

FLUCTUATIONS DEMOGRAPHIQUES AU COURS DU CYCLE DE VIE DU CaMV (Cauliflower mosaic virus). Estimation de la taille efficace des populations virales lors de la colonisation des feuilles de la plante hôte, évaluation de la multiplicité d'infection cellulaire au sein de ces feuilles, et estimation de la taille des goulots d'étranglement lors de sa transmission d'hôte à hôte par vecteur

Baptiste Monsion

▶ **To cite this version:**

Baptiste Monsion. FLUCTUATIONS DEMOGRAPHIQUES AU COURS DU CYCLE DE VIE DU CaMV (Cauliflower mosaic virus). Estimation de la taille efficace des populations virales lors de la colonisation des feuilles de la plante hôte, évaluation de la multiplicité d'infection cellulaire au sein de ces feuilles, et estimation de la taille des goulots d'étranglement lors de sa transmission d'hôte à hôte par vecteur. Génétique des populations [q-bio.PE]. Université Montpellier II - Sciences et Techniques du Languedoc, 2008. Français. ⟨NNT : ⟩. ⟨tel-00681500⟩

HAL Id: tel-00681500

https://theses.hal.science/tel-00681500v1

Submitted on 21 Mar 2012

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.



HAL Authorization

UNIVERSITE MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

T H E S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

Discipline : Biologie des populations et écologie

Ecole Doctorale : SIBAGHE

présentée et soutenue publiquement

par

Baptiste MONSION

le 28 Mai 2008

Titre :

FLUCTUATIONS DEMOGRAPHIQUES AU COURS DU CYCLE DE VIE DU CaMV *(Cauliflower mosaic virus)*

Estimation de la taille efficace des populations virales lors de la colonisation des feuilles de la plante hôte, évaluation de la multiplicité d'infection cellulaire au sein de ces feuilles, et estimation de la taille des goulots d'étranglement lors de sa transmission d'hôte à hôte par vecteur

JURY

Jean-Loup NOTTEGHEM	Professeur à l'Agro.M, Montpellier	Président
Christophe ROBAGLIA	Professeur à l'Université de la Méditerranée	Rapporteur
Olivier LE GALL	Directeur de Recherche à l'INRA, Bordeaux	Rapporteur
Stéphane BLANC	Directeur de Recherche à l'INRA, Montpellier	Directeur de thèse
Benoît MOURY	Directeur de Recherche à l'INRA, Avignon	Examinateur
Manuel PLANTEGENEST	Maître de conférence à l'ENSAR, Rennes	Examinateur

Résumé

Le CaMV (*Cauliflower mosaic virus*) est un virus de plante à ADN double brin transmis par des pucerons. Comme pour tous les virus avec un taux de mutation élevé, les larges fluctuations démographiques au cours du cycle de vie sont supposées avoir un fort impact sur l'évolution de la valeur sélective, du fait de la dérive génétique intense, et des effets associés du cliquet de Muller. Pourtant, de telles données démographiques sont rares.

Afin de suivre l'évolution des populations de CaMV nous avons construit 6 clones infectieux contenant chacun un marqueur génétique distinct à un même locus. Nous avons développé une nouvelle méthode d'analyse : Quantitative Single-letter Sequencing (QSS) permettant de déterminer leur présence/absence et de suivre l'évolution de leur fréquence au sein d'une plante infectée.

Tout d'abord, nous avons évalué la taille efficace des populations de CaMV durant l'invasion de l'hôte : plusieurs centaines à plusieurs milliers de génomes sont à l'origine de la colonisation de chaque feuille au cours de l'infection systémique. La diversité virale, au sein des populations du CaMV, est en conséquence distribuée de manière très homogène dans toutes les feuilles infectées de l'hôte. Cette évaluation est surprenante puisqu'elle est 10 à 100 fois supérieure à ce qui a été estimé auparavant chez tous les phytovirus étudiés. Elle démontre ainsi que la très faible taille efficace de population, et donc les larges fluctuations démographiques durant le cycle d'infection *in planta*, ne sont pas une règle générale chez les virus de plante.

Ensuite, nous avons déterminé la multiplicité d'infection cellulaire naturelle (MOI, multiplicity of infection) qui se définit ici comme le nombre moyen de génomes qui pénètrent dans (et infectent) chaque cellule de l'hôte. Cette étude est inédite dans le sens où ce paramètre n'avait jamais été établi chez un virus de plante, et où nous avons montré pour la première fois que la MOI n'est pas une constante mais qu'elle augmente au fil du développement de l'infection pour atteindre un plateau, à une valeur proche de 7 dans le cas du CaMV. Cette valeur est très élevée ; elle dépasse amplement les données disponibles dans la littérature, quelle que soit l'espèce virale considérée. Des indices indirects semblent notamment suggérer que la MOI d'autres espèces de phytovirus puisse être très faible. Les causes et/ou conséquences de ces comportements contrastés (MOI forte ou faible) sur la biologie du virus représentent un champ de perspectives considérable pour l'équipe dans les années à venir.

Enfin, à l'aide de la technique d'EPG, nous avons contrôlé le comportement alimentaire des pucerons vecteurs, et évalué l'impact de ce comportement sur le goulot d'étranglement induit sur la population virale lors de la transmission.

Mots clés : CaMV, Goulot d'étranglement, Taille efficace des populations, Transmission par puceron, Multiplicité d'infection

DEMOGRAPHIC FLUCTUATIONS IN THE LIFE CYCLE OF CaMV (Cauliflower mosaic virus)

Abstract

The *Cauliflower mosaic virus* (CaMV, a double-strand DNA plant virus) is non-circulatively transmitted by aphid vectors. As for many other viruses with a high mutation rate, ample demographic fluctuations along the life cycle, that are most often poorly defined, may have considerable impact on the evolution of fitness, due to intense genetic drift and associated Muller's ratchet.

In order to monitor the CaMV populations we constructed 6 infectious clones, each containing a unique genetic marker at the same locus. We have developed and implemented a new analysis method: Quantitative Single-letter Sequencing (QSS) to determine the presence/absence and monitor the frequency of the markers in an infected plant.

We first assessed the effective size of CaMV populations during host plant colonisation, and found that hundreds to thousands of genomes are founding the population in every single new leaf during systemic infection. This evaluation is 10 to 100 fold higher than that previously published for all other plant viruses, suggesting that ample demographic fluctuation during host infection is not a general rule for plant viruses.

Then, we addressed the natural multiplicity of infection of cells (MOI) in CaMV infected plants, a trait that has never been investigated before in plant viruses, and very rarely in other virus species. We showed for the first time that the mean MOI value is not a constant, and increases along the course of the host infection, reaching a maximum value close to 7 for CaMV, largely exceeding the rare estimates available in the literature on animal viruses or phages. Indirect hints collected in the literature suggest that other plant viruses may have an oppositely very low MOI (close to 1), and the causes and/or consequences of this contrasted situation on the biology of viruses is discussed in depth. This represents a new field of investigation to be developed in the group in the next future.

Finally, using the EPG technique, we have controlled the feeding behaviour of aphid vectors acquiring CaMV, and evaluated the impact of this vector behaviour on the size of the bottleneck induced on viral population during plant-to-plant transmission.

Key words : Cauliflower mosaic virus, Bottleneck, Effective population size, Aphid transmission, Multiplicity of infection.

Laboratoire d'accueil

Ce travail de thèse a été réalisé sur deux sites distincts suite au regroupement physique de l'ensemble du personnel de l'UMR BGPI au sein d'un bâtiment neuf construit à cette occasion sur le Campus International de Baillarguet :

2004-2008

Biologie et Génétique des Interactions Plante-Parasite (BGPI)[1]

INRA-CIRAD-SupagroM

TA A-54/K, Campus International de Baillarguet

Équipe CaGeTE

34398 Montpellier Cedex 5

France

2003-2004

Station de Pathologie Comparée

UMR 5087 INRA-CNRS-UMII

Équipe Vection

30380 Saint-Christol-lez-Alès

[1] l'UMR BGPI a été certifiée selon la norme ISO 9001 version 2000 pour les activités de "recherche appliquée et fondamentale, expertise, formation d'étudiants et de professionnels, dans le domaine des interactions plante - agent pathogène"

« Que nous soient donnés la pluie, la santé et l'enrichissement spirituel »

Brown, le Frontalier [2]

[2] Voir l'Annexe 3 pour plus de détails

à mes victimes

Suzym CASPERIE & Vibrecynore CABASSIER

que j'ai tant aimées [3]

[3] Simples anagrammes de bestioles sacrées : *Myzus persicae* & *Brevicoryne brassicae*

Remerciements

Sincères remerciements à tous les gens rattachés à mon *ka* qui se reconnaîtront et dont l'estime que j'ai pour eux n'aurait pas plus de valeur une fois leurs noms listés.

Bref, merci à tous ceux qui ont pu et su me rendre la vie plus agréable – où que ce soit de Caen à Madrid en passant par le Languedoc-Roussillon – malgré toutes les contraintes, coûts et douleurs qui leur ont été occasionnés de par leurs amours pour ma propre personne ou de par leur simples sympathies envers certaines de mes singularités.
Pour preuve de mon éternelle gentillesse légendaire, voici une aide probable au décryptage de ces remerciements : toute personne qui me connaît et qui trouve ces remerciements très impersonnels n'en fait certainement pas partie !
Merci à toute personne qui pense avoir été oubliée de se signaler !

Merci également à de nombreux artistes inconnus

Grand merci à Romane qui ne sait pas encore lire, d'où l'intérêt tout particulier qui me fait écrire son prénom comme l'unique de cette page.

Silence s'il vous plait, la petite dort et la lecture du manuscrit va bientôt commencer …

MERCI

Tables des matières

Résumé 2
Abstract 3
Laboratoire d'accueil 4
Epigraphe 5
Dédicace 6
Remerciements 7
Tables des matières 8
Liste des figures 12
Liste des tableaux 12
Liste des encadrés 13
Avant-propos 14

Introduction générale *16*

1 Le cycle viral du CaMV dans son hôte 19
1.1 L'infection au niveau de la cellule hôte 19
1.1.1 Le cycle cellulaire du CaMV 19
1.1.2 Fonctions des différentes protéines virales exprimées 21
1.1.2.1 Protéine exprimée à partir de l'ARN 19S : P6 21
1.1.2.2 Protéines exprimées à partir de l'ARN 35S : P1 à P5 23
1.1.2.2.1 P5 25
1.1.2.2.2 P1 25
1.1.2.2.3 P4 27
1.1.2.2.4 P3 29
1.1.2.2.5 P2 31
1.2 Colonisation de la plante hôte 33
1.2.1 Mouvement du CaMV depuis son site d'inoculation 33
1.2.1.1 Mouvement de cellule à cellule via les plasmodesmes 33
1.2.1.1.1 Rappel sur les différents modes de passage de cellule à cellule chez les phytovirus 34
1.2.1.1.2 Cas particulier du CaMV 36
1.2.1.1.3 Cheminement par les plasmodesmes jusqu'aux tubes criblés du phloème 37
1.2.1.2 Mouvement à longue distance et distribution du virus dans la plante 39
1.3 Symptomatologie & Interaction avec l'hôte 39
1.3.1 Symptomatologie et Pathogénicité 41
1.3.1.1 Symptômes et accumulation du virus dans le navet 41
1.3.1.2 Gamme d'hôte et Pathogénicité 43
1.3.2 Impact et gestion du RNA Silencing sur et par le CaMV 45
1.3.2.1 Un système de défense pour la plante 45
1.3.2.2 Un système contrôlé par le virus ? 47

2 La transmission des virus de plante 48
2.1 Les différents modes de transmissions des phytovirus 48
2.1.1 La transmission circulante multipliante (ou propagative) 49
2.1.2 La transmission circulante non-multipliante 50
2.1.3 La transmission non-circulante 50
2.1.3.1 Stratégie capside 52
2.1.3.2 Stratégie Helper 52
2.2 Précisions sur les phases d'acquisition et d'inoculation des virus non-circulants 52
2.2.1 Acquisition 56
2.2.2 Inoculation 56
2.3 La transmission du CaMV en détail 56
2.3.1 Stratégie Helper et modèle d'acquisition séquentielle 56
2.3.2 Acquisition – Rétention – Inoculation 58
2.3.2.1 Acquisition 58
2.3.2.2 Rétention 60
2.3.2.3 Inoculation 60
2.3.3 Interaction moléculaire Virus-Vecteur 60
2.3.3.1 Spécificité d'interaction (P2-récepteur) 60
2.3.3.2 Zone d'interaction et nature chimique du récepteur 62

3 Quelques aspects ciblés de génétique des populations virales 64
3.1 Description des notions de base 65
3.1.1 Variabilité chez les phytovirus 65
3.1.2 Description et propriétés de la quasi-espèce virale 66
3.1.3 Dérive génétique et cliquet de Muller 67
3.1.4 Cliquet de Muller et goulots d'étranglement 67
3.2 Connaissances chez le CaMV 67
3.2.1 Variabilité chez le CaMV 67
3.2.2 La quasi-espèce pour le CaMV 69
3.2.3 Impacts des goulots d'étranglement sur le CaMV 71
3.3 Objectif initial de la thèse 71

CHAPITRE I : Taille efficace des populations de CaMV intra hôte 74

1 Quantitative Single-letter Sequencing : a method for simultaneously monitoring numerous allelic variants in single DNA samples (Article 1) 76

2 Effective size of *Cauliflower mosaic virus* populations during host plant colonization (Article 2) 89

CHAPITRE II : Multiplicité d'infection (M.O.I.) 106

1 Introduction 107
1.1 La multiplicité d'infection : notions 107
1.1.1 La MOI et notre définition 107
1.1.2 L'infection multiple : co-infection 108
1.2 La MOI chez les virus : estimations disponibles 109
1.2.1 Virus de bactéries : phages 109
1.2.2 Virus d'animaux 109
1.2.2.1 Virus d'arthropodes : baculovirus 109
1.2.2.2 Virus de mammifères 110
1.2.2.2.1 Human immunodeficiency virus (HIV) 110
1.2.2.2.2 Cytomégalovirus (CMV) 111
1.2.3 Virus de plantes 112

2 Matériel & Méthodologie mise en œuvre 116
2.1 Principe général de la méthode développée 117
2.2 Variants viraux et amorces spécifiques utilisés 118
2.3 Quantification des variants et co-inoculation 118
2.4 Récolte des feuilles, protoplastation et fixation 120
2.4.1 Récolte des feuilles 120
2.4.2 Protoplastation 120
2.4.3 Fixation des protoplastes 122
2.5 Décontamination et isolement des protoplastes 123
2.5.1 Décontamination du surnageant 123
2.5.2 Isolement des protoplastes 123
2.6 Détection des variants à partir de protoplaste fixé 125
2.6.1 Libération de l'ADN viral et Nested-PCR-1 125
2.6.2 Nested-PCR-2 125
2.6.2.1 Par PCR classique 125
2.6.2.2 Avec le LightCycler® 480 : kit HRM 126
2.6.3 Vérification de l'efficacité la méthode de décontamination 126
2.7 Détermination des fréquences relatives de chaque variant dans une feuille 128
2.8 Pour le calcul de la MOI 129

3 Résultats 131
3.1 Prologue exploratoire du taux de co-infection 131
3.2 Estimation de la MOI 132
3.2.1 Fréquence des variants et taux de co-infection 132
3.2.2 Calcul de la MOI 136

4 Discussion 136
4.1 La MOI du CaMV *versus* autres virus d'hôtes pluricellulaires 136
4.2 Augmentation de la co-infection au cours du temps 136
4.3 Inoculation des variants sur une même feuille ou sur feuilles séparées 139
4.4 Deux stratégies semblent s'être établies 141
4.4.1 Faible MOI (≈1) 142
4.4.2 Forte MOI (≥3) 142

5 Perspectives 143

CHAPITRE III : Goulots d'étranglement imposés aux populations de CaMV lors de la transmission d'hôte à hôte par vecteur 145

1 Introduction 146

2 Matériels et Méthodes 149
2.1 Plantes sources, pucerons et plantes test 149
2.2 Electro-Penetro-Graphie (EPG) 150
2.2.1 Dispositif EPG (branchements) 150
2.2.2 Modalités des périodes d'acquisition du CaMV 151
2.3 Quantification des clones VIT dans les feuilles sources 151
2.4 Détection des clones VIT dans les plantes tests 151

3 Résultats 153
3.1 Taux de transmission selon la modalité d'acquisition 153
3.2 Fréquence de transmission de chaque clone VIT 154
3.3 Diversité transmise selon la modalité d'acquisition 156
3.4 Evaluation de la taille du goulot d'étranglement : nombre de génomes viraux efficacement transmis en fonction de la modalité d'acquisition *(analyses futures)* 159

4 Discussion 160
4.1 Neutralité des clones lors de la transmission 160
4.2 Cohérence et pertinence avec les études précédentes 161
4.3 Effet de la modalité d'acquisition ? 162

5 Perspectives 164
5.1 A brève échéance 164
5.2 A plus long terme 164
5.2.1 Effets de deux comportements alimentaires distincts du vecteur sur l'évolution du CaMV 164
5.2.2 Quantification de la HC-transcomplémentation 164
5.2.3 Taille du goulot d'étranglement chez les virus circulants 165

Conclusion générale 167

Matériels & Méthodes 172

1 Matériels 173
1.1 Matériel végétal 173
1.2 Souches virales 173
1.3 Espèces de pucerons 173
1.4 Souches bactériennes 173
1.5 Plasmides 174

2 Méthodes 174
2.1 Techniques relatives aux plantes 174
2.1.1 Culture des navets 174
2.1.2 Inoculation mécanique du CaMV 174
2.2 Techniques relatives aux pucerons-vecteurs 174
2.2.1 Conditions d'élevage 174
2.2.2 Tests de transmission classiques 174
2.2.2.1 A travers membrane de Parafilm® 174

2.2.2.2 Transmission de plante à plante 175
2.3 Techniques et stratégies utilisées pour le clonage 175
2.3.1 Amplification par PCR (Polymerase Chain Reaction) 175
2.3.2 Digestion par des enzymes de restriction 176
2.3.3 Déphosphorylation du plasmide 176
2.3.4 Purification de fragments d'ADN sur gel d'agarose 176
2.3.5 Ligation 177
2.3.6 Transformation et sélection 177
2.3.7 Criblage des colonies 177
2.3.7.1 Par digestion enzymatique 177
2.3.7.2 Par PCR sur des pools de colonies 178
2.3.8 Amplification et purification de plasmides recombinants 178
2.4 Techniques relatives aux virus 178
2.4.1 Purification de virus 178
2.5 Techniques relatives aux protéines 179
2.5.1 Production de protéines en systèmes hétérologues 179
2.5.1.1 Production de P2 et P3 en système baculovirus 179
2.5.1.2 Production de P3 en bactérie et purification 180
2.5.2 Analyse des protéines et quantification 181
2.5.2.1 Électrophorèse SDS-PAGE & Western Blot 181
2.5.2.2 Quantification densitométrique des virus 182

***Références Bibliographiques* 183**

***Annexes* 191**

1 Tests préliminaires de transmission par puceron 192
1.1 Evaluation de la neutralité des clones Div 192
1.2 Préparation d'une solution mixte : « Mix5Div » 192
1.3 Tests de transmissions avec *Myzus persicae* 194
1.3.1 Acquisition sur feuille infectée 194
1.3.2 Acquisition sur membrane 194

2 Programme « R » pour calculer la MOI 198

3 Détails sur l'épigraphe 200

Liste des figures

Figure 1 : Profils d'accumulation des différentes protéines virales au cours de l'infection de protoplastes par le CaMV 22
Figure 2 : Représentation schématique des domaines et motifs de la protéine P6 (TAV) du CaMV 22
Figure 3 : Représentation schématique de la protéine P5 26
Figure 4 : Représentation schématique de la protéine P1 26
Figure 5 : Représentation schématique du précurseur de la protéine de capside du CaMV et de ses produits de maturations 30
Figure 6 : Représentation schématique de la protéine P3 30
Figure 7 : Représentation schématique de la protéine P2 32
Figure 8 : Le corps clair se forme en drainant (accumulant) la P2 de toute la cellule infectée 32
Figure 9 : Durée nécessaire pour la colonisation des tissus vasculaires dans une feuille de navet inoculée par le CaMV 38
Figure 10 : Schéma d'un navet indiquant les paramètres qui influencent l'accumulation bilatérale et basipétale du CaMV dans les feuilles systémiques 42
Figure 11 : Relations génomiques entre les différentes espèces du genre Brassica 44
Figure 12 : Modèle pour la défense anti-virale basée sur le RNA silencing 46
Figure 13 : Corrélation entre phases de la ponction intracellulaire (pd), acquisition et inoculation de virus non-persistant 54
Figure 14 : La phase II-1 est une phase de salivation 55
Figure 15 : Le CaMV est préférentiellement acquis au niveau du phloème par son vecteur 59
Figure 16 : Taux de transmission des différents mutants du CaMV par 5 espèces de pucerons 61
Figure 17 : Caractérisation de la nature chimique du récepteur hypothétique du CaMV dans les stylets de son puceron vecteur 63
Figure 18 : Accumulation des mutations délétères au sein d'une population 68
Figure 19 : Variabilité de la quasi-espèce du CaMV 70
Figure 20 : Propriétés biologiques (phénotypiques) de génomes isolés composant la quasi-espèce du CaMV 70
Figure 21 : Distribution spatiale de virus en co-infection systémique dans N. benthamiana 113
Figure 22 : Position des amorces utilisées pour la Nested-PCR sur VIT1 et VIT3 119
Figure 23 : Inoculation en vis-à-vis pour quantifier le pouvoir infectieux 121
Figure 24 : Découpe des protoplastes fixés et déposés sur bandelette de plastique 124
Figure 25 : Validation de la méthode de décontamination des protoplastes fixés 127
Figure 26 : Symptomatologie provoquée par le CaMV au cours du temps 130
Figure 27 : Evolution du taux de co-infection au cours du temps 134
Figure 28 : Représentation graphique de l'estimation de la MOI et de son évolution au cours du temps 135
Figure 29 : Détection des clones VIT dans les plantes tests par PCR-RFLP-CE 152
Figure 30 : Distribution du nombre de marqueurs VIT détectés dans les plantes tests issues de chaque modalité d'acquisition 158

Figure A1 : Evaluation de la neutralité des clones Div 193
Figure A2 : Distribution du nombre de clones Div transmis par un puceron de plante à plante 195
Figure A3 : Distribution du nombre de clones Div transmis par un puceron (acquisition sur membrane) 195

Liste des tableaux

Tableau 1 : Dénombrement des lésions locales pour quantifier le pouvoir infectieux des préparations virales 121
Tableau 2 : Validation de la méthode de décontamination des protoplastes fixés 127
Tableau 3 : Evolution du taux de co-infection dans la plante au cours du temps 130
Tableau 4 : Evaluation de la fréquence des variants et du taux de co-infection cellulaire dans une feuille à 60 dpi 134
Tableau 5 : Evolution du taux de co-infection au sein d'une même feuille au cours du temps 138
Tableau 6 : Influence de la séparation spatiale de l'inoculation sur la co-infection 140
Tableau 7 : Taux de transmission en fonction de la modalité d'acquisition 155
Tableau 8 : Fréquence de transmission de chacun des marqueurs VIT 155
Tableau 9 : Nombre de marqueurs VIT transmis en fonction de la modalité 157
Tableau 10 : Test de différenciation par paires entre les modalités d'acquisition 158

Tableau A1 : Synthèse des tests de transmission réalisés à partir du Mix5Div 197

Liste des encadrés

Encadré 1
Structure génomique, réplication et cycle cellulaire du CaMV
(page 18)

Encadré 2
Les ORFs de l'ARN 35S et son épissage alternatif
(page 20)

Encadré 3
Traduction de l'ARN 35S par saut de ribosome
(page 24)

Encadré 4
Interactions entre les protéines virales du CaMV
(page 28)

Encadré 5
Le mouvement de cellule à cellule chez les phytovirus
(page 35)

Encadré 6
Le mouvement à longue distance selon le flux des photoassimilats
(page 40)

Encadré 7
Les différents modes de transmission des phytovirus par vecteur
(page 51)

Encadré 8
Technique d'électro-pénétrographie (EPG)
(page 53)

Encadré 9
La transmission du CaMV par vecteur
(page 57)

Avant-propos

Suite à une maîtrise en Biologie Cellulaire et Physiologie (option Physiologie végétale appliquée) qui a développé mon goût pour la virologie végétale, j'ai suivi le DEA de Ressources Phytogénétiques et Interactions biologiques afin d'effectuer un stage enthousiasmant sur le nombre de particules virales transmises par un insecte vecteur. Les résultats obtenus – bien que probants – se devaient d'être solidement validés.
La thèse, axée originellement sur la transmission du CaMV par puceron-vecteur – thème authentique et historique de l'équipe « Vection » – nous a entraîné vers l'étude de certains aspects relatifs à la génétique des populations, tels que les fluctuations démographiques que subissent les populations virales durant l'infection systémique des plantes hôtes, en plus de la transmission. Le paragraphe suivant décrit les travaux réalisés selon leur chronologie réelle qui, dans un souci d'homogénéité et de logique de présentation diffère légèrement de l'organisation des chapitres au sein du manuscrit.

Afin de suivre l'évolution des populations de CaMV nous avons construit 6 clones infectieux contenant chacun un marqueur génétique distinct à un même locus. Nous avons développé une nouvelle méthode d'analyse : Quantitative Single-letter Sequencing (QSS) qui nous a permis de déterminer la présence/absence et de suivre l'évolution de la fréquence des différents marqueurs au sein de chaque plante co-inoculée avec ces six variants.

Par cette méthode d'analyse, nous avons évalué la taille efficace (*Ne*) des populations de CaMV quand elles colonisent les nouvelles feuilles au cours de l'infection systémique. En effet, dans l'optique finale de caractériser l'échantillon viral prélevé (sur une plante infectée) et transmis par un puceron vecteur, il nous fallait préalablement vérifier que les clones ou variants viraux potentiellement transmis par le vecteur ne seraient pas éliminés par simple dérive génétique *in planta*, que ce soit avant leur

prélèvement ou après leur inoculation. Les valeurs estimées de *Ne* se sont révélées très élevées (une centaine à plusieurs milliers de génomes viraux fondant la population dans chaque feuille), indiquant une dérive négligeable durant l'infection systémique. Notre objectif final pouvait ainsi être poursuivi dans les meilleures conditions.

Les tests de transmission de plante à plante par puceron nous ont alors permis de démontrer l'existence d'un très fort goulot d'étranglement, spécifiquement à cette étape du cycle viral. En effet, sur les six variants présents dans les feuilles sources, seuls 2 à 3 en moyenne infectent chaque plante test, suite à l'inoculation par un insecte vecteur, et ceci indépendamment du fait qu'il ait acquis le CaMV dans une seule cellule du mésophylle de la feuille source, dans 5 cellules ou dans les vaisseaux du phloème.

De manière remarquable, dans les cas où le puceron-vecteur a pu acquérir le virus lors d'une seule ponction intra-cellulaire (contrôle EPG), 62% des plantes subséquemment infectées contenaient plusieurs variants génétiques. Cette observation nous a incité à suivre la « co-existence » des différents variants génétiques viraux dans la plante hôte, jusqu'au niveau intra-cellulaire, et ainsi à calculer la multiplicité d'infection (MOI) naturelle du CaMV durant la colonisation systémique de son hôte.

Il ne s'agissait pas de nous positionner sur des questions fondamentales de génétique des populations, ou plus généralement de biologie évolutive, en utilisant le CaMV comme modèle. Néanmoins, le cheminement résumé ci-dessus nous a entraîné à l'interface avec ces disciplines et, à chaque étape, nous avons essayé de replacer nos résultats dans ce contexte plus général afin de les valoriser au mieux. L'organisation globale du manuscrit est donc logiquement CaMV-centrée, les notions de biologie évolutive nécessaires à la valorisation de chaque « bloc » de résultats étant plus spécifiquement introduites dans les chapitres correspondants.

Introduction générale

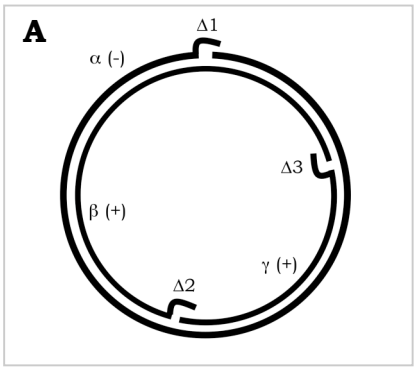
Le *Cauliflower Mosaic Virus* (CaMV) est un pararétrovirus – virus répliqué par transcription inverse, mais dont l'intégration du génome à celui de l'hôte n'est pas requise – de plante qui infecte naturellement les brassicacées et parfois les solanacées pour certaines souches particulières [1]. Le CaMV est le membre type du genre *Caulimovirus* et de la famille *Caulimoviridae*, il est constitué d'un génome circulaire à double brin d'ADN d'environ 8000 paires de bases (voir encadré 1A, [2]), ce qui lui a valu d'être le premier phytovirus séquencé [3]. Le génome est protégé par une capside pour former des particules virales icosaédriques de 52 nm [4].

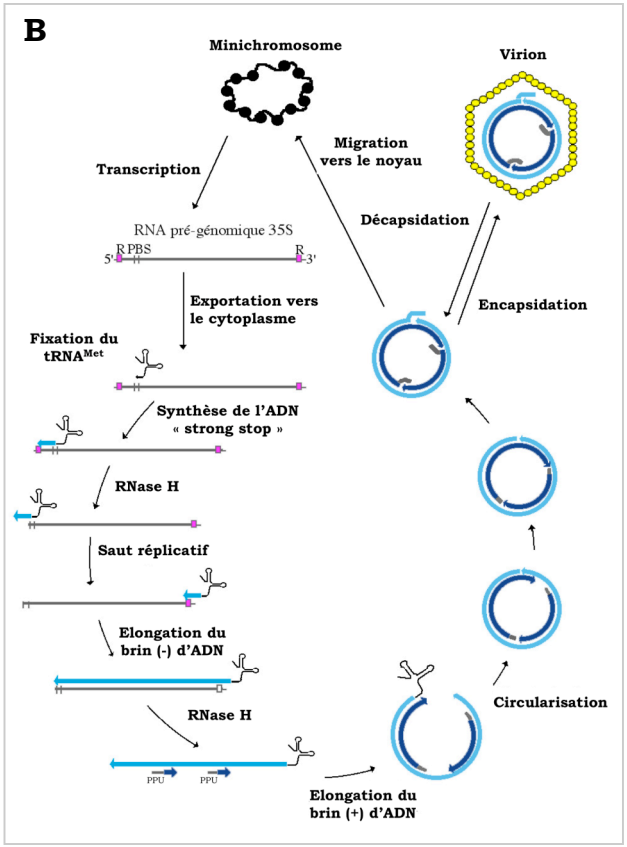
Ces virions sont naturellement transmis de plante à plante via les pucerons-vecteurs, comme de très nombreux autres phytovirus. La mobilité des vecteurs permet une large dissémination du virus dans les populations végétales sensibles. Le CaMV est à priori présent sur tous les continents, partout où hôtes et vecteurs potentiels coexistent [5]. L'impact de la maladie est généralement faible, mais peut occasionnellement être important sur certaines cultures de choux de Bruxelles, de choux pommés et de choux-fleurs, dans les zones de culture intensive. Le CaMV peut attaquer sévèrement les pommes (inflorescence) des choux cultivés, qui perdent ainsi leur valeur marchande. Les crucifères adventices et d'ornement peuvent aussi être atteintes. La lutte contre les pucerons vecteurs avec des insecticides est utilisée pour limiter la dissémination du virus et réduire les dégâts directs. Prévenir l'introduction initiale du virus dans les cultures en utilisant des aphicides est difficile, mais les méthodes culturales permettent de limiter les attaques de pucerons et l'établissement dommageable du virus [6].

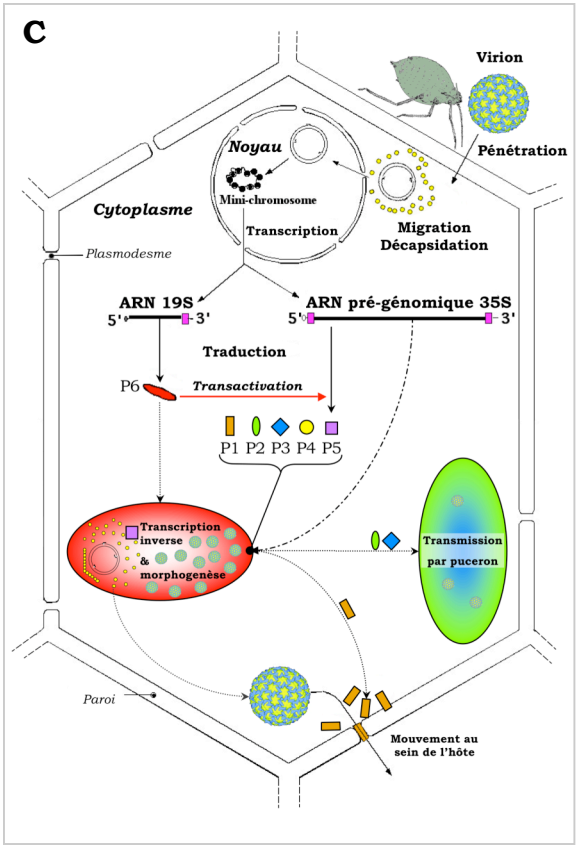
De part ses caractéristiques génomiques – offrant des commodités de travail indéniables en termes de biologie moléculaire afin de créer des mutants *in vitro* directement infectieux par exemple – et de son mode de transmission (mode non-circulant), commun à la majorité des phytovirus, le CaMV est un modèle incontournable en virologie végétale, du domaine de la biologie cellulaire à celui de l'évolution virale.

Encadré N°1

Structure génomique, réplication et cycle cellulaire du CaMV







A Structure génomique : le génome viral encapsidé consiste en un ADN double brin circulaire d'environ 8000 paires de bases qui comporte 3 interruptions de séquence résultant des différentes étapes du mécanisme de réplication (voir B).

B Réplication du génome viral : à partir de l'ARN 35S viral, et d'un ARN de transfert méthionine (ARNt) de l'hôte qui s'y lie au niveau d'un Primer Binding Site (PBS) de 14 pb, la transcriptase inverse (RT) initie la synthèse du brin d'ADN négatif jusqu'à l'extrémité 5' de l'ARN 35S. L'activité RNase H associée à la RT dégrade la matrice d'ARN dès qu'elle est rétro-transcrite. Le brin d'ADN alors en cours de synthèse par la RT est transféré en 3' de l'ARN 35S au niveau de la zone répétée (R : en rose) de 180 nucléotides qui est présente aux deux extrémités de l'ARN 35S. L'élongation de l'ADN reprend jusqu'à l'extrémité 5' du PBS, générant ainsi la future interruption de séquence Δ1. A cause de séquences PPU (polypurine), la RNase H laisse deux zones d'hétéroduplex non dégradées, à partir desquelles l'activité ADN polymérase ADN dépendante de la RT va amorcer la synthèse des brins d'ADN positifs complémentaires et engendrer la 2ième interruption (Δ2). Après un dernier transfert de brin de la RT, il y aura formation de la dernière interruption de séquence (Δ3), à la jonction de l'autre région PPU, avant l'encapsidation du génome viral.

C Cycle de multiplication du CaMV : les principales étapes du cycle viral sont: 1) pénétration du virus dans la cellule par puceron ou par mouvement de cellule à cellule, 2) migration du virion vers les pores nucléaires, 3) décapsidation et importation de l'ADN viral dans le noyau, 4) formation du minichromosome, 5) transcription de l'ADN viral, 6) traduction des ARN 19S et 35S, 7) réplication du génome viral par la transcriptase inverse (voir B), 8) morphogenèse des particules virales et 9) mouvement du virus de cellule à cellule à travers les plasmodesmes. Les étapes 6-8 semblent s'effectuer dans ou à proximité du corps d'inclusion dense (en rouge). En outre, le virus prépare la cellule hôte à la transmission par puceron par la formation d'une structure spécialisée (le corps d'inclusion clair : en bleu/vert) composée de P2, P3 et de quelques particules virales.

D'après [2].

1 Le cycle viral du CaMV dans son hôte

Comme tous les virus de bactéries, de plantes ou d'animaux, le CaMV est un parasite endocellulaire obligatoire. Il ne dispose pas de l'information génétique nécessaire pour synthétiser de manière autonome les systèmes métaboliques présents dans l'hôte qu'il exploite.

1.1 L'infection au niveau de la cellule hôte

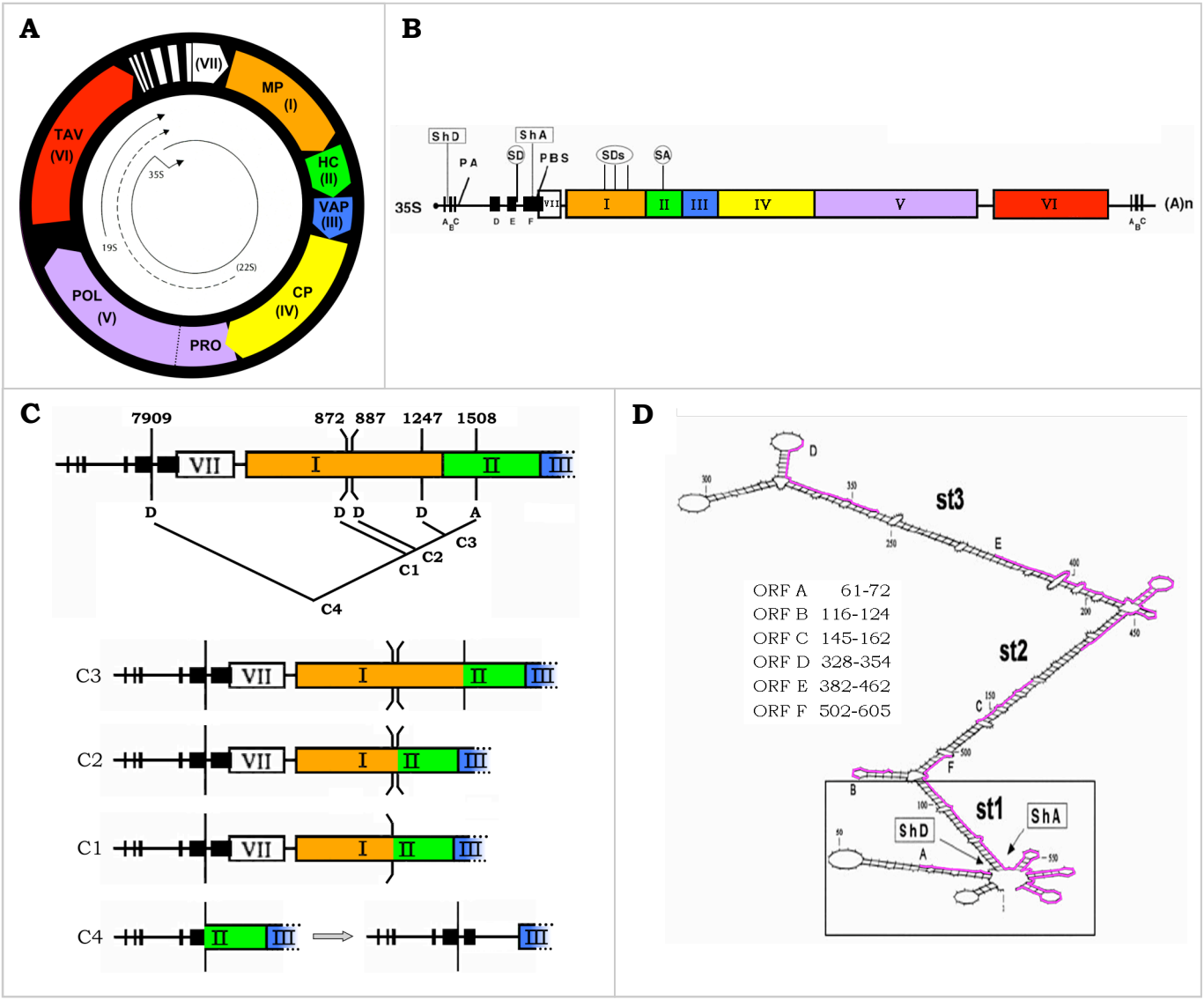
Dans la nature, un cycle de réplication viral s'enclenche dès qu'un ou des virion(s) rentre(nt) dans le cytoplasme d'une cellule végétale hôte, soit via les plasmodesmes qui assurent un continuum cytoplasmique entre cellules adjacentes, ou bien via les piqûres intracellulaires d'un puceron-vecteur.

