie C terminale de la Master Rep du FBNSV récemment obtenu sera bientôt testé.

ii) L'absence de détection du segment R dans certaines cellules où la protéine Master-Rep est présente peut être de nouveau liée au manque de sensibilité de la méthode FISH utilisée. Une étude visant à montrer la décorrélation d'intensité de fluorescence pour le couple segment R/proteine Master Rep ne serait pas satisfaisante étant donné que plusieurs raisons pourraient biaiser nos conclusions : seuil de réplication de R, timing d'expression et stabilité de la protéine Master-Rep.

iii) la protéine Master-Rep pourrait être co-encapsidée avec chaque segment comme c'est le cas chez certains virus à ARN simple brin négatif, expliquant ainsi que la Master Rep soit trouvée dans les cellules où le segment R n'y est pas forcément. Un moyen de tester cette hypothèse serait de vérifier que cette protéine se trouve également dans les purifications virales. Jusqu'à présent, nous n'avons pas réussi à obtenir de purifications assez « propres » pour pouvoir mener cette expérience.

Enfin, il n'est pas exclu que certaines protéines virales puissent trafiquer sans être toutefois fonctionnelles suite à leur dénaturation au cours de leur transport. S'il s'avérait qu'effectivement la protéine Master Rep pouvait trafiquer, la présence de cette protéine n'impliquerait donc pas forcément que celle-ci soit fonctionnelle et donc que les segments soient forcément répliqués. La présence de formes répliquatives à double brin ADN dans des cellules où le segment R est absent permettrait d'apporter une preuve des plus convaincantes de répliquions quand le segment R est absent. Cette expérience sera menée ultérieurement.

## 5. Perspectives à plus long terme

Plusieurs expériences restent donc à mener afin de vérifier que la coinfection des cellules par tous les segments génomiques n'est pas nécessaire au fonctionnement des virus multipartites et ne représente donc pas un coût majeur de ce système. Si, comme nos données le laissent présager, ce coût n'existe pas (en tout cas pas tel que considéré dans la littérature), plusieurs perspectives s'ouvriraient suite à ce travail.

Premièrement, il serait intéressant de connaître la nature des produits d'expression circulants : ARN ou protéines. Une étude FISH/immunofluorescence en microscopie confocale suivie d'une étude d'immunogold couplée à de la microscopie électronique – outil permettant de repérer avec plus de certitude les protéines et la nature des cellules observées – serait plus appropriée pour répondre à cette question. Cela pourrait être conduit pour l'ensemble des huit segments et de leurs produits d'expressions, afin de regarder si tous se comportent de manière identique ou si chaque segment ou produit de segment possède une stratégie propre. On peut effectivement imaginer que les produits d'expression de certains segments n'ont pas besoin d'être redistribués au sein des différentes cellules compagnes comme par exemple pour le segment N qui ne semble jouer un rôle qu'au moment de la transmission et ne pourrait être « utile » que dans les tubes criblés où les pucerons se nourrissent et acquièrent ce virus.

Deuxièmement, notre étude étant limitée aux pétioles, il serait intéressant de regarder si ces résultats sont généralisables à tous les tissus infectés par le virus. En effet, on peut imaginer que la MOI (Multiplicité d'infection cellulaire) peut être plus ou moins élevée selon le type cellulaire –comme cela a été montré par Gutierrez et al. pour un potyvirus, le *Turnip mosaic virus* (article en préparation) - ou encore selon l'organe considéré. Si la MOI était très élevée au sein de certains organes, tous les segments pourraient y coloniser chaque cellule permissive,

ce qui pourrait conduire à un fonctionnement optimal du virus. Ces organes constitueraient ainsi les lieux principaux de production virale, les organes étant infectés par seulement des sous-groupes de segments ne seraient ainsi que des lieux de production secondaire.

Troisièmement, une entrée aléatoire des segments au sein des cellules devrait engendrer une corrélation entre la fréquence relative des segments et le nombre de cellules qu'ils infectent. Un comptage du nombre de cellules infectées par chaque segment est donc prévu, comme mentionné ci-dessus.

Enfin, il serait intéressant de voir si ces résultats sont généralisables à d'autres virus multipartites mais aussi à d'autres types de virus, même des virus monopartites. De façon remarquable, des résultats obtenus par Déborah Conflon au sein de notre laboratoire sur l'association entre un géminivirus et un satellite semblent indiquer que d'autres systèmes viraux pourraient fonctionner grâce à une « complémentation à distance ».

**CHAPITRE IV :**

**INSERTION DES MARQUEURS GENETIQUES  
NEUTRES DANS LE CLONE INFECTIEUX DU  
FBNSV POUR L'ESTIMATION DES GOULOTS  
D'ETRANGLEMENT AU COURS DE SON CYCLE  
DE VIE**

# Chapitre IV : Insertion des marqueurs génétiques neutres dans le clone infectieux du FBNSV pour l'estimation des goulots d'étranglement au cours de son cycle de vie

---

## 1. Introduction

### **Préambule**

Toutes les études théoriques ayant trait à l'évolution des virus multipartites présentent la multiplicité d'infection cellulaire (MOI) comme un paramètre qui doit être extrêmement élevé pour que ces systèmes viraux puissent fonctionner. Il nous semblait donc incontournable de l'étudier et cette partie représentait l'un des objectifs initiaux de ma thèse. Cependant, la mise au point des outils nécessaires a été très longue. Nous avons eu un problème technique assez simple, mais dont nous nous sommes rendu compte très tard. Ce problème semble maintenant réglé, et les analyses de MOI pourront être menées à termes, bien qu'elles ne puissent en conséquence et malheureusement qu'être présentées comme des perspectives de ma thèse. Vu l'importance de la question, j'ai malgré tout tenu à décrire en détail ce travail inachevé, son contexte et sa portée pour la suite du programme sur le FBNSV.

Les résultats présentés dans le Chapitre III semblent montrer que le besoin de complémentation intracellulaire n'est pas absolu, et donc qu'une MOI élevée n'est pas obligatoire. Les virus multipartites, tout au moins les nanovirus, semblent effectivement avoir développé un mode de fonctionnement leur permettant d'éviter ce coût, en pouvant infecter même si tous les segments ne sont pas ensemble au niveau de chaque cellule individuelle. Si les résultats du Chapitre III sont confirmés, l'étude de la MOI d'un virus multipartite n'en restera pas moins intéressante, car jouant un rôle clef dans l'évolution virale. Une question importante qui demeure est celle d'une entrée aléatoire ou sélective des différents segments au sein des cellules. Dans le premier cas, la MOI de chaque segment devrait être dépendante de sa fréquence dans la plante infectée, alors que dans le second cas nous devrions observer une MOI spécifique à chaque segment, décorrélée de sa fréquence. Vu notre compréhension très limitée de la biologie de ces systèmes, il est aussi possible que certains segments génomiques aient besoin d'être ensemble au niveau intracellulaire (toutes les paires possibles n'ont pas été testées au Chapitre III), que d'autres

segments s'excluent mutuellement, ou encore que certains segments soient indifférents à la présence de certains autres (comme cela semble être le cas pour les paires testées au chapitre III). Les mesures de MOI seront un complément précieux sur ces interrogations.

### **Taille des populations et goulots d'étranglement**

La sélection naturelle et la dérive génétique sont deux forces majeures façonnant l'évolution virale. Si la sélection naturelle peut conduire à une adaptation rapide des virus à un nouvel environnement par la sélection des variants génétiques les plus adaptés, la dérive génétique peut au contraire ralentir cette adaptation en « échantillonnant » de manière aléatoire les variants génétiques qui ne sont pas forcément les mieux adaptés. Les effets de ces deux forces évolutives sont plus ou moins prononcés selon la taille des populations sur lesquelles elles agissent (Hartl & Clark, 1997). L'effet de la dérive génétique sur une population de petite taille, où seuls quelques rares individus contribuent à la génération suivante, sera effectivement plus important puisqu'elle pourra conduire rapidement à la perte de certains variants génétiques par simple effet du hasard. A l'inverse, dans une population de grande taille (idéalement de taille infinie), le hasard ne peut pas éliminer les variants les mieux adaptés, et ils envahiront à coup sûr la population. La taille des populations constitue donc un paramètre important pour déterminer quelle force évolutive prédominante façonne leur structure génétique et leur évolution. La taille totale d'une population est un paramètre qui peut être trompeur, puisque seuls les individus qui se reproduisent au sein de cette population participent à son évolution. La « taille efficace » des populations est plus communément utilisée et certainement plus pertinente. Elle peut être définie comme la taille d'une population idéale au sein de laquelle la dérive génétique aurait la même intensité que dans la population observée (Wright, 1931). Pour simplifier, et bien que cela ne soit pas absolument correct, il est le plus souvent possible d'assimiler la taille efficace d'une population au nombre réel d'individus de cette population qui passent leurs gènes à la génération suivante.

Même dans des populations de taille immense, telles que les populations virales, la taille efficace peut être extrêmement réduite. Par exemple, pour rester sur le cas des virus, si à partir d'une feuille contenant  $10^{10}$  particules virales, seuls 10 génomes initient l'infection dans la feuille suivante, alors la taille efficace de cette population peut être considérée de 10. De même, si l'on assimile la colonisation d'une plante par un virus à une génération virale (c'est très discutable mais c'est pour l'exemple) et si le nombre de génomes viraux transmis d'une plante

à une autre est de 2, alors la taille efficace de la population virale est de 2 ! Même si chaque plante contient des milliards de particules et de génomes viraux.

Ces réductions drastiques et temporaires de la taille des populations sont dénommées « goulot d'étranglement ». Les goulots d'étranglement populationnels ont deux effets majeurs. Premièrement, ils conduisent à une réduction de la diversité génétique, réduction qui sera d'autant plus marquée que le goulot d'étranglement est fort. Deuxièmement, ils déterminent le nombre de copies de génomes ou de gènes lors de l'initiation de l'infection, paramètre qui, comme évoqué précédemment, joue un rôle important dans la biologie des organismes.

Des goulots d'étranglement peuvent intervenir à différentes étapes du cycle viral : au cours de la colonisation de l'hôte et lors de la transmission. Les goulots d'étranglement à ces étapes sont largement influencés par ceux ayant lieu au niveau cellulaire (pour revue voir Gutiérrez et al., 2012).

### **Multiplicité d'infection cellulaire (MOI)**

L'étude du nombre moyen de génomes viraux – identiques ou non – qui participent efficacement à l'infection de chaque cellule individuelle durant la colonisation d'un hôte pluricellulaire est donc centrale. Ce paramètre est connu sous le nom de multiplicité d'infection cellulaire ou MOI. Il est important de noter que cette définition de la MOI n'est pas nécessairement équivalente au terme couramment utilisé lors des infections réalisées en cultures cellulaires. Dans ces infections en culture cellulaire, le terme MOI informe sur le nombre d'unités infectieuses apportées par cellule dans le milieu de culture, sans prendre en compte la possibilité que seule une partie des génomes viraux dans l'inoculum infecte réellement les cellules, ni la possibilité qu'une unité infectieuse puisse ne pas représenter qu'un seul génome viral. De plus, le terme MOI en culture cellulaire n'a pas de sens lorsque l'on considère l'infection d'un organisme pluricellulaire où la population virale progresse de façon potentiellement irrégulière à l'intérieur des tissus et organes avec une charge virale reflétant assez mal l'inoculum initial. La MOI est un paramètre majeur dans l'histoire de vie des virus car elle a un impact sur un nombre impressionnant de phénomènes liés à leur cycle biologique et à leur évolution. En effet, la MOI détermine en grande partie le potentiel d'interactions intracellulaires entre les différents variants génomiques d'une population virale. Parmi ces interactions intracellulaires, on peut citer la compétition, les échanges génétiques tels que le réassortiment et la recombinaison et, la complémentarité fonctionnelle.

## **Exclusion de la surinfection**

Les virus eux-mêmes semblent limiter leur MOI par des mécanismes moléculaires très diversifiés et communément regroupés sous les termes d'exclusion de la surinfection ou exclusion de l'infection secondaire (Gutiérrez et al., 2012). Ce processus est défini comme la capacité d'un virus déjà établi à empêcher la surinfection d'une cellule par un virus homologue. Pour les virus animaux, plusieurs mécanismes ont été décrits permettant de limiter le nombre de génomes entrant, et ce à différents stades du cycle viral : au moment de l'attachement du virus à la cellule, lors de la pénétration du génome viral dans la cellule cible et lors de la réplication (Huang et al., 2008; Lee et al., 2005; Nethe et al., 2005). Le *gene silencing* serait également un des mécanismes utilisé par les virus pour limiter leur MOI (Takeshita et al., 2004).

## **MOI chez les virus monopartites**

Comme déjà indiqué dans l'introduction, seules quelques études se sont intéressées à étudier la multiplicité d'infection cellulaire et, ce majoritairement pour des virus monopartites. A l'heure actuelle, les estimations formelles de MOI se limitent à un phage, un virus d'insecte, le VIH et trois virus de plante. Ces premières études ont montré que le nombre moyen de génomes infectant chaque cellule varie entre 1 et 13, indiquant que la MOI est limitée et varie selon les virus (Gutiérrez et al., 2012).

## **MOI et sMOI chez les virus multipartites**

Il est communément admis que la MOI des virus multipartites doit au contraire être élevée afin qu'il puisse y avoir complémentation intra-cellulaire entre les différents segments. Une seule étude a regardé la MOI d'un virus bipartite, le *Soil-borne wheat mosaic virus* (Miyashita and Kishino, 2010). Les auteurs de cet article ont trouvé qu'en moyenne 5 à 6 copies de l'ARN2 de ce virus pénétraient par cellule. Cette étude n'est cependant que partielle et ne s'adresse pas spécifiquement à la biologie des virus multipartites. En effet, elle s'est focalisée sur un seul des deux segments, l'ARN2, ne permettant de faire aucune prédiction sur le nombre de copies de l'ARN1 pénétrant par cellule, ni sur le lien possible entre la fréquence d'un segment au sein de la population virale et sa chance d'entrer dans une cellule. Ainsi, il demeure impossible de savoir si tous les segments d'un virus multipartite (bipartite dans ce cas précis), quelle que soit leur fréquence relative, ont la même chance d'entrer au sein d'une cellule.



Dans le cas des virus multipartites, il nous semble pourtant plus approprié de raisonner à l'échelle du segment qu'à l'échelle du génome viral, contrairement à ce qui a été fait précédemment. Nous proposons donc dans un premier temps de comparer la MOI du FBNSV sur un segment présent à forte fréquence au sein de la population virale de la fève, le segment N, et sur un segment présent à faible fréquence, le segment S. L'échelle du segment se justifie notamment au vu des résultats exposés dans le Chapitre III sur la répartition spatiale des segments du FBNSV au sein des tissus. On qualifiera ainsi de sMOI -pour MOI d'un segment- le nombre de copies entrant et se multipliant au sein d'une cellule pour un segment donné.

### **Importance de la sMOI**

Si les différents segments génomiques du FBNSV entrent effectivement de manière aléatoire au sein des cellules permissives, comme nos premières observations le laissent présager (Chapitre III), on s'attend à ce que le nombre moyen de copies pénétrant dans chaque cellule soit plus élevé pour les segments fréquents que pour les segments rares. Autrement dit, on s'attend à ce que la sMOI du segment N soit supérieure à la sMOI du segment S. Comme évoqué plus haut, cela sous-entendrait que le segment N serait plus soumis à la compétition, à la recombinaison et à la complémentation intracellulaire que le segment S. Il serait ainsi possible de prédire un taux d'évolution distinct pour différents segments, taux d'évolution qui serait lié à la fréquence relative des segments au sein de la population virale. Le processus d'exclusion de la surinfection pourrait au contraire empêcher l'entrée répétée d'un segment au sein de la même cellule et limiter ainsi ces interactions, induisant une sMOI non corrélée à la fréquence des segments et peut être égale pour certains, voire pour tous.