1.1.1 Le cycle cellulaire du CaMV

Dès qu'un virion a pénétré dans une cellule sensible, celui-ci migre vers le noyau, où l'ADN génomique (voir encadré 1C) est très probablement décapsidé au niveau de la membrane nucléaire avant d'y pénétrer [7]. A l'intérieur du noyau, le génome viral est converti en un mini-chromosome super enroulé par la machinerie cellulaire. Le génome du CaMV n'est probablement pas multiplié dans le noyau – comme le serait un épisome – mais il est transcrit par l'ARN polymérase II. Au moins deux ARN messagers fonctionnels : le 19S (subgénomique) et le 35S (de taille génomique) sont synthétisés puis exportés vers le cytoplasme sous forme polyadenylée ; un ARN subgénomique 22S a été décrit en 1985, mais sa fonctionnalité n'a pas été vérifiée [8] (voir encadré 2A, [9,10]). L'ARN 19S monocistronique est traduit en une protéine qui sert ensuite de trans-activateur indispensable à la traduction complète de l'ARN polycistronique 35S. En effet, les systèmes cellulaires eucaryotiques sains – qui ne possèdent généralement que des ARN messagers monocistroniques – sont incapables de réinitier la traduction sur un ARN messager polycistronique, après la fin du premier cadre ouvert de lecture [11]. Le 35S sert donc de (i) matrice pour la réplication du génome viral (voir encadré 1B), (ii) ARN messager pour l'expression des gènes viraux dont notamment, la transcriptase inverse et la protéine de capside, la réplication et l'encapsidation semblant s'opérer de manière concomitante.

Encadré N°2

Les ORFs de l'ARN 35S et son épissage alternatif



A Carte circulaire du génome de la souche Cabb S. Le génome du CaMV comprend 7 ORF majeurs (dont l'expression des 6 premiers a été vérifiée en plante : (I) protéine de mouvement, MP ; (II) helper component, HC ; (III) protéine associée au virion (VAP) ; (IV) capside, CP ; (V) protéase à acide aspartique, PRO, et polymérase, POL ; (VI) transactivateur de la traduction de l'ARN 35S et composant des corps d'inclusion denses, TAV) et 6 petits ORF (sORF: small ORF) dans la séquence leader en amont de l'ORF VII. Le génome viral comporte 3 promoteurs qui sont à l'origine des : (i) ARN pré-génomique 35S, (ii) ARN sub-génomique 19S et (iii) ARN sub-génomique putatif 22S d'après analyse *in silico*. L'ARN 19S code uniquement pour la protéine P6 tandis que l'ARN 22S putatif coderait pour la protéine P5. L'ARN 35S sert de messager polycistronique pour la synthèse des protéines P1 à P5 mais également de matrice pour la réplication par transcription inverse ; c'est pour cette raison qu'il est dit pré-génomique.

B Carte linéaire de l'ARN 35S coiffé en 5' et polyadénylé en 3'. Les ORFs I à VII (colorés) ainsi que les 6 sORF A à F (en noir) sont figurés. Les différents motifs ou sites sont indiqués : PA, signal de polyadénylation ; SD, site donneur pour l'épissage ; SA, site accepteur de l'épissage ; ShD, site de décollage pour le saut du ribosome ; ShA, site d'atterrissage pour le saut du ribosome ; PBS, site d'amorçage pour la transcription inverse; A(n), queue poly(A).

C Représentation schématique de l'épissage de l'ARN 35S. Les 6 sORF de la région leader sont représentés en noir et les ORF I, II et III sont en orange, vert et bleu, respectivement. L'ARN 35S subit un épissage alternatif dans 70% des cas. Il existe un site donneur d'épissage dans la séquence leader et trois sites donneurs dans l'ORF I. Par contre, il n'existe qu'un seul site accepteur d'épissage dans l'ORF II en position 1508. L'épissage alternatif produit des ARN où le début de l'ORF I est fusionné à la fin de l'ORF II en conservant le cadre de lecture – cas des produits C1, C2 et C3 représentant chacun 20% des ARN épissés – et un ARN (C4) où le début du sORF F est fusionné à la fin de l'ORF II sans conservation du cadre de lecture. Il se crée alors un nouvel sORF (G) en amont de l'ORF III. Les polypeptides codés par les ARN C1 à C3 ont été détectés dans les plantes infectées, sans qu'aucune fonction n'ait pu leur être attribuée.

D Organisation de la structure secondaire de la région « leader » de l'ARN 35S du CaMV. La structure en tige-boucle présente dans la séquence leader est très stable. Elle permet, entre autre, de rapprocher le sORF A de l'ORF VII. Elle contient trois tiges (st1 à st3) se différenciant par leur stabilité. Les sites « décollage » (ShD) et « atterrissage » (ShA) du saut du ribosomes sont indiqués. La structure minimale requise pour le saut du ribosome est encadrée (voir encadré n°3 pour le saut du ribosome).

D'après [9, 10].

Une fois que le cycle viral est bouclé – c'est-à-dire que les génomes viraux néo-synthétisés sont encapsidés – les particules virales peuvent soit (i) migrer de cellule à cellule par l'intermédiaire des plasmodesmes (voir section 1.2.1.1), et à longue distance en empruntant les réseaux vasculaires de la plante (voir section 1.2.1.2), soit (ii) rester dans la cellule infectée afin d'être ultérieurement prélevées et transmises à une autre plante par les pucerons-vecteurs (voir section 2.3).

1.1.2 Fonctions des différentes protéines virales exprimées

D'après les analyses *in silico,* l'ARN 35S (voir encadré 2B) possède 7 cadres ouverts de lecture (ORF ; Open Reading Frame) hypothétiques, susceptibles, au regard de leur taille, de coder pour des protéines biologiquement fonctionnelles. Cependant à ce jour, seules 6 protéines ont pu être détectées à partir de matériel végétal infecté (plantes ou protoplastes), et seules 5 d'entre-elles sont indispensables à l'infection systémique [12-15]. Ci après, seront brièvement décrites les différentes fonctions connues pour chacune des protéines virales, en respectant l'ordre correspondant à leur cinétique d'expression (figure 1) observée en protoplastes [16].

1.1.2.1 Protéine exprimée à partir de l'ARN 19S : P6

L'ARN monocistronique 19S comporte uniquement l'ORF VI qui code pour une protéine multifonctionnelle P6 (520 aa ; 62 kDa), exprimée de façon très précoce, par rapport à l'ensemble des autres protéines virales.

La justification majeure de cette précocité semble être le fait que des homodimères de P6 [17] constituent la matrice des corps d'inclusion denses, aussi appelés « viroplasmes » [18], à l'intérieur ou à la périphérie desquels elle transactive la traduction de tous les autres gènes sur l'ARN polycistronique 35S [19,20]. Pour ces propriétés, la protéine P6 est aussi appelée TAV, acronyme de « transactivator/viroplasmin ». L'expression des autres gènes à partir de l'ARN 35S s'effectue au travers d'un mécanisme de ré-initation du complexe de traduction au-delà d'un codon stop, où il serait libéré de l'ARN matrice sans l'action complexe de P6 [21,22]. Sous son contrôle, les facteurs d'initiations (notamment eIF3) semblent rester disponibles, même durant les phases d'élongation et de terminaison, et le

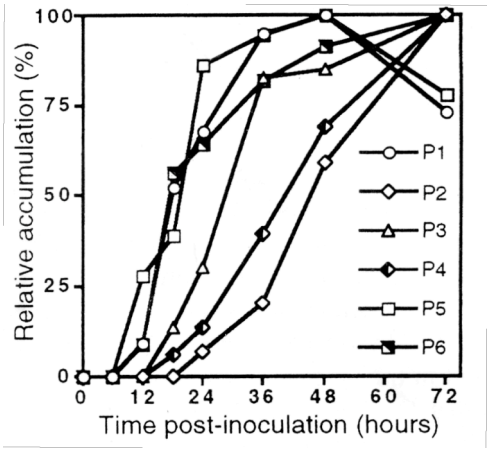


Figure 1 : Profils d'accumulation des différentes protéines virales au cours de l'infection de protoplastes par le CaMV. L'accumulation relative (en % de leur maximum) est déterminée à partir des intensités de coloration mesurées en western-immuno-blot. Des protoplastes de navet ont été transfectés avec des virions de la souche CM1841, et des aliquots ont été prélevés et analysés à des intervalles de temps réguliers durant 72h. *D'après [16].*

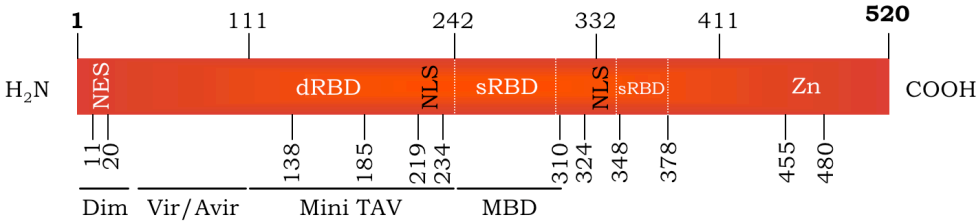


Figure 2 : Représentation schématique des domaines et motifs de la protéine P6 (TAV) du CaMV. La protéine P6 possèdent plusieurs domaines fonctionnels dont ceux impliqués dans sa propre dimérisation (Dim, aa 4-30 structurés en hélice alpha), et dans la virulence/avirulence (Vir/Avir) et la replication. Deux domaines sont importants pour l'activité de la transactivation traductionnelle de l'ARN 35S, le mini TAV (domaine minimal requis pour la transactivation) et le MBD (Multiple Binding Domain) qui se lie à un nombre important de cofacteurs cellulaires. Plusieurs motifs additionnels caractérisent la protéine TAV. Un motif d'export nucléaire (NES, Nuclear Export Signal) et deux motifs de localisation nucléaire (NLS, Nuclear Localization Signal) qui régulent le transport nucléocytoplasmique de TAV. TAV possède une activité de liaison aux ARN double brin (dsRBD, double strand RNA Binding Domain) et deux motifs de liaison aux ARN simple brin (ssRBD, single strand RNA Binding Domain). TAV contient aussi un motif en doigt de zinc (Zn). *D'après [24].*

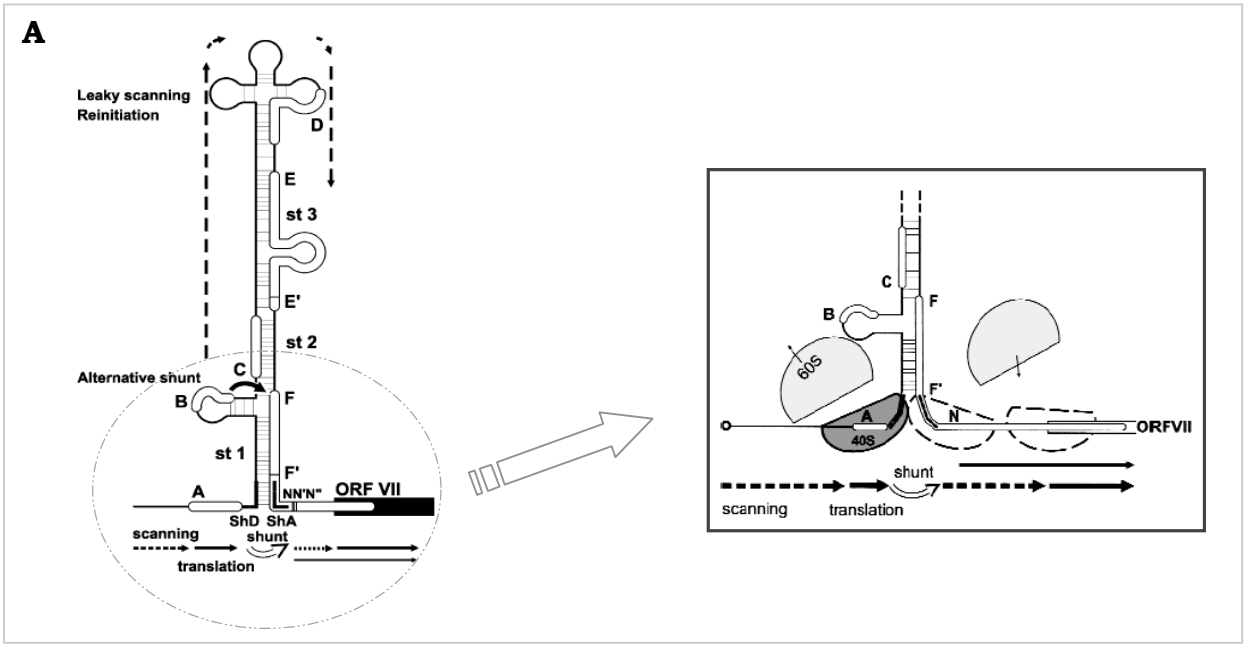
ribosome continue le scanning et reste compétent pour une ré-initiation au codon d'initiation suivant. Egalement durant les étapes précoces de la traduction, P6 participe au phénomène de ribosome shunt qui permet aux ribosomes ayant initié le scanning de l'ARN 35S au niveau de la coiffe d'effectuer un « saut » court-circuitant littéralement tout le scanning de la séquence 5' non traduite (leader), sur une longueur de plus de 500 nucléotides (voir éventuellement l'encadré 3 pour de plus amples détails).
Au-delà de son rôle déterminant dans le contrôle de la traduction au niveau des corps d'inclusion denses, P6 possède une activité de suppression du RNA silencing [23] mais si et seulement si, les deux signaux de localisation nucléaire (NLS) à l'origine d'un transport nucléo-cytoplasmique y sont préservés intacts [24]. Il est possible que P6 interfère avec la machinerie de PTGS (post-transcriptional gene silencing) au niveau du noyau puisque ces deux NLS sont essentiels pour l'infectivité du virus [24]. Un autre rôle important et encore hypothétique de P6 au niveau nucléaire serait le contrôle de l'épissage et de l'export de l'ARN 35S non-épissé vers le cytoplasme. Cette dernière possibilité semble cohérente avec l'existence, au sein de P6, de motifs protéiques de liaison aux ARN (simple et double brin) et d'export nucléaire (figure 2) [25]. En dirigeant l'ARN 35S non épissé vers le cytoplasme, P6 faciliterait : (i) la réplication du génome viral et (ii) la traduction des gènes I et II qui requièrent un ARN 35S non épissé.

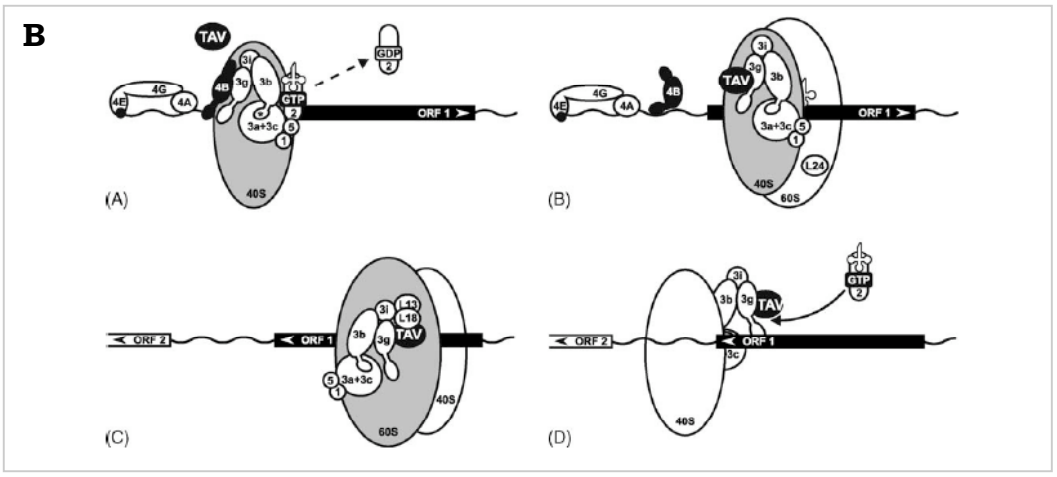
1.1.2.2 Protéines exprimées à partir de l'ARN 35S : P1 à P5

L'ARN 35S est le messager qui permet l'expression des ORFs I, II, III, IV et V. Bien que ces ORF se succèdent sur l'ARN 35S, le contrôle de leur cinétique d'expression reste totalement inconnu et ne semble pas suivre cette disposition linéaire (figure 1). Les protéines P1 à P5 seront décrites une à une en suivant leur cinétique d'apparition. Ce choix nous paraît plus réaliste ou judicieux car il permet d'appréhender au mieux la dynamique de l'infection au niveau cellulaire. Ceci nous amène à commencer avec les protéines P5 et P1 qui sont exprimées de manière précoce dans le cycle viral du CaMV, puis P3, P4 et enfin P2, puisque cette dernière semble s'accumuler de manière significative à la toute fin du cycle cellulaire.

Encadré N°3

Traduction de l'ARN 35S par saut de ribosome





A Représentation schématique des sauts de ribosome intervenant dans le leader du CaMV. La traduction de l'ARN 35S est initiée suite à un saut du ribosome durant lequel le complexe de lecture (scanning) est directement transféré d'un site « décollage » vers un site « atterrissage » sans lecture de la région intermédiaire. Cette zone est appelée séquence leader et possède 6 petits ORF (sORF) nommés de A à F. La structure secondaire de la région leader du CaMV suivi de l'ORF VII est figurée. Les traits continus noirs sur la structure indiquent les positions du site décollage (ShD) et du site atterrissage (ShA) pour le saut du ribosome. Les flèches figurent le déplacement des ribosomes au cours de leur scanning (recherche de site d'initiation ; flèches en pointillés), ou de la traduction (flèches pleines). L'initiation de la traduction peut avoir lieu au niveau de codons non-AUG (N, N', N"), mais la plupart des ribosomes initialisent au niveau de l'AUG de l'ORF VII. Stratégies d'initiation pour le CaMV : (i) mécanisme principal avec saut au niveau de la première tige (st1 ; arc de flèche ouvert) ; (ii) voie alternative par l'intermédiaire d'un saut à la seconde tige (st2 ; arc de flèche plein) ou alors (iii) sans aucun saut, par « leaky-scanning » : une migration linéaire des ribosomes le long de la séquence leader et ré-initiation (flèches noires discontinues). L'encart zoomé figure plus précisément la dissociation du complexe ribosomique (sous-unités 40S et 60S en grisées) suite à son passage sur le codon stop du premier cadre de lecture ouvert (sORF A) et leur réassociation préférentielle au niveau du codon AUG de l'ORF VII. *D'après [21, 22].*

B Modèle schématique de la fonction de transactivation de la traduction de l'ARN 35S polycistronique par P6 TAV. Les côtés externes des sous-unités ribosomiques (40S et 60S) sont représentées en grisé, et leurs côtés internes en blanc. Le modèle est vu de "face" avec le mouvement du complexe d'initiation, de gauche à droite pour (A) et (B), et de "dos" avec mouvement de droite à gauche pour (C) et (D). (A) eIF4B, de concert avec eIF4F (4E/4G/4A) ou eIFiso4F, interagit avec eIF3 lié au complexe de pré-initiation 48S. Le complexe scanne jusqu'à ce qu'il rencontre le premier codon start convenable durant l'étape de reconnaissance codon-anticodon. L'hydrolyse GTP dépendante d'eIF5 se produit et eIF2 est relarguée de la sous-unité 40S. EIF3 et ses sous-unités g de (3g), i, (3i), ter (3b), a et c (3a + 3c); eIF1 (1), eIF2 (2), eIF5 (5), ARNt et TAV sont représentées. (B) Au cours de la jonction avec la sous-unité 60S, eIF4B est déplacée du ribosome, alors que eIF3 est toujours associé au côté externe de la sous-unité 40S. TAV se lie au complexe eIF3/40S *via* la sous-unité EIF3 g, puis la traduction de l'ORF1 commence. L24 est positionné sur le côté interne de la sous-unité 60S. (C) Au cours du processus d'élongation, le complexe TAV/eIF3/80S est stabilisé, apparemment en partie par le transfert de TAV/eIF3 à la sous-unité 60S au travers d'une interaction de TAV avec L18 ou/et L13. (D) Le complexe TAV/eIF3 est vraisemblablement relocalisé sur la sous-unité 40S à la fin de l'ORF 1, ce qui est une des clés du succès de la réinitiation au codon AUG suivant. *D'après [22].*

1.1.2.2.1 P5

La protéine P5 est une polyprotéine de 78 kDa possédant une activité de protéinase à acide aspartique dans sa région N-terminale, et une activité de transcriptase inverse/RNAse H dans la région C terminale (figure 3). La protéinase (18 kDa) est relarguée par auto-clivage [26], et intervient ultérieurement dans la maturation de la protéine de capside P4 (voir paragraphe 1.1.2.2.3). Les fonctions de transcriptase inverse et de RNAse H quant à elles, interviennent lors de la réplication virale, pour synthétiser les nouveaux génomes à partir de l'ARN 35S pré-génomique (voir encadré 1B pour les différentes étapes de la réplication). La précocité de l'expression de P5 peut se comprendre par son implication dans la maturation du précurseur de la protéine de capside, aussi bien que par son rôle de réplicase pouvant immédiatement agir sur une fraction des ARN pré-génomiques.

1.1.2.2.2 P1

Le gène I code pour la protéine de mouvement du CaMV [27] composée de 327 acides aminés avec une masse moléculaire totale de 40 kDa. Cette protéine forme des tubules creux au travers des plasmodesmes, à l'intérieur desquels les virions migrent d'une cellule à une autre [28]. L'adressage de P1 en périphérie cellulaire (plus de détails à la section 1.2.1.1.2) est contrôlé par son domaine central [29], tandis que la formation des tubules requiert un très large domaine structural (acides aminés 1 à 282) [30] (figure 4). En outre, la partie C-terminale de P1 (acides aminés 295 à 327) qui est structurée en hélice α contrôle, via des interactions de type coiled-coil, (i) l'homo-trimérisation de la protéine et (ii) l'interaction *in vitro* avec le domaine N-terminal de P3 [31]. C'est précisément cette dernière interaction qui permettrait le(s) mouvement(s) intra et/ou intercellulaire des virions, la capside étant décorée de la protéine P3 (*cf.* encadré 4). De fait, les mutants de délétions de l'ORF III ou de la partie C-terminale de P1 sont incapables d'infecter une plante de manière systémique, bien qu'il se répliquent à l'intérieur de chaque cellule inoculée [30,32]. Enfin, P1 contient un domaine de fixation à l'ARN [33,34] entre les acides aminés 120 à 197 suggérant que l'ARN viral pourrait aussi transiter de cellule à cellule sous la forme d'un

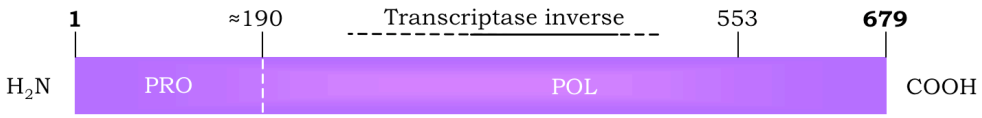


Figure 3 : Représentation schématique de la protéine P5. La protéine P5 est une polyprotéine de 78 kDa possédant respectivement les activités de protéinase à acide aspartique (PRO) et de polymérase (POL) dans les régions N et C terminales. La protéinase qui est libérée par auto-clivage intervient dans la maturation ultérieure de la protéine P4. La polymérase possède trois activités : (i) la transcriptase inverse qui est couplée à une activité ADN polymérase ADN dépendante et (ii) la RNAse H (motif DAS : aa 552-554) qui sont hypothétiquement séparées par un clivage supplémentaire. *D'après [26].*

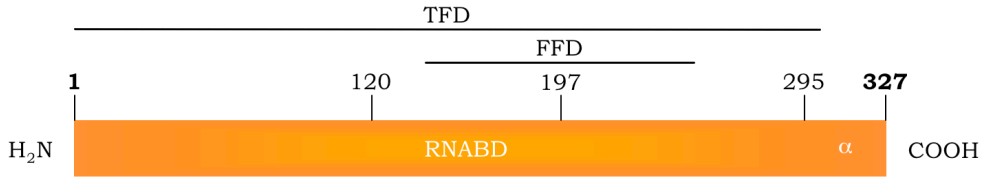


Figure 4 : Représentation schématique de la protéine P1. Constituée de 327 acides aminés pour un poids de 40 kDa, P1 est la protéine de mouvement du CaMV. Elle possède un domaine de fixation à l'ARN (RNABD) en son centre, et un domaine structuré en hélice α en C-terminal. Ce dernier permet une interaction de type coiled-coil avec la protéine P3. En outre, les domaines TFD et FFD, délimités par les traits horizontaux sont respectivement impliqués dans la formation des tubules au niveau des plasmodesmes et dans la formation de foci à la périphérie cellulaire. *D'après [31].*

complexe ribonucléoprotéique (voir encadré 5 [35]). Mis à part cette dernière possibilité, qui pourrait très tôt induire le mouvement d'un complexe ARN 35S-P1, la précocité de l'accumulation de cette protéine reste inexpliquée.

Les protéines suivantes (P4, P3 et P2) sont particulièrement importantes dans le cade de cette thèse puisqu'elles sont toutes trois impliquées dans la transmission du CaMV par pucerons-vecteurs.

1.1.2.2.3 P4

L'ORF IV du CaMV code pour le précurseur de(s) la protéine(s) de capside (P4) d'une masse moléculaire calculée de 57 kDa [36]. Cette protéine subit plusieurs clivages au niveau des domaines acides présents en N- et C-terminal, par l'intermédiaire de la protéinase à acide aspartique de P5 [26] et d'éventuelles autres protéases cellulaires non-identifiées, qui génèrent finalement trois polypeptides majoritaires de 42, 39 et 37 kDa (figure 5). Les polypeptides de 37 et 42 kDa sont les constituants majeurs de la capside [4], avec des proportions variables suivant les purifications virales. La capside du CaMV est formée de 420 sous-unités protéiques arrangées selon une symétrie d'ordre sept [37].

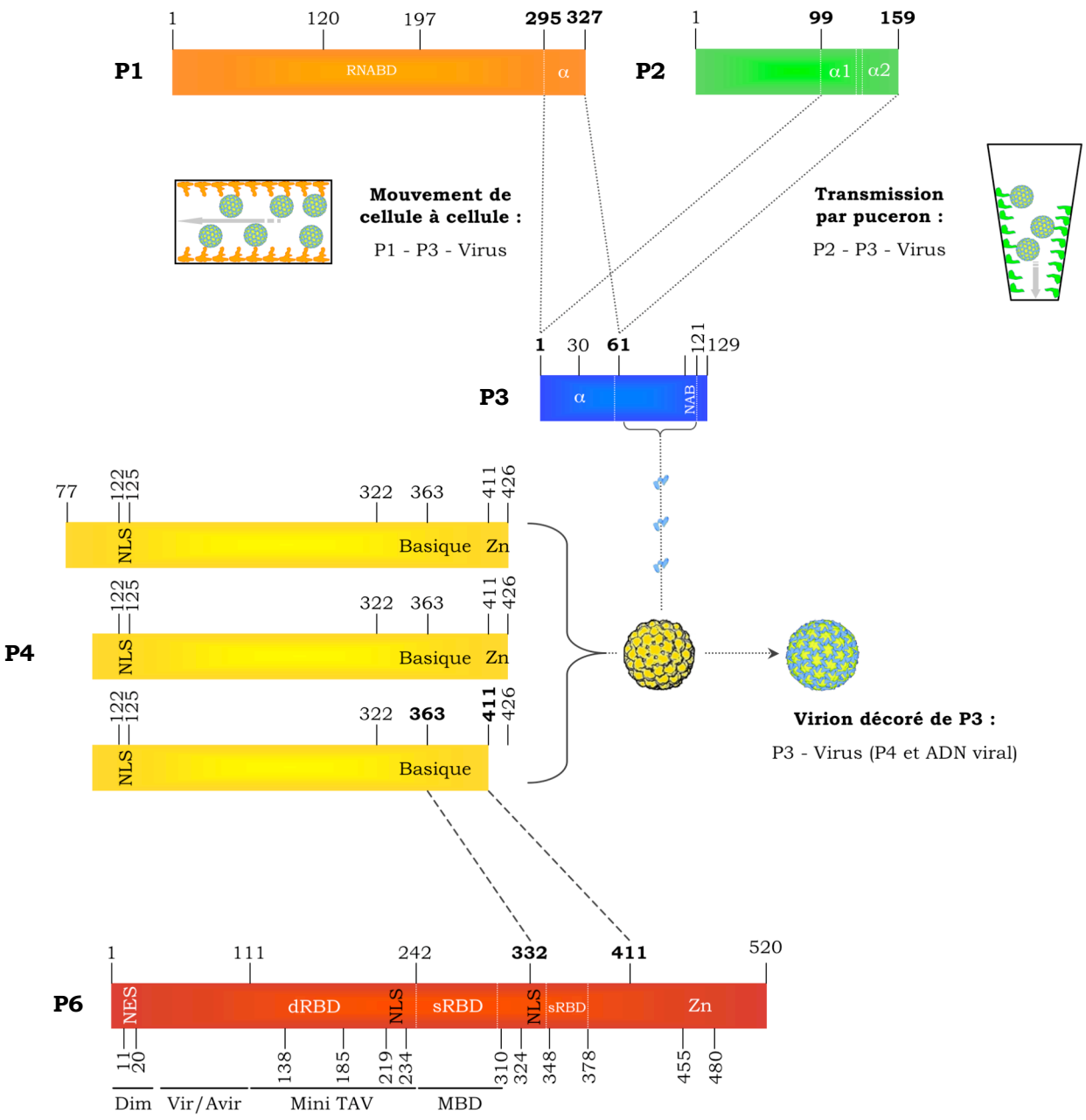
La protéine de capside subit deux autres types de modifications post-traductionnelles : (i) une glycosylation de ces polypeptides matures [38] dont le rôle biologique n'est pas connu et (ii) des phosphorylations aux extrémités N et C-terminales par une kinase de l'hôte (CKII : Casein Kinase II) qui sont indispensables à l'infectivité du virus [39,40].

Dans le domaine N-terminal de ces protéines de capsides, un NLS assure l'adressage des virions aux pores nucléaires, indispensable à l'infectivité du CaMV [41]. Il est à noter que la forme précurseur de 57kDa semble masquer ce NLS et inhibe le transfert vers le noyau. Ceci permettrait un contrôle de l'encapsidation des génomes viraux néo-synthétisés dans le cytoplasme à proximité des corps d'inclusions denses avant toute migration des virions [7].

La protéine P4 possède en C-terminal un motif « doigt de zinc » essentiel à l'infectivité, qui permet une interaction spécifique entre les différents polypeptides maturés (après un clivage minimal de la partie C-terminale

Encadré N°4

Interactions entre les protéines virales du CaMV



Les interactions entre différentes protéines du CaMV sont indiquées ci-dessus avec le code couleur suivant : P1 est représentée en orange, P2 en vert, P3 en bleu, P4 en jaune et P6 en rouge.

L'extrémité C-terminale de la protéine P3 interagit avec la particule virale – qui est composée des 3 produits résultant du clivage de la protéine P4 – alors que son extrémité N-terminale se lie à l'extrémité C-terminale des protéines P1 ou P2 afin d'opérer soit :

(i) au mouvement des virions via les plasmodesmes pour coloniser l'hôte (schématisation de ce mouvement dans un plasmodesme figuré ici en un rectangle)

(ii) à la transmission par puceron-vecteur suite à la rétention du complexe transmissible (P2-P3-Virus) au niveau de l' extrémité distale des stylets du vecteur (schématisée ici par un parallélépipède vertical) .

Par ailleurs, les protéines de capsides (P4) interagissent avec P6, dans le contexte des corps denses, via leur domaine basique en C-terminal,.

acide de la protéine de 57 kDa) et une région riche en purine présente au centre de la séquence leader de l'ARN 35S [42]. Cette association ARN 35S-protéine de capside semble nécessaire à la formation des particules virales, ce qui suggère que la réplication du CaMV, et donc la synthèse du génome ADN, est concomitante à l'encapsidation.
Enfin, un domaine riche en lysine (50% de lysine entre les résidus 363 et 411) permet une interaction de P4 avec P6 (encadré 4) [43], et ce même domaine est aussi impliqué dans la liaison aux acides nucléiques (ADN et ARN), indiquant un éventuel rôle additionnel de P6 vis-à-vis de l'encapsidation.

1.1.2.2.4 P3

La protéine P3 d'une masse moléculaire de 15 kDa (129 aa) est indispensable à l'infection systémique de la plante hôte [32], mais pas à la réplication du génome ni à la morphogenèse de particules virales dans un protoplaste transfecté [15]. Néanmoins, cette protéine est fréquemment appelée VAP pour « virion associated protein », car elle a d'abord été retrouvée avec les fractions enrichies en particules virales [44], puis identifiée sous la forme d'un réseau symétrique intimement associé à la particule virale (encadré 4) [45]. Dans ce réseau, la partie C-terminale sert d'ancrage dans la capside virale du CaMV, tandis que le domaine N-terminal en hélice α est exposé en surface. Il a été montré que c'est la région comprise entre les acides aminés 61 à 122 de P3 qui est responsable de l'interaction avec la capside virale (figure 6) [46]. Ce domaine est profondément enfoui dans la particule, et comporte une région riche en proline (aa 112 à 121) capable de fixer de manière non spécifique les acides nucléiques (figure 6) [32,47]. De multiples analyses par microscopie électronique et immuno-marquage semblent indiquer que les particules virales sont toujours associées à P3, où qu'elle se trouvent dans les plantes infectées. Ainsi, c'est ce complexe (illustré dans l'encadré 4) qui représenterait la forme mature et fonctionnelle de la particule du CaMV.
Le domaine N-terminal de P3, structuré en hélice α, s'associe via des interactions de type coiled-coil i) avec lui-même pour former un tétramère

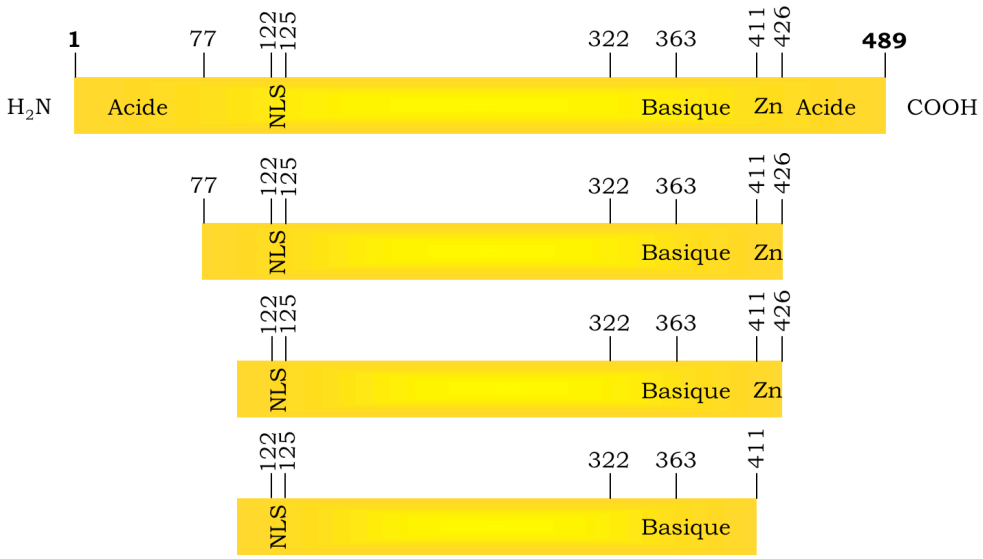


Figure 5 : Représentation schématique du précurseur de la protéine de capside du CaMV et de ses produits de maturation. Le précurseur de la protéine de capside (en haut) est clivé (entre autre, par la protéase à acide aspartique issue de P5) au niveau de ses extrémités acides N et C-terminales, afin de libérer les différents produits (p44, p39 et p37 de haut en bas) qui composent la capside des virions. Les domaines acide et basique, ainsi que le NLS et le motif en doigt de zinc (Zn) sont figurés. Les protéines de capside possèdent une activité de liaison aux acides nucléiques via un domaine aussi impliqué dans la dimérisation (aa 322-411). *D'après [40].*

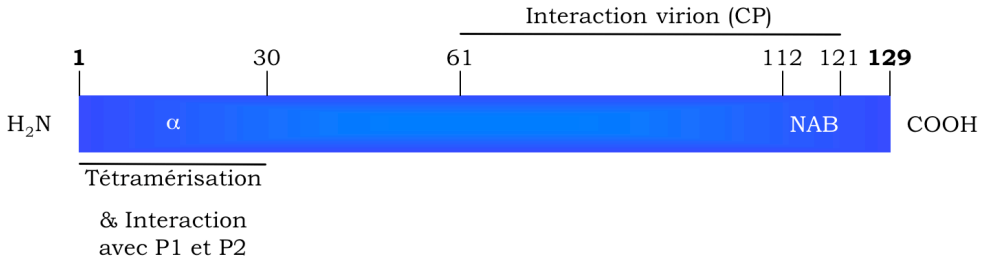


Figure 6 : Représentation schématique de la protéine P3. Comprenant 129 acides aminés pour un poids de 15 kDa, la protéine P3 est aussi dénommée VAP pour « Virion Associated Protein ». En effet, P3 interagit avec les virions par un large domaine de 60 acides aminés en C-terminal (trait noir supérieur) qui comporte aussi un motif pour la fixation aux acides nucléiques (NAB). Par ailleurs, les 30 premiers acides aminés de la protéine sont structurés en hélice α, ce qui offre à P3 la possibilité de former des homo-tétramères en solution ou bien d'interagir avec les protéines P1 ou P2 via des associations de type coiled-coil. *D'après [46].*

parallèle pour la forme libre de P3 [48,49], ii) avec la région C-terminale de P2, formant un complexe P2-P3-virion pour la transmission (voir ci-dessous) et iii) avec la partie C-terminale de P1, formant un complexe P1-P3-virion pour le mouvement *in planta* [31]. Le mécanisme spécifique contrôlant ces interactions particulières (encadré 4) et leurs possibles interférences est à ce jour totalement obscur.

Pour finir, une forme de P3 clivée en sa partie C-terminale (11kDa) par une protéinase à cystéine [50] est souvent retrouvée dans les extraits de plante infectée, cependant sa fonction biologique éventuelle n'est pas encore connue.

1.1.2.2.5 P2

L'expression de l'ORF II apparaît tardive, P2 ne s'accumulant qu'à la fin du cycle viral. Il s'agit d'une protéine non-structurale avec une masse moléculaire de 18 kDa [51]. La prédiction de sa structure secondaire montre deux domaines distincts séparés par une région repliée " au hasard " [52]: un domaine N-terminal en feuillet β et un domaine C-terminal (aa 99 à 159) composé de deux hélices α susceptibles de former une structure super-enroulée (coiled-coil) (figure 7 [53]). Blanc *et al.* [54] ont apporté la preuve directe que la protéine P2 est le facteur assistant de la transmission (FAT ou plutôt HC pour « helper-component ») car elle peut permettre la transmission par pucerons d'une souche délétée du gène II qui est naturellement non transmissible. Diverses expériences ont montré que P2 interagit d'une part avec le complexe P3-virus (via une interaction entre les hélices α des protéines P2 et P3), et d'autre part avec un site protéique spécifique au niveau de la cuticule des stylets du puceron-vecteur [55,56] (encadré 9A). Ces deux interactions créent un pont moléculaire entre la particule virale et les stylets du vecteur, qui est absolument indispensable au succès de la transmission [57,58] (encadré 4).

Il est connu de longue date que P2 n'est pas requise pour le déroulement du cycle infectieux dans la plante [12], ce qui fait de cette protéine un facteur viral spécialisé dans la transmission par vecteur [46]. En effet, P2 produite au niveau des corps d'inclusions denses migre (avec une partie de P3) à

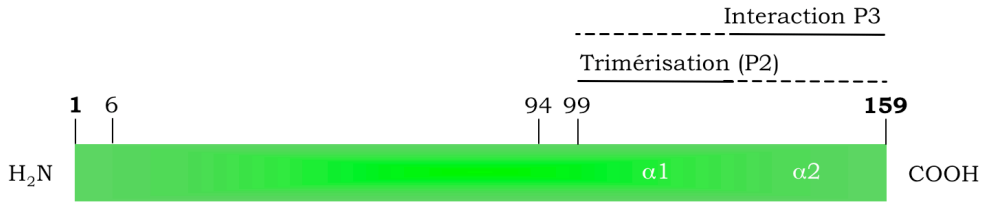


Figure 7 : Représentation schématique de la protéine P2. Constituée de 159 acides aminés pour un poids de 18kDa, la protéine P2 est le Helper Component du CaMV. Elle possède dans sa partie C-terminale deux régions structurées en hélice α (α1 : 99-128 ; α2 : 137-159) qui lui permettent d'interagir soit avec elle-même pour former majoritairement des trimères (*via* α1 essentiellement) ou soit avec P3 (*via* α2 très probablement). Les acides aminés 6 et 94 sont respectivement impliqués dans l'interaction avec le récepteur des pucerons-vecteurs et dans la formation du corps clair. *D'après [53, 55, 56 et 61].*

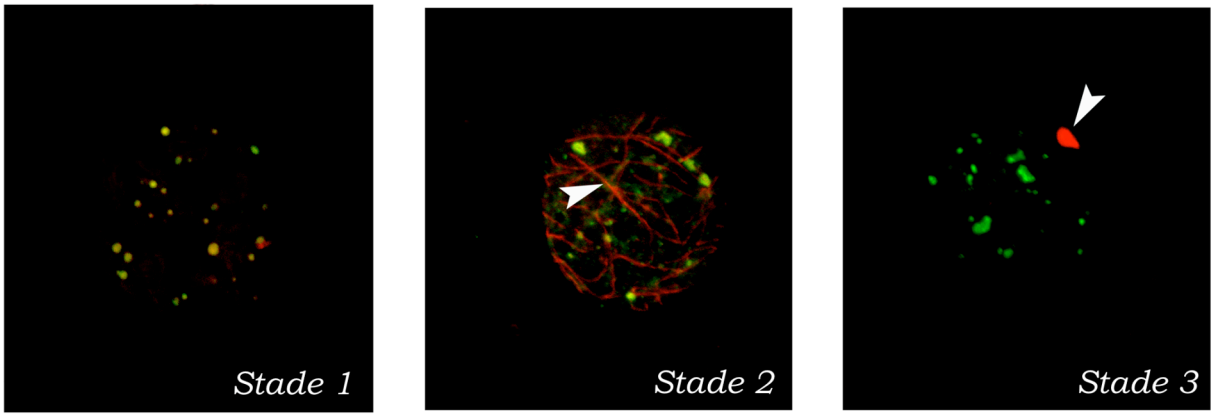


Figure 8: Le corps clair se forme en drainant (accumulant) la P2 de toute la cellule infectée. Sur ces photos, la protéine P6 (marqueur des corps denses) est immunomarquée en vert tandis que P2 est immunomarquée en rouge. Au stade 1, les protéines P2 et P6 co-localisent au niveau des corps denses (couleurs superposées = orange). Au second stade, la P2 qui a été produite au niveau des corps denses, semble les quitter et forme un réseau sur les microtubules (réseau signalé par la flèche blanche). Enfin au stade 3, P2 est accumulée dans une structure cellulaire unique, le corps clair (flèche blanche), et ne co-localise plus avec les corps denses. *D'après [59].*

l'aide du réseau de microtubules [59] (encadré 1C et figure 8) pour donner lieu à la formation d'un unique corps d'inclusion clair par cellule [60], à la structure extrêmement complexe dont la présence est essentielle et exclusivement dévouée à la transmission par vecteur [61] (voir section 2.3.1. pour plus de détails sur le corps clair).

1.2 Colonisation de la plante hôte

Lorsque le virus a terminé son cycle de réplication au sein d'une cellule hôte, il lui est essentiel de partir coloniser les cellules avoisinantes afin d'accroître sa descendance, d'envahir de nouveaux tissus et de rejoindre la sève élaborée qui permet à l'infection de se généraliser à l'ensemble de la plante et d'augmenter ses chances de transmission par vecteur.

1.2.1 Mouvement du CaMV depuis son site d'inoculation

Quel que soit le mode d'inoculation, naturel par un puceron virulifère ou bien expérimental par abrasion mécanique, les points d'entrée (ou sites d'infection primaire) sont généralement visualisables sous la forme de petites lésions locales chlorotiques. À partir de chaque point d'entrée, le virus va se répliquer, puis migrer de cellule à cellule, avant de gagner les faisceaux vasculaires, qui lui permettront un mouvement à longue distance.

1.2.1.1 Mouvement de cellule à cellule via les plasmodesmes

Les plasmodesmes (PD) ont été décrits en 1930 [62] comme de simples ponts cytoplasmiques reliant les cellules entre elles, définissant ainsi un domaine nommé symplasme. En fait, les plasmodesmes des plantes supérieures sont des complexes supramoléculaires constitués de membranes et de protéines. Sur le plan fonctionnel, ils sont capables de contrôler le seuil d'exclusion limite (SEL), pour la diffusion intercellulaire des métabolites et petites molécules, ainsi que de renforcer et de réguler le trafic intercellulaire des macromolécules telles que les acides nucléiques et les protéines [63]. Ainsi, de multiples événements essentiels au développement et au maintien d'une plante sont dépendants du trafic de molécules *via* les plasmodesmes.

1.2.1.1.1 Rappel sur les différents modes de passage de cellule à cellule chez les phytovirus

Les phytovirus disposent selon les espèces de 1 à 3 protéines dédiées à la régulation du mouvement dans la plante. Plusieurs mécanismes très divergents ont été décrits pour assurer les migrations d'unités infectieuses [64-67], certains de ces mécanismes pouvant co-exister au sein d'une même espèce virale (encadré 5).

- Passage de complexe MP-acide nucléique :

C'est le mode utilisé par le *Tobacco mosaic virus* (TMV, et d'autres espèces moins étudiées) avec un passage de cellule à cellule sous la forme un complexe ribonucléoprotéique composé de la protéine de mouvement (MP) et l'ARN génomique viral (pour revue voir [68]). Dans ce complexe, la molécule d'ARN est étirée en un brin fin de 2 nm de diamètre environ, structure plus adaptée au passage entre les fibres cloisonnant le cytoplasme dans le PD. En plus de fixer les acides nucléiques, il a été clairement démontré que les MP de certains virus – tobamo, tobra, dianthovirus – sont capables d'augmenter le SEL des plasmodesmes. A ce jour, aucune observation en microscopie électronique n'a pu montrer la présence de particules virales de TMV passant d'une cellule à une autre au travers des plamosdesmes.

- Passage de complexe MP-CP-acide nucléique :

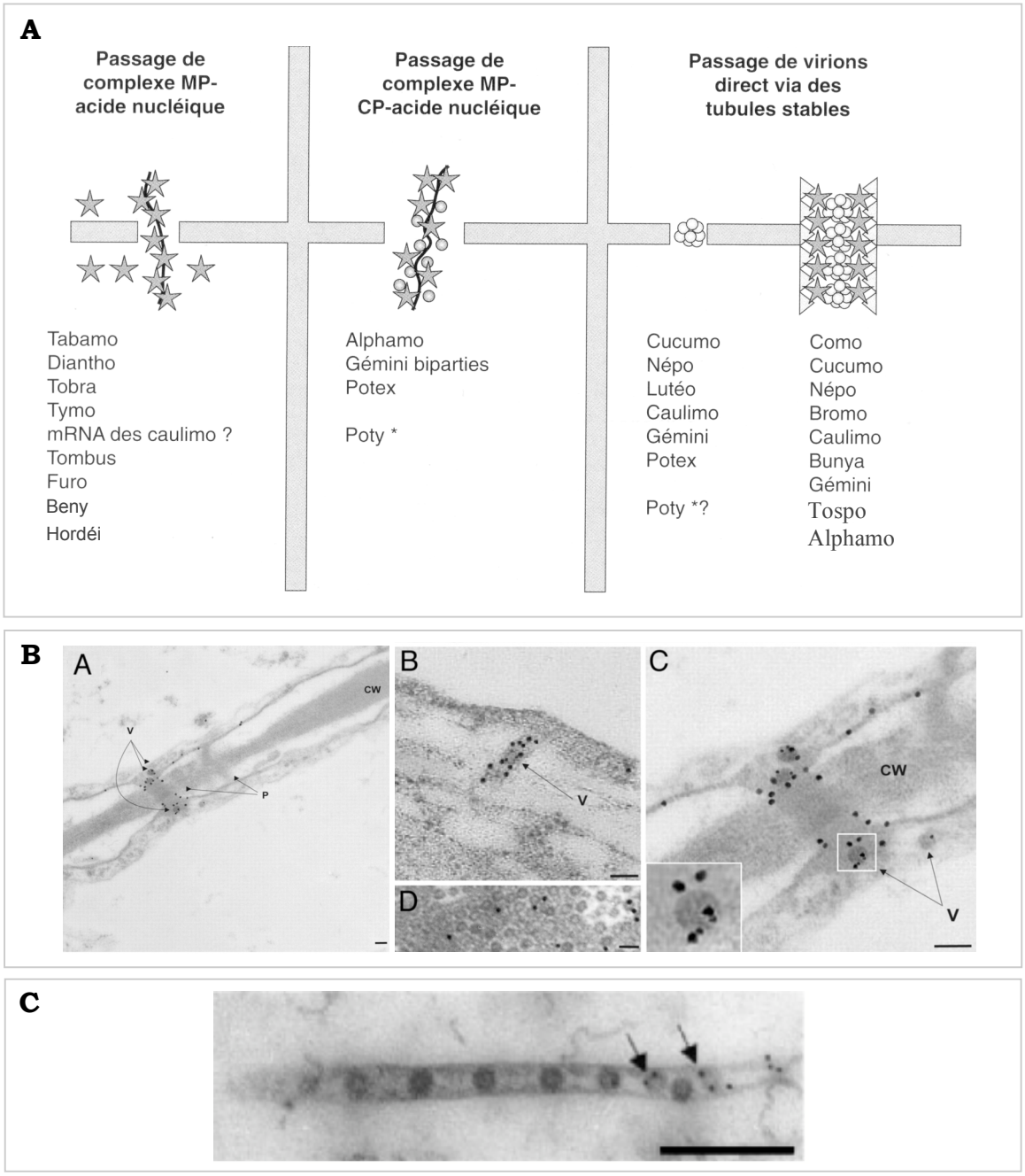
Ce mode de transport implique en plus de la (ou des) MP, la présence de la protéine de capside (CP) pour assurer le passage du matériel infectieux, sous la forme d'un complexe MP-ARN-CP sans toutefois que l'encapsidation du génome et donc la présence d'une capside mature ne soit requise. Pour faciliter le passage de ces complexes certaines MP (des alphamo, gémini, potex et potyvirus) peuvent accroîtrent le SEL des plasmodesmes. Le virus type pour ce mode de transport est l'*Alfalfa mosaic virus* (AMV), un alphamovirus appartenant à la famille des *bromoviridae* [69].

- Passage de particule virales matures (pour revue voir [70]) :

Ce type de passage correspond à la migration de génomes viraux encapsidés (voir encadré 5). Pour certaines espèces, dont les lutéovirus, les virions migrent tels quels de cellule à cellule au travers des PD. On peut alors observer directement les particules virales traversant les parois au niveau

Encadré N°5

Le mouvement de cellule à cellule chez les phytovirus



A Modélisation des différents mode de passage de cellule à cellule des phytovirus. Au travers des plasmodesmes, représentés ici par de simples trous dans la paroi (rectangles gris entouré de la membrane plasmique en noir) sont représentés : la protéine de mouvement (étoile), le génome viral (trait noir), la protéine de capside (disque gris), le virion (amas de disques blancs) et les éventuelles protéines cellulaires (triangle). Les listes au-dessous des schémas mentionnent les genres viraux concernés. Certains virus, tels que les Alphamo- et les Potexvirus, sont d'ailleurs capables de passer de cellule à cellule selon plusieurs modes. (*) la protéine HC-Pro et la capside des potyvirus jouent le rôle de protéine de mouvement. (?) suspicion faute de preuves suffisantes. *D'après [35].*

B Cas du CaMV : P3 et P1 co-localisent au niveau des plasmodesmes modifiés. Microscopie électronique sur coupes ultrafines de tissu infecté montrant deux cellules adjacentes. Le marquage immunogold de P3 (billes d'or de 6-nm) et de P1 (billes d'or de 18-nm) permet la localisation de chacune des protéines. (A) Deux plasmodesmes (P) traversant la paroi cellulaire (CW) sont indiqués ; l'un est modifié et contient des particules virales (V). (B) Coupe longitudinale d'un plasmodesme modifié comportant des particules virales (V) et montrant la distribution de P1. (C) Zoom sur le plasmodesme modifié vu en A ; un agrandissement dans l'encart permet de visualiser la localisation de P1 et P3 sur une particule virale. (D) Corps d'inclusion dense (edIB) contenant des particules virales et du marquage spécifique à P3 (billes d'or de 6-nm). Barre : 100 nm. *D'après [31].*

C Cas du CaMV : Passage de virions via des tubules stables. Image en microscopie électronique d'une coupe au niveau d'un plasmodesme (probablement entre une cellule compagne et une cellule du phloème). On peut voir la présence de virions passant d'une cellule à l'autre au travers d'une structure tubulaire. Les flèches noires indiquent P3 immunomarquée (couplage à des billes d'or de 10nm). Barre, 250 nm. *D'après [132].*

des PD en microscopie électronique. Pour d'autres espèces, dont l'exemple type est le *Cowpea mosaic virus* (CPMV, comovirus), des structures tubulaires apparaissent et traversent les parois au niveau des PD. Ces tubules très stables sont constitués en tout ou partie par la MP [71,72], et servent de « tunnel » dans lesquels s'engagent des « chapelets » de particules virales. Ce phénomène a été observé chez de nombreuses espèces virales (dont le CaMV) réparties dans des genres très disctincts : Caulimo, Como, Gémini, Népo et Tospovirus.

Parmi les protéines virales, la formation des tubules ne nécessite que la présence de la (ou des) MP : de telles tubules ont été obtenues dans des cellules de plantes transgéniques exprimant la (ou les) MP de plusieurs des virus cités ci-dessus, et ceci en l'absence d'infection virale. Il est par contre possible, et encore très peu documenté, que un ou plusieurs facteurs cellulaires participent à la formation de ces tubules.

Il faut cependant noter que, pour certains virus particuliers, l'attribution d'une ou plusieurs de ces catégories de mouvement reste ambiguë. C'est le cas de l'*Alfafa mosaic virus* (AMV), un alphamovirus dont la CP est nécessaire au mouvement de cellule à cellule, mais l'encapsidation n'est pas requise. Dans les protoplastes de *Vigna unguiculata* ou de *Nicotiana benthamiana* infectés, il y a formation de tubules remplis de particules virales. Par contre, dans les tabacs infectés, de rares tubules courts sont visibles exclusivement au front de virose, où les particules virales ne sont pas visibles bien que ces structures soient marquées par un anticorps dirigé contre la CP [73]. Dans ce cas particulier, les tubules ne seraient donc pas des structures stables comme dans le cas des Como et Népovirus, mais plutôt des structures très transitoires, avec peut-être un rôle spécifique encore inconnu.

Dans le cas des potyvirus, il reste à démontrer si le matériel infectieux migre sous forme d'un complexe CP-HCPro-ARN ou sous forme de virions associés à HC-Pro et peut-être à la protéine CI [74-77].

1.2.1.1.2 Cas particulier du CaMV

Il a été confirmé que les virions du CaMV passent de cellule à cellules via les plasmodesmes grâce à des tubules formés par la protéine de mouvement P1

[28], bien que ces tubules ne soient pas toujours directement observables (encadré 5B). Des études ont montré que P1 interagit avec au moins un partenaire cellulaire d'*Arabidopsis thaliana* (MPI7, une protéine membranaire associée aux vésicules (VAMP)) [78] et que sa localisation finale au niveau des plasmodesmes ne requiert pas le réseau de microtubules, mais un système endomembranaire fonctionnel [79]. Ainsi, d'après les interactions protéiques connues chez le CaMV, il est possible de proposer l'hypothèse que les complexes P3-virus sont pris en charge par la protéine de mouvement P1 (à un moment du cycle viral et en un lieu de la cellule non déterminés) afin d'être acheminé, au travers des tubules, dans les cellules adjacentes.
Par ailleurs, du fait que P1 et P4 possèdent un domaine de liaison à l'ARN, il est impossible d'exclure les éventuels passages de complexes MP-acides ribonucléiques (P1-ARN) ou de complexes MP-CP-acides ribonucléiques (P1-P4-ARN) *via* les plasmodesmes.

1.2.1.1.3 Cheminement par les plasmodesmes jusqu'aux tubes criblés du phloème

Les plasmodesmes se différencient (modifications du nombre et de la structure) de manière spécifique en fonction du domaine symplasmique auquel ils appartiennent et au cours du temps lors de la transition source-puits [80]. Ainsi, les plasmodesmes connectant les cellules du parenchyme périvasculaire au parenchyme phloémien ont des propriétés différentes de ceux qui interconnectent les cellules du mésophylle [67]. De même, dès le début de la différenciation cellulaire, des plasmodesmes spécialisés se forment entre les cellules destinées à devenir des cellules compagnes et les tubes criblés. Cette spécialisation correspond à un élargissement du canal de communication côté tube criblé et par la formation de ramifications plus étroites du côté des cellules compagnes [81,82]. Cette asymétrie au niveau des tissus sources où se situent les cellules exportatrices est à l'origine du transfert directionnel des sucres vers les tubes criblés.
Ces particularités anatomiques des plantes font que certains virus peuvent être ralentis par ces obstacles (« frontières » tissulaires) durant leur progression systémique, ou bien même complètement bloqués et restreint à un tissu particulier. En conséquence, la migration des virus depuis le

A

Days post-inoculation leaf removed	Systemic symptoms after leaf removed	Visible symptoms (chlorotic lesions) on inoculated leaf
0	−	−
1	−	−
3	−	−
5	+	−
7	+	+
9	+	+
11	+	+
13	+	+
15	+	+
17	+	+

B

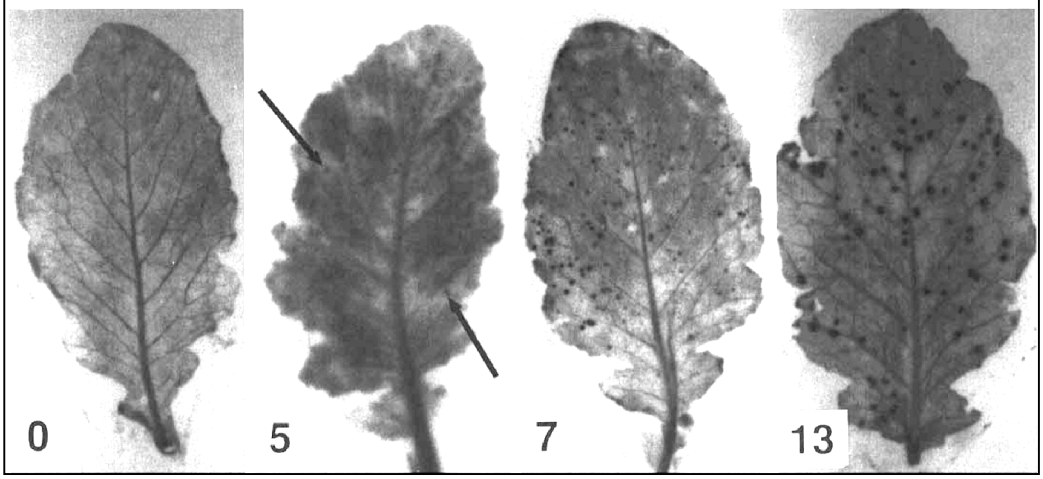


Figure 9 : Durée nécessaire pour la colonisation des tissus vasculaires dans une feuille de navet inoculée par le CaMV. Un lot de navet a été infecté mécaniquement au niveau d'une seule feuille par plante. Ensuite, ces feuilles inoculées ont été coupées et éliminées à des dates précises, et l'apparition ultérieure de symptômes a été surveillée sur les feuilles systémiques. (A) Tableau récapitulatif : les « + » signalent la présence de symptômes visibles sur les feuilles systémiques (1ère colonne) à 32 dpi et sur les feuilles inoculées (lors de leur élimination) au cours du temps. Le CaMV initie son mouvement à longue distance entre le 3ième et le 5ième jour après l'inoculation. Toutefois, la feuille inoculée ne montre pas de symptômes visibles avant le 7ième jour après son inoculation. (B) L'utilisation de l'hybridation moléculaire à l'aide d'une sonde d'ADN viral marquée au ^{32}P a permis grâce à sa sensibilité de mettre en évidence l'apparition des premiers foyers d'infection (signalés par les flèches) dès le 5ième jour.
D'après [83].

mésophylle vers le système vasculaire au travers des différents tissus ne se fait pas à une vitesse constante. Pour le CaMV, il a été montré qu'il faut un minimum de 3 à 5 jours sur plant de navet pour que le virus sorte de la feuille inoculée (figure 9) et initie une infection systémique [83], sans que le détail des vitesses de passage dans le mésophylle, dans les tissus périvasculaires et vasculaires au sein de cette feuille n'ait pu être analysé.

1.2.1.2 Mouvement à longue distance et distribution du virus dans la plante

La majorité des virus transmis par pucerons utilisent généralement le phloème pour leur transport à longue distance, mais il existe au moins une exception avec le *Blueberry shoestring virus* (BSSV) puisqu'il est détecté au niveau du xylème et du phloème [84]. En outre, plusieurs phytovirus transmis par coléoptères ou par champignons migrent à longue distance en empruntant le xylème [64,85]. Pour revue sur le transport systémique des virus voir [86].

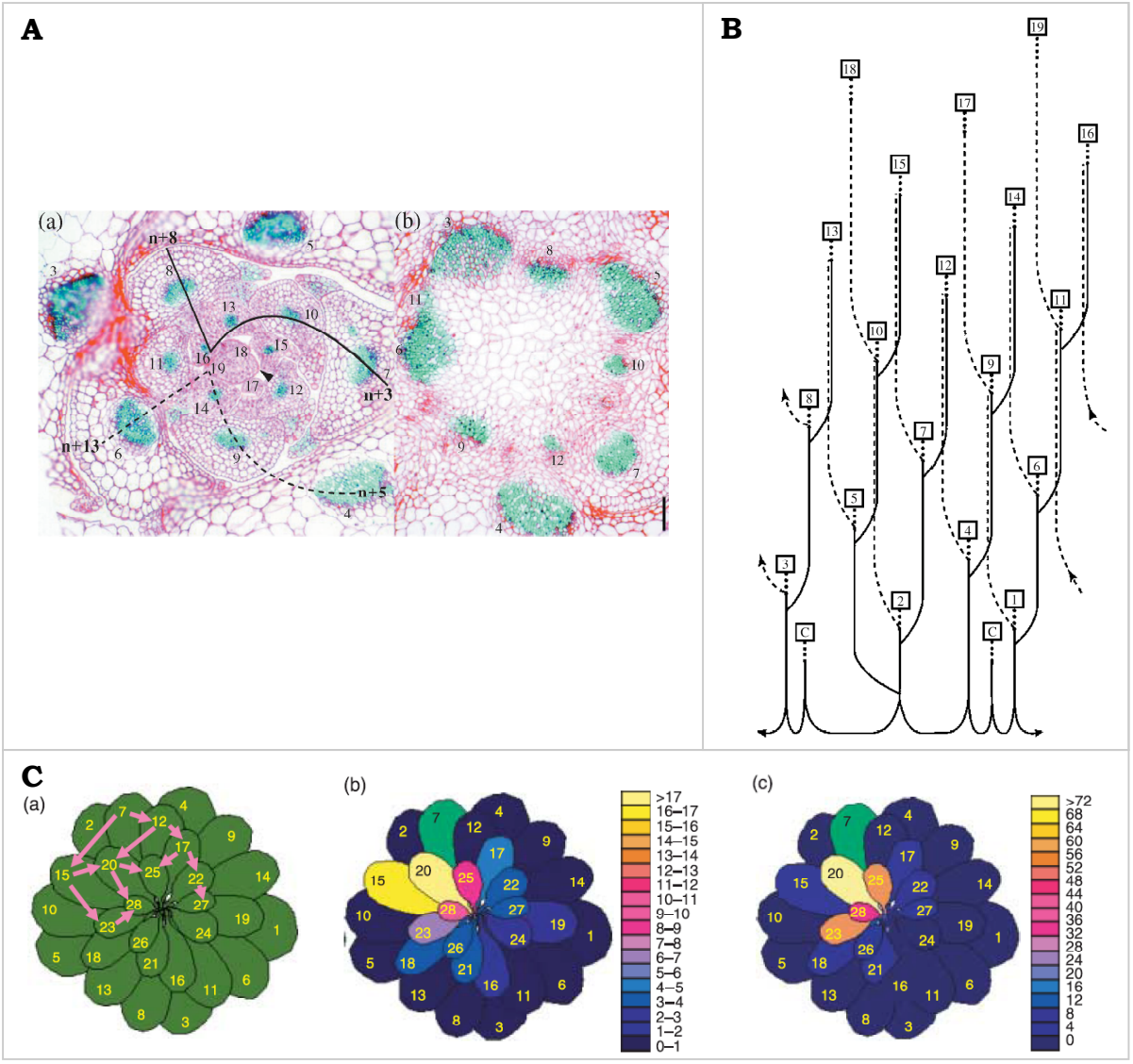
Le mouvement à longue distance du CaMV s'effectue *via* le phloème [83] sous la forme très probable de complexes P3-virions [58] qui sont chargés dans les tubes criblés à partir des cellules compagnes. Ensuite, la distribution de ces complexes P3-virions s'effectue selon le flux des photo-assimilats en fonction de la phyllotaxie [83]. C'est-à-dire qu'ils vont migrer à partir d'une feuille mature (source) pour coloniser principalement les jeunes feuilles (puits) qui sont à proximité immédiate en regard de l'insertion de leur pétiole. En d'autres termes, lorsque le virus emprunte la voie phloémienne, il se retrouve surtout emporté vers les feuilles puits (en croissance) qui présentent une connexion vasculaire directe avec la feuille source considérée [87]. Ainsi, pour les plantes en rosette le transport à longue distance s'effectue essentiellement en suivant les parastiches (motifs en spirale qui regroupent des feuilles non successives) (encadré 6, [88]).

1.3 Symptomatologie & Interaction avec l'hôte

Selon les souches virales du CaMV, les plantes utilisées (espèces ou écotypes), et les conditions environnementales, les symptômes peuvent varier [89,90]. Dans certains cas, l'infection peut même échouer en fonction

Encadré N°6

Le mouvement à longue distance selon le flux des photoassimilats



A Relation entre phyllotaxie et profil vasculaire au niveau du bourgeon végétatif apical d'*Arabidopsis thaliana*. (a) coupe transversale au niveau du méristème apical dans des Arabidopsis transgéniques avec un marquage spécifique du procambium (ATHB-8::GUS). La rosette de feuille est numérotée en spirale selon l'ordre de la chronologie ontogénique. Les lignes continues indiquent les contacts entre les parastiches orientés dans le sens des aiguilles d'une montre (n+3 et n+8) tandis que les lignes pointillées correspondent aux parastiches tournant dans le sens inverse des aiguilles d' une montre (n+5 et n+13). (b) Coupe transversale 66 µm sous le méristème apical. Les faisceaux vasculaires situent l'insertion individuelle des feuilles (3, 4, 5, 6 et 7) et des sympodia qui donneront naissance aux feuilles (8, 9, 10, 11, 12). Barre : 50 µm. *D'après [88].*

B Diagramme représentant les connexions vasculaire au sein de l'appareil végétatif d'*Arabidopsis thaliana*. Le système vasculaire est figuré en deux dimensions, comme si la rosette était déroulée à partir d'une fente pratiquée entre les feuilles 1 et 3. Le numéro de chaque feuille est placé dans un carré ; les cotylédons sont mentionnés de manière identique par des C . Les lignes continues représentent les sympodia vasculaires qui relient les feuilles selon le parastiche n+5, et les connexions entre les feuilles à l'extrémité distale de l'hypocotyle. Les lignes discontinues correspondent aux sympodia vasculaires qui relient les feuilles selon le parastiche n+8. Les pointillés symbolisent l'insertion des feuilles au-dessus du point de branchement. *D'après [88].*

C La distribution du CaMV se superpose au profil établi selon les relations source-puits. (a) schéma d'*Arabidopsis thaliana* avec 28 feuilles en rosette numérotées selon leur apparition. Les flèches indiquent les connexions vasculaires directes (n+5 et n+8) à partir de la feuille n°7. (b) Patron de l'allocation des sucres déterminé à partir de mesures d'accumulation de ^{14}C 5 jours après avoir appliqué la source de sucre marqué sur la feuille n°7 (en vert). Valeurs (moyennes sur 3 plantes) représentées pour chaque feuille suivant un code couleur (barre à droite) selon le pourcentage total de radioactivité et après normalisation par rapport à la surface foliaire totale. (c) Le profil d'apparition des symptômes après 31 jours d'infection. Les feuilles présentent un score selon l'ordre d'apparition des symptômes (4 pour la 1ère, 3 pour la 2nde, 2 pour la troisième et 1 pour la quatrième). Le schéma synthétise les résultats cumulés pour 27 plantes symptomatiques. Les feuilles sont colorées selon leur score cumulé en fonction du code couleur figuré sur la droite. La feuille inoculée (n°7) qui présente des lésions locales est colorée en vert. *D'après [87].*

du stade de développement de l'hôte au moment de l'inoculation. De plus, il a récemment été montré que le mouvement à longue distance du CaMV est dépendant de la voie de défense de l'éthylène [91]. De nombreuses interactions spécifiques entre le CaMV et ses plantes hôtes ont été découvertes et étudiées au cours des quinze dernières années.

1.3.1 Symptomatologie et Pathogénicité

Bien que les symptômes provoqués par le CaMV varient en fonction de l'hôte, la mosaïque et la marbrure des feuilles, l'éclaircissement des nervures et le long des nervures, ainsi que le rabougrissement et la déformation des plantes en constituent des caractéristiques principales très constantes.

1.3.1.1 Symptômes et accumulation du virus dans le navet

Dans un premier temps, on peut observer des taches chlorotiques (ou lésions locales) sur les feuilles inoculées, puis l'infection se généralise à toutes les feuilles néo-formées de manière systémique. Les premières feuilles qui montrent des symptômes systémiques développent très souvent un éclaircissement des nervures qui se limite à une petite région basipétale sur un seul côté du limbe [83]. Cette sectorisation s'expliquerait par le fait que ces feuilles – au moment où le virus arrive *via* le phloème – sont dans une phase de transition puits/source et que le CaMV ne pourrait donc pas remonter à contre-courant dans le phloème pour envahir un tissu source (figure 10).

Dans ces premières feuilles aux symptômes systémiques, le virus se concentre surtout le long des nervures (veines) et des veinelettes en comparaison avec les territoires qui se situent entre les veines. Par contre, dans les jeunes feuilles néoformées infectées de manière systémique, le virus s'accumule préférentiellement dans les îlots verts interveinaux plutôt que le long des nervures chlorotiques [92]. Des hybridations moléculaires sur plantes infectées suggèrent que le CaMV est présent (à des concentrations variables) dans la plupart des tissus végétatifs, mais on le retrouve de manière prépondérante dans les tissus vasculaires des pétioles, de la tige, de l'hypocotyle et des racines [92].

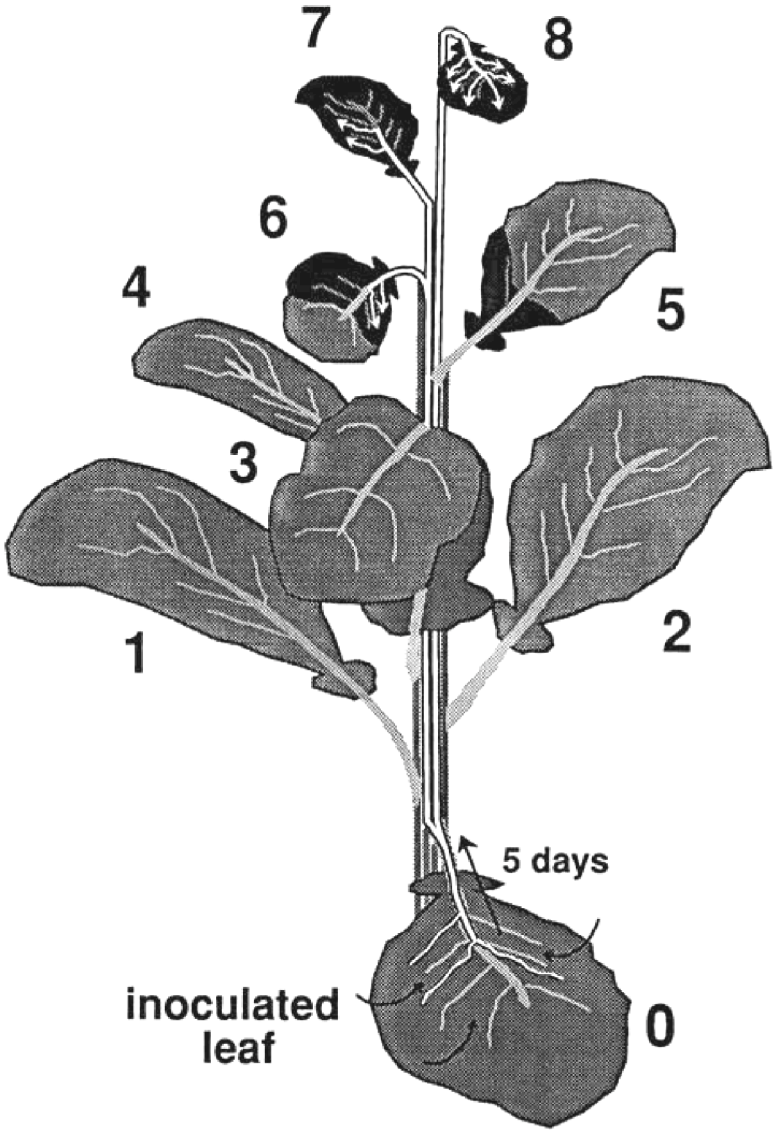


Figure 10 : Schéma d'un navet indiquant les paramètres qui influencent l'accumulation bilatérale et basipétale du CaMV dans les feuilles systémiques. La feuille inoculée est numérotée 0. Les lignes blanches indiquent les faisceaux vasculaires par lesquels le virus gagne les jeunes feuilles systémiques à partir de la feuille inoculée (au 5ième jour après inoculation). L'état puits ou source des feuilles est figuré par les couleurs (les feuilles en transition comportent donc 2 couleurs) : le gris correspond aux régions matures (source) produisant les photoassimilats, tandis que le noir symbolise les régions qui sont puits. La transition puits-source d'une feuille s'effectue progressivement à partir de sa partie distale, d'où l'accumulation de virus basipétale souvent observée sur les premières feuilles systémiques symptomatiques (feuilles 5 et 6). Par ailleurs, les jeunes feuilles (6 et 7) dont le point d'insertion du pétiole n'est pas adjacent à celui de la feuille inoculée (parastiches distincts), accumulent seulement le virus sur la partie de leur limbe la plus proche (flèches blanches) de la source virale (feuille 0). *D'après [83].*

1.3.1.2 Gamme d'hôte et Pathogénicité

Le CaMV a une gamme d'hôte assez étroite puisqu'il n'infecte généralement que la famille des *Brassicaceae*. Toutefois, certaines souches du CaMV (W260, S-Japan, NY8153, D4 et B29 [93]) sont capables d'infecter systémiquement quelques solanacées des genres *Datura* et *Nicotiana*, tandis que d'autres souches (comme CM1841) y induisent les systèmes de défense et restent exclusivement confinées à la feuille inoculée [1]. La souche W260 induit une réponse hypersensible (HR) sur les feuilles inoculées de *Nicotiana edwarsonii*, alors qu'elle déclenche la mort cellulaire systémique après les premiers symptômes sur *Nicotiana clevelandii*. Par contre, la souche D4 produit des symptômes systémiques non nécrotiques sur ces deux hôtes. Il existe aussi une situation particulière pour quelques espèces de brassicacées (par exemple : *Brassica napus* qui est un amphidiploide ; voir figure 11, [94]), où le virus se développe de manière systémique avec des symptômes caractéristiques, avant que la plante ne récupère un phénotype « sain » sans symptômes apparents. Ce phénomène est connu sous le nom de « récupération ».

L'analyse des symptômes produits par des CaMV chimériques (réalisés à partir de souches provoquant des symptômes distincts) sur différentes plantes a révélé que l'ORF VI est le déterminant majeur responsable de la spécificité d'hôte et de la sévérité des symptômes [95-98]. Par la suite, des plantes transgéniques d'*Arabidopsis thaliana* exprimant P6 sauvage ou des P6 chimériques ont réaffirmé le rôle de l'ORF VI dans la sévérité des symptômes [18,99]. Une étude du transcriptome d'*Arabidopsis thaliana* transgénique exprimant P6 a mis en évidence la surexpression ou la répression de plus de 30 gènes, et ces modifications sont très similaires à celles observées dans une plante infectée par le CaMV [100]. Il doit exister un mécanisme permettant ce dialogue (interactions directes ou indirectes) entre la P6 virale et la plante qui est à l'origine de l'apparition des symptômes.

Le produit de l'ORF VI est aussi responsable du déclenchement de la réponse hypersensible, dans les associations souche/hôte où elle a été décrite, et peut donc être considéré comme un gène d'avirulence [101]. En

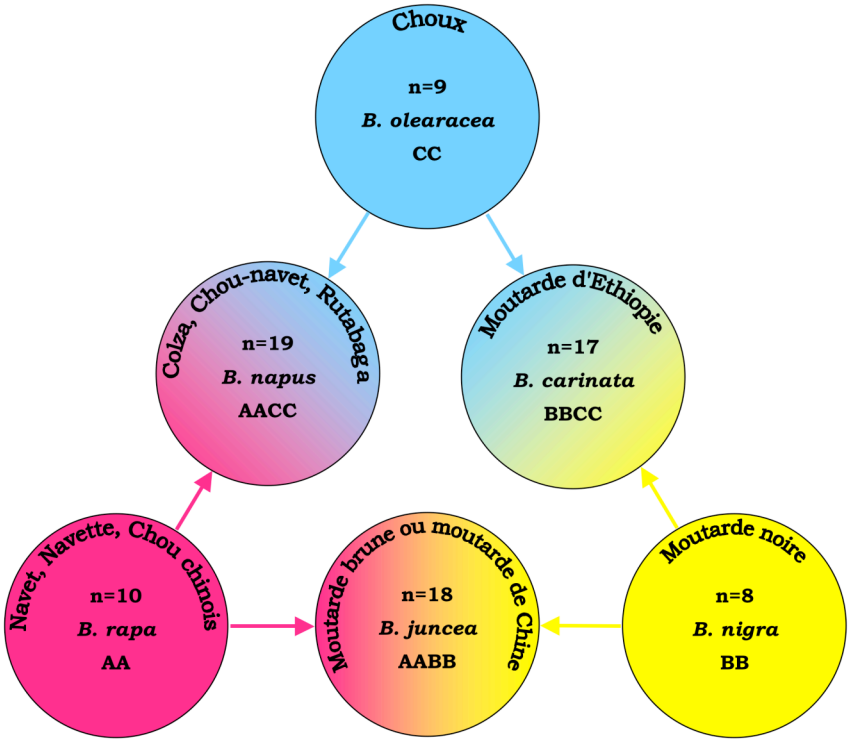


Figure 11 : **Relations génomiques entre les différentes espèces du genre Brassica.** Les espèces situées dans les angles du triangle sont les parents diploïdes à l'origine des descendants allotétraploïdes (amphidiploïdes) qui sont placés sur les côtés respectifs du triangle. Pour chaque espèce, des lettres (a, b et c) représentent la composition du génome. *D'après [94].*

effet, l'infection par le CaMV déclenche le « burst » oxydatif et active deux voies de défense chez *Arabidopsis thaliana* : celle de l'acide salicylique (SA), celle de l'acide jasmonique/éthylène [102].
Enfin, il a été signalé que les ORFs II et IV peuvent aussi influer sur la pathogénicité du CaMV, notamment dans les cas précis de souches infectant *Datura* et *Nicotiana* [103,104].