### **Calcul de la sMOI**

L'étude de la sMOI sera basée sur l'utilisation de deux variants stables et équicompetitifs pour chaque segment analysé, par exemple des variants Na et Nb, et des variants Sa et Sb. Par variants équicompetitifs, nous entendons qu'il n'y ait pas de sélection en faveur de l'un ou l'autre, et qu'en moyenne sur plusieurs répétitions de plantes infectées, chacun se maintienne au cours du temps et que sa fréquence par rapport à l'autre reste constante. Les deux variants équicompetitifs de ce segment seront inoculés avec les sept autres segments -non marqués- au sein des plantes. Un comptage en microscopie confocale du nombre de cellules infectées par un variant, par l'autre, ou par les deux sera réalisé. Ce comptage sera certainement effectué sur des cellules de pétioles au vu des résultats précédemment acquis. Le nombre moyen de copies de S

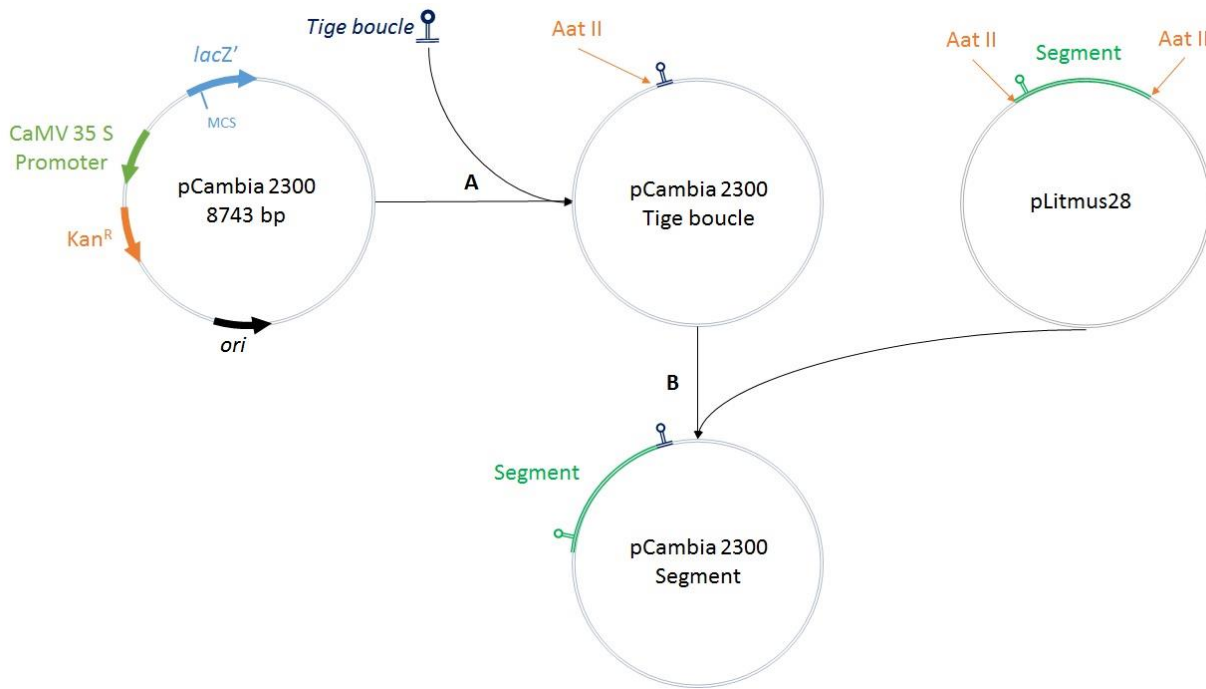
(la procédure sera la même pour N) entrant par cellule sera ensuite inféré par un calcul de maximum de vraisemblance basé sur (i) la fréquence relative des deux variants Sa et Sb mesurée dans chaque pétiole par qPCR et (ii) la proportion de cellules contenant Na, contenant Nb, et contenant les deux, pour chaque pétiole correspondant. Cette approche est similaire à celle déjà développée et éprouvée au laboratoire pour le calcul de la MOI du *Cauliflower mosaic virus* (CaMV, Gutiérrez et al., 2010). La première étape indispensable pour étudier la sMOI est donc d'obtenir des variants infectieux pour les segments ciblés, de vérifier que ces variants sont stables, puis finalement qu'ils sont équicompetitifs.

Nous n'avons pas réussi à aller plus loin que cette première étape, mais elle est vraiment déterminante pour la suite, et les deux approches différentes qui ont été utilisées pour y parvenir sont décrites ci-dessous.

## 2. Matériel et méthodes

### *2.1. Construction des clones « segment-tige boucle »*

Pour les virus à ssADN tels que les circovirus, geminivirus ou encore nanovirus, il est connu depuis très longtemps que pour que des clones dans un plasmide bactérien soient infectieux, la présence de deux copies du génome en tandem est nécessaire (Cheung, 2006). Il semble que la nécessité d'une recombinaison entre les deux répétitions, au niveau de l'origine de réplication (structure en tige boucle), permette l'excision et la circularisation du génome viral, et l'infection. La présence de deux copies en tandem dans des clones empêche l'utilisation des kits de mutagenèse basés sur des amplifications PCR (par exemple le kit « QuickChange »). Les seules séquences qui doivent être dupliquées dans un clone pour que celui-ci soit potentiellement infectieux sont les origines de réplication en tige-boucle. Notre laboratoire a mis au point une technique pour permettre le clonage d'une seule copie infectieuse d'un génome viral, adjacent à une copie de sa tige boucle déjà intégré au plasmide receveur. Cette technique permet alors l'utilisation des kits classiques de mutagenèse sur toute la séquence virale clonée (hormis celle de la tige boucle). Initialement développée pour le geminivirus monopartite TYLCV (Urbino et al., 2008), nous avons tenté de transférer cette approche pour nos clones des segments du FBNSV afin de pouvoir plus facilement y intégrer ensuite nos marqueurs génétiques (Figure 21).

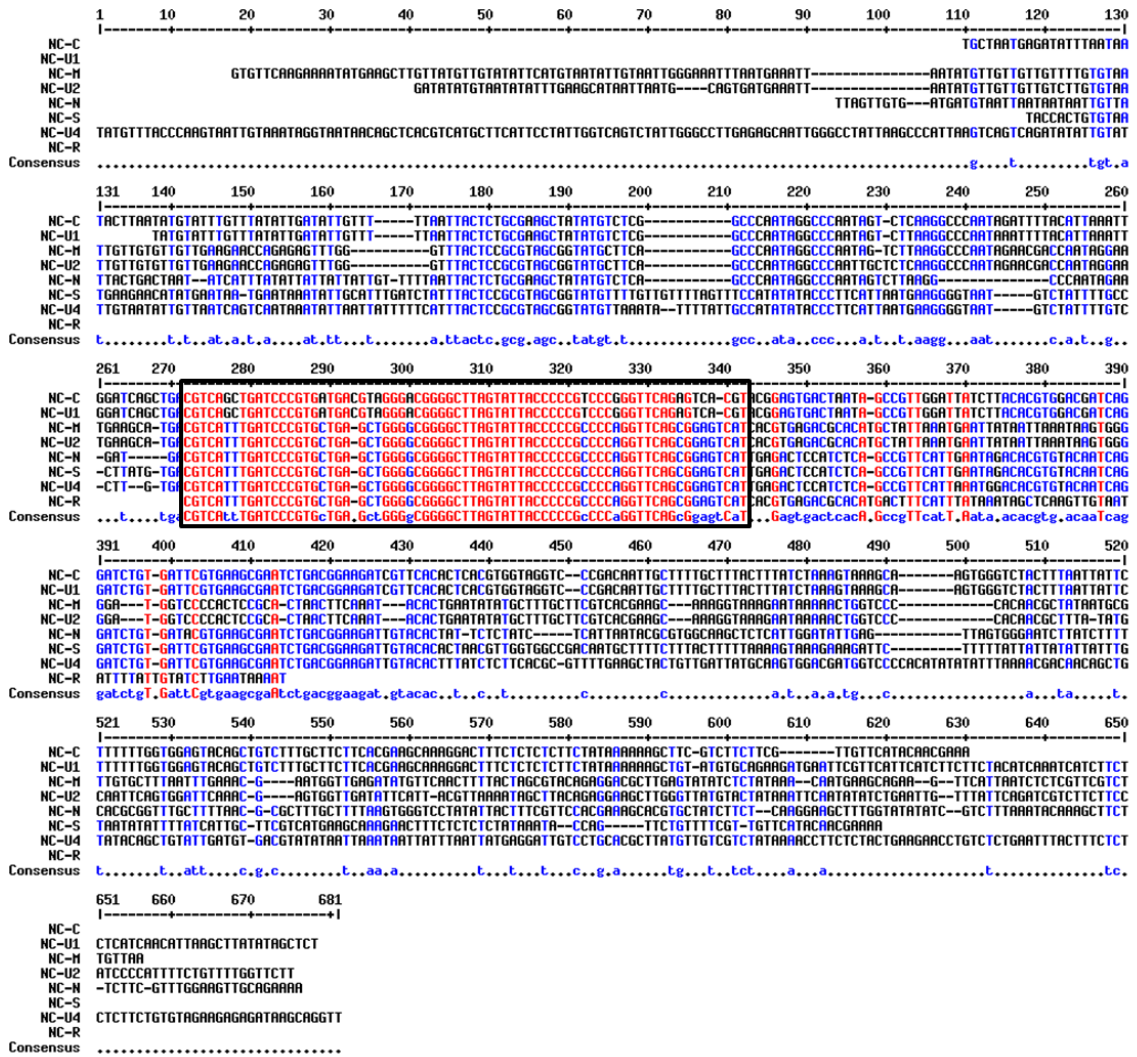


**Figure 21 : Étapes de construction des clones pCambia2300-segment-tige boucle**

(A) La région de l'origine de réplication des segments (séquence commune de type tige-boucle) est insérée au sein du plasmide pCambia2300. (B) Le clone obtenu est digéré par AatII. En parallèle, un segment est excisé du plasmide pLitmus28 par AatII et ligé dans le clone précédent. Les séquences complémentaires au plasmide pCambia2300 sont en lettres capitales tandis que les séquences complémentaires à la région de la tige-boucle sont en minuscules.

## - **Obtention des régions tige-boucle**

A l'instar des autres nanovirus, le FBNSV comporte une région commune de 78-81 nt flanquant une tige-boucle et contenant l'origine de réplication (CR-SL pour Common Region – Stem Loop). Cette région est identique à 100% entre les segments R, S, M, N, U<sub>2</sub> et U<sub>4</sub> et entre les segments C et U<sub>1</sub> mais ne présente que 83,6% d'homologie entre les deux groupes (Figure 22, Grigoras et al., 2010). Ainsi deux séquences tige-boucle ont dû être insérées au sein du vecteur pCambia 2300. Dans un premier temps, des « mégaprimers » contenant la séquence des tige-boucles, ainsi que les séquences adjacentes de la CR-SL et les séquences entourant le site de clonage du plasmide pCambia2300, ont été générés en suivant les instructions du kit de Quickchange Lightning Site-Directed Mutagenesis (Agilent Technologies). Les amorces utilisées pour générer ces deux mégaprimers sont détaillées dans le tableau 2. Les deux bandes obtenues à l'aide de ces mégaprimers et correspondant aux deux régions tige-boucle ont alors été purifiées après migration sur gel d'agarose à 2% (NuSieve GTG Agarose) selon les instructions du kit Wizard SV Gel and PCR Clean-Up system (Promega).



**Figure 22 : Alignement des séquences non-codantes des huit segments du Faba bean necrotic stunt virus mettant en évidence la région commune de la tige-boucle (encadrée en noir)**

La dernière ligne représente la séquence consensus. Les homologies strictes apparaissent en rouge. Cet alignement a été réalisé par « Multiple sequence alignment with hierarchical clustering » (Corpet, 1988).

Séquence cible	Direction	Séquence 5'-3'
pC_Tige boucle- CU1	sens	CTTTACACTTTTATGCTTCCGGCTtgacgtcagctgatcccgtg
	antisens	CACAATTCCACACAACATACGGgatccgtacgtgactctgaac
pC_Tige boucle- MNRSU2U4	sens	CTTTACACTTTTATGCTTCCGGCTgacgtcattgatcccgtgc
	antisens	CACAATTCCACACAACATACGGgatccgtgaacctgg

**Tableau 2 : Séquences des amorces utilisées pour l'obtention des « megaprimers »**

### - **Insertion des régions tige-boucle modifiées dans pCambia2300**

Les mégaprimers ainsi obtenus ont été insérés dans le plasmide pCambia2300 en suivant les instructions de Quickchange Lightning Site-Directed Mutagenesis pour l'insertion de mégaprimers. L'utilisation du plasmide pCambia 2300 se justifie par sa capacité à se multiplier à la fois dans *Escherichia coli* et dans *Agrobacterium tumefaciens* (plasmide binaire) et par son nombre de copies élevé au sein de *E.coli* et sa stabilité au sein d'*A.tumefaciens* ; ces avantages le rendent beaucoup plus simple à manipuler que pBin19, historiquement utilisé pour la construction des clones du FBNSV que nous possédons (Grigoras et al 2009). Les cartes des plasmides utilisés sont disponibles en annexe IV.

### - **Transformation et sélection des clones pCambia 2300-tigeboucle**

2µL de mix beta mercaptoéthanol contenant le pCambia 2300 avec la tige boucle, insérée à l'aide de nos mégaprimers dans l'étape précédente, ont alors été ajoutés aux 45µL de cellules XL10 Gold Ultracompetent cells (kit Quickchange Lightning Site-Directed Mutagenesis). Ce mélange a été placé dans de la glace pendant 30 minutes, puis incubé pendant 45s à 42 °C et immédiatement refroidi sur glace durant 2 min. L'ajout de 200 µl de milieu SOC (Super Optimal Broth with Catabolite repression) et une incubation d'une heure à 37°C sous agitation (250 rpm) ont permis d'initier la croissance des bactéries et d'induire la résistance à la kanamycine dans les cellules transformées. Les bactéries ont ensuite été étalées sur boîtes de Pétri de LB Agar contenant 80 µg/mL de X-gal et 20 mM d'IPTG et supplémentées de kanamycine. Les boîtes ont alors été placées durant une nuit à 37°C.

### - **Criblage des clones pCambia 2300-tigeboucle**

Un séquençage des plasmides pCambia 2300 ayant effectivement intégré la tige boucle sans autre modification artéfactuelle a été effectué sur des plasmides issus de colonies uniques et préalablement purifiés avec le Kit Promega Wizard Plus SV Minipreps. Les plasmides ainsi sélectionnés sont nommés pCambia 2300-tigeboucle. Il y en a deux, celui qui porte la séquence de la tige boucle commune aux 6 segments R, S, M, N, U<sub>2</sub> et U<sub>4</sub>, et celui qui porte celle commune aux segments C et U<sub>1</sub>.

## **Insertion des segments du FBNSV dans les plasmides pCambia2300-tigeboucle**

Les plasmides pCambia 2300-tigeboucle contenant la séquence attendue ont été linéarisés par AatII puis déphosphorylés par la Thermosensitive Alkaline Phosphatase (Promega). En parallèle, les segments génomiques du FBNSV ont été excisés, à l'aide de la même enzyme, à partir des plasmides pLitmus28 dans lesquels ils sont clonés sous forme monomérique (plasmides pLitmus28 fournis par T Timchenko). Les segments excisés ont été purifiés sur gel d'agarose à 1% en suivant les conditions du kit Wizard SV Gel and PCR Clean-Up system (Promega).

Les ligations destinées à insérer les segments du FBNSV dans les plasmides pCambia 2300-tige boucle ont été réalisées à température ambiante toute une nuit en présence d'une unité d'enzyme T4 DNA ligase (Promega), dans un volume final de 10 µl comprenant 100 ng de plasmide vecteur pour 3 fois plus d'insert (mol : mol). Une transformation bactérienne des produits de ligation a été menée suivant le protocole décrit précédemment avec des bactéries DH5α thermocompétentes.