1.3.2 Impact et gestion du RNA Silencing sur et par le CaMV

L'analyse de plantes présentant un phénomène de « récupération » a permis de démontrer l'implication du « gene silencing » dans la disparition tardive des symptômes [105,106]. La détection de vsRNA (virus-derived siRNAs) dérivant de la séquence du CaMV a irrémédiablement permis d'établir que ce virus est reconnu par la machinerie du RNA silencing, aussi dans *Arabidopsis thaliana* [107], et ceci malgré l'absence d'une étape d'ARN double brin lors de son cycle de réplication. Dans une classique course à l'armement entre hôte et parasite, il a récemment été démontré que la protéine virale P6 est un suppresseur de silencing [23], ce qui coïncide particulièrement bien avec le fait que dans la plupart des cas les suppresseurs de silencing sont des déterminants majeurs de la pathogénicité.
Pour revue sur la production et le transport du silencing en général chez les plantes voir [108,109] ainsi que la figure 12 [110].
Comme cela est précisé ci-dessous, il semble que le rôle du RNA silencing soit ambiguë pour le CaMV ; à la fois négatif et positif.

1.3.2.1 Un système de défense pour la plante

Lors de l'infection par le CaMV, des vsRNA de différentes tailles (21 à 25 nucléotides) s'accumulent au sein des tissus symptomatiques de l'hôte. Ils sont majoritairement produits à partir d'une structure secondaire fortement structurée en « tige-boucle » de la région « leader » de l 'ARN 35S. Les 4 produits des gènes *Dicer-like* (DCL1-4) d'*Arabidopsis thaliana* sont impliqués dans la biogenèse de ces vsRNA. Lors de l'infection par le CaMV, ce sont principalement les protéines DCL3 et DCL4 qui sont à l'origine de la production des vsRNA (24 et 21 nucléotides, respectivement) retrouvés dans

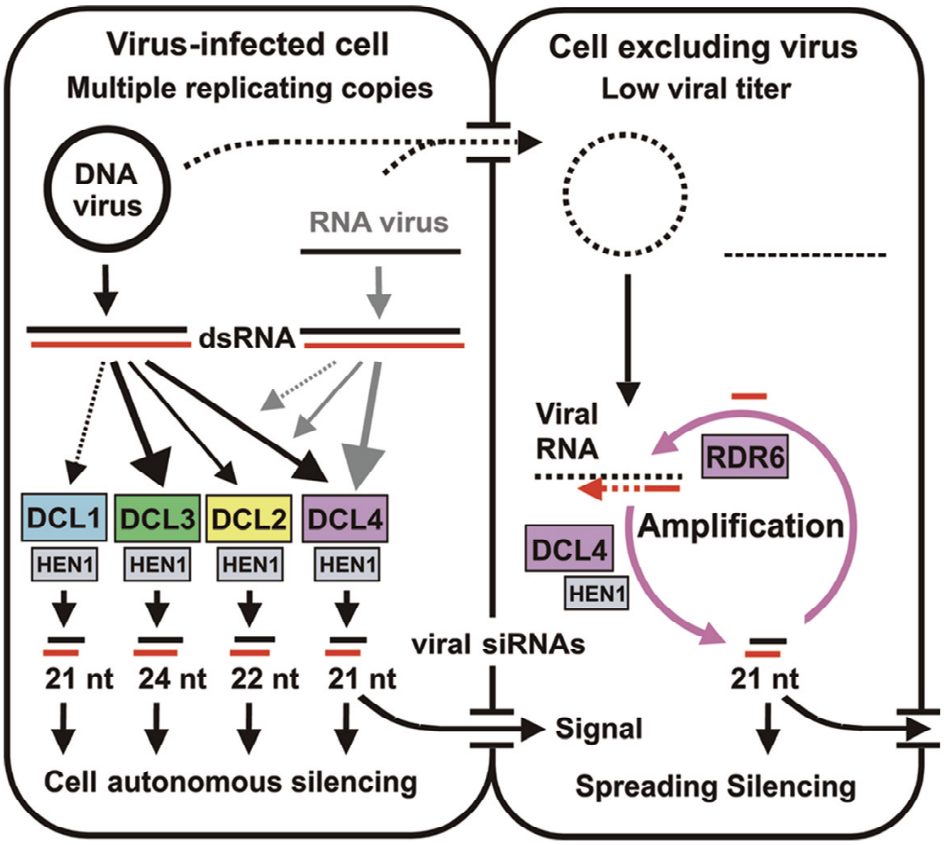


Figure 12 : Modèle pour la défense anti-virale basée sur le RNA silencing. Deux cellules végétales adjacentes sont schématisées. La cellule initialement infectée (à gauche) contient une forte concentration de virus à ADN ou ARN. L'ARN double brin (dsRNA) provient des génomes viraux indépendamment des ARN polymérases de l'hôte. L'importance des flèches reliant les dsRNA aux enzymes DCL représentent la contribution relative de chaque DCL dans la biogenèse des siRNA viraux. Pour les virus à ADN, toutes les DCL clivent les dsRNA en siRNAs de différentes tailles, avec DCL3, DCL4 et DCL2 agissant principalement dans cet ordre. Pour les virus à ARN, DCL4 est majoritairement impliquée, tandis que DCL2 et DCL3 compensent la défaillance éventuelle de DCL4 (par exemple due à une mutation ou à une suppression par le virus). Les fragments de 21 nt produits par DCL4 constituent potentiellement le signal requis pour la dispersion (diffusion) du VIGS. Les acides nucléiques (génomes) infectieux ainsi que les siRNA viraux migrent dans la cellule adjacente (à droite). Cependant le titre viral reste faible, puisque DCL4 et RDR6 amplifient le signal des siRNA entrants et clivent les transcrits viraux. Les siRNA viraux sont stabilisés par leur méthylation *via* HEN1. *D'après [110].*

la plante hôte, tandis que DCL2 présente une activité réduite. Toutefois, il semble exister une certaine redondance entre DCL2 et DCL4 puisque le niveau d'accumulation des vsRNA de 22 nucléotides s'accroît très fortement dans les mutants *dcl4* (absence de DCL4). Enfin, seule une très faible accumulation de vsRNA de 21 nucléotides – perceptible au travers du triple mutant *dcl2/dcl3/dcl4* – permet d'envisager un rôle pour DCL1 lors de l'infection de la plante hôte par le CaMV [107]. Ainsi le CaMV, virus à ADN double brin, est principalement ciblé par DCL4, DCL3 et DCL2, alors que le silencing dirigé contre les virus à ARN est majoritairement produit par DCL4 et DCL2 [111]. Les vsRNA de 24 nucléotides générés par DCL3 pourraient altérer la transcription du virus en méthylant spécifiquement le génome viral qui est condensé sous la forme de minichromosome dans le noyau [112]. Cette hypothèse est supportée par le fait que le génome viral du CaMV peut effectivement être méthylé [113], et que les siRNA endogènes de 24 nucléotides sont impliqués dans le processus de méthylation de l'ADN génomique chez *Arabidopsis thaliana* [114].

En tout état de cause, l'action de ces enzymes limite vraisemblablement le développement viral puisque l'accumulation du CaMV est accrue dans le triple mutant *dcl2/dcl3/dcl4* [107].

1.3.2.2 Un système contrôlé par le virus ?

La protéine P6 agit comme suppresseur de silencing à un niveau local et systémique. Toutefois, dans les plantes transgéniques (*Arabidopsis thaliana* exprimant P6) infectées par le CaMV, le transgène P6 ne réduit pas les quantités de vsRNA (21 et 24 nt) provenant de la séquence leader de l'ARN 35S [23].

Il a été démontré qu'un des vsRNA provenant de la séquence leader a une structure typique de miRNA, et réprime directement l'expression de 3 gènes d'*Arabidopsis thaliana* au sein des tissus infectés. En outre, des analyses bioinformatiques suggèrent que ce même vs/miRNA pourrait à lui seul interférer avec l'expression de 90 gènes cibles de cet hôte [107].

Le virus a-t-il détourné le RNA silencing pour son propre compte afin de modifier le métabolisme de son hôte ? Si oui, alors il est nécessaire au virus

de contrôler finement l'activité du RNA silencing en l'inhibant afin de contrôler un des systèmes de défense de l'hôte, tout en préservant une activité minimale nécessaire à la production des vs/miRNA permettant de moduler l'expression de certains gènes de l'hôte.

2 La transmission des virus de plante

Les plantes hôtes sont immobiles et ont généralement une durée de vie limitée. Une étape de transmission est donc indispensable au virus afin de se maintenir dans l'environnement. Cette transmission est le plus souvent assurée par des vecteurs *via* divers processus moléculaires, et constitue une des phases problématiques majeures du cycle viral. En effet, le virus endocellulaire est contraint de quitter son hôte et de passer par le milieu extérieur qui lui est défavorable, tout en restant infectieux.

Les différents modes de transmission par vecteur des phytovirus seront décrits dans cette section (pour revue voir [115]), avant de détailler les connaissances actuelles sur les interactions des différents partenaires moléculaires mis en jeu lors de la transmission du CaMV par puceron.

2.1 Les différents modes de transmissions des phytovirus

Le plus gros volume d'études sur la transmission par vecteurs concerne les virus transmis par des insectes dont l'adaptation de l'appareil buccal est de type « piqueur-suceur ». Ce type d'insecte correspond principalement aux trois ordres anciennement nommés : *Diptera*, *Heteroptera* et *Homoptera*. Ce dernier ordre contient les vecteurs responsables de la transmission de 58 % des virus de plantes [116], et regroupe les cochenilles, les cicadelles, les aleurodes, et surtout les aphides (pucerons). De manière générale, les pucerons ont un équipement sensoriel pour reconnaître leur plante hôte [117], tandis que les plantes perçoivent en retour leurs piqûres intra-phloémiques et déclenchent des systèmes de défenses contre lesquels le vecteur devra lutter pour se nourrir [118-120].

Quels que soient précisément le nombre et la nature de ces nombreuses interactions vecteurs-hôtes, l'étape de la transmission comporte toujours

trois phases : acquisition - rétention - inoculation. Historiquement, les mesures temporelles de ces trois phases chez des aphides vecteurs ont permis d'établir une première classification en trois catégories. Il y a tout d'abord (i) les virus « non persistants » dont la phase d'acquisition est de très courte durée (quelques secondes à quelques minutes) avec une fenêtre de rétention et un temps d'inoculation tout deux aussi limités à quelques minutes; puis (ii) les virus « semi persistants » qui peuvent survivrent plusieurs heures dans le vecteur – en dehors de leur plante hôte – et qui nécessitent des phases d'acquisition et d'inoculation pouvant s'étendre de quelques minutes à quelques heures; et enfin (iii) les virus dits « persistants » dont les phases d'acquisition et d'inoculation requièrent généralement plusieurs heures, et sont séparées par une longue période de latence durant laquelle l'insecte ne peut pas transmettre le virus. Une fois cette phase de latence terminée, les vecteurs maintiennent le virus inoculable durant une très longue période, souvent jusqu'à leur mort.
De plus récentes informations obtenues sur les mécanismes de l'interaction virus-vecteur (critères qualitatifs), permettent maintenant d'élaborer un système de classification applicable ou ajustable pour tous les virus de plantes *(pour revue voir [121]).*

2.1.1 La transmission circulante multipliante (ou propagative)

Le virus est ingéré avec la sève phloèmienne lors de la prise de nourriture du vecteur. Des virions traversent alors les cellules épithéliales de l'intestin et gagnent l'hémocoele. Ensuite, le virus diffuse par l'hémolymphe jusqu'à de nombreux organes dans lesquels il se réplique, notamment les glandes salivaires. Enfin, il peut y avoir inoculation lorsque le virus est excrété des cellules des glandes salivaires par exocytose et injecté dans la plante hôte avec la salive pendant une nouvelle piqûre.
L'acquisition de tels virus demande quelques minutes à quelques heures, selon la répartition du virus dans la plante hôte. Les particules virales doivent franchir plusieurs barrières membranaires du vecteur, ce qui implique des interactions moléculaires complexes. La phase de latence qui correspond au temps nécessaire au virus pour passer de la lumière

intestinale à la lumière des glandes salivaires se caractérise par une plus ou moins longue durée (en général plusieurs jours) pendant laquelle le vecteur n'est pas infectant. Le virus peut persister dans le vecteur jusqu'à sa mort et il contamine aussi ses descendants par transmission verticale trans-ovarienne : les descendants du vecteur ainsi contaminés seront capables de transmettre le virus.
Cette catégorie correspond à des espèces anciennement classées comme virus persistants.

2.1.2 La transmission circulante non-multipliante

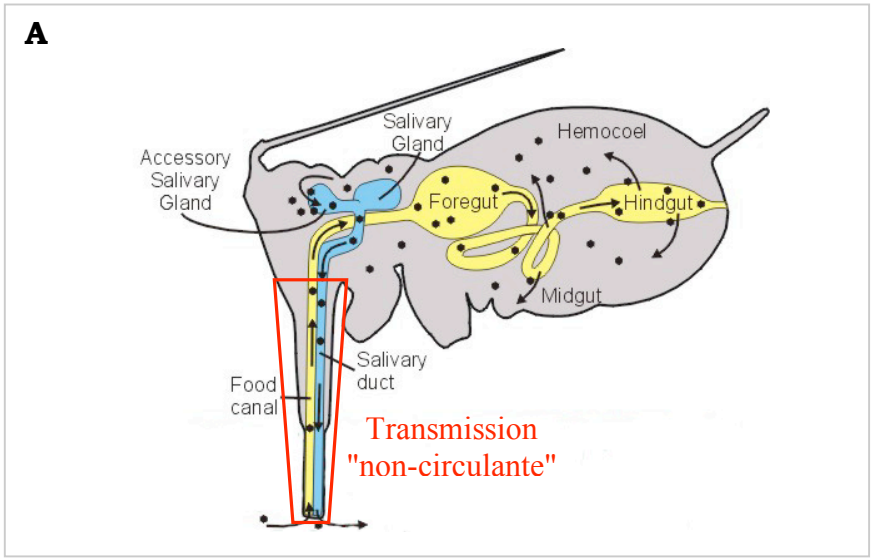
Les virus transmis selon ce mode sont transportés de façon interne comme dans le mode précédent, sauf qu'ils ne se répliquent jamais durant leur cycle dans le vecteur. La viabilité des particules virales dans l'hémolymphe est délicate, elle semble requérir l'apport stabilisant de protéines chaperonnes, comme la symbionine produite par les bactéries du genre *Buchnera*, endosymbiotiques de pucerons. Ces espèces étaient anciennement aussi considérées comme des virus persistants.

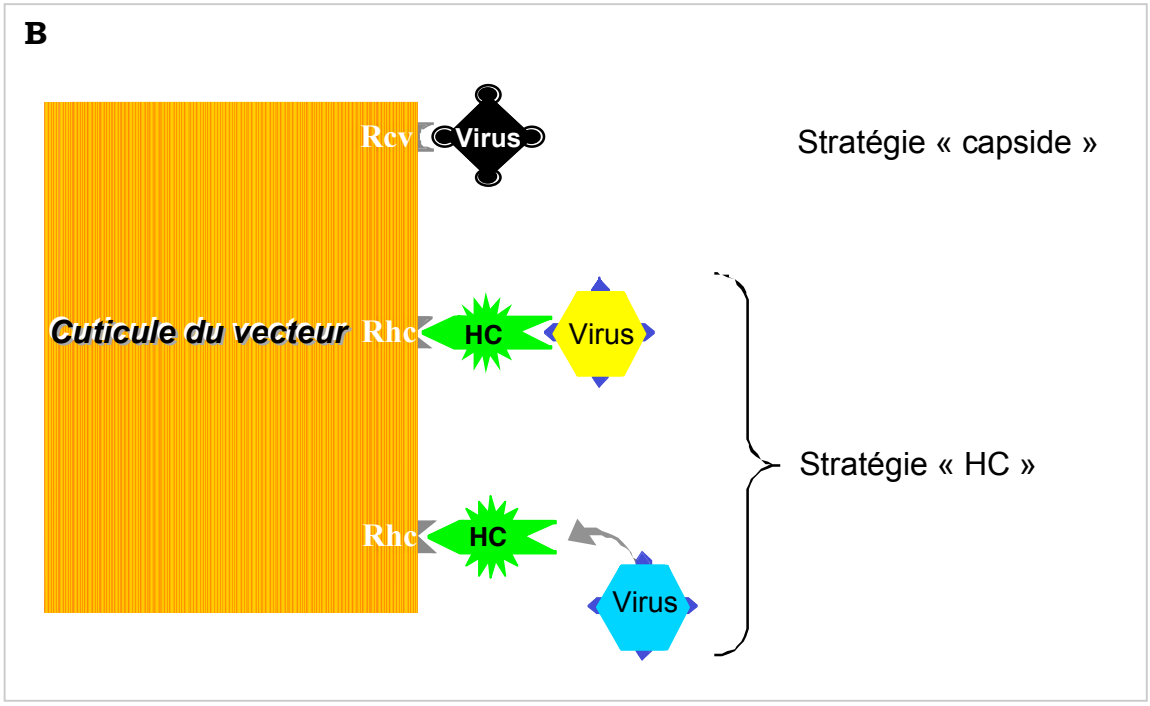
2.1.3 La transmission non-circulante

Cette catégorie regroupe aujourd'hui l'ensemble des virus anciennement catégorisés non- et semi-persistants, puisque cette dichotomie ne reflétait aucune différence qualitative démontrée.
Pour ce mode de transmission, l'association virus-vecteur est exclusivement externe (encadré 7A) et il n'y a pas de phase de latence. Lors de l'acquisition, des particules virales sont retenues au niveau de sites de rétention localisés sur la cuticule qui tapisse la face interne des canaux alimentaires et/ou salivaires des stylets, mais cette rétention ne dure que quelques minutes à quelques heures. Après acquisition, lors d'un repas ultérieur du vecteur sur une plante saine, quelques virions seront décrochés et inoculés au cours de la salivation [122]. De ce fait, seule une partie des particules virales ingérées par le vecteur va être inoculée ultérieurement à une plante.
Deux stratégies moléculaires distinctes permettent de gérer l'interaction virus-vecteur (encadré 7B):

Encadré N°7

Les différents modes de transmission des phytovirus par vecteur





A Circulation des virus au sein de leur vecteur selon leur mode de transmission. Lors des transmissions circulantes (multipliantes ou non), le virus effectue un cycle en passant par le milieu intérieur du vecteur. Il traverse les barrières intestinales et salivaires, avant d'être inoculé lors de phases de salivation. Au contraire, l'interaction virus/vecteur est exclusivement externe et restreinte à la zone encadrée en rouge pour le mode "non-circulant".

B Les deux stratégies moléculaires de la transmission non-circulante. Dans le cas de la stratégie capside, la particule virale interagit directement avec un récepteur hypothétique (Rcv, récepteur de la capside virale) dans les stylets du vecteur. Dans le cas de la stratégie « Helper », la particule virale interagit par l'intermédiaire d'une protéine non-structurale (Helper Component, HC) – ne participant pas à la structure de la particule virale – avec un récepteur hypothétique (Rhc). Les 2 récepteurs hypothétiques (Rcv et Rhc) pourrait être un seul et unique récepteur (R) ; les données actuelles ne permettent pas encore de trancher cette question. Enfin, le phénomène de HC-trans-complémentation est représenté en bas du schéma : un génome X (dans le virion jaune) code pour un HC (vert) qui va permettre à un génome Y (dans le virion bleu) d'être retenu dans les stylets et d'être ensuite transmis. La HC-trans-complémentation peut s'opérer au sein d'une espèce, et même parfois entre différentes espèces.

2.1.3.1 Stratégie capside

Cette stratégie implique une interaction directe entre un domaine d'une protéine structurale virale (la protéine de capside) et un récepteur hypothétique présent sur la cuticule du vecteur.

2.1.3.2 Stratégie Helper

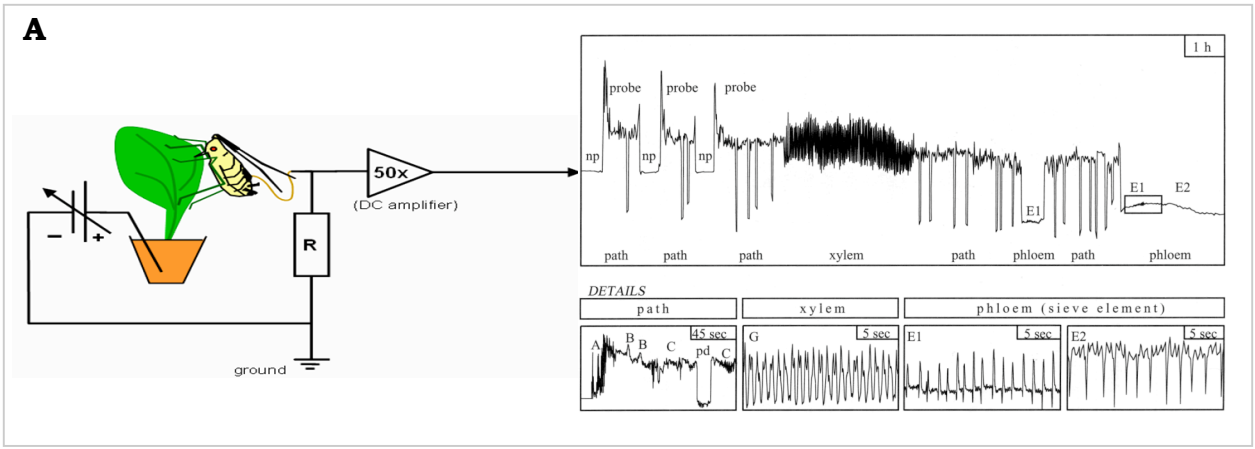
Cette stratégie de transmission fait intervenir un ou plusieurs composants additionnels dans l'interaction virus-vecteur. En effet, des particules virales purifiées ne peuvent pas être transmises, car la reconnaissance du récepteur chez le vecteur ne se fait pas *via* la protéine de capside. La transmission nécessite l'aide de protéines virales non-structurales (extracapsidaires), les Helpers Components (HC), qui sont capables d'interagir d'un côté avec le récepteur du vecteur et de l'autre avec la particule virale. Ils forment donc un pont moléculaire permettant la rétention du virus par le vecteur. La labilité intrinsèque de ce pontage moléculaire et/ou l'action de facteurs inconnus, libère(nt) des particules virales qui sont alors inoculées à une plante sensible. Le HC peut être acquis seul par le vecteur, avant l'acquisition des particules virales. Ainsi, le HC d'un virus transmissible est capable d'assister la transmission de souches non transmissibles (dont le HC est déficient ou absent) dans un phénomène défini comme la HC-transcomplementation [123] (encadré 7B).

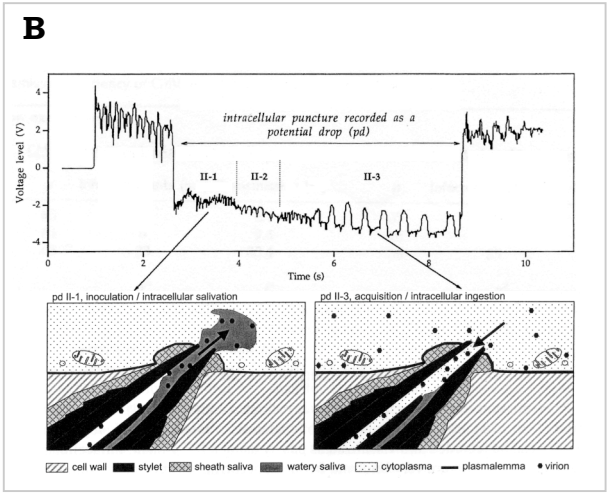
2.2 Précisions sur les phases d'acquisition et d'inoculation des virus non-circulants

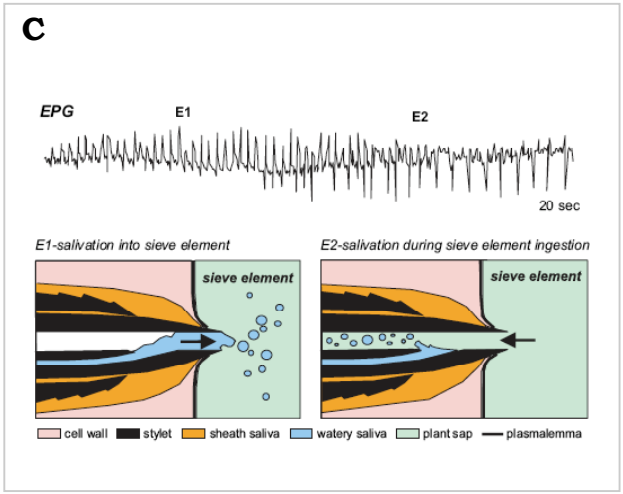
La technique d'Electro-Pénétro-Graphie (EPG) qui a vu le jour au début des années 60, permet d'analyser le comportement alimentaire d'insectes piqueurs suceurs, pour revue voir [124] et [116]. Les graphiques d'EPGs obtenus à partir de pucerons « monitorés » ont pu être décomposés en plusieurs phases ondulatoires caractéristiques, dont chacune a pu être corrélée à une étape du processus alimentaire (encadré 8) [125], grâce à l'appui d'images en microscopie électronique à transmission [126]. Ainsi pour les virus transmis suivant le mode non-circulant, les différentes phases impliquées dans l'acquisition et l'inoculation des virions ont pu être déterminées précisément.

Encadré N°8

Technique d'électro-pénétrographie (EPG)





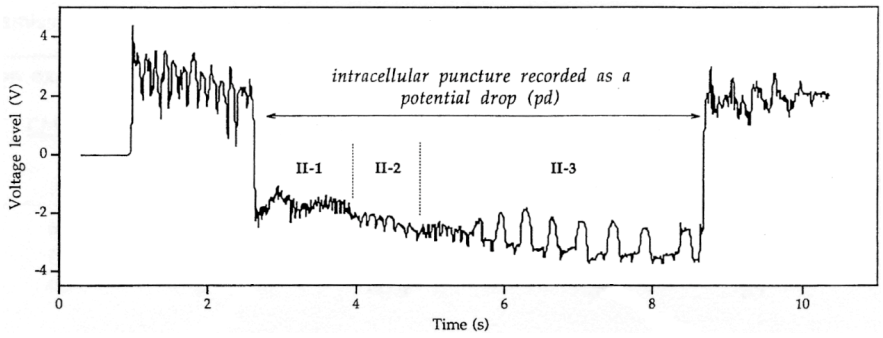


A Circuit EPG et enregistrement type du comportement alimentaire du puceron. Le circuit d'enregistrement pour l'EPG est représenté sur la gauche de la figure. La sortie de l'amplificateur (DC amplifier) est connectée à un appareil enregistreur – généralement un ordinateur – avec un écran de visualisation en temps réel et un disque dur pour le stockage des données. La partie droite de la figure schématise l'enregistrement sous EPG du comportement d'un puceron. Graphique supérieur : représente une heure d'enregistrement avec une alternance de phases entre piqûre (« probe ») ou non (« non-probing », np). Les deux premières piqûres dites « d'essai » comportent uniquement une phase de « pathway » (path) qui inclut des ponctions de cellules superficielles (épidermes et/ou mésophylle). La troisième piqûre comprend une phase xylémique et une phase phloémique. De nombreuse baisses de potentiels se produisent durant les périodes de pathway. La phase phloémique se décompose en 2 parties : une première de courte durée (E1) suivie par une seconde plus longue (E2). Le petit rectangle entoure la transition entre E1 et E2 (voir C pour plus de détails). Petits encarts inférieurs : détails de chaque phase avec leur forme d'onde caractéristique ; la phase de pathway se subdivise en plusieurs sous-phases : A, B, C – correspondant à une progression intercellulaire des stylets du puceron – et pd (potential drop) identifiée par une baisse de potentiel lors de brèves ponctions intracellulaires (voir B pour de plus amples détails). *D'après [125].*

B La ponction intracellulaire ou PD en détail. La partie supérieure expose une baisse de potentielle typique (pd) produite par *Myzus persicae* juste après le début d'une piqûre, ainsi que le détail de ses sous-phases intracellulaires (II-1, II-2 et II-3) en corrélation avec les mécanismes d'inoculation (en bas à gauche) et d'acquisition (en bas à droite) des virus non-persistants. Les schémas en coupe longitudinale– lors de la ponction intracellulaire d'une cellule d'épiderme ou de mésophylle – montrent les extrémités des stylets mandibulaires et des stylets maxillaires à l'intérieur desquels est figuré le canal commun (fusion des canaux alimentaire et salivaire). *D'après [128].*

C Fonctions associées aux sous-phases phloémiennes E1 et E2. La partie supérieure figure l'étape de transition entre les sous-phases E1 et E2. Les diagrammes inférieurs reflètent les activités liées à chacune des sous-phases. Durant E1, la valve cibariale localisée dans la tête de l'insecte ferme le canal alimentaire de sorte à ce que la salive excrétée aille dans le phloème. Pendant la sous-phase E2 il y a un phénomène de salivation continue mais cette salive n'atteint pas le phloème ; elle se retrouve mélangée avec la sève et remonte directement par le canal alimentaire. *D'après [119].*

(a)



(b) Acquisition experiment

	CMV transmitted by *Aphis gossypii* on melon			PVY transmitted by *Myzus persicae* on pepper		
Sub-phase	*n*	Infected plants (no.)	Transmission (%)	*n*	Infected plants (no.)	Transmission (%)
II-1	65	4	6.1	37	3	8.1
II-1+II-2	64	6	9.4	16	3	18.7
II-1+II-2+II-3	65	23	35.4	56	35	62.5

(c) Inoculation experiment

	CMV transmitted by *Aphis gossypii* on melon			PVY transmitted by *Myzus persicae* on pepper		
Sub-phase	*n*	Infected plants (no.)	Transmission (%)	*n*	Infected plants (no.)	Transmission (%)
II-1	53	24	45.3	12	6	50
II-1+II-2	48	28	58.3	13	5	38.5
II-1+II-2+II-3	60	27	45.0	27	9	33.3

Figure 13 : Corrélation entre phases de la ponction intracellulaire (pd), acquisition et inoculation de virus non-persistant. (a) Enregistrement EPG typique d'une ponction intracellulaire (ou potential drop, pd) à partir duquel il a été caractérisé trois sous-phases (II-1, II-2 et II-3) très brèves (quelques secondes tout au plus). (b) Acquisition des virus non-persistants par leur vecteur lors de la phase II-3. Des expérimentations d'acquisition contrôlée sous EPG sont menées avec deux systèmes vecteurs-virus (Virus : *Cucumber mosaic virus*, CMV et *Potato virus Y*, PVY), l'inoculation étant laissée libre durant plusieurs heures. Une seule piqûre intracellulaire est autorisée à chaque puceron. Celle-ci est interrompue artificiellement à la fin d'une phase II-1, II-2 ou II-3 lors de la première baisse de potentiel (pd), simplement en enlevant le puceron de sa plante. Les taux de transmission obtenus augmentent fortement et significativement – quelque soit le système vecteur-virus considéré – lorsque la phase II-3 est achevée durant l'acquisition. (c) Inoculation des virus non-persistants lors de la phase II-1. Le même dispositif EPG est utilisé cette fois-ci lors de l'inoculation du virus par le vecteur, l'acquisition étant constante et fixée à 5 minutes. Les résultats montrent que la phase II-1 est nécessaire et suffisante pour un relargage efficace des virions puisque les phases ultérieures ne permettent pas d'obtenir des taux de transmission significativement supérieurs. *D'après [128].*

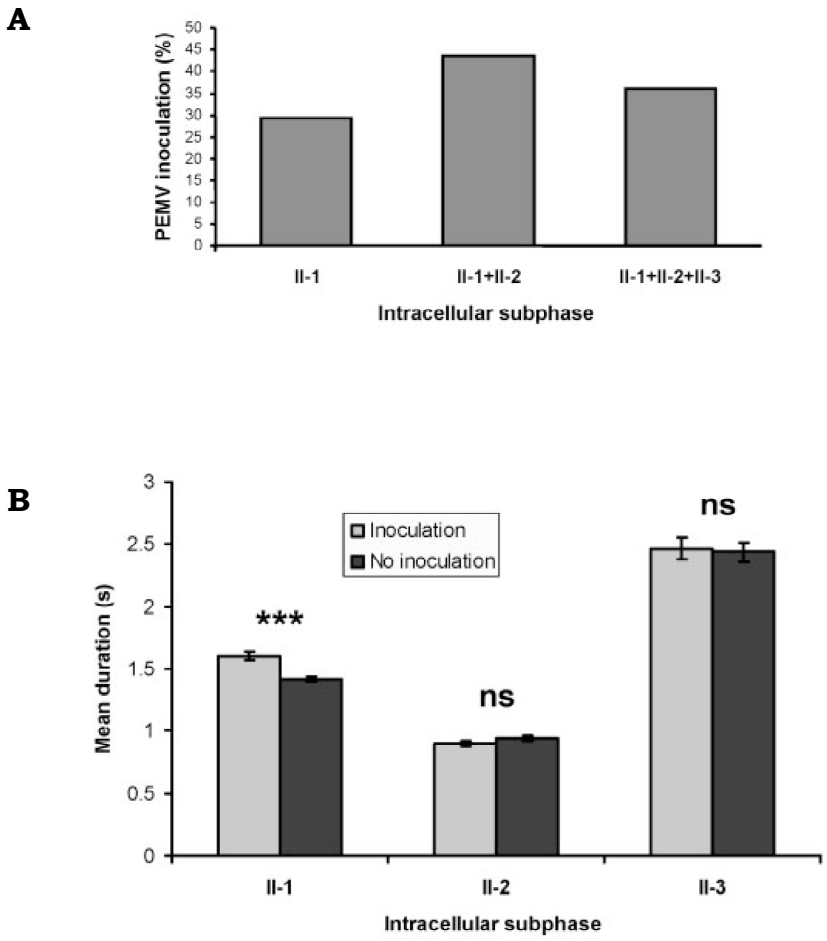


Figure 14 : La phase II-1 est une phase de salivation. Le *Pea enation mosaic virus* (PEMV) est transmis de manière circulante non-multipliante (non-propagative) ; il s'accumule dans les glandes salivaires et se retrouve inoculé lors de la salivation du vecteur (ici *Acyrthosiphon pisum*). (A) Efficacité d'inoculation du PEMV en fonction des différentes sous-phases intracellulaires. La ponction cellulaire effectuée par des pucerons virulifères est interrompue artificiellement lors de sous-phases II-1, II-2 ou II-3 (taille des échantillons : II-1, n=41 ; II-1+II-2, n=62 ; II-1+II-2+II-3, n=72). (B) Durée moyenne (± erreur standard de la mesure) des sous-phases intracellulaires, classées selon la réussite de l'inoculation du PEMV ou son echec. (***, P<0,001 ; ns, différence non significative ; résultats obtenus à partir de tests-*t*). *D'après [122].*

2.2.1 Acquisition

Une première étude d'EPG sur milieu artificiel (virus, HC et saccharose) à travers une membrane de parafilm® a montré que le *Tobacco etch virus* (TEV) – un potyvirus – est acquis par le vecteur lors de la sous-phase II-3 du Pd (Potential drop) [127]. Ensuite, il a été confirmé *in planta* avec deux virus transmis selon le mode non-circulant (le *Cucumber mosaic virus*, CMV, pour la stratégie capside, et le *Potato virus Y*, PVY, pour la stratégie « helper ») que cette sous-phase II-3, qui correspond à l'ingestion d'une partie du contenu cellulaire, est en effet majoritairement responsable de l'acquisition virale (figure 13) [128].

2.2.2 Inoculation

Dans la même étude que précédemment, les auteurs ont identifié la sous-phase II-1 comme étant impliquée dans le processus d'inoculation des virions [128]. L'activité du puceron liée à cette sous-phase fut débattue (entre salivation ou égestion) jusqu'en 2005 où des travaux l'ont associée à l'injection active de salive directement à l'intérieur du cytoplasme (figure 14) [122].

2.3 La transmission du CaMV en détail

Le virus de la mosaïque du chou-fleur est transmis par pucerons-vecteurs sur le mode non-circulant selon une stratégie « helper ».

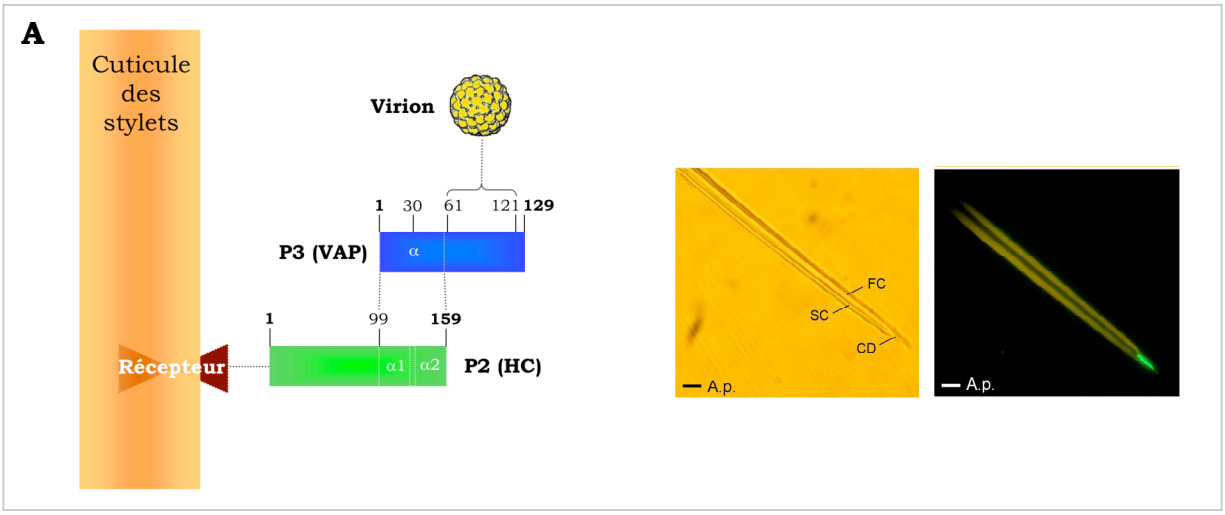
Parmi les 27 pucerons connus pour transmettre le CaMV [129], le puceron cendré du chou (*Brevicoryne brassicae*) et le puceron vert du pêcher (*Myzus persicae*) sont les deux principaux vecteurs [130]. Le puceron cendré du chou est un spécialiste des brassicacées alors que *Myzus percisae* est un généraliste (polyphage) qui s'alimente sur plus de 100 espèces végétales [131].