### **- Criblage des clones pCambia2300-segment**

Un premier criblage par PCR sur colonies a été mené avec une amorce sens se fixant sur pCambia2300 et les amorces antisens spécifiques de chaque segment répertoriées dans la supplementary table S6 de l'article 1 ; amorces permettant d'amplifier les clones lorsque l'insertion a eu lieu dans le bon sens (CR-SL de part et d'autre du reste de la séquence du segment, cf Figure 21). Les clones positifs ont été purifiés avec le Kit Promega Wizard Plus SV Minipreps puis envoyés à séquencer. Ces plasmides sont nommés pCambia 2300-segmentX, où X représente le nom du segment cloné. Ils ont finalement été électroporés dans la souche d'*A.tumefaciens* Cor308.

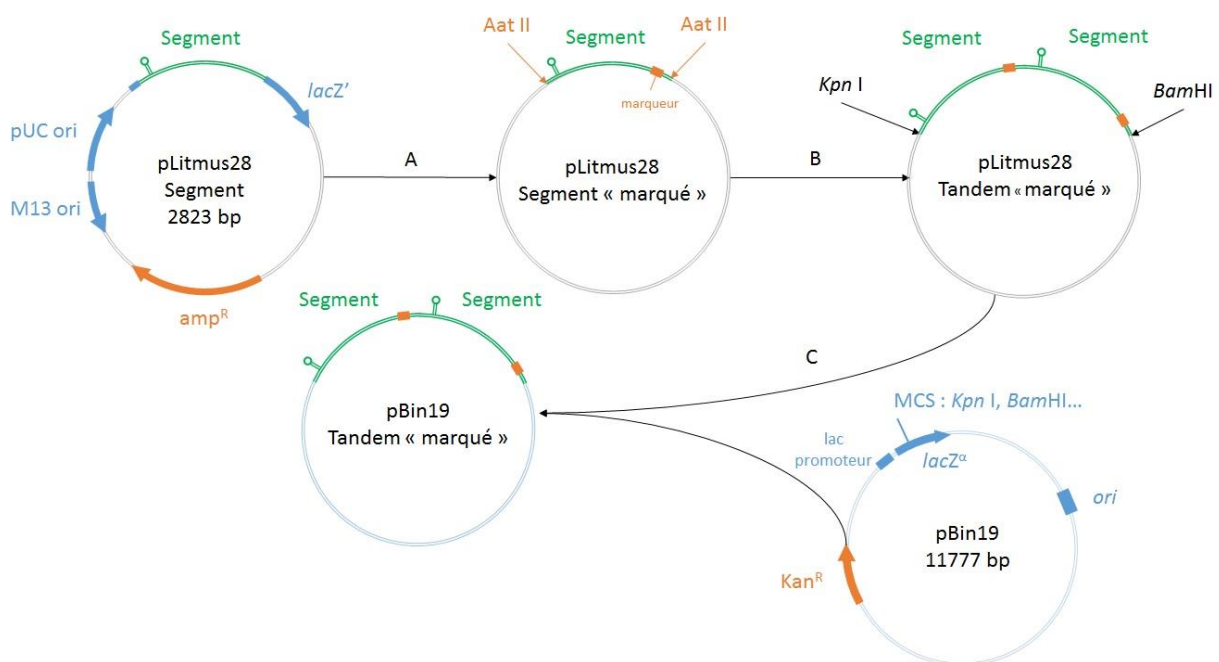
### **- Electroporation des clones pCambia2300-segmentX dans COR308**

100 ng d'ADN ont été ajoutés aux 40 µL de bactéries électrocompétentes Cor308, préalablement décongelées. Cette souche d'*A.tumefaciens* est celle dans laquelle les clones de FBNSV que nous utilisons en routine ont été introduits (Grigoras et al 2009). Les plasmides ont alors été électroporés par impulsion électrique dans ces bactéries (2.5 kV, 25 uFD et 400 Ohms ; Biorad Gene Pulser). L'ajout d'1 mL de LB et une incubation de trois heures à 28°C sous

agitation (250 rpm) ont permis d'initier la croissance des bactéries et d'induire la résistance. Les bactéries ont ensuite été étalées sur boîtes de Pétri contenant le milieu LB Agar supplémenté de kanamycine et de tétracycline. Les boîtes ont alors été placées durant deux à trois jours à 28°C.

## 2.2. Approche 2 : Construction de clones en tandem avec marqueurs génétiques

N.B. : Cette approche longue et fastidieuse, mais très classique et donc largement éprouvée dans plusieurs laboratoires, est schématisée dans la figure 23.



**Figure 23 : Etapes de la construction de clones des segments tandem avec marqueurs génétiques**

(A) Insertion du marqueur dans un clone contenant un monomère du segment d'intérêt. (B) Création d'un clone avec deux copies en tandem du segment marqué. (C) Clonage du tandem du segment marqué dans le plasmide binaire pBin19.



## - Insertion des marqueurs dans les segments

Cinq marqueurs génétiques différents ont été insérés en une position similaire située entre la fin de la séquence codante et le signal de polyadénylation dans les segments N et S. Pour le segment N, trois autres positions d'insertion ont également été testées : à deux sites différents après le signal de polyadénylation et à un site situé avant la séquence codante (Figure 24 ci-après). Ces insertions ont été effectuées au sein des segments clonés en monomères dans pLitmus 28. Une étape préliminaire d'alignement de séquences nous a permis de vérifier que la séquence des différents marqueurs (séquences de 22 nt provenant de *Xenopus laevis* (NC\_001573.1), toutes détaillées dans le Tableau 3) ne présentait pas une homologie trop forte avec la séquence des huit segments génomiques. Les marqueurs ont été insérés en suivant les instructions du kit de Quickchange Lightning Site-Directed Mutagenesis (Agilent Technologies). Les étapes sont similaires à celles décrites précédemment pour l'insertion de la tige-boucle au sein de pCambia 2300. Le criblage des clones pour la présence des marqueurs a été effectué par PCR sur colonies, suivies d'une vérification par séquençage.

Marqueur	Segment et position de l'insertion	Séquence 5'-3'
Mys 1	S-858, N-851, N- 899	ggcagttattatagtgagaagt
Mys 2	S-858, N-851	ggatagcaacagttgtataat
Mys 4	S-858, N-851	cgttgagttatcaggaatatgt
Mys 7	S-858, N-851, N-289	ccttataacttacagcactt
Mys 8	S-858, N-851	gagcaggaattatgacttatat
Inv1	N-949	cgtataacctaggtgactaag

**Tableau 3: Séquences des marqueurs insérés en différentes positions des segments S et N**

Les marqueurs Mys1, Mys2, Mys4, Mys7, Mys8 ont chacun été inséré dans des clones indépendants à la position 858 et 851 des segments S et N, respectivement. Les marqueurs Mys7, Mys1 et Inv1 ont aussi été insérés dans des clones indépendants à d'autres positions du segment N comme indiqué sur la figure 24.

## A.

5' - **GACGTC**ATTTGATCCCGTGCTGAGCTGGGGCGGGGCTTAGTATTACCCCGCCCCAGGTT**CAGCGGAGTC**ATTGAGACTC  
 81 CATCTCAGCCGTTTCATTGAATAGACACCGTGTACAATCAGGATCTGTGATTCGTGAAGCGAATCTGACGGAAGATTGTACA  
 161 CACTAACGTTGGTGGCCGACAATGCTTTTCTTTACTTTTAAAAAGTAAAGAAAGATTCTTTTTATTATTATATTATTGT  
 241 AATATATTTTATCATTGCTTCGTTCATGAAGCAAAGAACTTCTCTCTCT**TATAAA**TACCAGTTCGTTTTCGTTGTTTCATA  
 321 CAACGAAAA**ATGGT**GAGCAATTGGAAT**TGGT**CAGGCAAGAAAGGGTAGAAGAACACCTCGTTCGTGGTTACACCAGGCCTTT  
 401 CAAGTCAGCTGTTCC**TACA**ACCAGAGTCGTTCGTT**CAT**CAGTCAGCTGTATTGAAGAAAGATGACGTGCTCGGTAGTGAAA  
 481 TTAAGCCTGAAGGTGATGTTGCGCGTTACAAAATTCGTAAGGTGATGTTGAGTTGTA**CTCTG**AGGATGCGTCCGGGAGAG  
 561 TTGGTGAATTATCTGATCGTCAAGTGTAGTTCACCTATTGTTAA**TGGT**CTGCAGCTTT**CAC**AGCACCTGCTCTAATGGT  
 641 GAAAGAGAGTTGTCAAGATATGATTACTATCATAGGCAAGGTAAAGTATAATCGAATGGTGTAGCTGGATCAGATTGTA  
 721 CTA**AAAT**CATTC**AA**TAAAGTTCATCAGATATGAGCTGGTATTAGTCAAACACAACATTTGTACGTTGTAATGTATA**CCAGT**  
 801 **GAAGCTGTTAAGACTGTATTATAACATATAGTCTATATTGAAGTTTAA**TACCAC**TGTG****NNNNNNNNNNNNNNNNNNNNNN**  
 859 TAATGAAAAACATATGAAAAATG**AAT****AAA**TATTGCATTGATCCAA**TACTCCGCGTAA**CGGTATGTTTGTGTTT**TA**  
 939 GTTT**CCATATATACCCTTCATT AATGAAGGGSTAATGTCTATTTTG**CCCTTATGT**GACGTC** -3'

Légende (sens 5'-3'):

- AatII**
- Région commune de la tige-boucle
- Région Commune N-S-U4
- TATA BOX
- Région codante du segment (ORF)
- POLY A SIGNAL
- Région majeure commune
- Région commune S-U4

## B.

5' - **GACGTC**ATTTGATCCCGTGCTGAGCTGGGGCGGGGCTTAGTATTACCCCGCCCCAGGTT**CAGCGGAGTC**ATTGAGACTC  
 81 CATCTCAGCCGTTTCATTGAATAGACACCGTGTACAATCAGGATCTGTGATTCGTGAAGCGAATCTGACGGAAGATTGTACA  
 161 CTATTTCTATCTCATTAAATACGCGTGGCAAGCTCTCATTGGATATTGAGTTAGTGGGAATCTTATCTTTT**CAC**CGCGTT  
 241 TGCTTTTAAACGCGCTTTGCTTTTAAAGTGGGTCCTACATTACTTT**CGTTC****NNNNNNNNNNNNNNNNNNNNNN**CAGGAAAGC  
 299 ACGTGCTATCTTCTCAGGAAGCTTTGGT**ATATAT**CGTCTTTAAATACAAAGCTTCTTCTT**CGTT**TGGAAAGTGCAGAAA  
 379 **AATGGCAGATTGGTTTTCTAGT**CCTTT**GAAGACATGTACACAC**CGTCTGCGACTTCCCTTCTCTT**GCT**GGAAATCCACAGC  
 459 **AAGAGTTTATGTGTTGTGATAGCATGAAAGACAAGCTCAACGATTCAAGAAAGGTTCTGCTTGTAGTTGCAGTGTGAGC**  
 539 TTCAATGGT**AGTTTCTAT**GGTGAAACAGGAACGTTCTGGT**CAGTTGCAGATATCTATGAT**GGAAAGACGACGGT**GT**TTG  
 619 TCGACCGATTGGATATT**TACCAAT**GGTGGTTATCTGTATCATAATGATATGGCTATTACGAAAGT**CAGAGGACGTTTA**  
 699 ATCTCGACATCGAGTCTCAGTATT**GAAGAAAGATGAAGATTATAGAAGAAAATTTATAATTACCGTTCTTAATGAGAAC**  
 779 GGGTTAGATAGT**CAGTGTGATTGAAAATGTTTATTG**TACATTC**ACTCAGAA**TCAAAGT**GTAAT**TAGTTGTG**NNNNNNNN**  
 851 **NNNNNNNNNNNNNN**ATGATGTAAT**TAATAATAAT**TGTTATTACTGACTAATATCATTATAT**NNNNNNNNNNNNNNNN**  
 900 **NNNNN**ATTATTATTGGTTTTAAT**TACTCCGCGTAA**CGGTATGCTTCAGCCCAAT**NNNNNNNNNNNNNNNNNNNN**AGG  
 961 CCCAATAGTCTCAAGGCCAATAGAAGAT**GACGTC** - 3'

Figure 24 : Zones d'insertion des marqueurs au sein des segments S et N du FBNSV.

Les séquences complètes des segments S (A) et N (B) sont présentées, avec le site AatII aux extrémités qui est utilisé pour leur clonage. Toutes les caractéristiques des segments qui ont influencé notre choix pour la position de l'insertion des marqueurs sont indiquées avec un code couleur détaillé dans la figure.

La position des marqueurs est matérialisée par une série de « N » surlignés qui correspond au nombre de nucléotides insérés. La séquence exacte des nucléotides insérés est détaillée dans le Tableau 3.

(A) Un seul site d'insertion a été testé pour le segment S (position 858), mais cinq marqueurs de séquence différentes ont été introduits dans cinq clones indépendants.

(B) Quatre positions différentes ont été testées pour l'insertion de marqueurs au sein du segment N. Cinq marqueurs distincts ont été insérés à la position 851 dans des clones indépendants. Pour les positions 289, 899 et 949, un seul marqueur a été introduit.

NB : les marqueurs ne sont pas pris en compte dans le calcul des positions indiquées dans cette figure.

#### - **Obtention de tandems avec marqueurs**

Les monomères contenant les marqueurs d'intérêt ont été digérés par AatII puis purifiés sur gel. Les inserts ainsi obtenus ont été incubés pendant 4 min à 14°C avec la T4 DNA ligase et le tampon correspondant afin de favoriser l'émergence de tandems. Le vecteur pLitmus28 préalablement digéré par AatII et déphosphorylé a ensuite été ajouté selon un ratio insert : vecteur de 3 :1 (mol :mol). Une incubation d'une nuit à 4°C a ensuite permis cette ligation. Les produits de ligation ont été transformés par choc thermique dans la souche DH5α selon le protocole décrit précédemment. Le criblage des clones a été effectué par digestion en utilisant l'enzyme MluI pour le segment N et l'enzyme NdeI pour le segment S ; l'intérêt de ces enzymes est qu'elles permettent de vérifier à la fois la présence d'un tandem et son sens d'insertion. Une confirmation par séquençage a suivi ce premier criblage.

Les tandems ainsi obtenus ont alors été insérés dans le plasmide pBin19 après double digestion des inserts et du vecteur par les enzymes KpNI/BamHI. Les clones positifs ont finalement été électroporés selon le protocole décrit plus haut dans la souche d'*A.tumefaciens* Cor308.

#### - **Test d'infectiosité des clones**

Afin de tester si les clones obtenus étaient infectieux, chaque clone a été inoculé en présence des sept autres segments « sauvages » utilisés jusqu'à présent. Les agroinoculations ont été conduites suivant le protocole décrit dans le chapitre I. Leur infectiosité a ensuite été déterminée grâce à l'apparition de symptômes dans le cas des segments indispensables à l'infection de la fève. La présence des marqueurs a été vérifiée à la fois par qPCR en utilisant une amorce s'hybridant au marqueur testé et par séquençage.