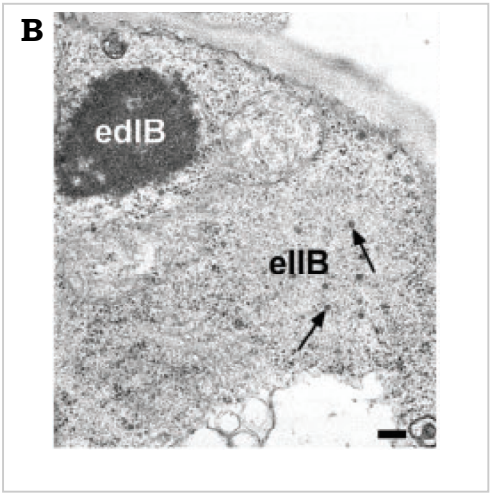
2.3.1 Stratégie Helper et modèle d'acquisition séquentielle

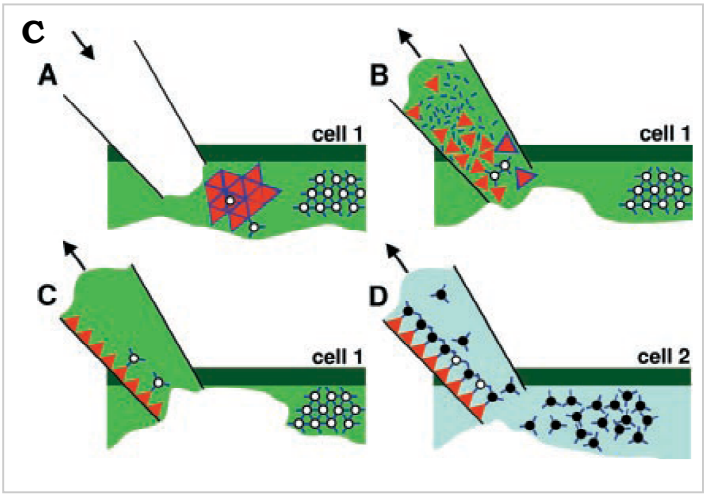
La protéine P2 du CaMV joue le rôle de HC en « pontant » le complexe P3-virus à un récepteur présent au niveau des stylets du puceron (voir encadré 9A). P2 est définie comme le HC, par opposition à P3, puisqu'elle peut être acquise seule par le puceron, puis permettre ultérieurement l'attachement

Encadré N°9

La transmission du CaMV par vecteur







A Le complexe transmissible et sa localisation dans les stylets du vecteur. Le complexe transmissible du CaMV – composé de P2 (HC, helper component), P3 (VAP, virus associated protein) et du virion (P4 et génome viral) – est retenu au niveau des stylets du puceron-vecteur *via* un récepteur protéique non identifié. Un domaine C-terminal (aa 61 à 121, comprenant un site de fixation aux acides nucléiques) de la protéine P3 lui permet de s'associer aux particules virales. Les protéines P2 et P3 interagissant par leur(s) hélice(s) α sont ici représentées sous la forme de monomère, toutefois leur forme active lors de la transmission pourrait être multimérique. Enfin, la partie N-terminale du HC est reconnue par le récepteur présent dans la cuticule des stylets. L'observation de stylets disséqués, puis incubés *in vitro* avec une P2 fusionnée à la GFP, a permis de mettre en évidence que le récepteur est localisé à l'extrémité distale (zone fluorescente verte) des stylets au niveau du canal commun (CD) (fusion entre le canal alimentaire (FD) et le canal salivaire (SC)). Image de gauche en lumière visible et image de droite sous U.V. Barre, 5µm. A.p., *Acyrthosiphon pisum. D'après [56].*

B Une cellule infectée présente deux types de corps d'inclusion distincts. L'observation en microscopie électronique à transmission d'une cellule infectée (ici par la souche B-JI) met en évidence la présence de deux corps d'inclusion : (i) edIB, corps d'inclusion denses, composé d'une matrice de protéine P6, et contenant 94% des particules virales et (ii) elIB, corps d'inclusion clairs comprenant les 6% des virions (flèches noires) restant et toute la P2. Un CaMV (mutant ponctuel dans P2) pour lequel la formation d'elIB est déficiente ne peut plus être transmis par son vecteur. Barre, 200nm. *D'après [61, 58].*

C Modèle d'acquisition séquentielle du CaMV par son puceron-vecteur. (A) dans une plante infectée, les composants viraux impliqués dans la transmission sont séparés spatialement dans l'elIB (à gauche) et les edIB (à droite). Alors que la majorité des virions (disques blancs) complexés à P3 (trait bleu) sont stockés dans les edIB, la totalité de P2 (triangle rouge), co-agrégée avec P3 et quelques virions, est localisée dans l'elIB. Quand un puceron perce la membrane plasmique avec ses stylets (en haut à gauche), il injecte de la salive dans la cellule végétale. (B) Ensuite, le puceron ingère une partie du contenu cellulaire contenant des corps d'inclusion viraux. Si un elIB est aspiré, il se désintègre et libère ses composants : P2, P3 et quelques rares complexe P3-virions. (C) Bien que la P3 libre soit perdue (avalée), les P2 et les complexes P2-P3-virions de l'elIB sont retenus dans les stylets du puceron. Ces derniers sont maintenant saturés de P2 et le puceron et donc compétant pour la transmission (inoculation) ainsi que pour l'acquisition de complexes P3-virions supplémentaires (disques noirs) au sein de la même cellule ou (D) lors d'une piqûre ultérieure dans une autre cellule. *D'après [58].*

de complexes P3-virus préformés. A l'opposé, l'acquisition de P3 ne peut pas se faire seule (ni même avec P2) lors d'un repas antérieur pour permettre une transmission efficace [58], d'où son autre nom de VAP employé dans la littérature. Au moins pour l'étape de la transmission, P3 se comporte littéralement comme une protéine de capside.

A la fin du cycle cellulaire viral, il y a deux types de corps d'inclusion cytoplasmiques : (i) les corps d'inclusion denses (edIBs, electron-dense Inclusion Bodies) formés d'une matrice de protéine P6 renfermant 94% des complexes P3-virus de la cellule, et (ii) le corps d'inclusion clair (elIB, electron-lucent Inclusion Body) qui contient une fraction de P3, toute la protéine P2, et les 6% de virions restants (encadrés 1C et 9B). P2 est ainsi spatialement éloignée (séparée) de la majorité des particules virales, en conséquence, il n'existe que relativement peu de complexes transmissibles P2-P3-virions au sein d'une cellule infectée. De ce fait, il est supposé que la reconstitution du complexe transmissible a majoritairement lieu dans les stylets du vecteur selon un modèle d'acquisition séquentielle [58] : lors d'une première piqûre, le puceron va absorber un elIB qui se désintègre en libérant le HC qui va se fixer sur les stylets avec quelques éventuels complexes P3-virus. Ensuite, ce puceron compétant pour la transmission peut piquer dans une autre cellule et gober un edIB, saturant ainsi tous les HC présents dans ses stylets. Ce modèle d'acquisition séquentielle (encadré 9C) permet de soutenir l'intérêt d'avoir deux structures distinctes, du fait qu'il favorise potentiellement l'acquisition d'une plus grande diversité génétique par le puceron-vecteur [58].

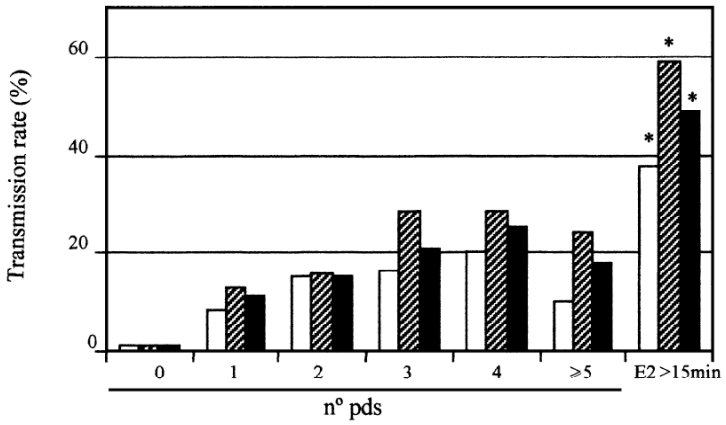
2.3.2 Acquisition – Rétention – Inoculation

2.3.2.1 Acquisition

Actuellement, il n'y a pas données expérimentales publiées sur la sous-phase impliquée dans l'acquisition du CaMV. Toutefois, il semble fort probable que l'acquisition se fasse lors de la sous-phase II-3 comme pour les autres virus transmis sur le mode non-circulant.

Par contre, il a été montré que le lieu de l'acquisition peut aussi être important pour le CaMV. En effet, lorsque le puceron s'alimente durant au

A



B

	no. of infected plants /total no. of test plants						
	no. of pds						E2≥ 15min
	0	1	2	3	4	5	
□ *M persicae*	0/29	12/145	5/33	5/31	3/15	4/38	9/24
▨ *B brassicae*	0/21	19/143	7/45	6/21	6/21	12/49	16/27
■ Totals	0/50	31/288	12/78	11/52	9/36	16/87	25/51

Figure 15 : Le CaMV est préférentiellement acquis au niveau du phloème par son vecteur. (A) Relation entre le nombre de ponctions intracellulaires (pds) ou l'ingestion dans le phloème produits par *Myzus persicae* (blanc) et *Brevicoryne brassicae* (hachuré), et le taux de transmission du CaMV ; l'ensemble des résultats pour les deux vecteurs est figuré en noir. * indique des différences significatives ($P<0,05$) lors de la comparaison des taux de transmission par paire entre les différentes classes, selon une analyse de χ^2, ou d'un test exact de Fisher quand les valeurs observées sont inférieures à 5. (B) Ratios entre le nombre de plantes infectées et le nombre total de plantes tests en fonction des différents traitements réalisés (0, 1, 2, 3, 4 à plus de 5 ponction(s) intracellulaire(s) ou bien 15 minutes minimum d'ingestion phloémienne). *D'après [132].*

moins 15 minutes au sein du phloème d'une plante virosée, le taux de transmission est significativement augmenté par rapport à la ponction d'une seule ou bien même de 5 cellules (figure 15) [132].

2.3.2.2 Rétention

Le CaMV fut, jusqu'à une époque encore très récente, catalogué comme un virus semi-persistant [133], du fait entre autre que *Brevicoryne brassicae* présente un fort taux de transmission suite à une période d'acquisition virale de 8h [134,135]. Toutefois, la vection du CaMV par certaines espèces de pucerons (ou même par divers clones de *Myzus persicae*) s'apparente parfois plus à celle d'un virus non-persistant du fait qu'une acquisition de 8h n'offre qu'un faible taux de transmission [134]. Les hypothèses actuelles pour expliquer ce phénomène particulier reposent sur une probable rétention accrue du complexe P2-P3-virus dans les stylets de *Brevicoryne brassicae,* et donc d'une éventuelle différence de composition et/ou quantité de salive aqueuse [136] qui serait nécessaire pour déstabiliser le complexe viral pour permettre l'inoculation, en comparaison avec les virus non-persistants « typiques ».

2.3.2.3 Inoculation

En accord avec l'hypothèse ci-dessus, en 2005, une étude d'EPG menée avec les deux principaux vecteurs du CaMV, n'a pas permis d'impliquer aussi clairement la salivation lors des premières piqûres superficielles dans les cellules de l'épiderme et du mésophylle, comme l'événement clé à l'origine de l'inoculation du virus [136,137].

2.3.3 Interaction moléculaire Virus-Vecteur

2.3.3.1 Spécificité d'interaction (P2-récepteur)

Une large gamme de pucerons transmet le CaMV, mais l'efficacité de la transmission se révèle variable en fonction des espèces et des clones étudiés [129,130,135]. Une possible explication à ce constat serait qu'il existe une interaction spécifique plus ou moins forte (ou labile) selon les vecteurs entre le récepteur putatif et le HC.

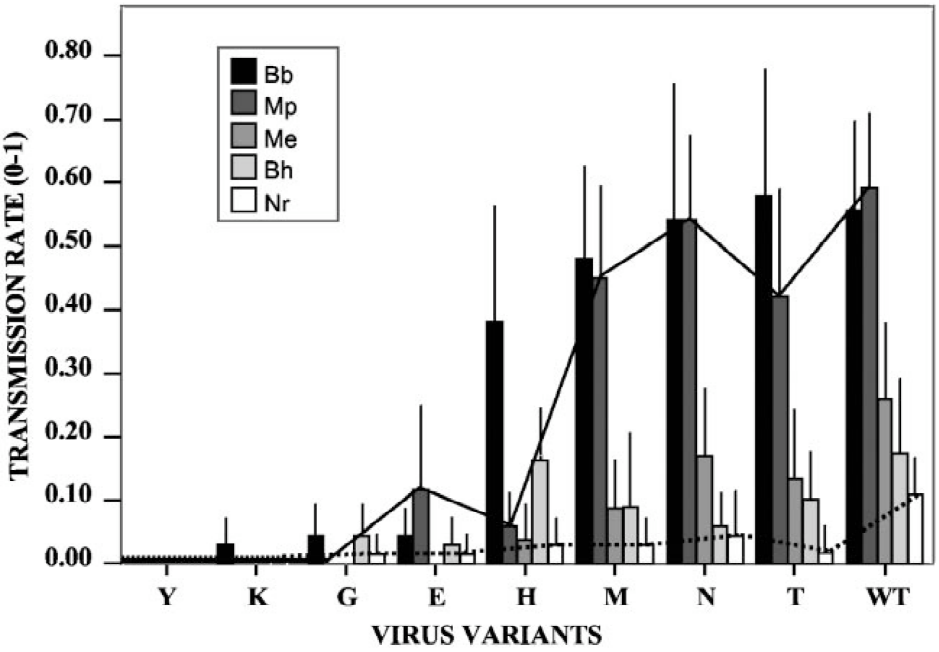


Figure 16 : Taux de transmission des différents mutants du CaMV par 5 espèces de pucerons. Le sixième acide aminé de la protéine P2 – naturellement Q (WT, wild type) – a été substitué par 8 autres (Y, K, G, E, H, M, N ou T) dans la souche CabbS du CaMV, avant de réaliser des tests de transmission par puceron à partir de plantes infectées. Les moyennes sont regroupées par mutant (notés au dessous), avec pour chacune une barre d'erreur standard de mesure. Bb, *Brevicoryne brassicae* ; Mp, *Myzus persicae* ; Me, *Macrosiphum euphorbiae* ; Bh, *Brachycaudus helichrysi* and Nr, *Nasonovia ribisnigri*. Il est a noter que *N. ribisnigri* (ligne en pointillés) transmet tous les mutants à des taux faibles mais similaires, alors que le taux de transmission par *M. persicae* (ligne continue) varie radicalement suivant le type de virus mutant testé. *D'après [55].*

Il a été découvert de manière fortuite, que le sixième acide aminé de la protéine P2 du CaMV est précisément impliqué dans l'interaction avec le puceron-vecteur, puisque le mutant Q6Y (substitution de Q – 6ième acide aminé – par Y) n'est plus transmissible par aucune espèce normalement vectrice. D'autres substitutions et l'utilisation de 5 espèces de pucerons ont mis en avant le rôle de cette position d'acide aminé dans la modification du spectre des vecteurs [55]. En effet (figure 16), certains mutants ne sont pas transmissibles par l'ensemble des 5 espèces de pucerons-vecteurs (cas de Q6K, Q6G et Q6E), tandis que d'autres mutants peuvent voir leur espèce vectrice préférentielle modifiée. Par exemple, l'espèce vectrice la plus efficace est *Myzus persicae* pour le mutant Q6E, alors qu'il s'agit de *Brevicoryne brassicae* pour le mutant Q6H. Il est remarquable de constater qu'aucun des mutants analysés n'a permis d'augmenter le taux de transmission maximum du CaMV « sauvage » et que toutes les séquences de souches naturelles disponibles sur GeneBank présentent le même acide aminé (Q), comme si cela correspondait à un optimum évolutif pour le virus.

2.3.3.2 Zone d'interaction et nature chimique du récepteur

Outre les travaux d'EPG mentionnés plus haut, des études encore plus anciennes corroborent l'idée que le site de rétention des virus non-persistants se situe au bout des stylets [138,139]. En effet, la transmission du *Potato virus Y* (PVY) par le puceron *Myzus persicae* est totalement inhibée lorsque l'extrémité des stylets du puceron-vecteur est traitée chimiquement au formaldéhyde.

Le développement récent, dans notre laboratoire, d'une technique de dissection et d'isolement des stylets mandibulaires et maxillaires a offert la possibilité d'incuber *in vitro* les stylets de 4 espèces de pucerons (dont 3 vectrices) avec une protéine de fusion P2-GFP. Ces expériences ont démontré que seule l'extrémité distale des stylets maxillaires des pucerons-vecteurs retenait la fluorescence, donc P2-GFP. Des études complémentaires en microscopie électronique à transmission ont aussi permis de découvrir, sur la même zone anatomique, des pseudo particules virales (VLPs, virus-like particles) dans les stylets de pucerons s'étant alimentés sur plante virosée.

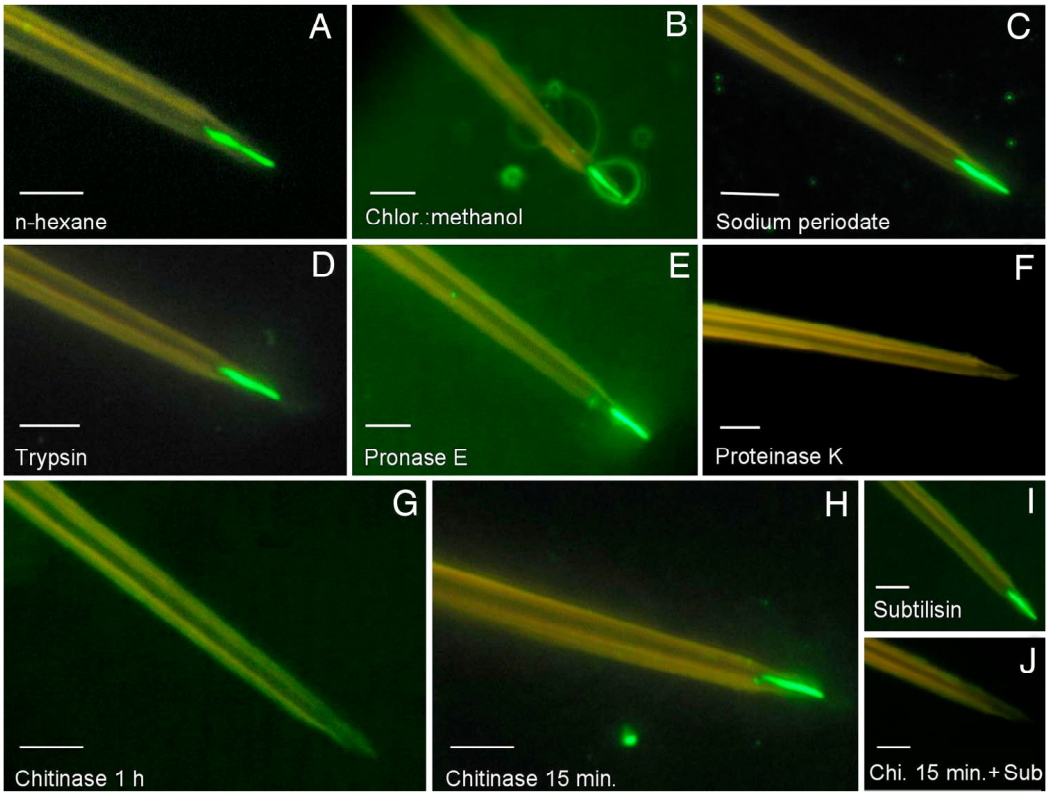


Figure 17 : Caractérisation de la nature chimique du récepteur hypothétique du CaMV dans les stylets de son puceron vecteur. Les stylets maxillaires sont pré-traités avec les composés indiqués avant incubation avec P2-GFP puis observation sous microscope à épifluorescence. Extraction des lipides polaires (A) et non polaires (B) avec du n-hexane et un mélange chloroforme :méthanol, respectivement. (C) Oxydation des carbohydrates avec du périodate de sodium. (D-F) Digestion avec diverses protéases comme indiqué. (G-I) Effet de la digestion de la chitine. Une digestion prolongée de la chitine provoque la perte du récepteur (G) alors qu'une digestion ménagée le maintient en place (H). Toutefois, la digestion modérée de la chitine semble déchausser suffisamment le récepteur pour le rendre accessible à la subtilisine (J), une protéase qui ne peut pas s'y attaquer quand il est pleinement enfoui (I). Barres : 5µm. *D'après [56].*

La zone de rétention correspond très précisément au fond du lit du canal commun, là où fusionnent le canal salivaire et le canal alimentaire (encadré 9A). Il a été vérifié que le mutant Q6Y (autrement appelé Rev5) ne se fixe pas aux stylets des espèces vectrices, suggérant que son défaut d'activité HC (mentionné plus haut) est lié à son incapacité à reconnaître le récepteur chez l'insecte. Des travaux supplémentaires avec différentes enzymes (chitinase et protéases) et solvants ont mis en évidence la nature protéique du récepteur (figure 17) qui se trouve enchâssé dans la matrice chitineuse des stylets [56].

3 Quelques aspects ciblés de génétique des populations virales

La transmission par vecteur est généralement perçue comme un moyen très efficace de dissémination virale. Or, cette efficacité apparente occulte plusieurs « difficultés » auxquelles les populations virales sont potentiellement soumises.

Un problème évident, et perceptible intuitivement, réside dans les fluctuations démographiques au cours de la transmission. Inévitablement, une population immense de particules virales (10^{12}-10^{14} virions) dans une plante infectée sera « échantillonnée », de manière plus ou moins aléatoire, par un ou des vecteurs qui ne prélèveront et ne transmettront que quelques rares génomes pour fonder la population future dans une nouvelle plante hôte. Il est facile de concevoir que ce phénomène puisse engendrer d'importantes modifications génétiques dans la population virale et donc influencer son évolution.

La composition de l'échantillon viral « fondateur », transmis à une nouvelle plante dépend : i) du nombre de vecteurs qui la visitent (densité de la population vectrice dans l'environnement), ii) du nombre de particules virales prélevées puis inoculées par chaque vecteur, iii) de la distribution de la diversité génétique de la population virale au sein de son hôte (accessibilité de cette diversité pour le vecteur), iv) du mécanisme de la formation du complexe transmissible dans la plante ou le vecteur (HC-trans-complémentation ou pas), et enfin v) du comportement alimentaire du vecteur (comment il échantillonne cette diversité).

Le projet de ma thèse se place précisément dans ce contexte, à un niveau d'analyse des interactions virus-vecteur où nous avons aussi dû intégrer des paramètres de l'interaction virus-hôtes et des concepts de génétique des populations et/ou de biologie évolutive. Cette troisième et dernière partie de l'introduction générale présente donc, le plus succinctement possible, les notions ou concepts de génétique des populations virales qui ont été à la base de nos questionnements et ont présidé à la définition de nos objectifs.

3.1 Description des notions de base

3.1.1 Variabilité chez les phytovirus

Les virus de plantes, tout comme les virus d'autres organismes, sont connus pour leur remarquable diversité génétique inter et intra-espèce [140]. Elle est principalement générée par les mutations et les recombinaisons, ce qui permet potentiellement une adaptation et une évolution rapide en réponse aux modifications de l'environnement.

Les mutations sont souvent produites par les ARN polymérases ARN dépendante (RdRp) et les transcriptases inverses qui ne disposent pas d'une activité de relecture (3'-5') capable de corriger les mésappariements. Leur taux d'erreur est estimé entre 2.10^{-5} et 1.10^{-4} par nucléotide et par événement de réplication [141].

Des évènements de recombinaison ont lieu chez les phytovirus (ADN et ARN). Ces recombinaisons homologues ou non peuvent générer de nouveaux variants, combiner sur un même génome des mutations apparues sur des génomes différents, ou encore régénérer des génomes fonctionnels à partir de génomes viraux comportant des mutations très délétères ou létales. En outre chez les virus segmentés et sans événement de recombinaison, les réarrangements entre fragments génomiques entraînent une augmentation de la diversité génétique virale. Il est à noter, que ces deux derniers événements à l'origine de variabilité sont possibles si et seulement si la cellule hôte est co-infectée par plusieurs variants ; un phénomène qui sera décrit et discuté en détail au cours du chapitre II.

3.1.2 Description et propriétés de la quasi-espèce virale

Dans une cellule infectée par un clone viral, la fréquence élevée de mutations/recombinaisons et le grand nombre de cycles de réplication aboutissent rapidement à une population virale très diversifiée [142-145]. On applique alors le terme de "quasi-espèce" à cette population qui est un ensemble hétérogène de variants apparentés [146-148]. La population virale ne peut en conséquence pas être assimilée à une séquence génomique unique, mais plutôt à un ensemble de séquences mutantes qui varient autour d'une séquence moyenne ou consensus ; la séquence consensus correspond à une séquence idéale dans laquelle est figurée pour chaque position, la base nucléique la plus fréquente dans la population ; il s'agit d'une séquence théorique qui peut même ne pas réellement exister dans la population des génomes viraux. L'ensemble des séquences théoriquement possibles (espace des séquences) n'est certainement jamais totalement représenté dans une quasi-espèce virale, puisque la diversité génétique générée est contrainte par des pressions de sélection diverses et variées, ainsi que par des phénomènes stochastiques comme la dérive génétique [149] *(pour revue sur les quasi-espèces voir [150-152]).*

Si un fort taux de mutation augmente théoriquement la taille du « nuage de variants» de la quasi-espèce, favorisant ainsi potentiellement l'apparition de mutations bénéfiques au sein de la population virale pour l'adaptation à de nouveaux environnements [153-155], il provoque aussi l'apparition de nombreux génomes non viables ou délétères [156-158]. Par ailleurs, il a été prédit par des modélisations mathématiques qu'une quasi-espèce virale n'est pas une simple collection de mutants co-existants, mais un groupe de variants dont les interactions multiples sont à l'origine de caractéristiques émergeantes au niveau de la population [146,152]. La sélection opérerait alors aussi sur cet ensemble de variants, et non pas uniquement de manière individuelle sur chaque génome mutant. Une coopération entre les différents variants d'une quasi-espèce d'un poliovirus a été démontrée comme nécessaire à l'infection généralisée de l'hôte, notamment au passage de la barrière hémato-encéphalique, ce qui suggère fortement que la sélection intervient parfois à un niveau populationnel [159]. Des propriétés

émergentes au niveau de la quasi-espèce ont aussi été suggérées et discutées lors de l'infection d'une plante hôte par le *Tobacco etch virus* [160].

3.1.3 Dérive génétique et cliquet de Muller

Dans une quasi-espèce, les mutations apparaissant très fréquemment au sein des génomes, la proportion de génomes non mutés est faible. Ainsi, en l'absence de recombinaison, la simple dérive génétique peut conduire à leur perte. On dit alors que le cliquet a passé un cran, chaque cran correspondant à la perte définitive de la classe de génomes les moins mutés (figure 18B [161]). Ce phénomène, nommé "cliquet de Muller", provoque une accumulation irréversible de mutations dans les populations de tailles finies en absence de recombinaison [162].

3.1.4 Cliquet de Muller et goulots d'étranglement

La plupart des mutations étant délétères, et le cliquet de Muller conduisant à leur accumulation, il devrait engendrer une diminution drastique de la valeur sélective moyenne des populations virales [163,164]. Les goulots d'étranglement, en réduisant temporairement la taille des populations, accentuent la dérive génétique et donc l'effet du cliquet de Muller (figure 18). Ainsi, pour des virus de bactéries [165,166], d'animaux [167-169] et plus récemment de plante [160], il est maintenant établi expérimentalement que les goulots d'étranglement génétiques répétés induisent une baisse très significative de la valeur sélective dans les populations.

3.2 Connaissances chez le CaMV

3.2.1 Variabilité chez le CaMV

Les substitutions et délétions constituent les deux sources principales de la variabilité décrite chez le CaMV. Les délétions sont créées notamment : (i) au cours de la transcription inverse, lors du transfert de brin provoquant des délétions à des sites spécifiques (région intergénique entre les gènes V et VI [170], milieu du gène II [12,171], origine de réplication [172]); (ii) lors des épissages de l'ARN 35S [172-174].

Le taux de substitutions dans la région intergénique (environ 700 pb comprises entre les gènes VI et VII) a été évalué à environ 4.10^{-4} à

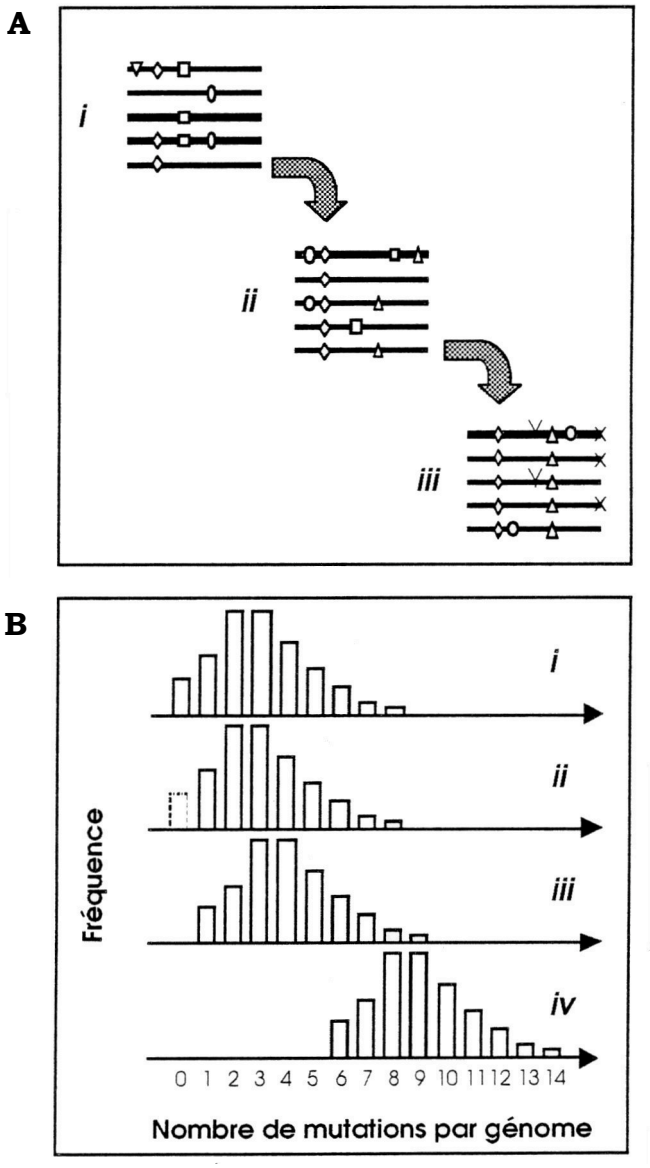


Figure 18 : Accumulation des mutations délétères au sein d'une population.

(A) par fixation allèlique au sein de la population. Une population de génomes tous différents est représentée par les différents traits et chaque symbole représente un allèle à un locus donné. Chaque flèche grise représente un goulot d'étranglement lors par exemple d'un événement de transmission.

(B) par cliquet de Muller. Un fort taux de mutation et la sélection naturelle contre les mutations délétères créent une distribution à l'équilibre du nombre de mutations au sein des génomes (i). Dans cette distribution, les génomes ayant le moins de mutations sont les mieux adaptés (avec une forte valeur sélective). Si les populations sont finies et que la dérive génétique a lieu, la classe d'individus-non-mutés peut être perdue aléatoirement (ii). Une fois que la classe d'individus-non-mutés est perdue, la population se déplace vers une autre distribution (iii) qui a la même distribution que la population ancestrale mais dont le nombre moyen de mutations est supérieur d'une mutation. Selon le nombre de « crans » du cliquet, le fardeau de mutations peut devenir de plus en plus important (iv). *D'après [161].*

6.10^{-4} substitutions par site et par cycle *in planta* (soit 3 à 5 substitutions par génome, [172]). Par ailleurs, l'analyse des 9 séquences génomiques virales (disponibles dans GeneBank), provenant de différentes plantes ayant été infectées par le CaMV en différentes localisations géographiques, a permis d'évaluer la distribution de la variabilité [175]. Les ORF IV et VI semblent être les régions les plus variables et la densité de substitutions silencieuses (ne changeant pas la séquence protéique) est globalement similaire selon les ORFs, bien qu'une plus grande densité semble présente dans l'ORF IV comparé à l'ORF II. La plupart des insertions/délétions seraient imputables au mauvais appariement des matrices durant la réplication par la transcriptase inverse, et les mutations ponctuelles seraient liées aux étapes de transcription et transcription inverse.

3.2.2 La quasi-espèce pour le CaMV

Dans l'équipe CaGeTE, un programme d'étude de la quasi-espèce du CaMV – notamment de la diversité génétique générée à chaque cycle *in planta* – a été initié à partir du clonage de 100 génomes viraux encapsidés, résultants de l'infection d'une plante à partir d'un seul clone viral initial. Le séquençage des 100 clones touche à sa fin, mais les données actuelles (sur 64 génomes complets) permettent d'ores et déjà de remarquer que 90% des génomes descendants arborent une ou plusieurs modifications par rapport à la séquence initiale (figure 19).

L'analyse systématique des propriétés biologiques de ces mutants a démontré que 47% des génomes viraux issus d'un cycle d'infection ne sont pas viables lorsqu'ils sont inoculés individuellement. En outre, l'évaluation préliminaire (visuelle) des propriétés biologiques de chaque variant de la quasi-espèce suggère que la pathogénicité observée pour une plante donnée est le reflet de la « moyenne » des effets des génomes qui la constituent. Pour exemple, deux de ces génomes isolés (clones QS-51 et QS-60) présentent une pathogénicité totalement différente de celle exprimée sur la plante d'où ils sont issus, l'un étant très sévère et l'autre atténué (figure 20) [176].

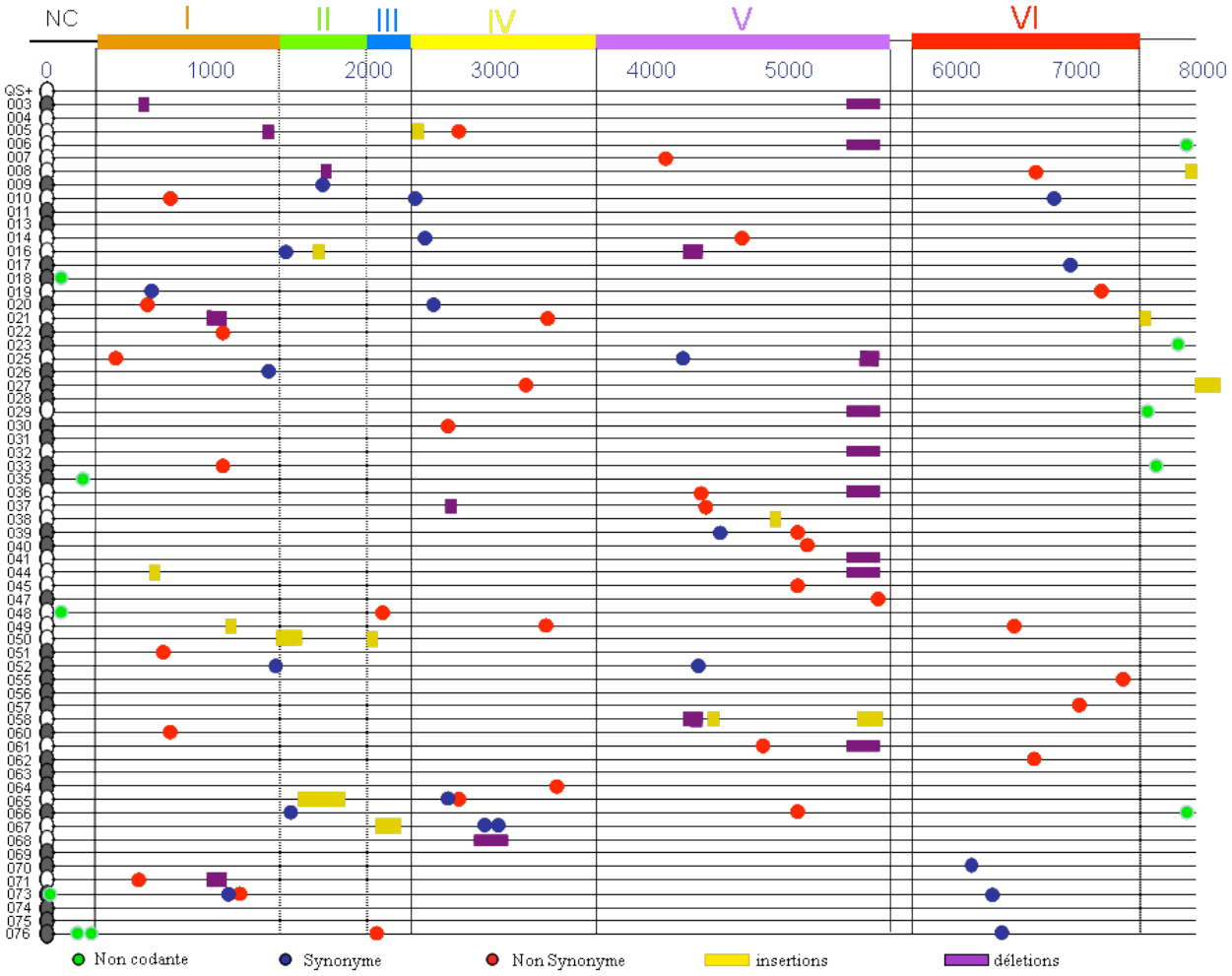


Figure 19 : Variabilité de la quasi-espèce du CaMV. L'étude de la quasi-espèce du CaMV présentée ici a été effectuée à partir du clonage de 64 génomes viraux résultants de l'infection d'une plante à partir d'une seule séquence (clone) initiale. Le séquençage de ces 64 génomes permet d'ores et déjà de remarquer la présence : (i) de délétions (en violet) dans les ORF I, II et V ; (ii) d'insertions (de séquence d'origine virale, en jaune) dans tout le génome à l'exception de l'ORF VI ; (iii) de mutations non synonymes (en rouge) dans tous les ORF sauf l'ORF II ; (iv) de mutations synonymes (en bleu) dans tout le génome mis à part l'ORF III [176]. La variabilité génétique générée en un seul passage dans l'hôte est donc très importante avec 90% des génomes qui arborent une ou plusieurs modifications par rapport à la séquence du clone initial. Les mutations présentent dans les régions non codantes sont figurées en vert. Les disques blancs ou noirs au début de chaque génome matérialisent le fait qu'ils soient viables ou non, respectivement, lors d'inoculation individuelle sur plante hôte.



Figure 20 : Propriétés biologiques (phénotypiques) de génomes isolés composant la quasi-espèce du CaMV. L'évaluation visuelle des propriétés biologiques de chacun des clones a permis de mettre en évidence des cas de symptômes atypiques extrêmes (très sévère : QS-51 et atténués : QS-60 en comparaison avec la plante dont ils sont issus : Cabb-S). Les variations de QS-51 et 60 sont toutes deux associées à une seule substitution non synonyme dans l'ORF I, à des positions relativement voisines, comme cela est visible dans la figure 19. *D'après [176].*

3.2.3 Impacts des goulots d'étranglement sur le CaMV

Rémy Froissart, dans l'équipe "Vection" (dorénavant CaGeTE) a historiquement initié un programme dans le but de suivre l'évolution *in planta* de la valeur sélective et de la virulence du CaMV en fonction de la "taille" des goulots d'étranglement génétiques imposés aux populations virales [177]. Lors de l'expérimentation, les goulots d'étranglement ont été produits de manière artificielle par de très nombreux clonages, de façon à contrôler totalement leurs tailles à chaque passage. Mais face à la charge monumentale que représentaient ces travaux et du fait de l'existence d'autres modèles viraux (non plante) s'exécutant plus aisément en cultures cellulaires, cette étude a été gelée faute de disposer d'un moyen confortable pour produire rapidement des goulots d'étranglement calibrés. L'utilisation de pucerons-vecteurs se présentait potentiellement comme un moyen rapide mais avec une efficacité inconnue vis à vis du nombre de génomes transmis.

3.3 Objectif initial de la thèse

L'objectif prioritaire de ma thèse découle précisément de ce dernier programme : la transmission par vecteur représente-t-elle un goulot d'étranglement sévère, et si oui, peut-on mesurer la taille de ce goulot d'étranglement et, pourquoi pas, le contrôler et le moduler ? Les chapitres qui composent le présent manuscrit s'articulent autour (ou dérivent) de cette question qui a clairement été la question centrale à l'origine de la thèse.
L'idée était de permettre l'acquisition de différents variants génétiques du CaMV par un puceron vecteur puis, suite à la transmission à une nouvelle plante hôte, de déterminer combien de ces variants ont effectivement participé à la nouvelle infection. Un certain nombre de connaissances préalables devaient être établies afin de pouvoir interpréter correctement les résultats produits par cette démarche.

1 - Tout d'abord, nous avons évalué la taille efficace d'une population de CaMV lors de l'infection systémique de son hôte, et l'existence de possibles goulots d'étranglement sévères à la colonisation de chaque feuille.

Outre l'information originale procurée par ce « sous-projet », il nous informe sur la disponibilité des variants génétiques du CaMV dans chaque feuille infectée qui sert de source virale pour le puceron vecteur. De plus, si la dérive génétique durant la colonisation de la plante hôte restait inconnue, il nous serait impossible d'estimer avec exactitude le nombre de variants transmis par un puceron. Nous serions en effet incapables de distinguer ceux perdus lors du transfert par le puceron, de ceux effectivement transmis mais ultérieurement perdus par dérive, lors de la colonisation systémique de la plante.

2 - Ensuite, nous avons déterminé la multiplicité d'infection cellulaire naturelle (MOI, multiplicity of infection) qui est un paramètre majeur dans le cycle de vie des virus. Ce paramètre est défini ici comme le nombre moyen de génomes du CaMV qui pénètrent dans et infectent chaque cellule de sa plante hôte. Encore une fois, ce sous-projet apporte des informations très originales en virologie et, dans le cadre de nos objectifs, il nous informe sur l'accessibilité des variants génétiques dans la population du CaMV, à l'intérieur même des feuilles, au niveau de chaque cellule individuelle. Il s'agit d'une information importante pour l'étude de la transmission, puisque les pucerons acquièrent souvent les virus non-circulants lors d'une ou quelques rares ponctions intra-cellulaires sur les plantes infectées .

3 - Finalement, avec ces informations en main, nous avons pu mettre en oeuvre notre protocole et estimer le nombre moyen de génomes du CaMV qui participent effectivement à l'infection d'une nouvelle plante hôte, suite à un événement de transmission par un puceron vecteur. A l'aide de la technique d'EPG, nous avons contrôlé le comportement alimentaire des pucerons vecteurs, et évalué l'impact de ce comportement sur la taille du goulot d'étranglement induit sur la population virale lors de la transmission.