### 3. Résultats

- **Construits segment-tige boucle**

La première approche adoptée pour obtenir différents variants d'un même segment visait à reproduire au préalable des clones « infectieux » des segments génomiques du FBNSV, mais qui seraient ensuite utilisables directement pour l'introduction de marqueurs génétiques par mutagenèse en une seule étape. Nous avons donc cherché à produire des clones « infectieux » contenant une duplication limitée à la séquence « tige-boucle » comme cela a été fait pour le *Tomato yellow leaf curl virus* (Urbino et al., 2008). Deux séquences de tiges-boucles différentes ont été insérées avec succès au sein du vecteur pCambia 2300, une pour recevoir les segments R, S, M, N, U<sub>2</sub> et U<sub>4</sub>, et une autre pour recevoir les segments C et U<sub>1</sub>. Les segments C, N, R, U<sub>1</sub> et U<sub>2</sub> ont pu ensuite être insérés avec succès, chacun au sein du plasmide contenant la séquence tige-boucle qui lui correspond (le clonage des segments M et R ont échoué et le clonage d'U<sub>4</sub> n'a pas été testé). Nos premiers tests d'agroinoculation visant à tester l'infectiosité de ces clones se sont avérés infructueux ; jamais le segment cloné sous une forme monomérique avec simple duplication de la région tige boucle n'a pu être retrouvé au sein de plantes inoculées. Lorsque nous testions un segment indispensable tel que R par exemple (les autres segments étant inoculés sous leur forme classique de clone en tandem), les plantes étaient tout simplement non infectées. Lorsque nous testions un segment non-indispensable tel que N, les plantes étaient bien infectées mais N indétectable par qPCR.

Ces résultats nous ont poussés à suivre la démarche classique des clones en tandem pour introduire les marqueurs génétiques au sein de nos clones infectieux du FBNSV. Même si cette méthode était *a priori* plus longue, elle avait peut-être plus de chance de procurer des clones marqués infectieux (Matériel et Méthodes, section 2.2).

Nous nous sommes rendus compte beaucoup plus (et malheureusement trop) tard que la non-infectiosité de nos clones « tige boucle-segment » n'était en fait pas liée à la non-fonctionnalité des clones eux-mêmes, mais due à la perte d'un plasmide (VirE) chez la souche bactérienne COR308 utilisée. Ce plasmide, au moins dans le cas de l'agro-inoculation des nanovirus, est particulièrement important pour un transfert efficace de l'ADN viral aux plantes. De récents tests conduits sur les clones pCambia 2300-segmentR et pCambia 2300-segment U1 transformés dans des bactéries COR308 qui contiennent VirE ont en effet prouvé l'infectiosité

de ces clones, qui pourront donc être utilisés dans la suite de ce programme de recherche pour l'introduction directe de marqueurs génétiques, ou pour tout autre besoin de mutagenèse, comme par exemple des approches de génétique inverse.

- **Insertion de marqueurs dans des tandems**

La « révélation » de la fonctionnalité de nos clones de type segment-tige boucle étant arrivée trop tard, nous nous étions donc tournés vers le protocole classique des clones en tandem décrit précédemment pour obtenir des segments marqués. Vu la lourdeur de ces clonages, seulement deux segments ont été choisis pour étudier la sMOI : les segments S et N, qui présentent respectivement des fréquences relatives faibles et élevées au sein de la population virale dans la fève. Cinq voire six marqueurs différents ont été introduits avec succès dans chacun de ces segments, clonés sous forme monomérique dans pLitmus28 (Figure 24 et Tableau 3). Cependant, sur le total des 8 constructions prévues pour le segment N (6 marqueurs au total, certains insérés en positions différentes dans le segment), seulement 5 ont pu être finalement obtenus sous forme tandems marqués. De même, sur les 5 constructions initialement prévues pour le segment S, seulement 3 tandems avec marqueurs ont été obtenus *in fine*, et ont donc pu être testés.

L'idée derrière l'insertion d'un nombre aussi élevé de marqueur dans chacun des segments, était que parmi eux certains risquaient d'être létaux, ou délétères. Nous espérons qu'avec ceux qui sont viables, des co-infections 2 à 2 révéleront au moins une paire équicompetitive pour chacun des segments N et S.

- **Infectiosité, stabilité et équicompetitivité des variants de N**

A ce jour, nous possédons 3 clones infectieux contenant chacun un marqueur différent pour le segment N ; ces 3 marqueurs sont positionnés au même endroit de la séquence du segment N (Marqueurs Mys2, Mys4 et Mys7, position 859, Figure 24). Un séquençage de type Sanger de l'apex des plantes infectées avec ces clones a montré la stabilité des marqueurs au cours de l'infection de la plante, où une délétion spontanée du ou des marqueurs n'a jamais été observée.

Des plantes ont été coinoculées avec deux de ces segments marqués en quantité équimolaire (N-Mys2/N-Mys7). Le trop faible nombre de plantes infectées au cours de cette expérience ne nous a pas permis de pouvoir tester leur équicompetitivité. Bien que nous puissions au moins

conclure que ces deux marqueurs co-existent, les mesures de leur quantité relative en qPCR sont trop peu nombreuses pour pouvoir mener un test statistique. Cette expérience devra être répétée.

Pour deux autres clones contenant les marqueurs Mys 1 et Inv1, respectivement en positions 899 et 949 du segment N (Figure 24), des plantes ont été inoculées et contenaient effectivement le segment N démontrant leur « infectivité ». La stabilité de ces marqueurs et leur équicompetitivité potentielle face aux autres marqueurs reste à évaluer.

- **Infectiosité, stabilité et équicompetitivité des variants de S**

Trois clones contenant trois marqueurs différents ont également été engendrés avec succès pour le segment S (Marqueurs Mys1, Mys7 et Mys8). Quelques plantes inoculées avec ces clones se sont révélées infectées, ce qui signe la fonctionnalité de ces trois segments S marqués. La stabilité des marqueurs et leur équicompetitivité (testée par paire) n'ont pas encore été évaluées.

## 4. Conclusion

Plusieurs tests restent à effectuer avant de pouvoir déterminer la sMOI des segments N et S avec les clones engendrés.

Une étape cruciale sera la démonstration que certains de nos segments marqués sont effectivement équicompetitifs quand ils sont co-inoculés à des plantes. Les méthodes statistiques de calcul de la sMOI sont en effet fondées sur une égale chance des deux segments marqués de pénétrer et de se répliquer au sein des cellules.

S'en suivront des comptages par FISH et microscopie confocale (hybridation de sondes spécifiques sur les marqueurs), pour déterminer la proportion de cellules co-infectées par les deux variants parmi l'ensemble des cellules d'un même pétiole contenant au moins un des deux variants. La mise au point de la technique de FISH présentée dans le chapitre précédent devrait permettre de limiter les difficultés techniques à ce niveau.

En fonction des résultats obtenus, une étude similaire sur les autres segments pourra être menée. L'utilisation des clones segment-tige boucle pourra faciliter le processus de clonage.

## **Perspectives à plus long terme**

Au-delà des études de sMOI à venir, qui restent une priorité de ce programme de recherche, nos clones marqués pourront permettre d'étudier les tailles efficaces des populations lors de la colonisation de la plante, de l'invasion des feuilles, et de la transmission par pucerons.

L'étude de la taille efficace lors de la transmission sera particulièrement intéressante, parce que le nombre d'unités infectieuses transmises par un insecte vecteur est un paramètre totalement inconnu à ce jour pour un virus transmis sur le mode circulant. Cette étude sur le FBNSV pourra ainsi élargir le champ de connaissances sur la taille efficace des populations lors de la transmission par insecte vecteur et plus particulièrement par puceron. La technique employée ici sera adaptée à partir de celle mise au point au laboratoire lors des études de la taille efficace des populations virales (Monsion et al., 2008) et de la MOI (Gutiérrez et al., 2010) sur le modèle CaMV.

L'étude des goulots d'étranglement génétique et populationnel des virus multipartites au cours de leur transmission par vecteur se limite à deux articles traitant du *Cucumber mosaic virus* (Ali et al., 2006; Betancourt et al., 2008). Outre la question de savoir si des segments génomiques qui s'accumulent à des fréquences différentes ont les mêmes chances d'être transmis, il sera intéressant de voir si ce qui est transmis correspond à une sorte de « core génome » (1 copie de chaque segment), ou si cela correspond à une formule génomique.

De même, l'étude de la taille efficace des populations du FBNSV au cours de la colonisation de son hôte permettra de savoir si ce virus a mis en place des mécanismes permettant de limiter les forts goulots d'étranglement, qui réduisent de manière drastique la taille des populations chez la plupart des virus étudiés lors de la colonisation de nouvelles feuilles (Gutiérrez et al., 2012).



## **CONCLUSION GENERALE**

## Conclusion générale

---

Les conclusions et perspectives spécifiques aux différents chapitres ayant été discutées au fil du manuscrit, cette dernière partie sera relativement brève.

Ce projet s'inscrivait dans une nouvelle thématique du laboratoire visant à pallier un manque dans la littérature relative aux virus multipartites. Le programme correspondant qui a été initié avec ma thèse (ou plutôt durant mon M2) avait pour objectif de mieux comprendre le fonctionnement de cette catégorie de virus en se basant sur l'étude d'un nanovirus, le *Faba bean necrotic stunt virus*.

La complexité à manipuler ce virus et le peu de connaissances à son égard ont rendu certaines expériences difficiles à réaliser et ne nous ont pas permis d'atteindre tous les objectifs initialement fixés. Notamment, l'étude de la multiplicité d'infection cellulaire et plus généralement des goulots d'étranglement au cours du cycle d'un virus multipartite n'a pas pu être menée à terme. Cependant, deux résultats principaux ont pu être dégagés : la description du phénomène de la formule génomique d'une part, et la localisation des différents segments au niveau cellulaire d'autre part.

### **Formule génomique**

Nous avons démontré que les segments du FBNSV présentaient des fréquences relatives différentes et que celles-ci étaient reproductibles au sein d'un hôte, quelles que soient les conditions à l'inoculation testées. Ces résultats nous ont amenés à définir un nouveau concept : celui de la formule génomique. Une première hypothèse a alors été émise : cette formule génomique serait le reflet d'une régulation différentielle du nombre de copies de gènes. L'impact des variations du nombre de copies de gènes sur l'expression génétique et le phénotype a été démontré chez de nombreux organismes à la fois procaryotes et eucaryotes (Myleiko 2008, Nair 2008, Hastings 2009, Almal 2012, Sjodin 2012). Parmi les perspectives immédiates de ce travail, il conviendra donc de confirmer que ceci est aussi le cas chez les virus multipartites.

i) Il sera intéressant de tester s'il existe un lien entre le nombre relatif de copies des segments du FBNSV dans une plante et la régulation de l'expression des gènes correspondant ainsi que

l'expression phénotypique. L'étude des quantités relatives des transcrits et des protéines semble particulièrement pertinente puisqu'*a priori* on pourrait penser que le nombre relatif de copies des segments est à l'origine d'un nombre différentiel d'ARN et de protéines et qu'ainsi, une corrélation positive existe entre ces quantités. *A contrario*, on peut imaginer que ces différences de copies visent à compenser une dégradation préférentielle ou une moindre production de produits d'expression des segments les plus fréquents et qu'ainsi une corrélation négative existe entre ces quantités. Cette hypothèse peut paraître étonnante mais pourrait expliquer le faible nombre relatif de copies du segment S. Ce segment code effectivement pour la protéine de capsid, et sa faible fréquence relative au sein de la population virale est assez surprenante étant donné que 60 protéines de capsides sont utilisées dans l'élaboration d'une seule particule virale. Cette analyse proposée est à considérer à deux niveaux. En effet les quelques lignes ci-dessus s'intéressent à la relation qui existe entre abondance moyenne d'un segment dans la population virale intra-hôte et abondance moyenne de sa protéine (ou son ARN) dans la plante infectée. L'autre niveau est celui de la relation directe qui pourrait exister entre la variation de la quantité d'un segment donné et la variation concomitante de ses produits d'expression. D'après les études théoriques sur le sujet cette relation n'est pas forcément positive ni même linéaire et tout cela sera à caractériser.

ii) L'impact de la formule génomique sur l'expression phénotypique constitue une deuxième perspective de ce travail visant également à comprendre son rôle biologique. Ainsi, l'analyse de l'accumulation virale et du taux de transmission pourra permettre de voir si la « setpoint genome formula »-formule génomique vers laquelle les fréquences de segments convergent-correspond à un stade optimal pour le virus. Les variations de formule entre des plantes infectées en parallèle permettront en outre de tester si des formules différentes engendrent des symptômes différents. Ce dernier point pourra aussi être analysé par génétique inverse, en essayant de créer des mutations dans le clone infectieux qui modifient la formule, et en suivant les modifications de symptôme associées.

La possibilité de modifier de manière différente le nombre de copies de chaque gène pourrait permettre au virus d'ajuster rapidement ce nombre, donc sa formule génomique, en fonction de l'environnement dans lequel il se trouve. On peut effectivement penser que selon son environnement, les « besoins » du virus diffèrent. Cela peut paraître particulièrement pertinent dans le cas du vecteur versus de l'hôte, au sein desquels les segments et/ou produits d'expression nécessaires au bon fonctionnement viral ne sont pas forcément les mêmes, ou en

même quantité. Dans le cas des pucerons, les modifications de formule génomique pourraient être fonctionnellement reliées à la transmission. Aucune de nos données ne nous permet pour le moment d'affirmer une telle conclusion, mais ces travaux seront poursuivis. De même, les interactions entre le virus et l'hôte pouvant différer selon les espèces végétales, un ajustement rapide du nombre de copies des gènes selon l'hôte pourrait être nécessaire au bon fonctionnement du virus. L'étude d'évolution expérimentale au sein de l'hôte *Medicago truncatula* présentée en complément de l'article 1 vise à répondre à cet aspect. Cet ajustement rapide du virus pourrait constituer un avantage unique aux virus multipartites.

Outre le rôle que la formule génomique peut jouer dans le fonctionnement du virus, une seconde question particulièrement intrigante est de savoir par quel processus, ou par quel(s) mécanisme(s) elle peut être régulée. Ces questions se posent aussi bien au sein des hôtes qu'au sein des vecteurs du FBNSV.

Nous avons émis deux grandes hypothèses pour l'établissement de la formule génomique *in planta*, dans l'article présenté au chapitre I, les hypothèses H1 et H2.

Pour commencer par la plus simple, H2, les segments seraient répliqués de manière différentielle ce qui les conduirait chacun à une fréquence spécifique. Ce mécanisme impliquerait cependant une boucle de rétrocontrôle afin que ces fréquences soient conservées et qu'ainsi les segments rares ne soient pas progressivement perdus. Le test de cette hypothèse pourrait se révéler délicat dans la mesure où il nécessite un système expérimental dans lequel aucun passage de cellule à cellule ne doit s'opérer, afin de mesurer le taux de répllication de chaque segment sans interférence qui serait liée au mouvement, ou à une quelconque pression de sélection au fil de la colonisation de l'hôte. La transfection de protoplastes pourrait être une solution intéressante à utiliser. La difficulté principale résiderait dans la transfection des huit segments au sein d'une cellule permissive. Cependant, cette approche a déjà été utilisée avec succès pour deux nanovirus différents pour lesquels une répllication au sein de protoplastes transfectés a été constatée (Chu et al., 1993; Franz et al., 1999) montrant ainsi la faisabilité de cette technique. La quantification du nombre de copies dans les protoplastes suite à une transfection en quantité équimolaire de chaque segment pourrait ainsi permettre de regarder si le taux de répllication des segments est effectivement différent. Il faudrait à ce stade vérifier que tous les segments ont la même probabilité d'entrer dans les différentes cellules afin que cela ne biaise pas les quantités finales observées. Mener cette expérience au cours d'une cinétique avec plusieurs points de temps pourrait régler toutefois ce problème.

La deuxième hypothèse émise, H1, postulait que les cellules infectées avec des proportions « optimales » de chaque segment participeraient de manière plus active à la production de nouveaux virus dans le milieu commun (la sève des tubes criblés) que celles infectées avec des fréquences non-optimales. Ainsi, petit à petit, les fréquences des segments convergeraient vers leurs fréquences optimales correspondant à la formule génomique observée. Au vu des résultats obtenus sur la distribution spatiale des segments, cette hypothèse pourrait très bien s'appliquer à un ensemble de cellules interconnectées et non à une cellule unique comme préalablement suggéré. Un ensemble de cellules infecté avec un nombre relatif optimal de copies de chaque segment fonctionnerait ainsi de manière optimale. Si au cours des comptages des cellules infectées avec des segments rares et de celles infectées avec des segments fréquents, une corrélation entre nombre de cellule infectées et fréquence du segment correspondant n'était pas démontrée, alors l'hypothèse H1 serait remise en cause.

Il est possible que les mécanismes à l'origine des changements de fréquence au sein des pucerons soient totalement différents de ceux *in planta*. En tout état de cause, à ce stade de nos connaissances, un tout premier objectif serait d'étudier les interactions cellulaires et moléculaires entre les pucerons et les virus. L'étude de la distribution spatiale des segments au sein de l'insecte pourrait par exemple permettre de savoir si tous les segments sont toujours ensemble et si une réplication au sein de l'insecte est envisageable (entrée des segments au sein de cellules nucléées, présence des segments dans les noyaux) ou complètement exclue.

L'importance du nombre de copies de gènes pour un virus multipartite n'avait jamais été jusqu'à ce jour invoquée et constitue pour le moment l'unique avantage ayant un impact direct sur la fréquence des segments. La formule génomique observée montre que certains segments sont rares et ont théoriquement plus de chance d'être perdus, augmentant le coût pour le système viral.

### **Etude de la distribution spatiale des gènes du FBNSV**

L'étude de la (co)localisation des différents segments du FBNSV s'adresse directement à l'existence du coût majeur considéré chez les virus multipartites, qui est la nécessité de mettre au moins une copie de chaque segment par cellule infectée. Comme je le comprends, ce coût est lié au fait que pour une grande proportion des cellules susceptibles cela ne sera pas le cas, et que le matériel viral qui y a pénétré se trouvera dans une impasse et sera perdu.

Je considère que l'étude expérimentale de ce coût constitue le deuxième résultat majeur de cette thèse et a été amplement discuté dans le chapitre III et en fin de revue bibliographique présentée en introduction. Nos premières investigations visant à étudier la répartition des segments au sein des cellules semblent montrer que le FBNSV aurait recours à un mécanisme lui permettant de limiter un tel coût. Nos observations semblent effectivement indiquer que la complémentation pourrait avoir lieu à une échelle intercellulaire. Bien qu'encore préliminaires, ces résultats pourraient révéler un nouveau mode de fonctionnement viral basé sur une communication intercellulaire permettant l'échange de produits d'expression des gènes viraux « à distance ». Après confirmation de ces résultats, il sera particulièrement intéressant de voir si un tel fonctionnement s'applique à d'autres virus multipartites –notamment à des virus non restreints au phloème- ainsi qu'à d'autres systèmes viraux.

S'ouvrira alors un pan quantitativement très important en termes de recherche qui s'attachera à comprendre ce fonctionnement au niveau moléculaire et cellulaire, c'est-à-dire à l'étude des propriétés de mouvement des ARNs et protéines chez ce type de virus.

Ces résultats sont assez novateurs et ouvrent de nombreuses perspectives qui seront abordées au sein d'un projet ANR qui fera suite à cette thèse.

# ANNEXES

# Annexes

---

## Liste des annexes

- Annexe I : Purification du FBNSV
- Annexe II : Calcul de la probabilité de perdre au moins un segment au cours de l'infection de la fève
- Annexe III : Phloème
- Annexe IV : Carte des plasmides utilisés au cours de cette thèse



## 1. Annexe I : Purification du FBNSV

Des feuilles symptomatiques ont été collectées à partir de fèves infectées par le FBNSV et homogénéisées dans un tampon citrate (10 mM; pH6.0) contenant 0.1% (v/v) Tween 20, 0.5% (w/v) de sulfite de sodium et 0.05% (w/v) de Cellulase Onozuka R10 (Yakult Pharmaceutical Industry Company Ltd., Japan). Après une nuit d'incubation à 4°C sous agitation, l'homogénat a été filtré par miracloth. Un volume de 25% de chloroforme-butanol (1:1) a été ajouté au filtrat. Après une heure d'agitation, ce mélange a été centrifugé pendant 20min à 10 000g et le surnageant a été récupéré. Les virions ont été précipités par ajout de polyéthylène glycol 6000 à 8% (w/v) et de 0.3M de chlorure de sodium suivi d'une centrifugation de 20 min à 15 000g. Le culot a été resuspendu dans un tampon phosphate (10mM, pH 7.5) et chargé sur un gradient discontinu de sucrose (à des concentrations respectives de 10, 15, 25, 30 and 45%) qui a été centrifugé pendant 4h à 35 000 rpm (Beckman rotor SW41 Ti). Les fractions contenant du virus ont été collectées à l'aide d'une pipette Pasteur, diluées trois fois dans du tampon phosphate et sédimentées pendant 3h à 40 000 rpm (Beckman rotor Ti 50.2). Les culots ont été finalement resuspendus dans du tampon phosphate et la présence de virions a été vérifiée par microscopie électronique à transmission (Jeol JEM 100CX II) opéré entre 60-80 kV.

## 2. Annexe II : Calcul de la probabilité de perdre au moins un segment au cours de l'infection de la fève

Sous l'hypothèse que l'infection des cellules s'effectue de manière aléatoire et indépendante pour les huit segments du FBNSV, le nombre de copies d'un segment donné entrant dans une cellule suit une loi de Poisson de paramètre  $\lambda p_i$ , avec  $\lambda$  le nombre de particules moyen entrant par cellule et  $p_i$  la fréquence relative du segment considéré au sein de la population virale dans une fève. Les particules sont ici considérées comme étant indistinguables quel que soit le segment qu'elles contiennent,  $\lambda$  correspond ainsi au nombre moyen total de particules entrant. La classe nulle de chaque distribution de Poisson correspond à la probabilité de ne pas être infectée par le segment correspondant :

$$P_i(0) = e^{-\lambda p_i}$$

Ainsi, la probabilité d'être infecté par au moins une copie de chaque segment est (l'indice  $c$  correspondant à un « génome complet »):

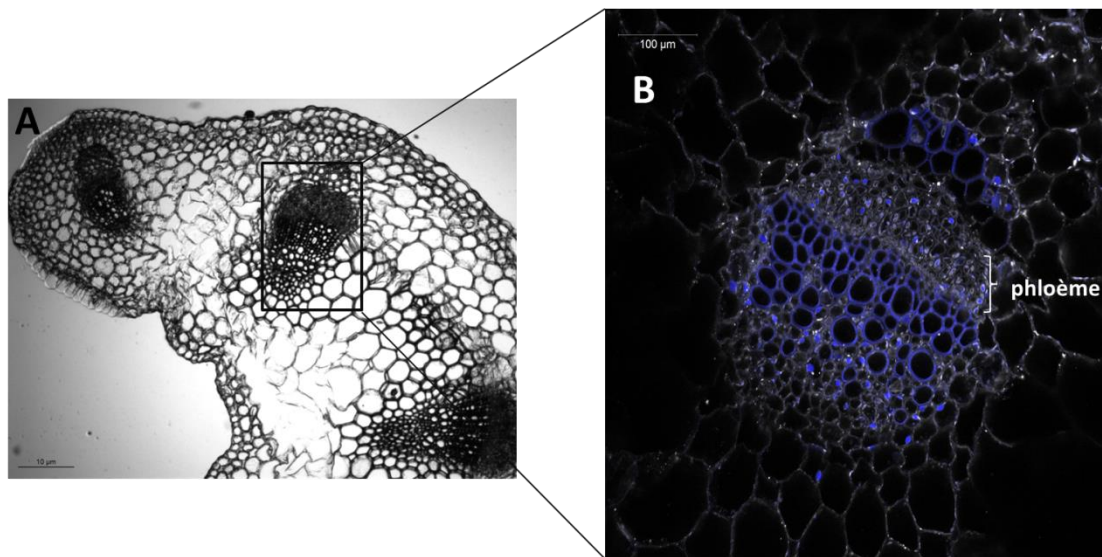
$$P(c) = \prod_i (1 - e^{-\lambda p_i})$$

On peut ainsi en déduire la probabilité qu'au moins un segment soit perdu lors de l'infection (l'indice  $inc$  correspondant à un « génome incomplet ») :

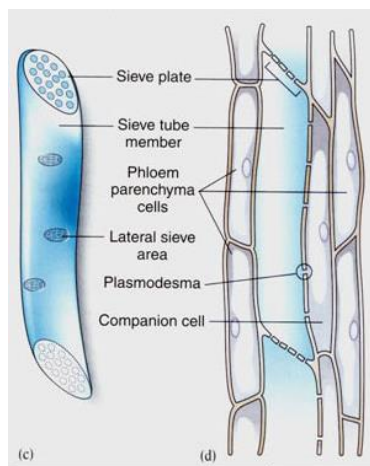
$$P(inc) = 1 - \prod_i (1 - e^{-\lambda p_i})$$

En fixant des valeurs de  $\lambda$ , on peut en déduire la probabilité de ne pas être infecté par un « génome entier ».

### 3. Annexe III : Phloème

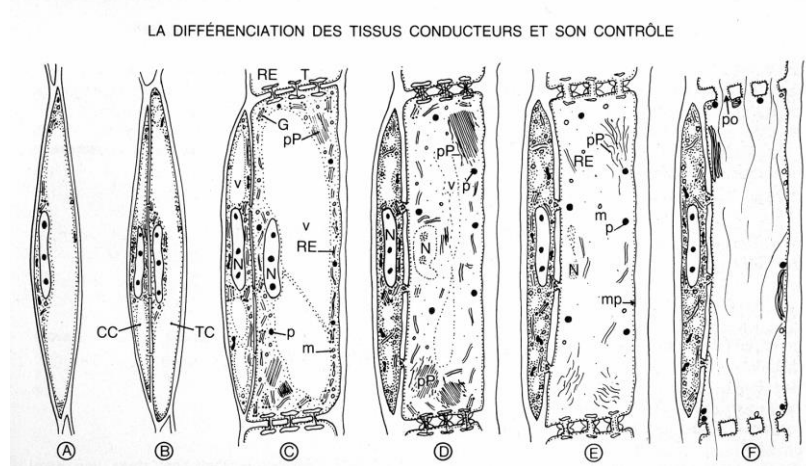


**Figure 25 : Localisation du phloème au sein d'un pétiole de feuille de fève saine.**



**Figure 26 : Schéma représentatif des éléments du phloème**

Les tubes criblés forment un syncytium de cellules vivantes « alimenté » par les cellules compagnes (<http://biochemunrated.wordpress.com/category/sieve-tube-element>)



**Figure 27 : Représentation schématique des étapes de différenciation des tubes criblés**

A partir d'une même cellule initiale qui va se diviser, la cellule compagne et l'élément criblé, cellule qui va donner le tube criblé vont se différencier de façon différente. La cellule criblée va progressivement perdre tous ces organites, puis les plasmodesmes situés dans l'axe longitudinal vont s'élargir pour former des pores, qui finiront par s'ouvrir pour laisser passer le flux de sève. CC: cellule compagne. TC: tube criblé. RE: réticulum endoplasmique. N: noyau. p: plastes. m: mitochondrie. PP: protéines P. v: vacuole. Po: pore. Mp: membrane plasmique. (D'après Atlas de biologie Végétale, Roland et al.2008)

4. Annexe IV : Carte des plasmides utilisés au cours de cette thèse

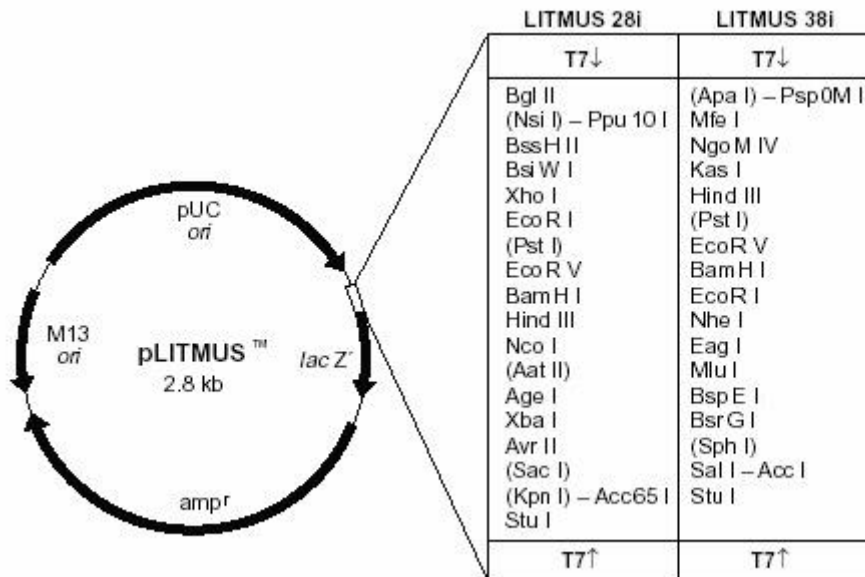


Figure 28 : Carte du plasmide pLitmus 28

(Image : <https://www.neb.com>)