NB : L'ensemble des méthodes nécessaires à l'obtention des résultats présentés dans ce manuscrit n'a pas été intégralement placé dans la section « Matériels et Méthodes ». En effet, de nombreuses techniques et protocoles

originaux ont dû être développés ; ils peuvent être considérés comme partie intégrante des résultats de la thèse, et sont en conséquence décrits au fil des chapitres correspondants. Seules les techniques de routine qui n'ont pas été mises au point pour nos questions spécifiques ont été rapidement listées et explicitées dans la section Matériels et Méthodes.

CHAPITRE I

Taille efficace des populations de CaMV intra hôte

Pour tenter d'évaluer la taille efficace des populations de CaMV durant l'infection systémique, nous avons choisi d'exploiter les variations quantitatives et aléatoires de la diversité génétique au sein de ces populations lorsqu'elles colonisent successivement les feuilles de leur hôte.

A cette fin, nous avons dû développer une méthodologie spécifique (QSS : Quantitative Single-letter Sequencing) qui nous permet d'identifier et de quantifier la fréquence relative de plusieurs variants génétiques, simultanément au sein d'une même population de génomes, donc dans un même échantillon d'ADN viral purifié.

Le chapitre I est donc organisé en deux parties :

- l'une décrivant la méthodologie sous forme d'un article récemment publié dans BMC Genomics

- l'autre décrivant les résultats, obtenus à l'aide de cette technique, à partir desquels nous avons estimé la taille efficace des populations du CaMV, *in planta*. Cette partie est présentée sous la forme d'un article soumis.

ARTICLE 1

Quantitative Single-letter Sequencing : a method for simultaneously monitoring numerous allelic variants in single DNA samples

Baptiste Monsion, Hervé Duborjal et Stéphane Blanc

Publié le 21 février 2008 dans BMC Genomics

BMC Genomics 2008, ***9:****85*
doi:10.1186/1471-2164-9-85

BMC Genomics

BioMed Central

Methodology article

Open Access

Quantitative Single-letter Sequencing: a method for simultaneously monitoring numerous known allelic variants in single DNA samples

Baptiste Monsion[1], Hervé Duborjal[2] and Stéphane Blanc*[1]

Address: [1]Biologie et Génétique des Interactions Plante-Parasite (BGPI), INRA-CIRAD-SupagroM, TA A-54/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France and [2]COGENICS GENOME Express SA, 38944 Meylan, France

Email: Baptiste Monsion - monsion@supagro.inra.fr; Hervé Duborjal - hduborjal@cogenics.com; Stéphane Blanc* - blanc@supagro.inra.fr

* Corresponding author

Published: 21 February 2008

BMC Genomics 2008, **9**:85 doi:10.1186/1471-2164-9-85

Received: 27 September 2007

Accepted: 21 February 2008

This article is available from: http://www.biomedcentral.com/1471-2164/9/85

© 2008 Monsion et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Pathogens such as fungi, bacteria and especially viruses, are highly variable even within an individual host, intensifying the difficulty of distinguishing and accurately quantifying numerous allelic variants co-existing in a single nucleic acid sample. The majority of currently available techniques are based on real-time PCR or primer extension and often require multiplexing adjustments that impose a practical limitation of the number of alleles that can be monitored simultaneously at a single locus.

Results: Here, we describe a novel method that allows the simultaneous quantification of numerous allelic variants in a single reaction tube and without multiplexing. Quantitative Single-letter Sequencing (QSS) begins with a single PCR amplification step using a pair of primers flanking the polymorphic region of interest. Next, PCR products are submitted to single-letter sequencing with a fluorescently-labelled primer located upstream of the polymorphic region. The resulting monochromatic electropherogram shows numerous specific diagnostic peaks, attributable to specific variants, signifying their presence/absence in the DNA sample. Moreover, peak fluorescence can be quantified and used to estimate the frequency of the corresponding variant in the DNA population.

Using engineered allelic markers in the genome of *Cauliflower mosaic virus*, we reliably monitored six different viral genotypes in DNA extracted from infected plants. Evaluation of the intrinsic variance of this method, as applied to both artificial plasmid DNA mixes and viral genome populations, demonstrates that QSS is a robust and reliable method of detection and quantification for variants with a relative frequency of between 0.05 and 1.

Conclusion: This simple method is easily transferable to many other biological systems and questions, including those involving high throughput analysis, and can be performed in any laboratory since it does not require specialized equipment.

Background

The need to analyze genetic variation within populations of various organisms has engendered a wide variety of techniques designed to identify genetic differences between related genomes [1,2]. The vast majority of methods currently available for detecting single or polynucle-

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

otide polymorphisms, as well as insertions and deletions of varying size, were originally developed for genotyping individual organisms within populations. By sampling and genotyping numerous separate individuals, the relative frequency of a given variant within a population can be estimated. However, the number of samples analyzed rapidly becomes a limiting factor, in terms of both time and cost considerations. As a consequence, a new generation of techniques is being developed that will allow more than one specific allele to be distinguished and quantified simultaneously [3-6] in a single sample containing nucleic acids pooled from several individuals [7,8]. The most successful techniques described so far have various advantages and drawbacks. On the one hand, techniques based on microarrays [9,10] or on mass spectrometry [11-14], although very efficient at sorting numerous variants within a single sample, require access to specialized equipment and are thus difficult and expensive to develop and implement. On the other hand, more amenable techniques based on real-time PCR or primer extension [15,16] often require multiplexing adjustments.

For pathogenic microorganisms, such as fungi, bacteria and especially viruses, populations infecting an individual host are often highly variable, intensifying the difficulties involved in distinguishing and accurately quantifying numerous alleles or variants co-existing in a single nucleic acid sample. This is best illustrated by considering the properties of viral populations, which exist as a swarm of related mutants commonly designated as a quasispecies [17]. One remarkable property of viral quasispecies is that the relative frequency and distribution of specific variants can vary at different phases of the virus life cycle [18]. In this situation, the ability to simultaneously monitor those specific variants can become critical in understanding the biology and evolution of the virus [19]. Similarly, monitoring changes in allele frequency is essential to the understanding of the evolution of viruses or other pathogens, as these changes reflect forces such as selection during adaptation [20] and genetic drift in phases of the life cycle where the effective population size is low [21,22].

We faced exactly this challenge when evaluating the effective population size of *Cauliflower mosaic virus* (CaMV) during systemic invasion of plant host tissues (to be described elsewhere). For this purpose, and because of concerns related to the available genotyping techniques mentioned above, we developed a novel analysis method based on classical detection of genetic variants by dideoxy fingerprinting [23,24]. This method, named Quantitative Single-letter Sequencing (QSS), allowed the simultaneous and accurate monitoring of six engineered allelic CaMV variants over time, determining in a single nucleic acid sample and in a single reaction process both the presence/absence and the relative frequency of each allele in the viral genome population.

Here, we describe the QSS method and its application to samples consisting of mixtures of purified plasmids or viral genomes extracted from a multiply infected host plant. We test the accuracy, reproducibility, robustness and limits of this method, and discuss its potential application to other biological systems in any standard laboratory setting. Since QSS applies to known sequences only, it will be useful for monitoring experimental populations, as well as natural populations where the different variants at the targeted locus have been previously identified.

Results

Single-letter sequencing distinguishes several markers in mixed plasmid solutions

Genetic markers were engineered into the CaMV genomes at exactly the same position. Thus, examining single letter (A) sequence traces from a mixed population should, in theory, reveal some common peaks at positions where more than one marker harbours a A, and discriminating peaks at positions where only one marker does so. Figure 1A shows an alignment of the marker sequences, highlighting the theoretically discriminating A-positions in green. Between one and three discriminating A-positions are expected for each of the 6 markers used in this study.

The actual observed positions of A-peaks from the amplified and single-letter-sequenced pCaVIT1-6 plasmids showed slight differences from their theoretical positions relative to a co-electrophoresed labeled DNA ladder (Figure 1B). Sequencing of each individual CaMV VIT clone was repeated three times, and the observed position of the A-peaks proved to be highly reproducible, with the position of each peak always occurring in the same place to within ± 0.2 bp (not shown). Therefore, we defined an experimental list of discriminating A-positions, using the actual observed peak positions rather than their original theoretically predicted positions, as explained and highlighted in red in Figure 1B.

This experimental list of discriminating A-positions was validated when all six markers co-existed in the DNA sample and were analyzed together. A mix of equal amounts of all pCaVIT1-6 plasmids was used as a template for PCR and subsequent single-letter sequencing. All experimentally-confirmed discriminating A-peaks were easily distinguished at their specific position on the electropherogram (Figure 1C), demonstrating that pooling plasmids in a single analysis did not lead to interference between the different markers. Hence, for each pCa-VIT plasmid, from one to three discriminating peaks can be reliably detected on the electropherogram and indicate the presence/

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

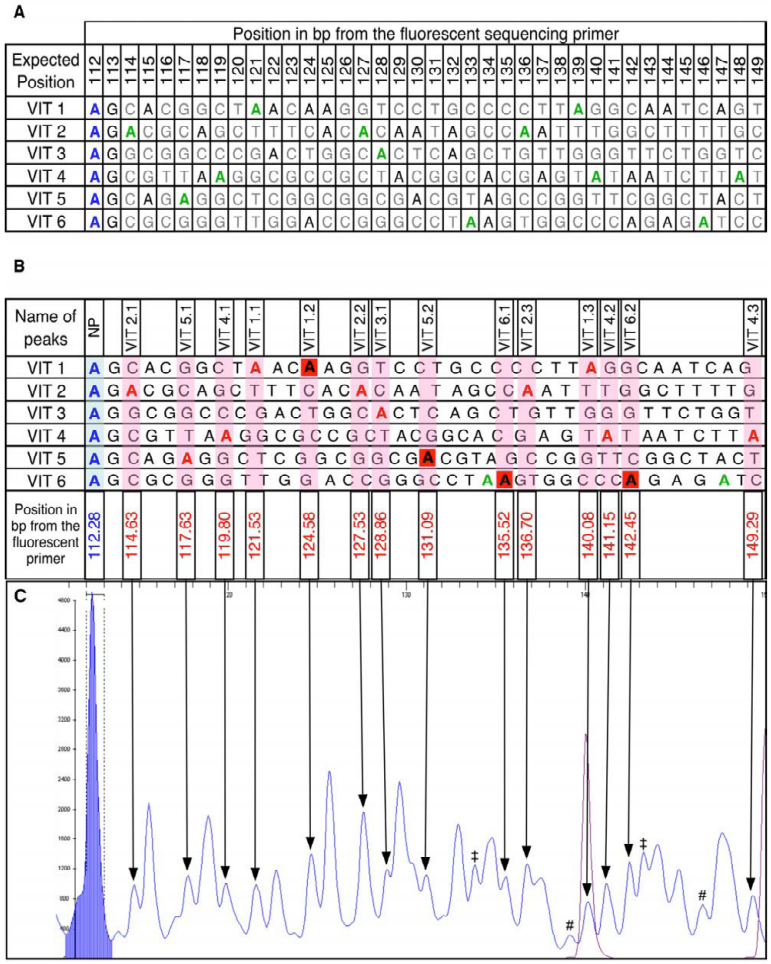


Figure 1

Sequence signatures identify different markers in a mixed DNA solution. (A) Sequences of the six genetic markers pCaVIT-1 to -6. The blue A on the left is the last common adenine shared between all pCa-VIT1-6 plasmids upstream of the polymorphic region. Adenine residues marked in green are theoretically expected to yield discriminating peaks; ther A residues are indicated in black. The expected position (in bp) of each base from the fluorescent sequencing primer is indicated at the top. (B) Experimentally observed positions of discriminating A-peaks (named VIT1.1 to VIT6.2 at the top) on sequence traces (not shown) from individually sequenced markers. Sequences of all six markers are slightly distorted in order to match the actual position of A peaks observed on individual electropherograms. The red A residues were each confirmed to yield a discriminating peak. Though not theoretically predicted to do so, the red-boxed As proved to yield discriminating peaks experimentally. The remaining green As failed to produce discriminating-peaks. Sequencing reactions of pCa-VIT1-6 were repeated three times, and the observed position of all peaks was highly reproducible: +/- 0.2 bp among repeats. At the bottom, the number in blue is the average position of the last A common to all sequences upstream of the differential markers. The numbers in red are the average positions of the observed discriminating peaks. (C) Single-letter Sequencing electropherogram from a mixed DNA solution containing all 6 markers. A mix of all 6 pCaVIT1-6 plasmids in equal amounts was used as a template for PCR and subsequent single-letter sequencing. All discriminating-peaks defined in B are indicated by arrows on the electropherogram. The shaded blue peak corresponds to the last common A. The scale at the top indicates the nucleotide position relative to that of the sequencing primer. The scale on the left is used for scoring the height of the peaks in arbitrary units provided by the STRand program. The two purple peaks correspond to molecular weight markers. # These peaks are artefacts. ‡ Though appearing as discriminating, these peaks actually overlap small artefactual peaks observed on at least one electropherogram from individually sequenced markers (not shown). They are thus not used further.

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

absence of the corresponding marker, determining the qualitative composition of the DNA sample.

In all subsequent experiments, only those peaks that perfectly matched their expected position (± 0.2 bp of the position shown in Figure 1B) were considered reliable and included in the analysis.

Standard curves for quantification from single-letter sequence traces

We used the last peak common to all sequences before the polymorphic region (shaded blue in Figure 1C) – called the normalizing-peak – to normalize the recorded data. The height of the normalizing-peak is contributed to by all the DNA molecules in the mixed solution, and is thus generated by 100% of the sequences in the analyzed sample. In contrast, the discriminating peaks are generated only by those sequences containing an A at the corresponding position. Each arrested DNA molecule produced by the sequencing process contains one single fluorochrome linked to the sequencing primer; hence the height of a discriminating-peak is proportional to the relative frequency of the corresponding marker in the DNA solution.

In order to prepare standard curves for the quantification of peak heights, we first prepared a series of mixed plasmid solutions containing pCa-VIT1 at a relative frequency of 5, 10, 15, 25, 50, 75 and 100%, maintaining equal amounts of the other five pCa-VIT plasmids in all but the 100% pCa-VIT solution. Similar mixed solutions were prepared for pCa-VIT2, pCa-VIT3, pCa-VIT4, pCa-VIT5 and pCa-VIT6 and processed for PCR amplification and single-letter sequencing as described in Methods.

To extract information from the electropherograms on the relative frequency of each marker present in a mixed solution, we assigned an intensity value (I) to each discriminating-peak, corresponding to its height relative to that of the normalizing-peak. For each given discriminating-peak, a standard curve was constructed by plotting the known frequencies of the corresponding marker against the observed I values (Figure 2). One standard curve for each discriminating-peak was established from 3 replicates at 100%, 2 replicates at 75, 50 and 25%, and 10 replicates at 15% of the corresponding marker in the mixed plasmid solution. In these replicate runs, the I values were highly reproducible and showed very limited variation (Figure 2). However, I values obtained at 10% and 5% were more variable and were not used in establishing the standard curves (mean standard deviations were 3 and 3.25%, respectively).

The polynomial regression function corresponding to each standard curve was calculated using Excel (Microsoft). The R^2 values obtained for all standard curves were remarkably high, and two examples (corresponding to curves with the highest and the lowest R^2 values) are shown in Figure 2.

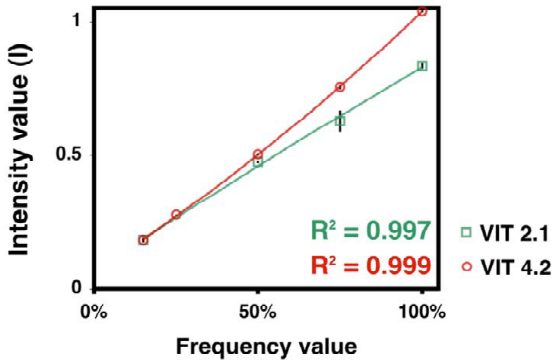


Figure 2
Standard curves for converting the intensity of discriminating peaks into frequency of corresponding markers. Standard curves were established for all 14 discriminating-peaks shown in Figure 1. Only those for VIT2.1 and VIT 4.2 are shown here, as they illustrate the worst and the best fit, respectively, between the recorded data and the polynomial regression function calculated using Excel (correlation coefficients R^2, are shown). Vertical bars represent standard deviation among repeats.

Accuracy of quantification of genetic markers within mixed DNA populations

The standard curves shown in Figure 2 were used to directly transform the observed I values into the frequency of the corresponding marker in test DNA solutions. Parallel estimates were possible when more than one discriminating-peak was available for a given marker. In such cases, the frequency was taken as the mean of the parallel estimates.

The accuracy of the quantification was assessed by comparing the measured frequencies to their expected values on plasmid mixes prepared from the reference stocks. Several mixes were tested (compositions listed in Table 1). Whether the mixes contained only 2 or all 6 markers, QSS yielded estimates remarkably close to the expected values for all markers tested (Table 2). Surprisingly, when the estimates for each marker present in a given solution were summed up, the total often deviated from 100%, ranging from about 90 to 120% (Table 2). This phenomenon was attributed to fluctuations in the baseline of sequence traces and was corrected by proportionally adjusting all marker frequencies to give a total of 100% (Table 3). This final correction yielded estimates even closer to the true

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

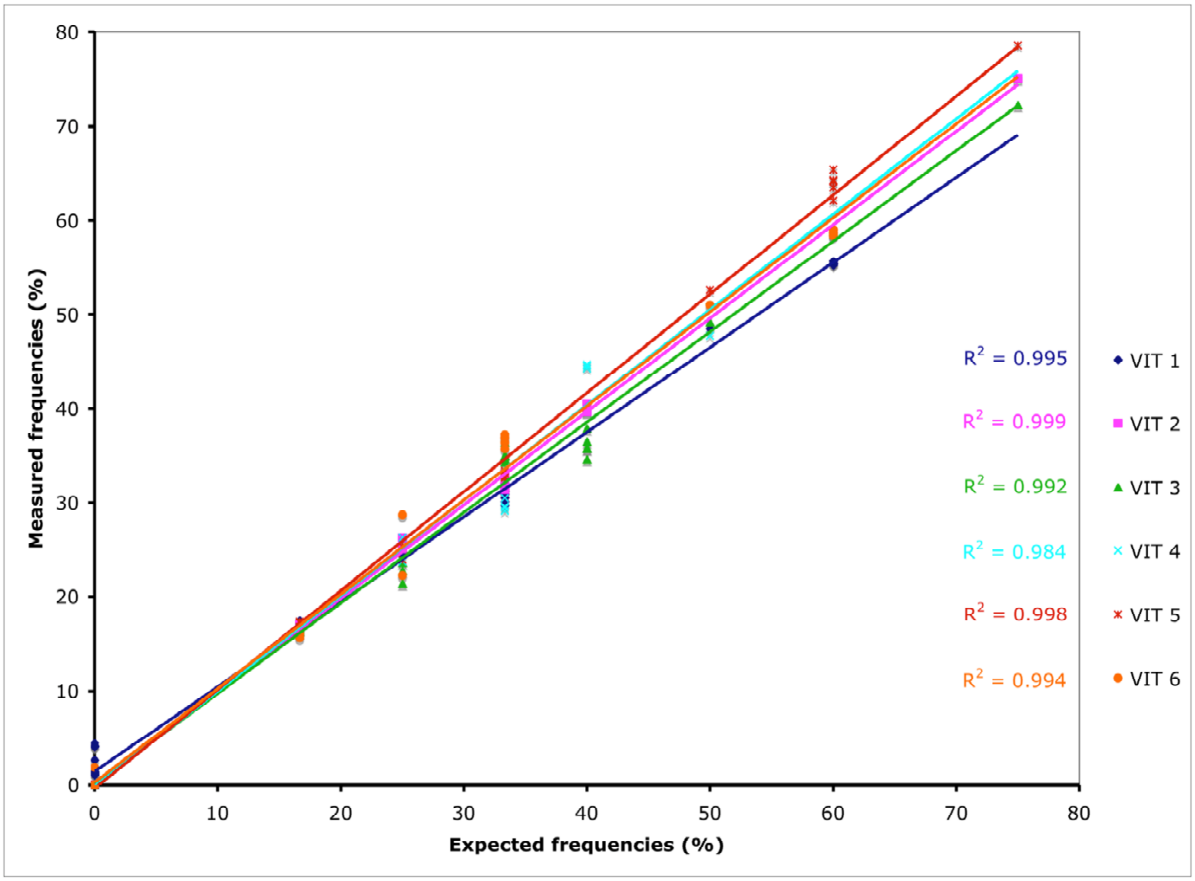


Figure 3
Accuracy of QSS measurements for the 6 genetic markers. Observed values are plotted against expected values for all measurements summarized in Table 3, plus four independent repeats of the analysis of mixes N°9 to 14. For each VIT marker, a linear regression function was deduced (colored lines). The near perfect scores for the R^2 correlation coefficient in each case illustrate the high degree of accuracy of QSS.

marker frequencies (discussed further below). The overall accuracy of the method is illustrated in Figure 3, which shows the linear regression function correlating expected values to the estimates. The R^2 correlation coefficient was excellent for all VIT markers, being between 0.984 and 0.999.

The name Quantitative Single-letter Sequencing (QSS) is used to describe the overall process of identifying discriminating peaks, recording their normalized height, and converting them into frequencies of genetic markers in a mixed DNA population.

Reproducibility of QSS in repeated analyses of plasmids and virus mixes

We performed five independent PCR-amplifications of plasmid mix N°9 (Table 1), containing equal amounts of all six markers, and further performed QSS analyses. In all five independent repeats, the estimated frequency of each marker (Table 4), ranging from 15.62 to 17.45%, was very close to the expected value of 16.6%. The standard deviation among repeats was very small (< 0.5%) for all markers considered (median of the 6 SD = 0.281%), confirming the high reproducibility of the process.

A similar reproducibility evaluation was also performed on a more realistic biological sample. A mixed population of viral DNA genomes was extracted from a single infected plant, originally inoculated with the CaMV Mix6VIT virus

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

Table 1: Composition of plasmid DNA mixes

	Percentage of each VIT					
Mix n°	*VIT1*	*VIT2*	*VIT3*	*VIT4*	*VIT5*	*VIT6*
1		75				25
2			75	25		
3			25		75	
4	50		25			25
5	25	25	50			
6	25			50	25	
7			25	25	50	
8		25			25	50
9	16.67	16.67	16.67	16.67	16.67	16.67
10	33.33				33.33	33.33
11		33.33	33.33	33.33		
12			40		60	
13		40				60
14	60			40		

Different mixes containing some or all pCa-VIT plasmids in the proportions indicated.

suspension as described in Methods. The extracted viral DNA was used in 19 independent PCR amplifications followed by QSS analysis (Table 5). Several conclusions can be drawn from the results of this experiment. i) QSS can be applied to "natural" samples, e.g. viral DNA solutions extracted from infected plants. ii) QSS is highly reproducible if genetic variants are sufficiently frequent in the population (above 5%), as the standard deviation among numerous repeats was very small. iii) The fact that the 6 markers were present in widely varying amounts allowed evaluation of the minimum frequency that can be reliably

Table 2: Accuracy of QSS on mixtures of plasmid DNA before corrections

	Percentage of each VIT						Total (%)
Mix n°	*VIT1*	*VIT2*	*VIT3*	*VIT4*	*VIT5*	*VIT6*	
1		74.8				22.2	97
2			78.02	28.5			106.52
3			19.4		71.1		90.5
4	50		23.5			29.6	103.1
5	27.3	29.1	54.6				111
6	27.8			52.9	28.1		108.8
7			23.3	23.4	51.7		98.4
8		22.7			21.6	46.2	90.5
9	17	16.9	16.6	16.8	16.7	16.5	100.5
10	30.3				32.1	34.6	97
11		36.4	40.1	33.5			110
12			36.5		59.8		96.3
13		39.2				58.3	97.5
14	66.8			53.1			119.9

Marker frequency quantification, obtained from the standard curves shown in Figure 2. The sum of frequencies in each mix is shown on the right.

Table 3: Accuracy of QSS on mixtures of plasmid DNA after final corrections

	Percentage of each VIT						Total (%)
Mix n°	*VIT1*	*VIT2*	*VIT3*	*VIT4*	*VIT5*	*VIT6*	
1		77.1				22.9	100
2			73.3	26.7			100
3			21.4		78.6		100
4	48.5		22.8			28.7	100
5	24.6	26.2	49.2				100
6	25.6			48.6	25.8		100
7			23.7	23.7	52.6		100
8		25.1			23.9	51	100
9	16.9	16.8	16.6	16.7	16.6	16.4	100
10	31.2				33.1	35.7	100
11		33.1	36.4	30.5			100
12			37.9		62.1		100
13		40.2				59.8	100
14	55.7			44.3			100

Frequency values after proportional correction to give a total of 100% in each mix.

measured by QSS. When considering the standard deviation calculated for each marker among the 19 repetitions, it is clear that the method can efficiently detect and reliably quantify genetic variants present in the population at a frequency above 5% (Markers VIT1-4). This 5% frequency appears as the threshold for QSS reliability, both because corresponding variance between repeated measures proved too high when establishing the standard curves (see above), and because rarer markers (VIT5 and VIT6) yielded discriminating peaks close to the sequencing baseline and could not be exploited unequivocally. iv) Finally, any variation observed for a given marker among the repeated experiments was only slightly affected by the quality of sequence traces (data not shown), an observation indicating the "robustness" of the method as specifically evaluated in the next section.

Robustness of QSS on samples of sub-optimal quality

In the analysis of viral samples from the CaMV Mix6VIT-infected plant (Table 5), repetitions 1–11 and 12–19 were performed using 2 ng and 5 ng of viral DNA matrix, respectively, at the initial PCR amplification step. Presumably due to differences in the efficiency of the PCR, or for other unknown reasons, these two subsets yielded final trace sequences of different quality (not shown). Higher baselines in the former subset seemingly resulted in biased sums of marker frequency, which totalled between 116 and 137% before final correction (Table 6), whereas such biases were smaller in the latter subset, with total frequencies ranging from 110 to 117%. However, the final corrected estimates from both data subsets showed only very small differences in the mean frequency of markers, as well as in the standard deviation observed between

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

Table 4: Reproducibility of QSS on mixtures of plasmid DNA

	Percentage of each VIT						Total (%)
Replicates of Mix N°9	*VIT1*	*VIT2*	*VIT3*	*VIT4*	*VIT5*	*VIT6*	
1	16.92	16.78	16.55	16.72	16.59	16.43	100
2	17.45	16.99	17.14	16.12	16.19	16.12	100
3	17.17	16.58	16.81	16.74	16.79	15.92	100
4	17.06	16.47	16.32	16.68	17.33	16.14	100
5	17.19	16.36	17.03	16.60	17.21	15.62	100
Mean	17,16	16,63	16,77	16,57	16,82	16,05	100
SD	0.195	0.251	0.336	0.261	0.465	0.301	

Reproducibility of QSS as evaluated by the standard deviation (SD) on 5 replications of the analysis of Mix n°9. Proportional correction of the frequency of all markers was applied to give a total of 100% in each mix.

repeats within subsets (Table 7). We thus conclude that, even on samples yielding trace sequences of sub-optimal quality, QSS still provides accurate and reproducible information on the genetic composition of a DNA population.

Discussion

This report describes the development of a novel technique – Quantitative Single-letter Sequencing (QSS) – allowing simultaneous quantification of the frequencies of numerous individual allelic genetic markers in a mixed DNA population without the need for multiplexing adjustments. Although several techniques for the simultaneous detection and quantification of multiple variants at a single locus have been described previously [25-27], the QSS method has several advantages making it an interesting alternative. Implementation of QSS requires no specialized equipment, and it involves only a single PCR step and one sequencing reaction, allowing high throughput analysis. Two sequencing plates (96 wells each) were sufficient to establish all the standard curves in this study, and we were able to subsequently process hundreds of individual biological samples within a week. Due to these practical advantages, our method competes very well with easily accessible techniques based on real-time PCR [28], primer extension [26], quantitative sequencing [28], or RFLP analysis [29]. Indeed, the simultaneous monitoring of up to 6 CaMV variants afforded by QSS represents a significant progress, as other techniques require complex multiplexing setups that have thus far precluded efficient monitoring of more than 2 or 3 allelic variants in a single reaction tube. A concomitant study, recently published by another research group, used complex (four-letter) sequence electropherogram and statistical modeling to extract qualitative and quantitative information from bacterial DNA samples containing several allelic variants [30]. Remarkably, this study and QSS together confirm that methods based on quantitative sequencing can be applied to a wide variety of microorganisms. The two methods however have essential differences. QSS uses simple single-letter sequence traces, allowing the direct quantification of variants, exclusively from discriminating peaks, with no statistical modeling. Moreover, the use of fluorescent sequencing primer in QSS theoretically allows

Table 5: Accuracy and reliability of QSS on viral DNA extracted from an infected plant

	Replicate n°																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Average	SD
VIT 1	42	40	42	41	41	41	40	41	41	44	42	44	44	43	43	42	43	42	43	42.0	1.3
VIT 2	5.4	5.6	5.5	5.8	5.9	6.6	6.0	5.8	6.1	4.9	5.6	5.2	5.2	5.1	5.3	6.6	5.8	5.3	5.4	5.6	0.5
VIT 3	39	36	40	38	37	39	39	40	39	41	39	41	42	42	42	38	40	42	42	39.9	1.8
VIT 4	11	12	11	11	12	11	11	11	11	9	11	9	8	9	9	11	9	9	9	10.3	1.2
VIT 5	1.4	3.3	1.1	1.9	2.5	1.3	2.2	1.3	1.2	1.1	1.4	0.6	ND	0.5	0.6	2.8	1.4	1.1	1.1	1.5	0.7
VIT 6	1.1	2.9	1.0	1.9	1.7	1.4	2.1	1.2	1.0	ND	0.7	ND	ND	ND	ND	ND	ND	ND	ND	1.5	0.6

Viral DNA was extracted from a plant infected with the CaMV Mix6VIT, and processed independently 19 times for PCR amplification and QSS analysis. Average final estimates (proportionally corrected to give a total of 100% in each case) and standard deviation (SD) among repeats are shown on the right.

Peaks yielded by low-frequency markers VIT 5 and VIT6, when identified, emerged just above the base line of the sequence traces. As indicated in the text, the corresponding frequency estimates reported in this table fall below the threshold of QSS accuracy, as confirmed by the high SD associated with VIT5 and VIT6, and by numerous repeats where estimates could not be obtained (ND), because the corresponding peaks were not clearly distinguishable on the electropherogram.

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

Table 6: Sum of the frequency estimates of all the markers before final correction

Replicate n°																		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
125	137	120	128	129	123	125	122	122	116	121	111	110	111	110	117	110	112	111	**Total amount (%)**

The sum of the frequency estimates of all markers before final correction is given here as an indicator of the quality of the sequence traces. Two subsets of differing quality can be distinguished (see text): replicates 1–11 (sums range from 116–137%), replicates 12–19 (sums range from 111–117%).

multiplexing in future development of the method. For instance, when two distinct polymorphic regions are present in the DNA sample, the use of two sequencing primers labeled with different fluorochromes can be envisaged, each targeting a specific locus. Such multiplexing, with two distinct unlabeled sequencing primers is not even theoretically possible with the use of dye-terminators (as in [30]), as the deciphering of overlapping sequences from the two distinct loci would be impossible.

The remarkable accuracy of the estimates obtained (MD and R^2) and the very small standard deviation (SD) between repeats, make QSS equally, if not more, reliable than other techniques for which we could extract similar parameters from the literature (Table 8). We also tested the method on DNA samples of different quality (purified plasmid mixes versus viral DNA extracted from plants) and used variable amounts of initial template for PCR, and found that even in sub-optimal conditions yielding sequence traces with relatively high background or baseline, estimates of the genetic composition of the population were only slightly affected.

During development of the QSS method, we observed an intriguing phenomenon (already mentioned above and presented in Table 2). When summing up the estimated relative frequency of all variants within a DNA sample, a value different from 100% was most often observed. We attribute this phenomenon, at least partly, to the fact that in most cases the baseline of the sequence traces is not exactly zero (not shown). Normalization of the height of the discriminating peaks to that of a peak common to all sequences (the normalizing peak) was intended to correct for this kind of fluctuation, but obviously did not completely do so when the baseline deviated significantly from zero. Moreover, attempts at normalizing discriminating peaks to any other common peak, positioned before or after the marker region, or even with an averaged height of several of these common peaks, did not significantly ameliorate this phenomenon (not shown). One explanation could be that the baseline is the same for both the normalizing peak and the discriminating peaks. Since the former is contributed to by all sequences within the population and the latter solely by a fraction thereof, the proportion of the total height of different peaks attributable to the baseline will be different. We believe that this leads to a slight overestimation of the discriminating peak, yielding sums above 100% when the baseline is above zero, with a possible distortion inversely correlated to the relative frequency of the genetic variant within the population. Although it may be possible to correct this undesirable artifact by further development of the QSS method, we show here that a simple proportional adjustment of all values to give a total of 100% is sufficient to provide an excellent approximation of true marker frequencies.

Table 7: Robustness of QSS on samples of sub-optimal quality

	Average		SD	
	1–11 set	12–19 set	1–11 set	12–19 set
VIT 1	41.2	43.0	1.1	0.7
VIT 2	5.8	5.5	0.4	0.5
VIT 3	38.8	41.3	1.3	1.3
VIT 4	11.1	9.2	0.7	0.6
VIT 5	1.7	1.0	0.7	0.8
VIT 6	1.5	ND	0.6	ND

Mean frequency estimates and SD for all markers within the two subsets shown in Table 6.

Conclusion

QSS is very well suited to monitoring allele frequency changes in populations of pathogens such as viruses, and probably fungi or bacteria, in either single host extracts or pooled extracts from numerous hosts, provided that the polymorphism at the locus under surveillance is previously identified. It could also be applied to estimation of homo- or heterozygosity in di- or polyploid organisms, as well as to pools of DNA genomes extracted from numerous individuals to determine the relative frequency of variants within populations.

The genetic markers introduced into the CaMV genome in this study were not designed specifically for the development of QSS, and this has two theoretical implications. On the one hand, it suggests that designing markers specifically for QSS (i.e. several discriminating peaks per variant with appropriate spacing between each) would likely avoid partially overlapping peaks (such as those visible in Figure 1C) and thus allow monitoring of even more co-existing variants in a population. On the other hand, it shows that QSS can be applied to any non-specifically

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

Table 8: Compared performance of QSS and other available methods

Methods and associated references	MSD[a] (%)	MD[b] (%)	R^2
QSS	0.281 to 0.950	1.255	0.984 to 0.999
BAMPER [38]	3.8*	ND	0.9999
Micro-Array [39, 40]	3.5 to 4.1	2.4	0.971 to 0.9921
Micro-Satellite [41]	ND	ND	0.97
Mass Spectrometry [29]	1.55	ND	ND
PE+DHPLC [42]	1.4	1.2	0.977
Pyrosequencing [29, 43, 44]	0.07* to 1.9	ND	0.979 to 0.996
Quantitative Sequencing [28]	4.2	1.44	ND
RFLP [29]	2.8	ND	ND
RFMP [25]	ND	ND	0.992
Single base extension [26, 29]	0.27 to 1.75	1.5 to 2.15	ND
SYBR Green [28, 45]	1.65 to 6.47	1 to 1.12	0.997
TaqMan Probe [28, 29, 46]	0.75 to 3.18	1.47	0.9984

[a] Reproducibility is evaluated by the median of standard deviations (MSD) among repeats on different variants, and is expressed as a frequency in %. For QSS, the overall MSD is 0.401%; the values shown here, 0.281 and 0.950% are calculated from Table 4 (plasmid mix) and Table 5 (viral DNA population extracted from plant), respectively.

[b] Accuracy is estimated i) by the median deviation (MD), being the median of all differences between observed and expected allele frequencies, and ii) by the correlation coefficient between observed and expected frequencies (R^2).

ND: Not-Determined

(*) these values refer to standard deviation (SD), because the data available in the literature do not allow calculation of the MSD.

designed sequence variation, such as natural variants, again with the requirement that polymorphism at the corresponding locus is previously known. The method proved very efficient even with marker VIT3, where only one discriminating peak was available. Thus, the polymorphic region can either be very diverse, or bear only minimal sequence changes generating only a single discriminating peak.

Methods

Engineering markers in full-length CaMV clones

CaMV is the type member of the genus *Caulimovirus* and has a circular dsDNA genome of around 8000 bp, depending on the isolate. We used plasmid pCa37 [31], containing the full-length genome of the Cabb-S isolate of CaMV, to amplify viral gene II with primers P2Spe5' (5'-GGACTAGTATGAGCATTACGGGACAACCG-3') and P2Spe3'Killer (5'-AGCTCCTAGGTTAGCCAATAATATTCTTTA-3'). The resulting PCR product was digested with SpeI and AvrII, and cloned into plasmid pΔII-S at the unique *Spe*I restriction site separating ORFs I and III [32]. Reintroducing the PCR-amplified gene II into pΔII-S generated the full-length clone pCa-VIT0 retaining a unique SpeI restriction site between ORFs I and II.

Six different short dsDNA sequences of 40 bp each (generating the six distinct genetic markers shown in Figure 1A), with SpeI cohesive ends, were then inserted into the SpeI site in pCa-VIT0 (the sequences of the oligonucleotides used for generating these dsDNA markers are available upon request). The six clones obtained were named pCa-VIT1 to pCa-VIT6. All plasmids were purified using a Midiprep Kit (Qiagen), verified by sequencing, and stored in distilled water at -20°C until use.

Quantifying pCa-VIT plasmids for PCR template preparation

Purified pCa-VIT1-6 plasmids stocks were thawed and carefully quantified by measuring the UV absorption at 260 nm of two dilutions for each plasmid (1 and 1/10), and duplicating each analysis in two different spectrophotometers (Varian Cary 50; Shimadzu UV-160A). A 1 mL solution, designated as the reference stock, was immediately prepared for each plasmid at a concentration of 50 $ng \cdot \mu L^{-1}$, and the concentration was verified again using the same two spectrophotometers before storage at -20°C. All single or mixed plasmid solutions subsequently used as PCR templates in the present study were prepared from these reference stocks.

Infectivity and stability of CaMV-VIT clones

Turnip plants (*Brassica rapa* var. "Just Right") were maintained in an insect-proof greenhouse under controlled conditions (25/19°C day/night with a photoperiod of 16/8 hours day/night).

Seedlings at the three-leaf stage were mechanically inoculated with *Sal*I-digested plasmid (pCa37 or one of the pCa-VIT1-6 plasmids; 2 µg in 1× TE buffer per plant), as previously described [33]. Symptoms characteristic of systemic CaMV infection appeared between 14 and 21 days post inoculation (dpi) in all cases, indicating that the infectivity of the marker-containing CaMV-VIT1-6 clones was comparable to that of the wild type clone.

BMC Genomics 2008, **9**:85 http://www.biomedcentral.com/1471-2164/9/85

At 21 dpi, each CaMV-VIT1-6 virus clone was transferred to new healthy plantlets by mechanical sap inoculation. After two successive plant-to-plant passages (21 days each), viral DNA was extracted from single plants as described below, and the integrity of all six genetic markers was verified by sequencing, confirming their stability.

Inoculation of plants with a mixed CaMV-VIT population

Leaves systemically infected with CaMV viral clones were collected at 21dpi, ground in buffer 1 (200 mM Tris, pH 7, 3v/w), and stirred overnight with 1.5 M urea and 2% triton-X-100 (final concentration). These crude extracts were then clarified by centrifugation at 3,000 × g for 15 min., and the supernatants were further ultra-centrifuged through a 15% sucrose cushion at 200,000 × g for 2 hours. The virus particles in the pellet were suspended in 600 µl of buffer 2 (100 mM Tris, pH7; 2.5 mM $MgCl_2$), submitted to final clarification at 15,000 × g for 20 min. and stored at -20°C until use.

Equal volumes (200 µl) of virus particle preparations from plants infected with each of the 6 CaMV-VIT clones were pooled to produce a mixed population designated Mix6VIT. Note that we were aware that the different CaMV-VIT clones are not necessarily present in equal amounts in this mix at this point. Young healthy plantlets were then mechanically inoculated by carefully rubbing 20 µl of Mix6VIT on the total surface of three young leaves previously powdered with the abrasive carborundum. Symptoms indicative of CaMV infection appeared within 8 to 11 days, and all plants were considered systemically infected at 13 dpi.

For unknown reasons, CaMV symptoms always appear much earlier on plants inoculated with viral particles than on those inoculated with plasmid DNA, as in the previous section.

Purification of viral DNA and PCR amplification

CaMV genomic DNA was purified from infected plants at 13 dpi according to the protocol described previously [34], plus an additional cleaning step using the Wizard DNA Clean-up kit (Promega).