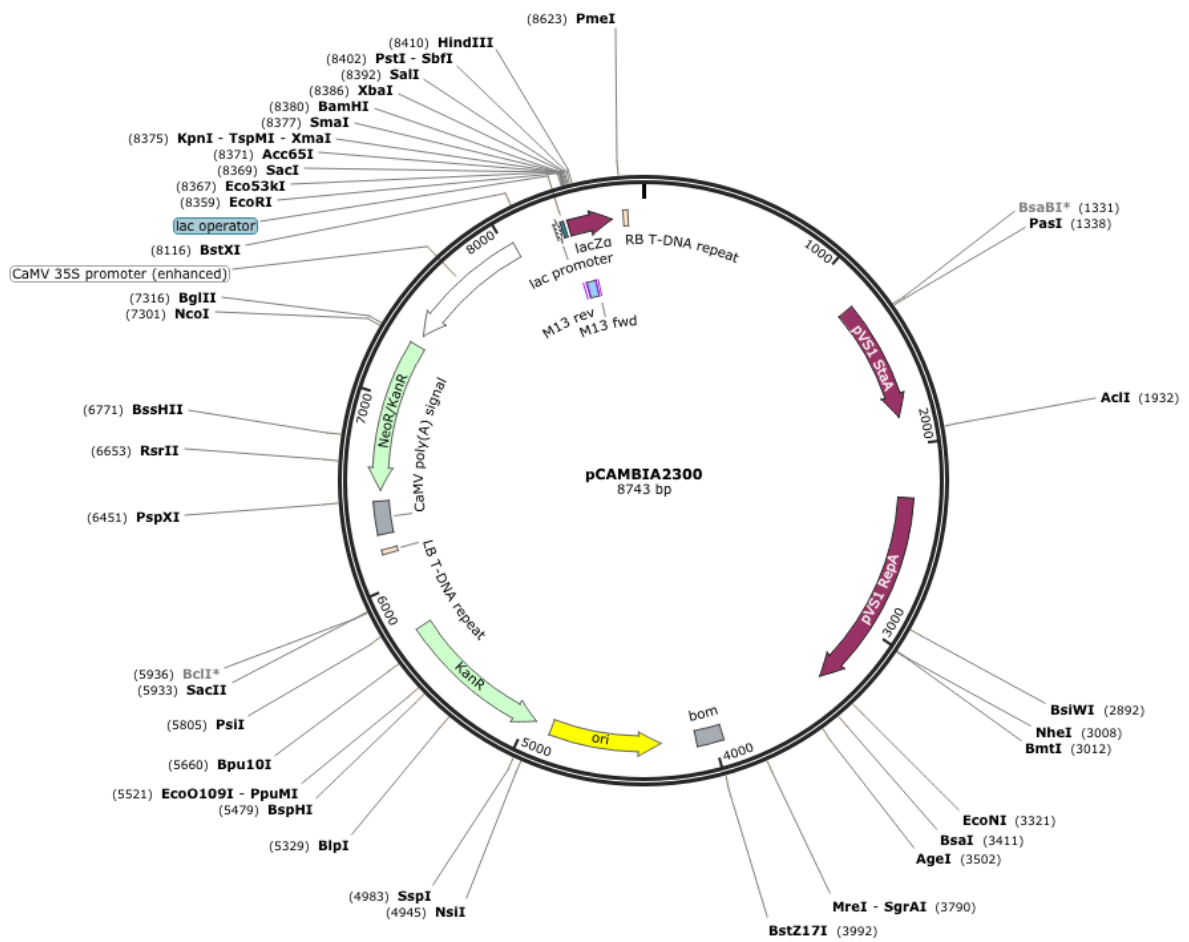
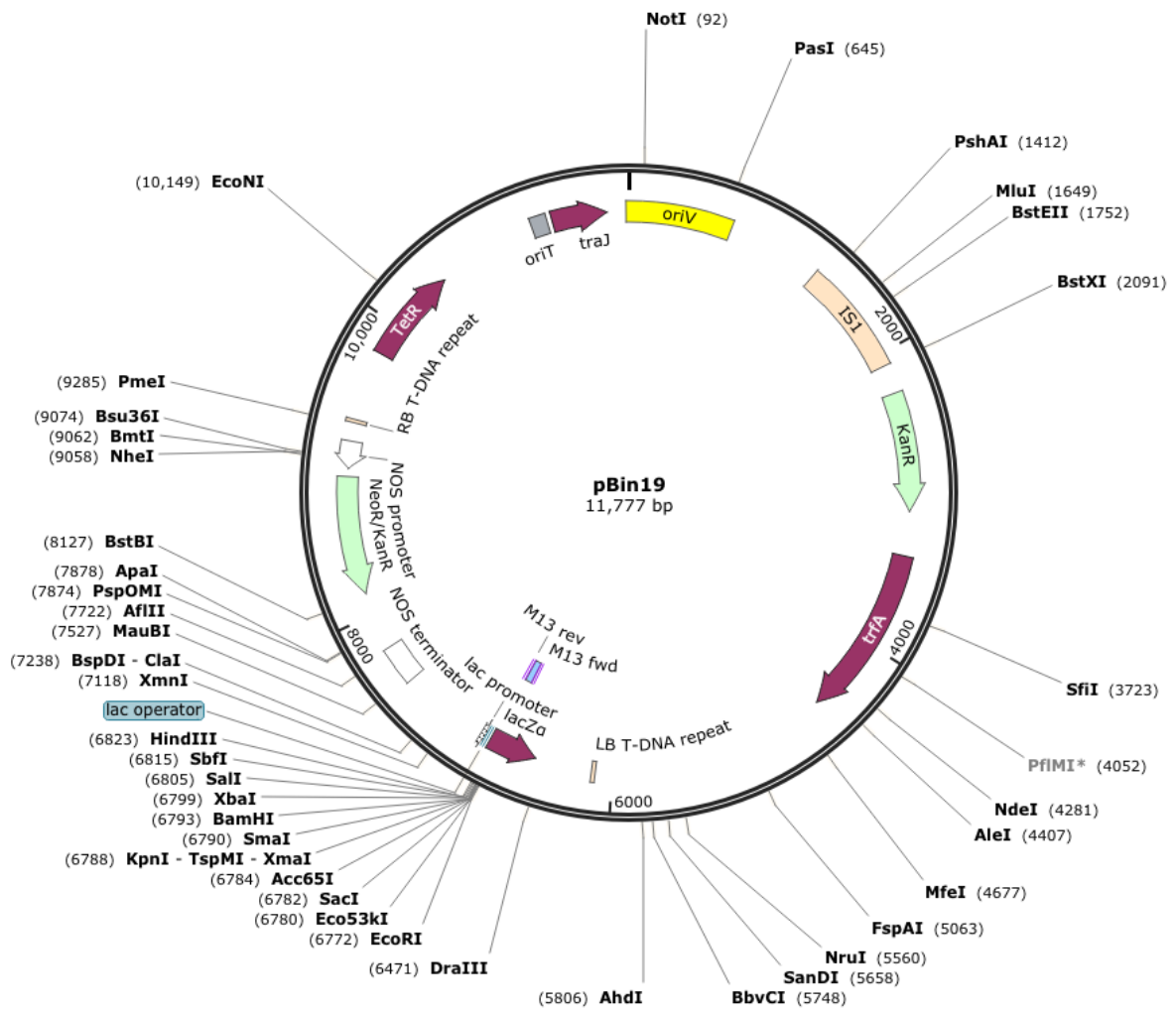


Figure 29 : Carte du plasmide pCambia2300

(Image : <http://www.snapgene.com>)



**Figure 30 : Carte du plasmide pBin19**

(Image : <http://www.snapgene.com>)

## **REFERENCES BIBLIOGRAPHIQUES**

## Références bibliographiques

---

- 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), pp. 37-46.
- 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), pp. 1045-1053.
- Ahlquist P., Dasgupta R., Kaesberg P., 1984. Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. *Journal of molecular biology*, 172 (4), pp. 369-383.
- Ali A., Li H., Schneider W.L., Sherman D.J., Gray S., Smith D., Roossinck M.J., 2006. Analysis of Genetic Bottlenecks during Horizontal Transmission of Cucumber Mosaic Virus. *Journal of Virology*, 80 (17), pp. 8345-8350.
- Anhalt M.D., Almeida R.P.P., 2008. Effect of temperature, vector life stage, and plant access period on transmission of Banana bunchy top virus to banana. *Phytopathology*, 98 (6), pp. 743-748.
- Annamalai P., Rao A.L.N., 2008. RNA encapsidation assay. *Methods in Molecular Biology (Clifton, N.J.)*, 451, pp. 251-264.
- 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), pp. 90-100.
- Bald J.G., 1950. *Measurement of concentration of plant virus suspensions. Viruses 1950.*, pp. 17-29.
- Bald J.G., 1937. The use of numbers of infections for comparing the concentration of plant virus suspensions: dilution experiments with purified suspensions. *Annals of Applied Biology*, 24 (1), pp. 33-55.
- Basnayake V.R., Sit T.L., Lommel S.A., 2006. The genomic RNA packaging scheme of Red clover necrotic mosaic virus. *Virology*, 345 (2), pp. 532-539.
- Basnayake V.R., Sit T.L., Lommel S.A., 2009. The Red clover necrotic mosaic virus origin of assembly is delimited to the RNA-2 trans-activator. *Virology*, 384 (1), pp. 169-178.
- Bergua M., Zwart M.P., El-Mohtar C., Shilts T., Elena S.F., Folimonova S.Y., 2014. A Viral Protein Mediates Superinfection Exclusion at the Whole-Organism Level but Is Not Required for Exclusion at the Cellular Level. *Journal of Virology*, 88 (19), pp. 11327-11338.



- Betancourt M., Fereres A., Fraile A., Garcia-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), pp. 12416-12421.
- Blanc S., Drucker M., Uzest M., 2014. Localizing Viruses in Their Insect Vectors. *Annual Review of Phytopathology*, 52 (1), pp. 403-425.
- Blomme G., Ploetz R., Jones D., De Langhe E., Price N., Gold C., Geering A., Viljoen A., Karamura D., Pillay M., Tinzaara W., Teycheney P.-Y., Lepoint P., Karamura E., Buddenhagen I., 2013. A historical overview of the appearance and spread of *Musa* pests and pathogens on the African continent: highlighting the importance of clean *Musa* planting materials and quarantine measures: A historical overview of the appearance and spread of *Musa* pests and pathogens in Africa. *Annals of Applied Biology*, 162 (1), pp. 4-26.
- Bonnet J., Fraile A., Sacristán S., Malpica J.M., García-Arenal F., 2005. Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus. *Virology*, 332 (1), pp. 359-368.
- Brault V., Herrbach É., Reinbold C., 2007. Electron microscopy studies on luteovirid transmission by aphids. *Micron*, 38 (3), pp. 302-312.
- Bressan A., Watanabe S., 2011. Immunofluorescence localisation of Banana bunchy top virus (family Nanoviridae) within the aphid vector, *Pentalonia nigronervosa*, suggests a virus tropism distinct from aphid-transmitted luteoviruses. *Virus Research*, 155 (2), pp. 520-525.
- Brown I.H., Harris P.A., McCauley J.W., Alexander D.J., 1998. Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. *Journal of General Virology*, 79 (12), pp. 2947-2955.
- De Bruyn A., Villemot J., Lefeuvre P., Villar E., Hoareau M., Harimalala M., Abdoul-Karime A.L., Abdou-Chakour C., Reynaud B., Harkins G.W., Varsani A., Martin D.P., Lett J.-M., 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 (1), p. 228.
- Bujarski J.J., Kaesberg P., 1986. Genetic recombination between RNA components of a multipartite plant virus. *Nature*, 321 (6069), pp. 528-531.
- Bull J.C., Godfray H.C.J., O'Reilly D.R., 2001. Persistence of an Occlusion-Negative Recombinant Nucleopolyhedrovirus in *Trichoplusia ni* Indicates High Multiplicity of Cellular Infection. *Applied and Environmental Microbiology*, 67 (11), pp. 5204-5209.
- Carluccio A.V., Zicca S., Stabolone L., 2014. Hitching a ride on vesicles: cauliflower mosaic virus movement protein trafficking in the endomembrane system. *Plant Physiology*, 164 (3), pp. 1261-1270.
- Chao L., 1988. Evolution of sex in RNA viruses. *Journal of Theoretical Biology*, 133 (1), pp. 99-112.

- Chao L., 1992. Evolution of sex in RNA viruses. *Trends in Ecology & Evolution*, 7 (5), pp. 147-151.
- Chao L., 1991. Levels of selection, evolution of sex in RNA viruses, and the origin of life. *Journal of theoretical biology*, 153 (2), pp. 229–246.
- Cheung A.K., 2006. Rolling-Circle Replication of an Animal Circovirus Genome in a Theta-Replicating Bacterial Plasmid in *Escherichia coli*. *Journal of Virology*, 80 (17), pp. 8686-8694.
- Choi Y.G., Rao A.L., 2000. Molecular studies on bromovirus capsid protein. VII. Selective packaging on BMV RNA4 by specific N-terminal arginine residuals. *Virology*, 275 (1), pp. 207-217.
- Choi Y.G., Rao A.L.N., 2003. Packaging of Brome Mosaic Virus RNA3 Is Mediated through a Bipartite Signal. *Journal of Virology*, 77 (18), pp. 9750-9757.
- Chou Y. -y., Vafabakhsh R., Doganay S., Gao Q., Ha T., Palese P., 2012. One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proceedings of the National Academy of Sciences*, 109 (23), pp. 9101-9106.
- Chu P.W.G., Helms K., 1988. Novel virus-like particles containing circular single-stranded DNAs associated with subterranean clover stunt disease. *Virology*, 167 (1), pp. 38-49.
- Cohen S., 1989. Acquisition, Interference, and Retention of Cucurbit Leaf Curl Viruses in Whiteflies. *Phytopathology*, 79 (1), p. 109.
- Comas-Garcia M., Cadena-Nava R.D., Rao A.L.N., Knobler C.M., Gelbart W.M., 2012. In Vitro Quantification of the Relative Packaging Efficiencies of Single-Stranded RNA Molecules by Viral Capsid Protein. *Journal of Virology*, 86 (22), pp. 12271-12282.
- Czosnek H., Ghanim M., Morin S., Rubinstein G., Fridman V., Zeidan M., 2001. Whiteflies: Vectors, and victims (?), of geminiviruses. In : *Advances in Virus Research*. Elsevier, pp. 291-322. vol.57.
- Delatte H., Reynaud B., Granier M., Thornary L., Lett J.M., Goldbach R., Peterschmitt M., 2005. A new silverleaf-inducing biotype Ms of *Bemisia tabaci* (Hemiptera: Aleyrodidae) indigenous to the islands of the south-west Indian Ocean. *Bulletin of Entomological Research*, 95 (01).
- DíAz-Pendón J.A., CañIzares M.C., Moriones E., Bejarano E.R., Czosnek H., Navas-Castillo J., 2010. Tomato yellow leaf curl viruses: ménage à trois between the virus complex, the plant and the whitefly vector: Tomato yellow leaf curl viruses. *Molecular Plant Pathology*, 11 (4), pp. 441-450.
- Ding B., 2010. Viroids: self-replicating, mobile, and fast-evolving noncoding regulatory RNAs. *Wiley Interdisciplinary Reviews: RNA*, 1 (3), pp. 362-375.
- Drake J.W., 1993. Rates of spontaneous mutation among RNA viruses. *Proceedings of the National Academy of Sciences*, 90 (9), pp. 4171–4175.

- Druett H.A., 1952. Bacterial invasion. *Nature*, 170 (4320), p. 288.
- Duffy S., Holmes E.C., 2009. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *Journal of General Virology*, 90 (6), pp. 1539-1547.
- Edwards K., Johnstone C., Thompson C., 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic acids research*, 19 (6), p. 1349.
- Elde N.C., Child S.J., Eickbush M.T., Kitzman J.O., Rogers K.S., Shendure J., Geballe A.P., Malik H.S., 2012. Poxviruses Deploy Genomic Accordions to Adapt Rapidly against Host Antiviral Defenses. *Cell*, 150 (4), pp. 831-841.
- Feng J.-L., Chen S.-N., Tang X.-S., Ding X.-F., Du Z.-Y., Chen J.-S., 2006. Quantitative Determination of Cucumber Mosaic Virus Genome RNAs in Virions by Real-Time Reverse Transcription-Polymerase Chain Reaction. *Acta Biochimica et Biophysica Sinica*, 38 (10), pp. 669-676.
- Feng J., Zeng R., Chen J., 2008. Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. *BioTechniques*, 44 (7), pp. 901-912.
- Filée J., 2013. Route of NCLDV evolution: the genomic accordion. *Current Opinion in Virology*, 3 (5), pp. 595-599.
- Flores R., Serra P., Minoia S., Di Serio F., Navarro B., 2012. Viroids: From Genotype to Phenotype Just Relying on RNA Sequence and Structural Motifs. *Frontiers in Microbiology*, 3.
- Fontes E.P., Gladfelter H.J., Schaffer R.L., Petty I.T., Hanley-Bowdoin L., 1994. Geminivirus replication origins have a modular organization. *The Plant Cell*, 6 (3), pp. 405-416.
- Franz A., Makkouk K.M., Vetten H.J., 1998. Acquisition, retention and transmission of faba bean necrotic yellows virus by two of its aphid vectors, *Aphis craccivora* (Koch) and *Acyrtosiphon pisum* (Harris). *Journal of Phytopathology*, 146 (7), pp. 347-355.
- Franz A.W., van der Wilk F., Verbeek M., Dulleman 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), pp. 210-219.
- Froissart R., Roze D., Uzest M., Galibert L., Blanc S., Michalakis Y., 2005. Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. *PLoS biology*, 3 (3), p. e89.
- Fulton R.W., 1962. The effect of dilution on necrotic ringspot virus infectivity and the enhancement of infectivity by noninfective virus. *Virology*, 18 (3), pp. 477-485.
- Gallagher K.L., Sozzani R., Lee C.-M., 2014. Intercellular Protein Movement: Deciphering the Language of Development. *Annual Review of Cell and Developmental Biology*, 30 (1), pp. 207-233.

- Garcia-Arriaza J., Manrubia S.C., Toja M., Domingo E., Escarmis C., 2004. Evolutionary Transition toward Defective RNAs That Are Infectious by Complementation. *Journal of Virology*, 78 (21), pp. 11678-11685.
- Garret A., Kerlan C., Thomas D., 1996. Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, *Myzus persicae* Sulz. *Archives of Virology*, 141 (7), pp. 1279-1292.
- Ghanim M., Morin S., Czosnek H., 2001. Rate of Tomato yellow leaf curl virus translocation in the circulative transmission pathway of its vector, the whitefly *Bemisia tabaci*. *Phytopathology*, 91 (2), pp. 188-196.
- Gildow F.E., 1993. Evidence for Receptor-Mediated Endocytosis Regulating Luteovirus Acquisition by Aphids. *Phytopathology*, 83 (3), p. 270.
- Gladfelter H.J., Eagle P.A., Fontes E.P.B., Batts L., Hanley-Bowdoin L., 1997. Two Domains of the AL1 Protein Mediate Geminivirus Origin Recognition. *Virology*, 239 (1), pp. 186-197.
- Godefroy-Colburn T., THIVENT C., PINCK L., 1985. Translational discrimination between the four RNAs of alfalfa mosaic virus. *European journal of biochemistry*, 147 (3), pp. 541-548.
- Goldhill D.H., Turner P.E., 2014. The evolution of life history trade-offs in viruses. *Current Opinion in Virology*, 8C , pp. 79-84.
- Gonzalez-Jara P., Fraile A., Canto T., Garcia-Arenal F., 2009. The Multiplicity of Infection of a Plant Virus Varies during Colonization of Its Eukaryotic Host. *Journal of Virology*, 83 (15), pp. 7487-7494.
- Gopinath K., Kao C.C., 2007. Replication-independent long-distance trafficking by viral RNAs in *Nicotiana benthamiana*. *The Plant Cell*, 19 (4), pp. 1179-1191.
- Gray S., Gildow F.E., 2003. L UTEOVIRUS -A PHID I NTERACTIONS \*. *Annual Review of Phytopathology*, 41 (1), pp. 539-566.
- Grigoras I., Ginzo A.I. d. C., Martin D.P., Varsani A., Romero J., Mammadov A.C., Huseynova I.M., Aliyev J.A., Khey-Pour A., Huss H., Ziebell H., Timchenko T., Vetten H.-J., Gronenborn B., 2014. Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *Journal of General Virology*, 95 (Pt\_5), pp. 1178-1191.
- Grigoras I., Gronenborn B., Vetten H.J., 2010. First Report of a Nanovirus Disease of Pea in Germany. *Plant Disease*, 94 (5), pp. 642-642.
- Grigoras I., Timchenko T., Grande-Perez A., Katul L., Vetten H.-J., Gronenborn B., 2010. High Variability and Rapid Evolution of a Nanovirus. *Journal of Virology*, 84 (18), pp. 9105-9117.

- Grigoras I., Timchenko T., Katul L., Grande-Perez A., Vetten H.-J., Gronenborn B., 2009. Reconstitution of Authentic Nanovirus from Multiple Cloned DNAs. *Journal of Virology*, 83 (20), pp. 10778-10787.
- Gronenborn B., 2004. Nanoviruses: genome organisation and protein function. *Veterinary Microbiology*, 98 (2), pp. 103-109.
- Grylls N., Butler F., 1959. Subterranean clover stunt, a virus disease of pasture legumes. *Australian Journal of Agricultural Research*, 10 (2), p. 145.
- Gutiérrez S., Michalakis Y., Blanc S., 2012a. Virus population bottlenecks during within-host progression and host-to-host transmission. *Current Opinion in Virology*, 2 (5), pp. 546-555.
- Gutiérrez S., Michalakis Y., Blanc S., 2012b. Virus population bottlenecks during within-host progression and host-to-host transmission. *Current Opinion in Virology*, 2 (5), pp. 546-555.
- Gutiérrez S., Yvon M., Thébaud G., Monsion B., Michalakis Y., Blanc S., 2010a. Dynamics of the Multiplicity of Cellular Infection in a Plant Virus. *PLoS Pathogens*, 6 (9), p. e1001113.
- Gutiérrez S., Yvon M., Thébaud G., Monsion B., Michalakis Y., Blanc S., 2010b. Dynamics of the Multiplicity of Cellular Infection in a Plant Virus. *PLoS Pathogens*, 6 (9), p. e1001113.
- Habili N., Francki R.I., 1974. Comparative studies on tomato aspermy and cucumber mosaic viruses. *Virology*, 60 (1), pp. 29-36.
- Hafke J.B., Furch A.C.U., Reitz M.U., van Bel A.J.E., 2007. Functional Sieve Element Protoplasts. *PLANT PHYSIOLOGY*, 145 (3), pp. 703-711.
- Hafner G.J., Harding R.M., Dale J.L., 1997. A DNA primer associated with banana bunchy top virus. *Journal of General Virology*, 78 (2), pp. 479-486.
- Hafner G.J., Harding R.M., Dale J.L., 1995. Movement and transmission of banana bunchy top virus DNA component one in bananas. *Journal of General Virology*, 76 (9), pp. 2279-2285.
- Harding R.M., Burns T.M., Dale J.L., 1991. Virus-like particles associated with banana bunchy top disease contain small single-stranded DNA. *The Journal of General Virology*, 72 (Pt 2), pp. 225-230.
- Hastings P.J., Lupski J.R., Rosenberg S.M., Ira G., 2009. Mechanisms of change in gene copy number. *Nature Reviews Genetics*, 10 (8), pp. 551-564.
- Hayakawa T., Kojima K., Nonaka K., Nakagaki M., Sahara K., Asano S., Iizuka T., Bando H., 2000. Analysis of proteins encoded in the bipartite genome of a new type of parvo-like virus isolated from silkworm—structural protein with DNA polymerase motif. *Virus research*, 66 (1), pp. 101-108.

- Hicks A.L., Duffy S., 2014. Cell Tropism Predicts Long-term Nucleotide Substitution Rates of Mammalian RNA Viruses. *PLoS Pathogens*, 10 (1), p. e1003838.
- Hipper C., Brault V., Ziegler-Graff V., Revers F., 2013. Viral and Cellular Factors Involved in Phloem Transport of Plant Viruses. *Frontiers in Plant Science*, 4.
- Hogenhout S.A., Ammar E.-D., Whitfield A.E., Redinbaugh M.G., 2008. Insect Vector Interactions with Persistently Transmitted Viruses\*. *Annual Review of Phytopathology*, 46 (1), pp. 327-359.
- Hollox E.J., Hoh B.-P., 2014. Human gene copy number variation and infectious disease. *Human Genetics*, 133 (10), pp. 1217-1233.
- Horser C.L., Karan M., Harding R.M., Dale J.L., 2001. Additional Rep-encoding DNAs associated with banana bunchy top virus. *Archives of virology*, 146 (1), pp. 71–86.
- Huang I.-C., Li W., Sui J., Marasco W., Choe H., Farzan M., 2008. Influenza A Virus Neuraminidase Limits Viral Superinfection. *Journal of Virology*, 82 (10), pp. 4834-4843.
- Hughes A.L., 2004. Birth-and-death evolution of protein-coding regions and concerted evolution of non-coding regions in the multi-component genomes of nanoviruses. *Molecular Phylogenetics and Evolution*, 30 (2), pp. 287-294.
- Hu J.-M., Fu H.-C., Lin C.-H., Su H.-J., Yeh H.-H., 2007. Reassortment and Concerted Evolution in Banana Bunchy Top Virus Genomes. *Journal of Virology*, 81 (4), pp. 1746-1761.
- Hull R. (Éd.), 2014. *Plant virology*. Fifth edition. Amsterdam ; Boston, Elsevier/AP, Academic Press is an imprint of Elsevier, 1104 p.
- Ince W.L., Gueye-Mbaye A., Bennink J.R., Yewdell J.W., 2013. Reassortment Complements Spontaneous Mutation in Influenza A Virus NP and M1 Genes To Accelerate Adaptation to a New Host. *Journal of Virology*, 87 (8), pp. 4330-4338.
- Iranzo J., Manrubia S.C., 2012. Evolutionary dynamics of genome segmentation in multipartite viruses. *Proceedings of the Royal Society B: Biological Sciences*, 279 (1743), pp. 3812-3819.
- Jaspars E.M.J., 1975. Plant viruses with a multipartite genome. *Advances in virus research*, 19, pp. 37-149.
- Josefsson L., King M.S., Makitalo B., Brannstrom J., Shao W., Maldarelli F., Kearney M.F., Hu W.-S., Chen J., Gaines H., Mellors J.W., Albert J., Coffin J.M., Palmer S.E., 2011. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proceedings of the National Academy of Sciences*, 108 (27), pp. 11199-11204.
- Jung A., Maier R., Vartanian J.-P., Bocharov G., Jung V., Fischer U., Meese E., Wain-Hobson S., Meyerhans A., 2002. Recombination: Multiply infected spleen cells in HIV patients. *Nature*, 418 (6894), pp. 144–144.

- Van Kammen A., 1968. The relationship between the components of cowpea mosaic virus. *Virology*, 34 (2), pp. 312-318.
- Katju V., Bergthorsson U., 2013. Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Frontiers in Genetics*, 4.
- Katul U., Vetten H.J., Maiss E., Makkouk K.M., Lesemann D.-E., Casper R., 1993. Characterisation and serology of virus-like particles associated with faba bean necrotic yellows. *Annals of Applied Biology*, 123 (3), pp. 629–647.
- King R.W., Zeevaart J.A.D., 1974. Enhancement of phloem exudation from cut petioles by chelating agents. *Plant Physiology*, 53 (1), pp. 96–103.
- Kumari S.G., Rodoni B., Vetten H.-J., Loh M.H., Freeman A., Van Leur J., Bao S., Wang X., 2010. Detection and Partial Characterization of Milk vetch dwarf virus Isolates from Faba Bean (*Vicia faba* L.) in Yunnan Province, China: Milk vetch dwarf virus in China. *Journal of Phytopathology*, 158 (1), pp. 35-39.
- Kuroda T., Natsuaki T., Okuda S., Teranaka M., 2005. Analysis of natural populations and possible natural reassortment of Cucumber mosaic virus. *Journal of General Plant Pathology*, 71 (3), pp. 221-229.
- Kwon C.S., Chung W.-I., 2000. Differential roles of the 5' untranslated regions of cucumber mosaic virus RNAs 1, 2, 3 and 4 in translational competition. *Virus research*, 66 (2), pp. 175–185.
- Kwon S.-J., Seo J.-K., Rao A.L.N., 2012. Detection and Quantification of Viral and Satellite RNAs in Plant Hosts. In: Jin H., Gassmann W. (Éd.), *RNA Abundance Analysis*. Totowa, NJ, Humana Press, pp. 131-141. vol.883.
- Lafforgue G., Tromas N., Elena S.F., Zwart M.P., 2012. Dynamics of the Establishment of Systemic Potyvirus Infection: Independent yet Cumulative Action of Primary Infection Sites. *Journal of Virology*, 86 (23), pp. 12912-12922.
- Lee Y.-M., Tscherne D.M., Yun S.-I., Frolov I., Rice C.M., 2005. Dual Mechanisms of Pestiviral Superinfection Exclusion at Entry and RNA Replication. *Journal of Virology*, 79 (6), pp. 3231-3242.
- Lin H.-X., Rubio L., Smythe A.B., Falk B.W., 2004. Molecular Population Genetics of Cucumber Mosaic Virus in California: Evidence for Founder Effects and Reassortment. *Journal of Virology*, 78 (12), pp. 6666-6675.
- Lister R.M., 1968. Functional Relationships between Virus-specific Products of Infection by Viruses of the Tobacco Rattle Type. *Journal of General Virology*, 2 (1), pp. 43-58.
- López-Bueno A., Villarreal L.P., Almendral J.M., 2006. Parvovirus variation for disease: a difference with RNA viruses? *Current Topics in Microbiology and Immunology*, 299 , pp. 349-370.
- Lot H., Kaper J.M., 1976. Further studies on the RNA component distribution among the nucleoproteins of cucumber mosaic virus. *Virology*, 74 (1), pp. 223-226.

- Lot H., Marchoux G., Marrou J., Kaper J.M., West C.K., Van Vloten-Doting L., Hull R., 1974. Evidence for three functional RNA species in several strains of cucumber mosaic virus. *Journal of General Virology*, 22 (1), pp. 81–93.
- McDonald S.M., Patton J.T., 2011. Assortment and packaging of the segmented rotavirus genome. *Trends in Microbiology*, 19 (3), pp. 136–144.
- Medina R.A., García-Sastre A., 2011. Influenza A viruses: new research developments. *Nature Reviews Microbiology*, 9 (8), pp. 590–603.
- Mehta P., Wyman J.A., Nakhla M.K., Maxwell D.P., 1994. Transmission of Tomato Yellow Leaf Curl Geminiviruses by *Bemisia tabaci* (Homoptera: Aleyrodidae). *Journal of Economic Entomology*, 87 (5), pp. 1291–1297.
- Melcher U., 2000. The ‘30K’ superfamily of viral movement proteins. *Journal of General Virology*, 81 (1), pp. 257–266.
- Mileyko Y., Joh R.I., Weitz J.S., 2008. Small-scale copy number variation and large-scale changes in gene expression. *Proceedings of the National Academy of Sciences*, 105 (43), pp. 16659–16664.
- Mills D.R., Peterson R.L., Spiegelman S., 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 58 (1), pp. 217–224.
- Mindich L., 2004. Packaging, replication and recombination of the segmented genomes of bacteriophage  $\Phi 6$  and its relatives. *Virus Research*, 101 (1), pp. 83–92.
- 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), pp. 1828–1837.
- Monsion B., Froissart R., Michalakis Y., Blanc S., 2008. Large Bottleneck Size in Cauliflower Mosaic Virus Populations during Host Plant Colonization. *PLoS Pathogens*, 4 (10), p. e1000174.
- 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*, 104 (45), pp. 17891–17896.
- Muhire B.M., Golden M., Murrell B., Lefeuvre P., Lett J.-M., Gray A., Poon A.Y.F., Ngandu N.K., Semegni Y., Tanov E.P., Monjane A.L., Harkins G.W., Varsani A., Shepherd D.N., Martin D.P., 2014. Evidence of Pervasive Biologically Functional Secondary Structures within the Genomes of Eukaryotic Single-Stranded DNA Viruses. *Journal of Virology*, 88 (4), pp. 1972–1989.
- Nee S., 1988. Deleterious mutation and genetic recombination. *Nature*, 331 (6154), p. 308.
- Nee S., 1987. The evolution of multicompartmental genomes in viruses. *Journal of molecular evolution*, 25 (4), pp. 277–281.



- Nethe M., Berkhout B., Van der Kuyl A.C., 2005. Retroviral superinfection resistance. *Retrovirology*, 2 (1), p. 52.
- Neumann G., Noda T., Kawaoka Y., 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature*, 459 (7249), pp. 931-939.
- Ni P., Vaughan R.C., Tragesser B., Hoover H., Kao C.C., 2014. The Plant Host Can Affect the Encapsidation of Brome Mosaic Virus (BMV) RNA: BMV Virions Are Surprisingly Heterogeneous. *Journal of Molecular Biology*, 426 (5), pp. 1061-1076.
- Noda T., Kawaoka Y., 2012. Packaging of influenza virus genome: Robustness of selection. *Proceedings of the National Academy of Sciences*, 109 (23), pp. 8797-8798.
- Nouri S., Arevalo R., Falk B.W., Groves R.L., 2014. Genetic Structure and Molecular Variability of Cucumber mosaic virus Isolates in the United States. *PLoS ONE*, 9 (5), p. e96582.
- Ojosnegros S., García-Arriaza J., Escarmís C., Manrubia S.C., Perales C., Arias A., Mateu M.G., Domingo E., 2011. Viral Genome Segmentation Can Result from a Trade-Off between Genetic Content and Particle Stability. *PLoS Genetics*, 7 (3), p. e1001344.
- De Paepe M., Taddei F., 2006. Viruses' Life History: Towards a Mechanistic Basis of a Trade-Off between Survival and Reproduction among Phages. *PLoS Biology*, 4 (7), p. e193.
- Palukaitis P., 2014. What has been happening with viroids? *Virus Genes*, 49 (2), pp. 175-184.
- Pérefarres F., Thébaud G., Lefeuvre P., Chiroleu F., Rimbaud L., Hoareau M., Reynaud B., Lett J.-M., 2014. Frequency-dependent assistance as a way out of competitive exclusion between two strains of an emerging virus. *Proceedings. Biological Sciences / The Royal Society*, 281 (1781), p. 20133374.
- Pierrugues O., Guilbaud L., Fernandez-Delmond I., Fabre F., Tepfer M., Jacquemond M., 2007. Biological properties and relative fitness of inter-subgroup cucumber mosaic virus RNA 3 recombinants produced in vitro. *The Journal of General Virology*, 88 (Pt 10), pp. 2852-2861.
- 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), pp. 264-268.
- Poranen M.M., Tuma R., 2004. Self-assembly of double-stranded RNA bacteriophages. *Virus Research*, 101 (1), pp. 93-100.
- Pouwels J., 2003. Identification of distinct steps during tubule formation by the movement protein of Cowpea mosaic virus. *Journal of General Virology*, 84 (12), pp. 3485-3494.
- Pressing J., Reanney D.C., 1984. Divided genomes and intrinsic noise. *Journal of molecular evolution*, 20 (2), pp. 135-146.
- Qiu W., Moyer J.W., 1999. Tomato Spotted Wilt Tospovirus Adapts to the TSWV N Gene-Derived Resistance by Genome Reassortment. *Phytopathology*, 89 (7), pp. 575-582.