PCR amplification of viral sequences was performed in a total volume of 100 µl, using VENT® DNA Polymerase (New England Biolabs) according to the supplier's recommendations. Samples (2–5 ng) of plasmid mix or purified viral DNA were amplified by 35 PCR cycles, using the CaMV-specific primers F748 (5'-CTTGGAGCGGTCAAAATATTG-3') and Ris7 (5'-GTTGGGTACCTAAGGCTTCTAATATCTC-3'), generating a 1301 bp fragment spanning 601 and 660 bp upstream and downstream, respectively, of the introduced genetic marker. The primers were designed relatively far away from the polymorphic region to alleviate putative problems related to allele specificity during PCR amplification. The amplified DNA fragment was finally purified with a double phenol/chloroform extraction and ethanol precipitation, suspended in 45 µl water and processed for single-letter sequencing as described below.

Single letter sequencing

The single-letter sequencing reactions were carried out by COGENICS Genome Express according to a proprietary protocol initially developed for DACS© technology [35]. Briefly, the quality and quantity of each PCR product were assessed by agarose gel electrophoresis. The PCR products were diluted as required and sequencing reactions were performed with 13 fmol of template DNA. The dye-primer single-letter sequencing reaction was performed using the oligonucleotide 6FAM-1236bis (5'-ccttcaaataggtaacagtgc-3' linked to a 6FAM fluorochrome at its 5' end) and ddATP (0.1 mM final concentration) as the sole terminator to halt strand elongation [36], thus labeling only fragments terminated by adenine (A). [Note – In this study, we produced trace sequences only for the targeted base adenine (A), although theoretically any other nucleobase could be similarly targeted.] After sequencing, the reactions were purified by ethanol precipitation and suspended in formamide containing a labeled DNA ladder (GenScan 500 Liz; Applied Biosystems, Foster City, CA, USA). The single-letter sequencing signatures were analyzed on a capillary electrophoresis ABI 3730 DNA analyzer (Applied Biosystems) according to the manufacturer's instructions for DNA fragment analysis.

The 6FAM-1236bis sequencing primer was positioned 112 bases upstream of the marker region – a compromise between too long a distance engendering a decrease in fluorescence sequencing signals, and too short a distance increasing the risk of putative allele specificity problems. In our case, the electropherogram was exploitable without significant decrease in signal intensity from 75 to 200 bp downstream of the fluorescent sequencing primer.

Extracting data from single-letter sequencing electropherograms

From the scanned images of the single-letter sequencing signatures, electropherograms were produced and analyzed using STRand – *Nucleic Acid Analysis Software* (freely available at [37]). The position of peaks as well as their height (directly provided by STRand in arbitrary units for the latter) were recorded for marker identification and quantification, respectively, as described in detail in Results.

Authors' contributions

BM performed research except single-letter sequencing reactions which were carried out by HD. SB had the idea

to exploit single-letter sequencing for distinguishing and quantifying several virus variants in a mixed population. BM, HD and SB extracted data from single-letter sequencing electropherograms. HD helped to draft the manuscript. BM and SB wrote the manuscript. All authors have read and approved the final version of this manuscript.

Acknowledgements

Turnip seeds were kindly provided by Takii Seed Compagny (Japan). We thank the CaGeTE group and Orebokech for critical reading of the manuscript. This work was funded by the French National Institute for Agronomic Research: Grant INRA-SPE-JE77 to SB, and by the Region Languedoc-Roussillon: Grant INRA-LR to BM.

References

1. Mir KU, Southern EM: **Sequence variation in genes and genomic DNA: methods for large-scale analysis.** *Annu Rev Genomics Hum Genet* 2000, **1:**329-360.
2. Gibson NJ: **The use of real-time PCR methods in DNA sequence variation analysis.** *Clin Chim Acta* 2006, **363(1–2):**32-47.
3. Wittwer CT, Herrmann MG, Gundry CN, Elenitoba-Johnson KS: **Real-time multiplex PCR assays.** *Methods* 2001, **25(4):**430-442.
4. Rudi K, Zimonja M, Hannevik SE, Dromtorp SM: **Multiplex real-time single nucleotide polymorphism detection and quantification by quencher extension.** *Biotechniques* 2006, **40(3):**323-329.
5. Lindroos K, Sigurdsson S, Johansson K, Ronnblom L, Syvanen AC: **Multiplex SNP genotyping in pooled DNA samples by a four-colour microarray system.** *Nucleic Acids Res* 2002, **30(14):**e70.
6. Archak S, Lakshminarayanareddy V, Nagaraju J: **High-throughput multiplex microsatellite marker assay for detection and quantification of adulteration in Basmati rice (Oryza sativa).** *Electrophoresis* 2007, **28(14):**2396-2405.
7. Norton N, Williams NM, O'Donovan MC, Owen MJ: **DNA pooling as a tool for large-scale association studies in complex traits.** *Ann Med* 2004, **36(2):**146-152.
8. Sham P, Bader JS, Craig I, O'Donovan M, Owen M: **DNA Pooling: a tool for large-scale association studies.** *Nat Rev Genet* 2002, **3(11):**862-871.
9. Gilad Y, Borevitz J: **Using DNA microarrays to study natural variation.** *Curr Opin Genet Dev* 2006, **16(6):**553-558.
10. Ng JK, Liu WT: **Miniaturized platforms for the detection of single-nucleotide polymorphisms.** *Anal Bioanal Chem* 2006, **386(3):**427-434.
11. Ragoussis J, Elvidge GP, Kaur K, Colella S: **Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research.** *PLoS Genet* 2006, **2(7):**e100.
12. Edwards JR, Ruparel H, Ju J: **Mass-spectrometry DNA sequencing.** *Mutat Res* 2005, **573(1–2):**3-12.
13. Corona G, Toffoli G: **High throughput screening of genetic polymorphisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.** *Comb Chem High Throughput Screen* 2004, **7(8):**707-725.
14. Jurinke C, Oeth P, van den Boom D: **MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis.** *Mol Biotechnol* 2004, **26(2):**147-164.
15. Wada K, Kubota N, Ito Y, Yagasaki H, Kato K, Yoshikawa T, Ono Y, Ando H, Fujimoto Y, Kiuchi T, *et al.*: **Simultaneous quantification of Epstein-Barr virus, cytomegalovirus, and human herpesvirus 6 DNA in samples from transplant recipients by multiplex real-time PCR assay.** *J Clin Microbiol* 2007, **45(5):**1426-1432.
16. Wu JH, Liu WT: **Quantitative multiplexing analysis of PCR-amplified ribosomal RNA genes by hierarchical oligonucleotide primer extension reaction.** *Nucleic Acids Res* 2007, **35(11):**e82.
17. Wilke CO: **Quasispecies theory in the context of population genetics.** *BMC Evol Biol* 2005, **5:**44.
18. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R: **Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population.** *Nature* 2006, **439(7074):**344-348.
19. Watzinger F, Ebner K, Lion T: **Detection and monitoring of virus infections by real-time PCR.** *Mol Aspects Med* 2006, **27(2–3):**254-298.
20. Saccheri I, Hanski I: **Natural selection and population dynamics.** *Trends Ecol Evol* 2006, **21(6):**341-347.
21. Edwards CT, Holmes EC, Wilson DJ, Viscidi RP, Abrams EJ, Phillips RE, Drummond AJ: **Population genetic estimation of the loss of genetic diversity during horizontal transmission of HIV-1.** *BMC Evol Biol* 2006, **6:**28.
22. de la Iglesia F, Elena SF: **Fitness Declines in Tobacco Etch Virus upon Serial Bottleneck Transfers.** *J Virol* 2007, **81(10):**4941-4947. Epub 2007
23. Langemeier JL, Cook RF, Issel CJ, Montelaro RC: **Application of cycle dideoxy fingerprinting to screening heterogeneous populations of the equine infectious anemia virus.** *Biotechniques* 1994, **17(3):**484-486. 488, 490.
24. Soucek P, Skjelbred CF, Svendsen M, Kristensen T, Kure EH, Kristensen VN: **Single-track sequencing for genotyping of multiple SNPs in the N-acetyltransferase 1 (NAT1) gene.** *BMC biotechnology* 2004, **4:**28.
25. Kim YJ, Kim SO, Chung HJ, Jee MS, Kim BG, Kim KM, Yoon JH, Lee HS, Kim CY, Kim S, *et al.*: **Population genotyping of hepatitis C virus by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis of short DNA fragments.** *Clin Chem* 2005, **51(7):**1123-1131.
26. Norton N, Williams NM, Williams HJ, Spurlock G, Kirov G, Morris DW, Hoogendoorn B, Owen MJ, O'Donovan MC: **Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools.** *Hum Genet* 2002, **110(5):**471-478.
27. Li ZG, Chen LY, Huang J, Qiao P, Qiu JM, Wang SQ: **Quantification of the relative levels of wild-type and lamivudine-resistant mutant virus in serum of HBV-infected patients using microarray.** *J Viral Hepat* 2005, **12(2):**168-175.
28. Wilkening S, Hemminki K, Thirumaran RK, Bermejo JL, Bonn S, Forsti A, Kumar R: **Determination of allele frequency in pooled DNA: comparison of three PCR-based methods.** *Biotechniques* 2005, **39(6):**853-858.
29. Shifman S, Pisante-Shalom A, Yakir B, Darvasi A: **Quantitative technologies for allele frequency estimation of SNPs in DNA pools.** *Mol Cell Probes* 2002, **16(6):**429-434.
30. Trosvik P, Skanseng B, Jakobsen KS, Stenseth NC, Naes T, Rudi K: **Multivariate analysis of complex DNA sequence electropherograms for high-throughput quantitative analysis of mixed microbial populations.** *Appl Environ Microbiol* 2007, **73(15):**4975-4983.
31. Franck A, Guilley H, Jonard J, Richards K, Hirth L: **Nucleotide sequence of cauliflower mosaic virus DNA.** *Cell* 1980, **21:**285-294.
32. Froissart R, Uzest M, Ruiz-Ferrer V, Drucker M, Hebrard E, T H, Blanc S: **Splicing of the Cauliflower mosaic virus 35S RNA serves to downregulate a toxic gene product.** *J Gen Virol* 2004, **85(pt 9):**2719-26.
33. Melcher U, Steffens DL, Lyttle DJ, Lebeurier G, Lin H, Choe IS, Essenberg RC: **Infectious and non-infectious mutants of cauliflower mosaic virus DNA.** *J Gen Virol* 1986, **67(Pt 7):**1491-1498.
34. Gardner RC, Shepherd RJ: **A procedure for rapid isolation and analysis of cauliflower mosaic virus DNA.** *Virology* 1980, **106:**159-161.
35. De Leeuw M, Mouret J, JP I: **WO03068987 (A3): Discriminative analysis of clone signature.** *WO: Cogenics Genome Express SA, 11 Chemin des Prés, F-38944 Meylan, France* 2003:7.
36. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors.** *Proc Natl Acad Sci USA* 1977, **74(12):**5463-5467.
37. **STRand – Nucleic Acid Analysis Software** [http://www.vgl.ucdavis.edu/informatics/download_strand.php]
38. Zhou G, Kamahori M, Okano K, Chuan G, Harada K, Kambara H: **Quantitative detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER).** *Nucleic Acids Res* 2001, **29(19):**E93.
39. Bang-Ce Y, Peng Z, Bincheng Y, Songyang L: **Estimation of relative allele frequencies of single-nucleotide polymorphisms in different populations by microarray hybridization of pooled DNA.** *Anal Biochem* 2004, **333(1):**72-78.

40. Meaburn E, Butcher LM, Schalkwyk LC, Plomin R: **Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans.** *Nucleic Acids Res* 2006, **34(4):**e27.
41. Schnack HG, Bakker SC, van 't Slot R, Groot BM, Sinke RJ, Kahn RS, Pearson PL: **Accurate determination of microsatellite allele frequencies in pooled DNA samples.** *Eur J Hum Genet* 2004, **12(11):**925-934.
42. Giordano M, Mellai M, Hoogendoorn B, Momigliano-Richiardi P: **Determination of SNP allele frequencies in pooled DNAs by primer extension genotyping and denaturing high-performance liquid chromatography.** *J Biochem Biophys Methods* 2001, **47(1–2):**101-110.
43. Gruber JD, Colligan PB, Wolford JK: **Estimation of single nucleotide polymorphism allele frequency in DNA pools by using Pyrosequencing.** *Hum Genet* 2002, **110(5):**395-401.
44. Lavebratt C, Sengul S, Jansson M, Schalling M: **Pyrosequencing-based SNP allele frequency estimation in DNA pools.** *Hum Mutat* 2004, **23(1):**92-97.
45. Germer S, Holland MJ, Higuchi R: **High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR.** *Genome Res* 2000, **10(2):**258-266.
46. Mattarucchi E, Marsoni M, Binelli G, Passi A, Lo Curto F, Pasquali F, Porta G: **Different real time PCR approaches for the fine quantification of SNP's alleles in DNA pools: assays development, characterization and pre-validation.** *J Biochem Mol Biol* 2005, **38(5):**555-562.

ARTICLE 2

Effective size of *Cauliflower mosaic virus* populations during host plant colonization

Baptiste Monsion, Rémy Froissart, Anne Philip, Yannis Michalakis et Stéphane Blanc

Soumis pour publication en avril 2008

Effective size of *Cauliflower mosaic virus* populations during host plant colonization

Baptiste Monsion[1], Rémy Froissart[1, 2], Anne Philip[1], Yannis Michalakis[2], Stéphane Blanc[1*]

[1] UMR BGPI, INRA-CIRAD-SupAgroM, TA A-54/K, Campus International de Baillarguet, 34398 Montpellier Cedex 5. France.

[2] UMR 9926, IRD, 911 Avenue Agropolis, B.P. 64501, 34394 Montpellier Cedex 5. France.

*To whom correspondence should be addressed:
Stéphane Blanc
Tel: + 33 (0)4 99 62 48 04
Fax: + 33 (0)4 99 62 48 22
E-mail: blanc@supagro.inra.fr

Running title: Effective size of CaMV populations

Key words: Plant virus, population size, population bottleneck, CaMV, systemic infection

Abstract:

The effective size (*Ne*) of populations determines whether selection or genetic drift predominantly shapes their genetic structure and evolution. This trait of foremost importance in highly variable and rapidly evolving organisms is poorly documented experimentally in viruses, particularly viruses of plants. All available studies, however, have demonstrated the existence of huge within-host demographic fluctuations, drastically reducing *Ne* upon systemic invasion of different organs and tissues. Notably, extreme bottlenecks have been detected at systemic leaf colonization in all viral species investigated so far, sustaining the general idea that some unknown obstacles would represent barriers unfailingly imposed to the development of plant viruses. This idea has important implications as it places genetic drift as a constant major force in plant virus evolution, and *de facto* precludes viral adjustments and trade-offs on important traits requiring elevated *Ne*.

By co-inoculating several genetic variants of *Cauliflower mosaic virus* (CaMV) into a large number of replicate host plants, and by monitoring their relative frequency within the viral population over the course of the host systemic infection, only minute stochastic variations could be detected. This indicated a very limited role for genetic drift during plant systemic infection by this virus, and allowed the estimation of CaMV *Ne* during successive leaves colonization to several hundreds of viral genomes, a value at least a hundred fold higher than that reported for all other plant viruses investigated so far.

These results suggest that severe bottlenecks and extremely small *Ne* do not apply to all viruses infecting plants. They demonstrate that barriers generating bottlenecks in some viral species might well be bypassed by others, and thereby suggest that *Ne* is a trait that can be directly or indirectly modulated by plant viruses.

Author summary:

Whether directional selection or stochastic genetic drift is the major force driving the evolution of viruses largely depends on the size of viral populations, the former being predominant in large exponentially growing populations, and the latter taking over when population sizes are transiently or durably reduced. This question is intensively debated in both plant and animal viruses, as demographic fluctuations along their life cycle are poorly understood. In plants, an extremely small population size, down to a few founder genome units colonizing each leaf, could be formally estimated in two instances, and all other virus species investigated so far have consistently been shown to undergo extreme demographic bottlenecks during the systemic invasion of their host. This situation is conveying the general idea that all viruses are confronted to comparable “barriers” in plants, imposing repeated transient decrease in their population size, and setting the genetic drift as a major constant driver of their evolution. Through the example of *Cauliflower mosaic virus*, we here mitigate this general idea by showing that at least one virus species can overcome the putative limiting barriers and massively invade leaves with hundreds to thousands of founding genome units.

Introduction

The main forces driving changes of the frequencies of alleles within populations are selection (deterministic) and genetic drift (stochastic). While both forces obviously act concomitantly in most cases, the balance of their respective action is governed by a number of factors, one of the most important being the effective population size (Ne) [1-3]. This trait, commonly defined as the size of an ideal population which would drift at the same rate as the observed population [4], is widely documented in the literature both theoretically [5-8], and experimentally [9-11] in a broad variety of organisms. Ne significantly informs whether the evolution of a given population might be better described by a deterministic or a stochastic evolution model [12-14]. In brief, when *Ne* is large, competition between genetic variants is fully acting, with no or little interference of random processes and selection shapes the genetic composition of corresponding populations. Conversely, in populations with a small *Ne*, various processes resulting in stochastic "sampling" of genetic variants that will engender the next generation are prevalent, compensating for, or even annihilating, the effect of selection.

Experimental evolution using viruses as biological models has developed enormously during the past two decades. A series of converging reports on viruses of bacteria, animals and plants confirmed that protocols maintaining constant large population sizes lead to selection of the fittest variants, thus augmenting the mean fitness of viral populations [15-17]. On the other hand, the repeated experimental imposition of severe bottlenecks (temporarily reducing Ne) is associated with relaxed selection and strong drift augmenting the Muller's Ratchet effect, thus dramatically reducing the mean fitness in the population [18-21]. Viruses are intuitively perceived as having extremely large population sizes during host invasion. While this is true for the census population size, i.e. the number of viral genomes accumulating in single hosts for example, it does not indicate what fraction thereof is actually actively replicating and propagating, or is efficiently progressing into new organs and tissues, yielding subsequent generations upon host colonization. Estimating viral Ne during systemic infection of a host has proved experimentally difficult, as illustrated by the impressive number of studies published, and still debated, on HIV [22]. Consequently, virus species for which Ne has been formally evaluated during invasion of various organs and tissues of the host remain extremely few in both plants and animals. The question of whether deterministic or stochastic models better explain the evolution of viral populations is thus pending further investigation of Ne.

In plants, several specific barriers could impose severe bottlenecks on populations of invading viruses. The level at which the existence of such bottlenecks has been most

documented is the colonization of new leaves, not only because of the obvious practical ease of comparing the genetic content of different leaves from the same host, but also because this level results from long-distance movement of the virus population, which is loaded into the vascular system from source leaves and unloaded into sink leaves (for review see [23]). As discussed later, other phenomena could also generate population bottlenecks but virus movement within the host plant through plasmodesmata is most often considered as a major putative obstacle to exponential expansion.

While not directly estimating *Ne*, two experimental evolution studies demonstrated the existence of viral population bottlenecks during leaf colonization. With a mixture of 12 engineered genetic variants of *Cucumber mosaic virus* (CMV) co-inoculated into tobacco plants, Li and Roossinck [24] showed that diversity decreased randomly but continuously in the viral population, as increasing numbers of variants were lost when systemic infection progressed into upper, newly-formed leaves. Even more illustrative was the spatial distribution of the genetic diversity of a Plum pox virus (PPV) population, maintained in a perennial host tree for over 13 years [25]. While a large number of related PPV variants could be distinguished and detected in various organs and tissues, distinct subpopulations were shown to be isolated in different branches. Extreme population bottlenecks were further evidenced when the virus progressed into newly formed leaves, which were all colonized by one single viral genotype.

The effective size of virus populations during systemic infection of host plants has been evaluated more formally in two instances. Populations of Wheat streak mosaic virus (WSMV) invading a wheat tiller [26], and of Tobacco mosaic virus (TMV) invading a tobacco leaf [27], were shown to be founded by as few as 4 and 2-20 virus particles, respectively, hence again resulting from severe bottlenecks.

All the studies cited above converged to the conclusion that the population size of plant viruses fluctuates dramatically and can be temporarily remarkably small, i.e. of the order of one to a few genome units founding the population that subsequently develops to billions of genomes in each systemically infected leaf. Such a demographic regime suggests that genetic drift is a major force in plant virus evolution, as proposed and discussed by several authors [24-28]. The fact that extreme population bottlenecks are consistently described for unrelated virus species infecting mono- or di-cotyledonous hosts appealingly suggests that all plant viruses might be subject to the same phenomenon [24,27-29], perhaps related to the unavoidable physical barriers that hamper the systemic movement of all viruses in plants. However, this tempting generalisation definitely requires closer inspection as, if proven true, it would illustrate the impossibility of viral adjustments and trade-offs on important traits regulating Ne. In contrast, a single counter example would demonstrate that viruses can

indeed overcome such putative barriers, of whatever form, and possibly regulate *Ne,* or other important traits impacting *Ne*. Further research on more diverse virus species, with different replication strategies and different life cycles, is still needed in order to assume extremely small *Ne* as a general rule during single host plant infection.

In this report, we assessed the existence of bottlenecks during systemic host colonization in populations of *Cauliflower mosaic virus* (CaMV), a DNA virus whose biological properties differ largely from those of the RNA viruses investigated in the studies cited above. Monitoring the frequency of several engineered allelic variants within leaves of single host plants revealed remarkably small stochastic fluctuations in the genetic structure of numerous CaMV populations, even over considerable periods of time. Beyond demonstrating that CaMV populations are not subject to intense genetic drift, and hence do not undergo severe demographic bottlenecks, the slight stochastic fluctuations detected were exploited to infer the effective size of CaMV populations upon systemic leaf colonization. Our estimates consistently indicate that several hundreds of genome units founded the viral populations in all leaves analyzed (originating from different plants). This value, which is ~100-fold higher than values previously reported for other plant viruses, demonstrates that extremely small *Ne* is not a general rule for plant viruses, and opens up the possibility of trade-offs on this trait that could be directly or indirectly regulated by the virus.

Materials and Methods

Engineered CaMV variants

The six plasmids (pCa-VIT1 to pCa-VIT6) used in this study to generate the six CaMV allelic variants have been described in detail and characterized previously [30]. All are infectious full-length clones of the CaMV Cabb-S isolate [31], where a specific genetic marker (a dsDNA sequence of 40 bp) has been inserted between CaMV ORFs II and III. When inoculated individually into turnip plants, each CaMV clone (CaMV-VIT1 to CaMV-VIT6) induces symptoms similar to CaMV wild type, and all six genetic markers have been shown to be stably maintained in the viral genome even after three successive passages in plants [30].

Plant growth conditions and inoculation with mixed CaMV-VIT populations

All turnip plants (*Brassica rapa* var. “Just Right”) were maintained in an insect-proof greenhouse under controlled conditions: 25/19°C day/night with a photoperiod of 16/8 hours day/night.

Plants infected for 21 days with one of the CaMV-VIT1 to -VIT6 clones (1 plant per clone) were used to prepare virus particle-enriched fractions as previously described [30]. Equal volumes of each of the six virus particle-enriched fractions were pooled to produce a mixed inoculum (designated Mix6VIT) containing all six allelic variants of CaMV in undetermined proportions. Fifty healthy young plantlets, at the three-leaf stage, were mechanically infected by rubbing 20µL of Mix6VIT solution onto the entire surface of the two larger leaves, previously powdered with abrasive carborundum. First symptoms indicative of CaMV infection appeared on non-inoculated leaves within 7 to 9 days, and all plants proved systemically infected at 11 dpi.

Harvesting of leaf samples

At 13 dpi, i.e. very soon after systemic infection had developed, the inoculated leaves were discarded and all large expanded leaves were collected. In all cases, 2 to 3 young expanding or emerging leaves in the centre of the rosette were left unsevered in order to allow continuous growth of the plants for further sampling as described below. The viral population extracted from each plant at this early stage was considered as the initial population.

Thirty-two days later (at 45 dpi) the same plants had grown continuously and produced 10 to 15 newly expanded leaves. At this stage, one single leaf from each plant, randomly chosen between the 5^{th} and the 12^{th} newly formed leaves, was collected. The viral population extracted from each single leaf at this late stage was considered as the final population. All leaf samples were stored at -20°C until further viral DNA purification.

The rationale of this sampling protocol is further discussed below.

Purification of viral DNA and genetic composition analysis

Purification of viral DNA from the harvested leaf samples, as well as the analysis of the genetic structure of the corresponding viral genome populations (i.e. determination of the presence/absence of various genetic markers and quantification of their relative frequency in the genome population) were performed exactly as described previously, using the QSS (Quantitative Single-letter Sequencing) method [30].

Statistical analysis

Changes in the relative frequencies of the six alleles (VIT1 to VIT6) between the initial and final populations, sampled as described above, were precisely monitored in 50 replicate test plants.

In order to calculate the number of founder genomes initiating each final population in a single leaf, all parameters accounting for the transition from P (initial frequency of a given marker) to P' (final frequency of the same marker) must be evaluated precisely. For any

given plant the difference between P and P' (∆P) may be potentially attributed to genetic drift, and to selection:

∆P = ∆P Drift + ∆P Selection (1)

There is no a priori reason to assume that any selection affecting the frequency of the markers is heterogeneous across plants (e.g. favouring a marker in a plant and selecting against it in another plant). Under this assumption, selection will not influence the variance among P' values in the final CaMV populations. Hence, the variance of the estimates of P and P' between plants can be written as:

Var (P') = Var (P) + Var (Drift) (2)

The variance due to drift is equal to

Var (Drift) = *p(1-p)/Ne* (3)

Where *p* is the frequency of a given marker in initial populations. From equations (2) and (3) we obtain :

$$N_e = \frac{p(1-p)}{Var(P') - Var(P)} \qquad (4)$$

To provide a confidence interval for this estimate we used a resampling technique. We bootstrapped over plants: for each bootstrap we randomly drew a sample of 50 plants with replacement, calculated Var(P'), Var(P) and p, and from that estimated Ne. We repeated this procedure 1000 times and constructed a distribution of Ne. The 95% confidence interval corresponds to the 2.5 and 97.5 percentiles of this distribution. In several bootstraps the final variance was smaller than the initial variance, yielding negative values of Ne. In such cases our estimation method does not work. Such cases would correspond to conditions leading to convergence of marker frequencies across plants, i.e. conditions where drift is negligible, and thus Ne can be regarded as infinite. We conservatively discarded such cases from the calculation of the upper limit of the confidence interval.

Results

Genetic structure of initial CaMV populations in whole plants

Following systemic spread of infection in plants inoculated with a mixture of 6 CaMV genetic variants (Mix6VIT), populations of viral genomes in each plant were purified from all fully expanded and systemically infected leaves, excluding inoculated leaves. DNA samples from each of 50 replicate plants were considered as the initial populations and were submitted to QSS analysis to detect markers in CaMV genomes and quantify their relative frequency as described in Materials and Methods. The mean frequency of each marker and the variance among the 50 CaMV populations collected at this stage is given in Table 1 (full dataset with details for each plant is provided in the online supplement; Table S1).

The six CaMV-VIT genotypes were found in widely unequal ratios in all 50 cases, with mean relative frequencies being close to 0 for CaMV-VIT5 and VIT6, and up to around 40% for CaMV-VIT1 and VIT3. However, a striking observation was that each marker was found with a very similar frequency in all 50 replicate CaMV populations (Table S1), as demonstrated by the very low standard deviation among repeats (Table 1). This result clearly indicates that little stochastic variation is generated by the inoculation process, or between the inoculation process and the sampling of initial populations. The frequency differences between markers being very consistent in all 50 repeats, they are certainly due to a non-stochastic phenomenon (discussed below).

Genetic structure of final CaMV populations in single leaves

At late infection stages, 32 days after collecting the initial populations, a single systemically infected leaf per individual plant, randomly chosen between the 5^{th} and the 12^{th} new expanded leaves, was harvested. The CaMV populations extracted from these sampled leaves are considered as “final populations”, and their genetic structure (Table 1 and S1) was evaluated in exactly the same manner as that of the initial populations.

Remarkably, in all 50 repeats, the final population resembled the corresponding ancestor population despite the different sampling process (individual leaves for final populations versus pools of leaves for initial populations), demonstrating that very little change had occurred over this considerable time period, whatever the position or age of the analyzed leaf [see $E(\Delta p)$ values in Table 1, and compare P and P’ values in Table S1 for each of the 50 plants]. Consequently, as between initial populations, there were only small variations between the final populations, although the standard deviation was slightly higher in the latter case (Table 1).

Table 1: Genetic structure of initial and final CaMV populations

	Initial populations[a] P[b] (SD[c] x10^4)	***Final populations***[a] P'[b] (SD[c] x 10^4)	E(Δp)[d]
VIT 1	0.416 (3.52)	0.458 (4.27)	0.042
VIT 2	0.042 (1.90)	0.033 (1.55)	-0.009
VIT 3	0.425 (4.19)	0.409 (5.08)	-0.016
VIT 4	0.102 (2.88)	0.092 (3.19)	-0.010
VIT 5	0.017 (1.89)	0.019 (1.61)	0.002
VIT 6	0.017 (0.98)	0.036 (1.70)	0.019

The genetic structure of the 50 sampled CaMV populations is determined by the different markers present in each, as well as by their respective relative frequency. For reasons of clarity, only mean frequency values and standard deviations among the 50 repeats are shown here. The full dataset with details of the 50 CaMV populations is available as an online supplement (Table S1).

[a] Sampling protocol for initial and final populations is described in the Material and Methods

[b] Mean frequency of each marker over the 50 replicates and expressed as a percentage of the total CaMV genomes

[c] Standard deviation of marker frequencies across the 50 replicates.

[d] Mean difference between initial and final frequencies for each marker.

Table 2: Estimates of the effective size of CaMV populations during systemic infection of a single leaf

Marker[a]	**Var(P') x10^4**	**Var(P) x10^4**	**Ne**[b]	**95% confidence interval**
VIT1	18.13	12.39	423.00	180 – 5034
VIT3	25.79	17.60	298.45	122 – 3219
VIT4	10.19	8.30	484.61	166 – 5193

[a] only markers VIT1, VIT3 and VIT4 were used for estimating *Ne*, for reasons explained in the text

[b] *Ne* is calculated according to equation 4 (see Materials and Methods), where p is estimated from values given in Table

Taken together, these results argue for the absence of large stochastic variations in the genetic structure of a CaMV population upon progression of the infection into newly formed leaves. This suggests that the effective size (*Ne*) of CaMV populations during host systemic colonization is likely to be surprisingly large (compared to data previously published on other plant viruses) as specifically evaluated in the next section.

Estimation of CaMV effective population size during systemic colonization of leaves

Because the average relative frequency of markers CaMV-VIT2, VIT5 and VIT6 was close to, or even below, the limit of detection of the QSS method for accurate analysis [30], their quantification was considered poorly reliable, and hence these markers were not used for further *Ne* analysis. We thus used only the three markers VIT1, VIT3 and VIT4 to estimate *Ne,* as described in the Materials and Methods.

The estimates of *Ne* from these three markers yielded remarkably high values, corresponding to several hundreds of viral genomes founding the population contained in every single leaf (Table 2). The lower limit of the confidence interval was >100 in all cases.

Discussion

This report evaluates the effective size of CaMV populations during systemic invasion of plant leaves. Several previous studies have used distinct experimental protocols to tackle similar questions with various viral species; the specifics and rationale of the protocol used here are discussed below. To eliminate variations in the genetic structure of viral populations that could be related to the inoculation process, the virus population present in the whole plant (excluding the inoculated leaves) soon after systemic spread of infection was considered as the starting point of the experiment, putative subsequent changes thus occurring only under the influence of within-plant processes. At this initial stage, the virus population uploaded into, and circulating within, the vascular system is most likely best represented by the overall content of the systemically infected leaves, as they have either received virus from the vasculature, exported virus into it, or both. Expanded infected leaves were therefore harvested for initial analysis, carefully preserving 2 to 3 young newly expanding leaves on the still-growing plant. The infected plants were left to grow for a period of 32 days, during which the virus population successively colonized 10 to 15 emerging and expanding new leaves. Regardless of where the virus population originates from during this process (vascular system and roots at the beginning, then increasing numbers of leaves later on), analysis of single leaf (between positions 5 and 12 above the initial harvest) contents collected 32 days later should reveal the existence of any putative bottlenecks at any stage of the systemic infection. Indeed, final CaMV populations result largely from the successive

leaf-to-leaf passages that occur sequentially when young sink leaves are infected, become sources and subsequently export virus into new sinks.

All genetic variants (CaMV VIT1 to VIT6) were co-inoculated at similar locations and at the same time point onto the two first true leaves of turnip plantlets. The reasons why the relative proportions of the six variants are highly unequal later in systemically infected plants is not known. However, all 50 test plants contained similar ratios of the six variants, indicating that there was very little stochastic variation during or after the inoculation process. Whether the initial Mix6VIT inoculum contained drastically unequal amounts of the variants or whether some were much less infectious than others was not investigated further. The alternative hypothesis that variants VIT5 and 6 were rapidly eliminated due to a negative effect of the marker sequence on their relative fitness could not be directly assessed with the presented dataset because, in all cases, these variants were present at a frequency below the accuracy threshold of the QSS quantification technique. Neutrality tests applied to VIT1, 2, 3 and 4, demonstrated that only VIT1 and 3 might be slightly selected for and against, respectively (data not shown). Nevertheless, both these variants were used for estimating *Ne* because, as indicated in the results, the calculation method is not affected by equal selection in the 50 repeats, and because they were the most frequent in all plants, providing highly reliable monitoring of changes in their relative frequency over time.

The values of P' were estimated from individual leaf samples, each collected on a different test plant at a random position (between leaf position 5 and 12 above the initial harvest point). The remarkable observation that only relatively small variations in P' are recorded in leaves from different plants (see standard deviation in Table 1), strongly indicates that very little variation occurs between leaves of the same plant. We therefore conclude that a similarly large founding population of several hundred viral genomes colonizes each newly formed leaf in plants systemically infected by CaMV.

The number of founder CaMV genomes in systemically infected leaves appears to be at least a hundred fold higher than that determined in previously published studies for various RNA viruses [25-27]. Although there are clear differences in the inoculation and sampling protocols described in different studies, the presence of severe bottlenecks in WSMV [26], TMV [27], CMV [24], PPV [25], and their obvious absence in CaMV cannot be explained by artefacts of experimental design. Indeed, at least in the present study and in three of the cited examples [24-26], viruses similarly moved not only from inoculated to systemic tissues, but also from systemically infected tissues to newly formed leaves or tiller. As mentioned in

the Introduction, the previous repeated demonstration that different viruses undergo extreme bottlenecks during systemic infection of their host plant suggested that unavoidable barriers (e.g. connections between cells, etc.) may exist similarly for all plant viruses. The results presented here clearly demonstrate that such putative limiting barriers can be surmounted by some viruses and that the size of the viral population circulating *in planta* might thus be controlled and needs to be evaluated for each virus individually.

The situation described here for CaMV re-opens the question as to what phenomena actually generate bottlenecks during virus infection. A reasonable hypothesis explaining the presence or absence of severe demographic bottlenecks would be the regulation of multiple infection of cells by several genomes of the viral population; in other words, regulation of the multiplicity of infection (MOI) of host cells. Indeed, it was proposed recently that bottlenecks seen in a PPV population infecting a Prunus tree could stem from the fact that viral genomes cannot secondarily invade tissues that have already been infected by a closely related genome [25], thus preventing extensive mixing of genetic variants within host. Consistently, two isogenic PPV genomes, but expressing fluorescent proteins of different colors, revealed their mutual exclusion within infected tissues, with apparently rare co-infection of single cells [32], a phenomenon also reported for other virus species known to undergo severe bottlenecks during host infection, such as WSMW [26], TMV [33], and a few other RNA viruses [34]. This phenomenon could logically induce a very low MOI, in turn engendering small Ne (and bottlenecks), due to competition for host “territories” between variants of the same population. Accordingly, that the larger Ne found for CaMV would suggest a higher MOI is totally consistent with the remarkably high within-host recombination rate described for this virus [35], recombination being detectable solely in multiply infected cells. We believe that explaining, at least in part, variations in Ne in different virus species by their different capacity to multiply infect host cells is a very appealing hypothesis that will stimulate further novel research, the “natural” MOI during infection of a multi-cellular host being virtually unknown in viruses of both animals and plants.

Acknowledgments

Turnip seeds were kindly provided by the Takii Seed Compagny (Japan). This work was funded by INRA grant JE77, and a PhD grant from Region-Languedoc-Roussillon/INRA-SPE. YM acknowledges support from CNRS and IRD. We are thankful to the CaGeTE group for critical reading and discussion of the manuscript.

References

1. Nei M, Maruyama T, Chakaraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29: 1-10.
2. Gherman A, Chen PE, Teslovich TM, Stankiewicz P, Withers M, et al. (2007) Population Bottlenecks as a Potential Major Shaping Force of Human Genome Architecture. PLoS Genet 3: e119.
3. Kimura M, Maruyama T, Crow JF (1963) The Mutation Load in Small Populations. Genetics 48: 1303-1312.
4. Wright S (1931) Evolution in Mendelian Populations. Genetics 16: 97-159.
5. Criscione CD, Blouin MS (2005) Effective sizes of macroparasite populations: a conceptual model. Trends Parasitol 21: 212-217.
6. Nei M, Tajima F (1981) Genetic Drift and Estimation of Effective Population Size. Genetics 98: 625-640.
7. Tajima F, Nei M (1984) Note on genetic drift and estimation of effective population size. Genetics 106: 569-574.
8. Crow JF, Kimura M (1972) The effective number of a population with overlapping generations: a correction and further discussion. Am J Hum Genet 24: 1-10.
9. Araki H, Waples RS, Blouin MS (2007) A potential bias in the temporal method for estimating Ne in admixed populations under natural selection. Mol Ecol 16: 2261-2271.
10. Wang J (2005) Estimation of effective population sizes from data on genetic markers. Philos Trans R Soc Lond B Biol Sci 360: 1395-1409.
11. Johnson JA, Bellinger MR, Toepfer JE, Dunn P (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. Mol Ecol 13: 2617-2630.
12. Rouzine IM, Rodrigo A, Coffin JM (2001) Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. Microbiol Mol Biol Rev 65: 151-185.
13. Jain K, Krug J (2007) Deterministic and stochastic regimes of asexual evolution on rugged fitness landscapes. Genetics 175: 1275-1288.
14. Wahl LM, Krakauer DC (2000) Models of experimental evolution: the role of genetic chance and selective necessity. Genetics 156: 1437-1448.
15. Weinreich DM, Chao L (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution Int J Org Evolution 59: 1175-1182.
16. Willi Y, Fischer M (2005) Genetic rescue in interconnected populations of small and large size of the self-incompatible Ranunculus reptans. Heredity 95: 437-443.
17. Clarke DK, Duarte EA, Moya A, Elena SF, Domingo E, et al. (1993) Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. J Virol 67: 222-228.
18. Chao L (1990) Fitness of RNA virus decreased by Muller's ratchet. Nature 348: 454-455.
19. de la Iglesia F, Elena SF (2007) Fitness Declines in Tobacco Etch Virus upon Serial Bottleneck Transfers. J Virol.