- Rager M., Vongpunsawad S., Duprex W.P., Cattaneo R., 2002. Polyploid measles virus with hexameric genome length. *The EMBO journal*, 21 (10), pp. 2364–2372.
- Reguera J., Malet H., Weber F., Cusack S., 2013. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proceedings of the National Academy of Sciences*, 110 (18), pp. 7246–7251.
- Ren X., Tao Y., Cui J., Suo S., Cong Y., Tijssen P., 2013. Phylogeny and evolution of porcine parvovirus. *Virus Research*, 178 (2), pp. 392–397.
- Ribi re M., Olivier V., Blanchard P., 2010. Chronic bee paralysis: A disease and a virus like no other? *Journal of Invertebrate Pathology*, 103 , pp. S120–S131.
- Ritzenthaler C., 2011. Parallels and distinctions in the direct cell-to-cell spread of the plant and animal viruses. *Current Opinion in Virology*, 1 (5), pp. 403–409.
- Roossinck M.J., 2002. Evolutionary History of Cucumber Mosaic Virus Deduced by Phylogenetic Analyses. *Journal of Virology*, 76 (7), pp. 3382–3387.
- 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), pp. e45–e45.
- Sakai Y., Kiyotani K., Fukumura M., Asakawa M., Kato A., Shioda T., Yoshida T., Tanaka A., Hasegawa M., Nagai Y., 1999. Accommodation of foreign genes into the Sendai virus genome: sizes of inserted genes and viral replication. *FEBS letters*, 456 (2), pp. 221–226.
- Sanchez-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), pp. 10805–10815.
- Sanjuan R., Nebot M.R., Chirico N., Mansky L.M., Belshaw R., 2010. Viral Mutation Rates. *Journal of Virology*, 84 (19), pp. 9733–9748.
- Sarker S., Patterson E.I., Peters A., Baker G.B., Forwood J.K., Ghorashi S.A., Holdsworth M., Baker R., Murray N., Raidal S.R., 2014. Mutability Dynamics of an Emergent Single Stranded DNA Virus in a Na ive Host. *PLoS ONE*, 9 (1), p. e85370.
- Sarkies P., Miska E.A., 2014. Small RNAs break out: the molecular cell biology of mobile small RNAs. *Nature Reviews Molecular Cell Biology*, 15 (8), pp. 525–535.
- Savory F.R., Ramakrishnan U., 2014. Asymmetric patterns of reassortment and concerted evolution in Cardamom bushy dwarf virus. *Infection, Genetics and Evolution*, 24 , pp. 15–24.
- Scholthof H.B., 2005. Plant virus transport: motions of functional equivalence. *Trends in Plant Science*, 10 (8), pp. 376–382.

- Scholthof H.B., Scholthof K.-B.G., Kikkert M., Jackson A.O., 1995. Tomato bushy stunt virus spread is regulated by two nested genes that function in cell-to-cell movement and host-dependent systemic invasion. *Virology*, 213 (2), pp. 425–438.
- Shirasawa-Seo N., 2005. Characteristics of the promoters derived from the single-stranded DNA components of Milk vetch dwarf virus in transgenic tobacco. *Journal of General Virology*, 86 (6), pp. 1851-1860.
- 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 .
- Simon-Loriere E., Holmes E.C., 2011. Why do RNA viruses recombine? *Nature Reviews Microbiology*, 9 (8), pp. 617-626.
- Sinisterra X.H., 2005. Differential transcriptional activity of plant-pathogenic begomoviruses in their whitefly vector (*Bemisia tabaci*, Gennadius: Hemiptera Aleyrodidae). *Journal of General Virology*, 86 (5), pp. 1525-1532.
- Sivakumaran K., Bao Y., Roossinck M.J., Kao C.C., 2000. Recognition of the Core RNA Promoter for Minus-Strand RNA Synthesis by the Replicases of Brome Mosaic Virus and Cucumber Mosaic Virus. *Journal of Virology*, 74 (22), pp. 10323-10331.
- Solovyev A.G., Savenkov E.I., 2014. Factors involved in the systemic transport of plant RNA viruses: the emerging role of the nucleus. *Journal of Experimental Botany*, 65 (7), pp. 1689-1697.
- Spiegelman S., Mills D.R., Kramer F.R., 1975. *The extracellular evolution of structure in replicating RNA molecules. Stability and origin of biological information.*
- Spiegelman Z., Golan G., Wolf S., 2013. Don't kill the messenger: Long-distance trafficking of mRNA molecules. *Plant Science*, 213 , pp. 1-8.
- Stace-Smith R., Tremaine J.H., 1973. Biophysical and biochemical properties of tomato aspermy virus. *Virology*, 51 (2), pp. 401-408.
- Stainton D., Kraberger S., Walters M., Wiltshire E.J., Rosario K., Halafihi M., Lolohea S., Katoa I., Faitua T.H., Aholelei W., Taufa L., Thomas J.E., Collings D.A., Martin D.P., Varsani A., 2012. Evidence of inter-component recombination, intra-component recombination and reassortment in banana bunchy top virus. *Journal of General Virology*, 93 (Pt\_5), pp. 1103-1119. [Consulté le 07/08/2014].
- Stranger B.E., Forrest M.S., Dunning M., Ingle C.E., Beazley C., Thorne N., Redon R., Bird C.P., de Grassi A., Lee C., Tyler-Smith C., Carter N., Scherer S.W., Tavare S., Deloukas P., Hurles M.E., Dermitzakis E.T., 2007. Relative Impact of Nucleotide and Copy Number Variation on Gene Expression Phenotypes. *Science*, 315 (5813), pp. 848-853. [Consulté le 12/10/2014].

- Takeshita M., Shigemune N., Kikuhara K., Furuya N., Takanami Y., 2004. Spatial analysis for exclusive interactions between subgroups I and II of Cucumber mosaic virus in cowpea. *Virology*, 328 (1), pp. 45-51. [Consulté le 08/08/2014].
- Tilsner J., Oparka K.J., 2012. Missing links? - The connection between replication and movement of plant RNA viruses. *Current Opinion in Virology*, 2 (6), pp. 705-711.
- Timchenko T., 2006. Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus. *Journal of General Virology*, 87 (6), pp. 1735-1743.
- Timchenko T., Bernardi F., 2007. *Les nanovirus, petits virus de plantes : similitudes et différences avec les géminivirus*.
- 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), pp. 189-195.
- 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), pp. 10173-10182. [Consulté le 07/08/2014].
- Tromas N., Zwart M.P., Lafforgue G., Elena S.F., 2014. Within-Host Spatiotemporal Dynamics of Plant Virus Infection at the Cellular Level. *PLoS Genetics*, 10 (2), p. e1004186.
- Turgeon R., Wolf S., 2009. Phloem Transport: Cellular Pathways and Molecular Trafficking. *Annual Review of Plant Biology*, 60 (1), pp. 207-221.
- Turner P.E., Burch C.L., Hanley K.A., Chao L., 1999. Hybrid frequencies confirm limit to coinfection in the RNA bacteriophage phi6. *Journal of Virology*, 73 (3), pp. 2420-2424.
- Ueki S., Citovsky V., 2011. To Gate, or Not to Gate: Regulatory Mechanisms for Intercellular Protein Transport and Virus Movement in Plants. *Molecular Plant*, 4 (5), pp. 782-793.
- Urbanowicz A., Alejska M., Formanowicz P., Blazewicz J., Figlerowicz M., Bujarski J.J., 2005. Homologous Crossovers among Molecules of Brome Mosaic Bromovirus RNA1 or RNA2 Segments In Vivo. *Journal of Virology*, 79 (9), pp. 5732-5742.
- 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 (1), p. 135.
- Vaughan R., Tragesser B., Ni P., Ma X., Dragnea B., Kao C.C., 2014. The Tripartite Virions of the Brome Mosaic Virus Have Distinct Physical Properties That Affect the Timing of the Infection Process. *Journal of Virology*, 88 (11), pp. 6483-6491.
- Van Vloten-Doting L., Jaspars E.M.J., 1977. *Plant covirus system: three component systems*. *Comprehensive Virology 11*, pp. 1-53.

- Van der Walt E., Martin D.P., Varsani A., Polston J.E., Rybicki E.P., 2008. Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. *Virology Journal*, 5 (1), p. 104.
- Wang Y.J., Yao Q., Chen K.P., Wang Y., Lu J., Han X., 2007. Characterization of the genome structure of Bombyx mori densovirus (China isolate). *Virus Genes*, 35 (1), pp. 103-108.
- Watanabe S., Bressan A., 2013. Tropism, compartmentalization and retention of banana bunchy top virus (Nanoviridae) in the aphid vector Pentalonia nigronervosa. *Journal of General Virology*, 94 (Pt\_1), pp. 209-219.
- Wetering F., 1996. Tomato Spotted Wilt Tospovirus Ingestion by First Instar Larvae of *Frankliniella occidentalis* Is a Prerequisite for Transmission. *Phytopathology*, 86 (9), p. 900.
- Wichgers Schreur P.J., Oreshkova N., Moormann R.J.M., Kortekaas J., 2014. Creation of Rift Valley Fever Viruses with Four-Segmented Genomes Reveals Flexibility in Bunyavirus Genome Packaging. *Journal of Virology*, 88 (18), pp. 10883-10893.
- Wright S., 1931. Evolution in Mendelian Populations. *Genetics*, 16 (2), pp. 97-159.
- Yang X., Zhou M., Qian Y., Xie Y., Zhou X., 2014. Molecular variability and evolution of a natural population of tomato yellow leaf curl virus in Shanghai, China. *Journal of Zhejiang University SCIENCE B*, 15 (2), pp. 133-142.
- Zagórski W., 1978. Translational Regulation of Expression of the Brome-Mosaic-Virus RNA Genome in vitro. *European Journal of Biochemistry*, 86 (2), pp. 465-472.
- Zaitlin M., Palukaitis P., 2000. Advances in understanding plant viruses and virus diseases. *Annual Review of Phytopathology*, 38 (1), pp. 117-143.
- Zandi R., van der Schoot P., 2009. Size Regulation of ss-RNA Viruses. *Biophysical Journal*, 96 (1), pp. 9-20.
- Zwart M.P., Daròs J.-A., Elena S.F., 2011. One Is Enough: In Vivo Effective Population Size Is Dose-Dependent for a Plant RNA Virus. *PLoS Pathogens*, 7 (7), p. e1002122.
- Zwart M.P., Tromas N., Elena S.F., 2013. Model-Selection-Based Approach for Calculating Cellular Multiplicity of Infection during Virus Colonization of Multi-Cellular Hosts. *PLoS ONE*, 8 (5), p. e64657.



## Résumé

Les virus multipartites sont caractérisés par un génome divisé en au moins deux segments, chacun étant encapsidé individuellement. Bien que ces virus représentent près de 30% des genres de virus de plantes connus, les coûts et avantages engendrés par cette organisation génomique restent mal connus. Basée sur l'étude d'un *Nanovirus*, le *Faba bean necrotic stunt virus* (FBNSV), ma thèse vise à mieux comprendre les dynamiques des populations d'un virus multipartite au cours de l'infection de son hôte et de sa transmission par pucerons.

Dans un premier temps, nous avons démontré que chacun des huit segments composant le génome du FBNSV s'accumulait de manière reproductible à une fréquence relative spécifique, certains segments représentant près de 30% de la population virale au sein de la fève et d'autres n'excédant pas 2%. En changeant de plante hôte, nous avons pu montrer que les fréquences relatives des différents segments étaient hôte-spécifiques. Nous avons par ailleurs observé des changements de fréquences relatives au sein des pucerons vecteurs, changements révélant des interactions intimes entre le FBNSV –précédemment perçu comme traversant simplement les barrières cellulaires des pucerons- et son vecteur.

Cette plasticité et l'ajustement correspondant de la composition génétique de la population virale au sein de l'hôte pourrait permettre aux virus multipartites de s'adapter rapidement à un nouvel environnement grâce à un changement spécifique du nombre de copies de ses gènes. Cependant, ces différences de fréquences relatives pourraient également constituer un coût additionnel pour le virus si, comme il est généralement admis, tous les segments ont besoin de pénétrer au sein de la même cellule pour garantir l'infection. En tenant compte des segments rares au sein des populations virales, ce coût apparaît énorme et nous avons donc décidé de vérifier l'existence d'un tel coût expérimentalement.

Une étude de la distribution des différents segments génomiques au sein des cellules végétales est donc en cours et ce, dans le but de regarder si les huit segments sont toujours ensemble au sein des cellules ou s'ils peuvent être physiquement séparés. Cette question est d'importance majeure non seulement en ce qu'elle permet de réévaluer le coût potentiel des virus multipartites mais aussi en ce qu'elle permet de tester le concept du cycle de réplication « cellule-autonome » chez les virus.

Enfin, le maintien de l'intégrité génomique des virus multipartites reposant sur de relativement faibles goulots d'étranglement tout au long du cycle viral, des outils ont été mis en place pour quantifier ces goulots d'étranglement.

## Abstract

Multipartite viruses are characterized by a genome divided into two or more segments, each encapsidated individually into separated virus particles. Although they represent more than 30% of the known plant virus genera, little is known on how such viral systems actually work and on the advantage(s) and cost(s) of their peculiar genome organization. Based on the study of a *Nanovirus*, the *Faba bean necrotic stunt virus* (FBNSV), my thesis aimed at better understanding the dynamics of a multipartite virus population during within host disease development and during aphid transmission. We have first demonstrated that each of the eight segments composing the FBNSV genome reproducibly accumulates at a specific relative frequency, some segments representing up to 30 % of the total viral DNA within an infected plant and others not exceeding 2%. By changing the host plant species, we could further show that the relative frequency of the different segments is host-specific. Besides, we observed changes in segment frequencies within the vectors, thus revealing intimate interactions between the FBNSV –which was previously thought to simply traverse cellular barriers- and its vectors.

This plasticity and the corresponding adjustment of the genetic composition of the within host viral population could enable the multipartite virus to rapidly adjust to a novel environment, by specifically changing the copy number of its different genes. However, it could also constitute an additional cost for the virus if, as generally assumed, all segments need to enter the same cell to secure infection. When considering the rare segments in FBNSV populations, the cost appear enormous and we have decided to experimentally address this question.

We are currently looking at the distribution of the different genome segments within individual cells of host plants in order to empirically assess whether all 8 segments are always together within individual cells or whether they can be physically separated. This question is of major importance, not only for re-evaluating the possible costs in multipartite viral systems, but also for challenging the concept of a cell-autonomous replication cycle in viruses.

Finally, as the genomic integrity preservation of multipartite viruses is based on rather small bottlenecks during the whole viral life cycle, tools have been developed to quantify these bottlenecks.