20. Novella IS, Quer J, Domingo E, Holland JJ (1999) Exponential fitness gains of RNA virus populations are limited by bottleneck effects. J Virol 73: 1668-1671.
21. Yuste E, Sanchez-Palomino S, Casado C, Domingo E, Lopez-Galindez C (1999) Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J Virol 73: 2745-2751.
22. Kouyos RD, Althaus CL, Bonhoeffer S (2006) Stochastic or deterministic: what is the effective population size of HIV-1? Trends Microbiol 14: 507-511.
23. Waigmann E, Heinlein M (2007) Viral Transport in Plants; Robinson DG, editor: Springer-Verlag Berlin Heidelberg. 188 p.
24. Li H, Roossinck MJ (2004) Genetic bottlenecks reduce population variation in an experimental RNA virus population. J Virol 78: 10582-10587.
25. Jridi C, Martin JF, Marie-Jeanne V, Labonne G, Blanc S (2006) Distinct viral populations differentiate and evolve independently in a single perennial host plant. J Virol 80: 2349-2357.
26. French R, Stenger DC (2003) Evolution of Wheat streak mosaic virus: dynamics of population growth within plants may explain limited variation. Annu Rev Phytopathol 41: 199-214.
27. Sacristan S, Malpica JM, Fraile A, Garcia-Arenal F (2003) Estimation of population bottlenecks during movement of tobacco mosaic virus in tobacco plants. J Virol 77: 9906-9911.
28. Garcia-Arenal F, Fraile A, Malpica JM (2003) Variation and evolution of plant virus populations. Int Microbiol 6: 225-232.
29. French R, Stenger DC (2005) Population structure within lineages of Wheat streak mosaic virus derived from a common founding event exhibits stochastic variation inconsistent with the deterministic quasi-species model. Virology 343: 179-189.
30. Monsion B, Duborjal H, Blanc S (2008) Quantitative Single-letter Sequencing: a simple method for simultaneously monitoring numerous allelic variants in single DNA samples (in press). BMC Genomics.
31. Franck A, Guilley H, Jonard J, Richards K, Hirth L (1980) Nucleotide sequence of cauliflower mosaic virus DNA. Cell 21: 285-294.
32. Dietrich C, Maiss E (2003) Fluorescent labelling reveals spatial separation of potyvirus populations in mixed infected Nicotiana benthamiana plants. J Gen Virol 84: 2871-2876.
33. Giritch A, Marillonnet S, Engler C, van Eldik G, Botterman J, et al. (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. Proc Natl Acad Sci U S A 103: 14701-14706.
34. Takahashi T, Sugawara T, Yamatsuta T, Isogai M, Natsuaki T, et al. (2007) Analysis of the Spatial Distribution of Identical and Two Distinct Virus Populations Differently Labeled with Cyan and Yellow Fluorescent Proteins in Coinfected Plants. Phypathology 97: 1200-1206.
35. Froissart R, Roze D, Uzest M, Galibert L, Blanc S, et al. (2005) Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. PLoS Biol 3: e89.

Supplemental data on line:

Table S1: Whole dataset from the 50 replicate infected plants

	Plant N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
P	VIT 1	46.1	38.8	45.7	44.7	43.2	41.1	48.7	43.2	41.6	44.9	42.9	45.9	44.6	37.5	40.5	41.8	41.7	35.1	39.2	45.4	41.5	43.5	38.9	40.5	39.0
	VIT 2	4.7	3.7	0.7	3.3	5.0	3.9	4.3	8.6	2.3	4.6	3.0	5.3	4.5	3.9	3.8	4.1	3.8	5.4	2.1	3.2	4.0	3.2	12.1	2.2	4.9
	VIT 3	39.1	44.6	46.8	37.1	37.7	39.9	40.4	34.8	47.4	39.6	42.6	42.3	42.5	48.7	44.7	40.9	44.2	50.4	49.7	42.2	44.1	44.7	33.7	44.1	42.3
	VIT 4	10.2	10.4	6.7	13.2	11.6	11.9	5.3	12.2	8.5	9.4	10.2	6.2	8.1	8.6	9.4	11.5	10.3	9.1	9.1	9.2	8.9	8.6	8.1	13.2	12.5
	VIT 5	-	-	-	-	-	2.7	-	-	-	-	-	0.3	0.3	1.3	0.6	0.2	-	-	-	-	-	-	7.1	-	-
	VIT 6	-	2.5	-	1.7	2.4	0.5	1.3	1.2	0.4	1.5	1.3	-	-	-	1.1	1.6	-	-	-	-	1.4	-	-	-	1.3
P'	VIT 1	56.2	38.1	46.3	46.1	49.1	49.5	50.3	53.5	42.0	55.4	45.7	53.7	46.7	43.7	44.3	50.0	44.6	49.9	48.6	46.3	37.9	45.3	43.1	44.2	45.6
	VIT 2	-	2.8	3.0	-	5.5	1.4	4.5	-	-	-	2.5	2.2	1.4	2.8	4.7	2.3	3.7	2.4	2.1	5.8	5.4	3.6	6.6	-	-
	VIT 3	39.1	51.9	38.0	47.1	35.2	37.9	37.9	32.0	48.1	38.5	40.8	38.8	40.6	42.8	40.5	41.5	43.3	42.0	45.0	31.3	39.2	40.0	40.5	44.7	41.8
	VIT 4	4.7	7.1	12.1	6.8	10.2	11.1	7.3	9.6	5.7	6.1	11.0	5.2	6.8	7.3	10.5	6.2	8.3	5.6	4.2	9.7	12.0	9.2	7.8	10.6	10.9
	VIT 5	-	-	-	-	-	-	-	4.9	1.6	-	-	-	0.2	-	-	-	-	-	-	1.8	4.1	1.9	2.1	0.6	1.7
	VIT 6	-	-	0.7	-	-	-	-	-	2.5	-	-	-	4.2	3.4	-	-	-	-	-	5.1	1.4	-	-	-	-

	Plant N°	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
P	VIT 1	42.8	37.6	31.9	40.1	37.9	39.5	40.5	44.0	42.7	45.4	43.8	43.9	43.7	41.1	37.3	40.0	36.4	47.6	41.1	43.5	42.1	42.0	48.2	37.9	34.2
	VIT 2	1.3	3.0	6.2	4.5	4.8	0.4	3.1	3.2	6.8	5.5	3.6	2.6	4.9	3.9	6.4	6.3	4.2	4.1	3.7	6.1	2.8	2.7	4.9	3.7	6.0
	VIT 3	50.0	37.4	39.1	45.0	43.4	54.3	44.6	40.8	40.1	40.1	43.4	45.0	41.0	42.9	45.8	43.4	43.6	35.7	40.1	39.8	48.2	43.2	35.2	41.5	39.0
	VIT 4	5.8	20.4	20.8	9.7	11.3	5.8	9.5	11.3	10.4	8.9	9.1	8.5	9.3	10.5	10.1	10.0	11.3	9.9	12.3	10.5	6.9	10.2	9.5	14.4	12.6
	VIT 5	-	-	-	0.5	2.6	-	-	0.7	-	-	-	-	1.1	-	-	0.3	-	-	-	0.2	-	-	2.1	2.5	4.9
	VIT 6	-	1.6	2.0	-	-	-	2.4	-	-	-	0.1	-	-	1.6	0.3	-	4.5	2.6	2.8	-	-	1.9	-	-	3.3
P'	VIT 1	40.7	45.7	37.2	44.1	39.1	45.3	41. 8	50.4	48.3	46.7	46.5	48.2	48.3	44.9	42.4	42.8	41.3	48.3	42.8	46.8	51.0	42.7	42.1	41.1	43.2
	VIT 2	-	4.0	5.3	0.4	3.8	-	2.2	3.6	4.3	1.6	2.4	1.6	4.9	0.6	3.3	5.3	-	1.7	1.6	5.5	2.6	4.4	2.9	-	-
	VIT 3	52.8	28.8	41.5	42.4	41.9	50.4	49.6	37.2	39.7	35.0	43.6	42. 8	31.4	44.4	45.2	39.3	41.3	41.9	45.3	33.1	38.4	37.4	41.6	45.9	36.4
	VIT 4	6.5	20.0	15.9	9.1	10.9	4.3	6.3	8.8	7.3	10.2	7.4	7.0	13.2	9.3	9.2	11.7	14.7	6.4	9.9	14.5	7.6	13.1	7.6	9.7	13.8
	VIT 5	-	1.6	0.1	-	4.3	-	-	-	0.4	1.0	-	0.4	2.2	0.8	-	0.9	2.7	1.8	0.4	-	0.4	2.4	-	3.4	6.5
	VIT 6	-	-	-	3.9	-	-	-	-	-	5.5	-	-	-	-	-	-	-	-	-	-	-	-	5.9	-	-

P represents the marker relative frequency in initial populations

P' represents the marker frequency in final populations

All values are indicated as % of the viral genome population

- marker not detected

CHAPITRE II

Multiplicité d'infection (M.O.I.)

1 Introduction

Du fait que plusieurs centaines de génomes du CaMV fondent la population d'une feuille néoformée, il nous a semblé pertinent d'affiner l'analyse de la distribution des variants viraux jusqu'au niveau intracellulaire. En effet, d'après le comportement alimentaire des pucerons et les caractéristiques de transmission des virus non-persistants (brève période d'acquisition), il semble particulièrement utile de connaître la composition génétique virale d'une ou des quelques cellules qui sont susceptibles d'être ponctionnées par le vecteur.

Par ailleurs, le nombre de variants viraux présents dans une seule cellule hôte et même plus encore la MOI (Multiplicity Of Infection) sont très peu étudiés expérimentalement chez les virus de plantes et d'animaux. Il s'agit pourtant d'une question générique fascinante et porteuse sur le plan de la biologie et de l'évolution des virus, puisque la MOI est à l'origine de processus importants comme la compétition intracellulaire, la complémentation fonctionnelle et la recombinaison.

1.1 La multiplicité d'infection : notions

1.1.1 La MOI et notre définition

Le terme MOI est couramment employé en culture cellulaire afin de définir le ratio du nombre d'unités infectieuses virales sur le nombre de cellules cibles ; dans ce cas la MOI correspond à une valeur numérique fixée par l'expérimentateur (par exemple 10). C'est à dire que la quantité d'unités infectieuses dans un inoculum est généralement titrée en « Plaque Forming Units » (PFU/mL), d'après le nombre de plages de lyse générées pour un volume d'inoculum donné. L'inoculum est alors dilué de manière à ce qu'au moment de l'inoculation il y ait – par exemple – 10 fois plus d'unités infectieuses (PFU) que de cellules dans la culture (MOI=10).

Cette valeur de MOI « théorique », qui détermine donc un nombre d'unités infectieuses disponibles dans le milieu environnant par cellule aussi présente dans le même milieu, ne correspond pas nécessairement au nombre réel d'unités infectieuses pénétrant et se multipliant dans chaque

cellule. En effet, il peut exister toutes sortes de mécanismes de régulation pour l'entrée de ces unités infectieuses dans la cellule, les limitant à un nombre restreint, ou encore de mécanismes d'exclusion ne permettant pas la co-existence de 2 unités infectieuses dans la même cellule. Sous l'influence de l'ensemble de ces possibles phénomènes, le nombre des unités infectieuses qui infectent effectivement chaque cellule peut être considéré comme la MOI vraie qui n'est souvent pas considérée, ni comparée à la MOI théorique dans les infections de cultures cellulaires.
Le paramètre qui nous intéresse est précisément la valeur réelle de la MOI et ceci en condition d'infection « naturelle » dans un hôte pluricellulaire.

1.1.2 L'infection multiple : co-infection

Une cellule peut se retrouver infectée de manière multiple (MOI>1) de deux façons : soit (i) par des infections synchrones, soit (ii) par une primo-infection suivie d'un ou plusieurs événement(s) de surinfection (infections asynchrones). La MOI que nous analysons s'explique potentiellement par ces deux possibilités, nous regardons seulement ce qu'il y a dans les cellules à un moment donné. Il est impossible de distinguer, par la méthode employée, si ce que l'on détecte provient de virions rentrés simultanément ou les uns après les autres.
Par ailleurs, la nature même des unités infectieuses est sujette à varier, c'est à dire que les éléments en présence dans une même cellule peuvent provenir d'une seule population ou quasi-espèce virale, ou de plusieurs populations co-infectant la plante, chacune pouvant correspondre à des isolats, souches ou même espèces différentes. Tout objet d'étude (plante, tissu ou cellule) peut donc être qualifié comme co-infecté dès que plusieurs unités infectieuses y sont entrées et détectées.
Néanmoins, dans ce chapitre, lorsque nous parlerons de MOI, c'est à l'infection multiple des cellules de l'hôte par plusieurs génomes d'une même population ou quasi-espèce, et durant le développement normal ou « naturel » de l'infection systémique, que nous ferons référence.

1.2 La MOI chez les virus : estimations disponibles

1.2.1 Virus de bactéries : phages

Le phage ϕ6 est un virus de bactérie dont le génome est un ARN double brin. Il utilise le pilus de son hôte bactérien (par exemple *Pseudomonas syringae*) afin d'y pénétrer. Il y a une vingtaine d'années, un laboratoire s'intéressant justement à l'entrée du phage dans son hôte a mesuré le nombre maximal de particules virales adsorbées sur la bactérie ainsi que le nombre de phages y entrant effectivement. Par des incubations avec des inoculums contenant diverses proportions de phages marqués radioactivement au ^{14}C, suivies de centrifugations, rinçages à froid, et élimination des phages adsorbés en surface par du Triton X-100, les mesures de la radioactivité incorporée dans les bactéries ont démontré que plusieurs particules de ϕ6 sont capables de pénétrer une seule cellule, la moyenne se situant aux alentours de 3 virions [178]. Les mêmes auteurs ont également montré que de 10 à 50 virions peuvent s'adsorber sur le pilus bactérien en fonction de la souche utilisée. Ils proposent donc qu'il doit exister un mécanisme limitant le nombre d'entrée : soit au niveau du pilus bactérien qui pourrait se rétracter suite à la fusion de la membrane virale, soit que les bactéries infectées par un grand nombre de phages meurent prématurément sous l'effet des enzymes virales.

1.2.2 Virus d'animaux

1.2.2.1 Virus d'arthropodes : baculovirus

Les baculovirus sont des virus à ADN double brin circulaire qui infectent les insectes [179]. Lors de l'infection d'un hôte, deux formes de virions sont produites : (i) la forme incluse (plusieurs particules virales) dans une matrice dense de protéine, appelée ODV (occlusion-derived virus), qui résiste dans l'environnement jusqu'à l'ingestion par de nouveaux hôtes sains, et (ii) la forme libre appelée BV (budded virus) qui passe seulement de cellule à cellule à l'intérieur d'un hôte. Ainsi, c'est la forme ODV qui va être à l'origine de l'infection primaire d'un nouvel hôte. La raison ayant poussé les auteurs à étudier la MOI chez ce virus est le fait que la diversité génétique virale présente dans cet inoculum (ODV), et donc transmise d'hôte en hôte, dépend directement de la MOI de la cellule qui l'a produit.

La MOI chez ce baculovirus a été estimée à l'aide d'un mutant de délétion du gène impliqué dans la formation d'ODV. Ce mutant peut se propager à un nouvel hôte si, et seulement si, il co-infecte une cellule avec le virus sauvage et peut être inclus dans les ODV par complémentation. La co-inoculation des deux mutants et leurs passages successifs d'hôte en hôte a permis de déterminer que plus de 90% des cellules infectées renferment plus d'un génome au moment de la mort de la larve d'insecte. La valeur de MOI moyenne déduite a été estimée à 4,3 génomes par cellule [180].
Cette valeur de MOI élevée était envisageable car les populations naturelles de baculovirus comportent souvent de multiples variants génétiques, essentiellement des mutants de délétions, ce qui suggérait l'existence de complémentations fonctionnelles probablement à un niveau intracellulaire.

1.2.2.2 Virus de mammifères

1.2.2.2.1 Human immunodeficiency virus (HIV)

Le HIV appartient à la famille des rétrovirus qui possèdent un génome à ARN simple brin. Pour ce virus, qui évolue très rapidement, il est primordial d'apprécier la MOI, ou tout au moins le taux de co-infection des cellules par différents variants génétiques, afin de mesurer le potentiel en terme de recombinaison.
Chez deux patients infectés, l'étude de cellules de rate a révélé la présence de 1 à 8 provirus (génomes intégrés) dans chaque cellule infectée [181]. Plus de 75% des cellules infectées renferment au moins deux provirus avec une moyenne de 3,2 provirus par cellule. Toutefois, cette valeur serait sous-estimée, car la méthode utilisée (FISH) a permis exclusivement la détection des formes intégrées. Or, dans une autre étude, publiée par le même laboratoire, l'utilisation de la PCR sur ces mêmes cellules a aussi mis en évidence la présence de formes circulaires non intégrées [182].
L'infection multiple du HIV est-elle un événement purement stochastique ou bien existe-t-il des mécanismes qui pourraient réguler la MOI ?
Dans une étude expérimentale en cultures cellulaires (293 T et $CD4^+$ T), l'inoculation de deux virus recombinants (exprimant chacun une protéine distincte) illustre clairement que le taux de co-infection observé est

supérieur à ce qui serait attendu de manière totalement aléatoire (selon la probabilité théorique de rencontre, calculée d'après les fréquences initiales de chacun des deux virus) [183]. Toutefois, les auteurs suggèrent que la seule présence de cellules non-infectables, ainsi que de sous populations de cellules plus ou moins susceptibles à l'infection par le HIV, pourraient être à l'origine de cette différence observée.

1.2.2.2.2 Cytomégalovirus (CMV)

Les cytomégalovirus sont des virus d'animaux appartenant à la sous-famille des betaherpèsvirus, dont les génomes de grandes tailles (180 à 250 kb) sont à ADN double brin linéaire [179]. Ils sont capables de co-infecter leurs hôtes pluricellulaires, et la présence de multiples variants dans un hôte est corrélée à l'élévation de la charge virale et de la virulence. Il a été établi *in vivo* que cette co-infection est aussi présente au niveau cellulaire [184]. Pour parvenir à cette démonstration, les auteurs ont utilisé deux mutants de cytomégalovirus murin (MCMV), un exprimant la recombinase Cre et l'autre portant une GFP sous un promoteur inductible par Cre, dans le but d'estimer la co-infection au travers du nombre de cellules fluorescentes.
De plus, par la combinaison du virus sauvage et d'un virus atténué (croissance et dissémination plus lente), ils ont pu appréhender par hybridation *in situ* à 2 couleurs (2C-ISH) sur des cellules du foie très peu susceptibles à l'infection (0,013 %) quel bénéfice pourraient retirer ces virus de l'infection multiple d'une cellule. Au final, cette approche a permis de constater que le taux de co-infection observé (4,3%) est 70 fois plus élevé qu'attendu par rapport à un processus stochastique. Les auteurs suggèrent que des phénomènes d'encapsidation multiples et de migration de cellules co-infectées pourraient être à l'origine de ce taux de co-infection supérieur à l'attendu. Les auteurs postulent que la co-infection peut influencer positivement la valeur sélective du virus, grâce à une *trans*-complémentation fonctionnelle (le virus atténué bénéficiant de protéines produites par le sauvage) plutôt que par la recombinaison génétique qui n'a pas pu être détectée. Quoi qu'il en soit, les 4,3% de cellules co-infectées ne témoignent pas de l'existence d'une MOI moyenne élevée lors de l'infection de l'hôte par ce virus.

Au final, il y a très peu d'études qui vont jusqu'à déterminer la MOI naturelle durant l'infection d'un hôte pluricellulaire par un virus. La majorité des travaux publiés (y compris les plus détaillés mentionnés ci-dessus) se limitent à des informations sur le taux de co-infection et/ou le nombre moyen de variants détectés au sein d'une cellule. Il est ici nécessaire d'insister sur la distinction entre taux de co-infection et MOI. Si les deux paramètres sont évidemment liés, et si le taux de co-infection est directement observable, le calcul de la MOI demande en plus la connaissance de la probabilité d'infection par chacun des variants, afin de déduire le nombre moyen de génomes qui doit rentrer dans chaque cellule pour expliquer le taux de co-infection observé.
Ainsi pour les virus d'animaux, la seule MOI disponible est celle estimée chez les baculovirus. Qu'en est-il chez les virus de plante ?

1.2.3 Virus de plantes

Chez les phytovirus la MOI réelle, ou naturelle, durant l'infection systémique des plantes n'a jamais été formellement évaluée, quelle que soit l'espèce considérée. Toutefois, diverses informations fragmentaires et indirectes peuvent être déduites à partir de travaux s'intéressant à d'autres questions. Ces informations permettent de supposer que certains virus de végétaux ont une MOI naturelle très proche de 1.
La co-infection de plants de *Nicotiana benthamiana* par deux potyvirus (soit deux clones du *Plum pox virus* (PPV), soit deux clones du *Tobacco vein mottling virus* (TVMV), soit un clone PPV et un clone TVMV), exprimant des protéines fluorescentes différentes (GFP et DsRed), montre qu'il existe une séparation spatiale entre des virus de la même espèce ou du même genre, chacun occupant des territoires distincts matérialisés par des plages cellulaires soit verte, soit rouge, avec une limite nette à leur conjonction. Seule une petite frange à la frontière entre ces territoires présente des cellules possiblement co-infectées, puisque contenant les deux couleurs de fluorescence (figure 21, [185]). Toutefois, l'utilisation de protéines fluorescentes solubles et non adressées vers un compartiment cellulaire particulier peut entraîner la diffusion passive de celles-ci vers les cellules

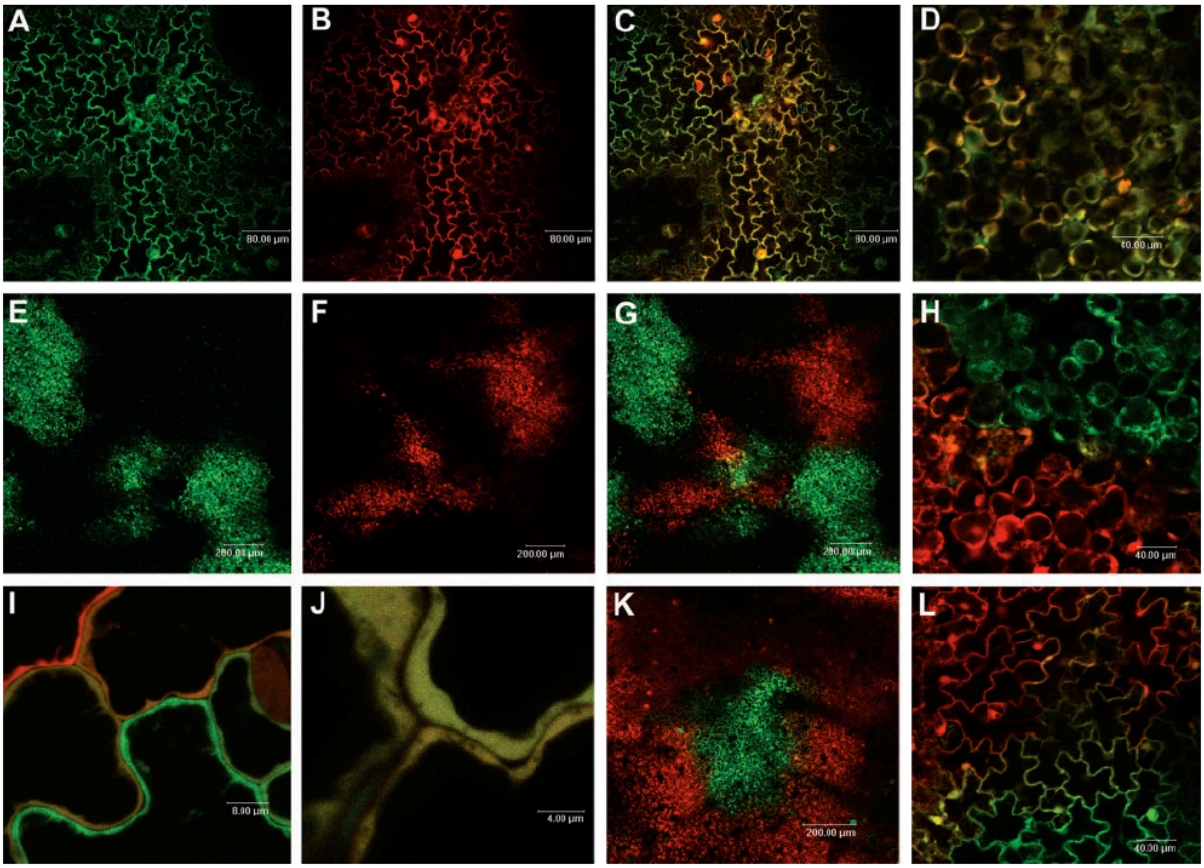


Figure 21 : Distribution spatiale de virus en co-infection systémique dans *N. benthamiana*. La microscopie confocale sur une plante co-inoculée par PPV-NAT-*AgfpS* (A) et PVX201-*optRed* (B) révèle la présence de nombreuses cellules épidermiques co-infectées, comme le montre la superposition des images (C). (D) Superposition d'images similaires à (C) qui montre des cellules du mésophylle co-infectées. La double infection d'une plante par PPV-NAT-*AgfpS* (E) et PPV-NAT-*red* (F) se traduit par une séparation des deux populations virales dans des « territoires » distincts de l'épiderme, comme illustré par la superposition des images (G). (H) Cette séparation spatiale est confirmée dans les cellules du mésophylle. (I) Gros plan issu de (G) au niveau de la jonction entre groupes de cellules différemment colorées. Les cellules co-infectées apparaissent en jaune alors que les cellules avoisinantes ne comportent que le PPV-NAT-*AgfpS* (en vert) ou le PPV-NAT-*red* (en rouge). (J) Gros plan issu de (C) montrant des cellules co-infectées par PPV-NAT-*AgfpS* et PVX201-*optRed,* arborant une couleur jaune pâle uniforme due au mélange rouge/vert. (K) Plante co-inoculée par PPV-NAT-*red* et TVMV-*gfp* affichant une séparation spatiale entre les deux espèces de potyvirus. (L) Gros plan issu de (K) révélant que dans ce cas aussi, une frange de cellules co-infectées pourrait exister au niveau des zones de jonction. *D'après [185].*

adjacentes, au travers des plasmodesmes modifiés pour faciliter le passage du virus. Cette possibilité n'a pas été envisagée et elle remet en cause la réelle co-existence de deux variants colorés même au sein de cette étroite frange de cellules.

A l'opposé, la co-infection des plantes avec n'importe quelle combinaison de l'un de ces potyvirus et un potexvirus (*Potato virus X,* PVX ; marqué aussi par une protéine fluorescente) aboutit à l'observation de nombreuses cellules co-infectées [185], indiquant que l'exclusion réciproque des virus apparentés décrits ci-dessus n'est pas effective pour des virus non apparentés.

Un travail semblable, effectué à partir de marquage fluorescent d'un autre potyvirus (*Bean yellow mosaic virus,* BYMV) et d'un cheravirus (*Apple latent spherical virus*, ALSV), a donné des résultats similaires [186]. C'est à dire que des clones viraux d'une même espèce ne sont pas retrouvés au sein d'une même cellule (mis à part éventuellement au niveau des frontières entre leurs aires de distribution respectives), tandis que l'infection mixte avec deux espèces de genres distincts montre la superposition des aires de répartition des deux virus.

Ces observations, réalisées à partir de virus ayant une protéine fluorescente codée par le génome, sont renforcées par les résultats d'une étude ayant analysée 230 feuilles issues d'un arbre infecté par le PPV depuis 13 ans. En effet, ces résultats révèlent la présence d'un seul variant génétique majoritaire dans chacune des feuilles analysées, tandis que l'ensemble de l'arbre comporte un total de 33 haplotypes [187]. Les auteurs de cette étude suggèrent que l'absence d'infection multiple au sein d'une feuille puisse être liée à la colonisation rapide, par les génomes d'un unique variant, des quelques cellules initiant la formation du bourgeon, suivie de l'inhibition de la surinfection par un autre variant.

Enfin, il semble aussi fort probable que deux variants de *Tobacco mosaic virus* (TMV) ne réussissent que très rarement à infecter une même cellule. En effet, la co-expression de deux protéines fluorescentes, introduites chacune dans un même vecteur d'expression viral issu du TMV, n'est observée que dans 4 à 5 % des cellules de la plante co-inoculée [188].

Bien qu'aucune ne vise directement à évaluer la MOI, toutes ces données semblent converger vers le fait que les phytovirus à ARN positif simple brin semblent peu à même d'infecter les cellules de manière multiple. Ceci, en tout cas, pour des infections multiples de génomes très apparentés.

En ce qui concerne notre modèle biologique, il a été établi que le CaMV possède un taux de recombinaison remarquablement important, aboutissant à plus de 50% de génomes recombinants dans une population virale suite à un seul passage dans une plante hôte [189]. Ce résultat suppose que la MOI du CaMV soit relativement élevée, puisqu'une condition *sine qua non* de la détection d'un événement de recombinaison est qu'au moins deux génomes contenant des marqueurs distincts aient co-infecté une même cellule. Sur la base de ce résultat, il nous est alors apparu très stimulant d'entreprendre l'évaluation formelle de la MOI naturelle du CaMV, en particulier dans le cadre de ce programme de thèse, pour les raisons suivantes :

i) la MOI informe sur la diversité potentielle de l'échantillon viral qui est accessible aux pucerons dans une seule cellule
ii) la MOI naturelle n'a jamais été formellement évaluée chez les phytovirus et une unique fois chez un virus d'animaux
iii) les données disponibles pour les phytovirus (mentionnées ci-dessus) suggèrent une faible MOI, notre modèle biologique pourrait aller à contresens et démontrer que plusieurs situations très contrastées peuvent exister suivant les espèces virales
iv) la distribution de la diversité génétique virale au niveau des cellules, à l'intérieur d'une feuille, est un complément d'information élégant vis à vis de notre étude sur la taille efficace (*Ne*) présentée au chapitre précédent. Démontrer une forte MOI pour le CaMV nous paraîtrait en parfaite cohérence avec une *Ne* supérieure à celle décrite pour les autres virus de plantes (à ARN simple brin positif notamment) (*cf.* dernier paragraphe de la discussion de l'article 2). Eventuellement, une corrélation positive entre MOI et *Ne* pourrait être proposée (*cf.* discussion du chapitre).

NB : Ce travail a été initié et développé en collaboration très étroite et constante tout d'abord avec Michel Yvon (TR INRA), puis aussi avec Serafin Gutiérrez (Post-doc INRA) qui prendra en charge la rédaction imminente de l'article correspondant, et le développement des perspectives de ce projet.

2 Matériel & Méthodologie mise en œuvre

Nous avions tout d'abord pensé mettre au point, et adapter à des plantes infectées par le CaMV, des techniques liées à l'hybridation *in situ*. Ceci aurait théoriquement permis de détecter simultanément deux variants distincts du CaMV, par exemple VIT1 et VIT3 utilisés dans les chapitres précédents, à l'aide de sondes fluorescentes de différentes couleurs.

Cette mise au point a été un échec pour des raisons qui devront être élucidées, car ces techniques pourraient bien devenir indispensables à l'avenir (*cf.* Discussion et Perspectives). Nous avons tenté deux approches différentes.

i) Le FISH (Fluorescent In Situ Hybridization) pour lequel nous n'avons pas réussi à clairement identifier de structures virales marquées spécifiquement dans les cellules. Il est possible que cette difficulté provienne d'un défaut d'accessibilité des génomes viraux, enchâssés dans des virions, eux-mêmes enfouis au cœur des corps denses.

ii) Le PRINS (Primed in situ labelling) basé sur l'amorçage spécifique et *in situ* d'un oligonucleotide sur une matrice cible, afin d'élonguer l'amorce en incorporant des nucléotides fluorescents. Si cette technique est particulièrement bien adaptée au marquage de régions chromosomiques sur des caryotypes, elle ne s'est pas avérée facilement adaptable à notre matériel viral. En particulier nous n'avons pas réussi à inhiber l'amorçage non spécifique sur l'ADN chromosomique et mitochondrial dans des protoplastes infectés.

Peut-être une plus grande persévérance aurait elle été nécessaire, cependant, en parallèle, Michel Yvon a eu l'idée (et a démontré la faisabilité)

d'une détection spécifique des variants du CaMV par nested-PCR, sur cellules isolées. Techniquement, il s'agit d'une première en virologie végétale qui paraissait jusqu'alors irréalisable. Nous nous sommes par la suite logiquement concentrés sur cette alternative particulièrement performante.

2.1 Principe général de la méthode développée

Le protocole original développé pour cette partie du travail est basé sur le principe suivant :

Deux variants génétiques du CaMV, distinguables par PCR, sont co-inoculés à une plante. Après, ou au cours du développement de l'infection systémique, une feuille est prélevée et protoplastée. Environ 3000 protoplastes sont utilisés pour déterminer les fréquences relatives de chacun des deux variants dans l'ensemble de cette feuille, tandis que d'autres protoplastes sont isolés et servent individuellement de matrice pour détecter chacun des deux variants génétiques. De ces PCR sont déduits à la fois le taux de cellules infectées, et le pourcentage infecté par l'un, par l'autre, ou par les deux variants (donc le taux de cellules co-infectées). Un calcul statistique nous permet alors d'inférer le nombre de génomes qui infectent en moyenne chaque cellule, donc d'inférer la MOI.

Ce calcul statistique repose sur le postulat que les deux variants ont une capacité identique à pénétrer et infecter une cellule. En effet, sous cette hypothèse, la probabilité que chacun y pénètrent, donc la probabilité qu'ils co-infectent, ne dépendra que de leur fréquence dans la population virale (donnée connue) et du nombre qui rentre dans chaque cellule (MOI, valeur inconnue). Il est alors possible, à partir des fréquences observées pour chacun des variants, de prédire un taux théorique de cellules co-infectées en fonction de la MOI, et ainsi d'évaluer la valeur de MOI qui explique le mieux les résultats observés (le taux de co-infection observé).

L'ensemble de cette technique a demandé des ajustements très fins de chacun des protocoles, qui ne correspondent pas à des procédures classiques utilisées en routine. C'est pour cette raison qu'ils sont décrits en détail dans ce chapitre et non pas dans le Matériels et Méthodes général.

2.2 Variants viraux et amorces spécifiques utilisés

Les variants génétiques du CaMV utilisés ici sont VIT1 et VIT3, déjà décrits dans l'article 1 [190]. La fréquence relative de ces deux variants génétiques a été montrée comme étant très stable durant le développement de l'infection systémique et le passage de la population de feuille en feuille, même après une période de plus de 30 jours. A l'intérieur d'une feuille individuelle, ces deux variants peuvent donc être considérés neutres, la variation de leur fréquence relative étant négligeable (*cf.* Article 2).

Les couples d'amorces utilisés (positions figurées sur la figure 22) sont les suivants:

Big One :

F 5'-CACGATGCCACAGGTAAGG-3'

R 5'-CTATTAGTTGCTCTTTAGGTTGGG-3'

Net 1 :

F 5'-TAACAAGGTCCTGCCCCTTA-3'

R 5'-ACTCCGAAGGGTCTTTGCTT-3'

Net 3 :

F 5'-CTCAGCTGTTGGGTTCTGGT-3'

R 5'- TCCTCTTTAGTTAAGGGCTCTGG-3'

Tiny 3 :

F 5'-TAGCCCAAAACAAAGCCTTATTAG-3'

R 5'-AATGCTCATACTAGACCAGAACC-3'

VLight :

F 5'-AGGTAACAGTGCTTCATCC-3'

R 5'-GTTGTCCCGTAATGCTCA-3'

2.3 Quantification des variants et co-inoculation

Suite aux purifications des particules virales VIT1 et VIT3, nous avons estimé leurs concentrations par spectrophotométrie dans le but de réaliser leur co-inoculation avec un ratio 1/1 (voir Matériels et Méthodes en fin de manuscrit).

Nous avons alors vérifié et quantifié le pouvoir infectieux de ces deux préparations virales – sur un lot de 15 plantes – en effectuant des

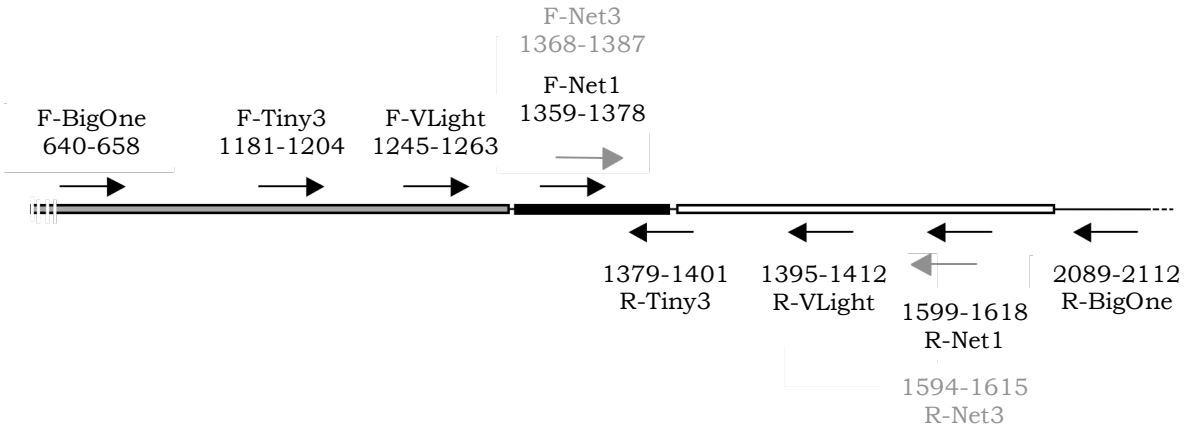


Figure 22: **Position des amorces utilisées pour la Nested-PCR sur VIT1 et VIT3.** Sur le trait horizontal, représentant une partie du génome viral du CaMV, sont figurées les positions des différents couples d'amorces utilisés lors des Nested-PCR. La région noire (1348-1392) sur le génome viral correspond à l'insertion (entre les gènes I et II, respectivement représentés en gris et blanc) des marqueurs spécifiques de VIT1 et de VIT3. Le couple *BigOne* est utilisé pour la première amplification (Nested-PCR-1). Lors de la seconde étape (Nested-PCR-2), les amorces *F-Net1*, *F-Net3* et *R-Tiny3*, localisées dans la région d'insertion des marqueurs, permettent spécifiquement de détecter la présence d'un VIT donné. Pour la Nested-PCR-2, le kit HRM (High Resolution Melting master, Roche) associé à l'appareil LightCycler® 480 permet à l'aide du seul couple d'amorce *(VLight)* de différencier les protoplastes mono-infectés par VIT1 ou VIT3, des protoplastes co-infectés par VIT1 et VIT3.

inoculations mécaniques (à des concentrations de 100 et 50 ng/µL) avec chacun des clones VIT, disposé en vis à vis de part et d'autre de la nervure centrale (2 feuilles par plante). A 9 jours post-inoculation, nous avons effectué un dénombrement des lésions locales provoquées par chacun des deux clones VIT, sous lumière ultraviolette (figure 23). Le tableau 1 montre que ces deux préparations de particules virales n'avaient pas le même pouvoir infectieux, la différence étant évaluée à VIT1 = 1,65 VIT3. Nous avons alors ajusté les concentrations respectives de VIT1 et VIT3 afin d'obtenir autant de lésions locales pour chacun des variants à volume d'inoculum égal. Pour préparer les inoculations ultérieures, les solutions de VIT1 et VIT3 utilisées ont donc été respectivement fixées à 75 et 123,75 ng/µL, et un volume total de 20 µL par feuille a été utilisé.

2.4 Récolte des feuilles, protoplastation et fixation

2.4.1 Récolte des feuilles

Toutes les feuilles récoltées et analysées sont des feuilles ayant développé des symptômes systémiques. A chaque date (21, 42, 60 et 84 jours post-inoculation ; voir le protocole d'échantillonnage dans la partie Résultats), les feuilles choisies sont celles en cours d'élongation ayant atteint un stade d'expansion compatible avec une bonne efficacité de protoplastation, elles correspondent en général à des feuilles en position 4 ou 5 en descendant depuis le centre de la rosette. Les feuilles prélevées sont donc de plus en plus « haut » dans la rosette, avec un intervalle de 4 à 6 feuilles entre chaque prélèvement.

2.4.2 Protoplastation

Pour éviter de peler péniblement l'épiderme inférieur des feuilles prélevées, celles-ci sont immergées (durant 2 à 5 minutes selon l'âge de la feuille) dans une solution d'eau distillée additionnée de 6% de DOMESTOS Original, puis soumises à des rinçages successifs à l'eau milliQ jusqu'à disparition complète de la mousse. À l'aide d'une seringue munie d'un filtre 0,22 µm, 12,5 mL de tampon de lyse (0,5 M de D-Mannitol avec 0,5% de cellulase et 0,05% de Macérozyme en poids/volume) sont déposés dans des boîtes de

	1 VIT par feuille				2 VIT sur chaque feuille			
VIT n°	1		3		1	3	1	3
[inoculum] ng/µL	50	100	50	100	50	50	100	100
Nombre de lésions locales (LL) sur chaque moitié de feuille					11	15	216	196
	56	180	10	35	33	27	82	40
	127	173	8	18	26	16	25	12
	70	136	109	95	42	33	50	21
	24	57	16	128	20	8	82	54
	9	41	7	23	6	2	369	132
					256	154	64	45
					69	35	79	70
Moyenne	57	117	30	60	58	36	121	71
Rapport LL 100/50	2,05		1,99					
Rapport VIT1/VIT3					1,60		1,70	

Tableau 1 : Dénombrement des lésions locales pour quantifier le pouvoir infectieux des préparations virales. Dans le but de nous assurer de la qualité des purifications virales, des inoculations mécaniques ont été réalisées à 2 concentrations (50 et 100 ng/µL). Tout d'abord, le même variant a été inoculé à deux concentrations différentes sur les deux moitiés d'une même feuille, afin de vérifier que la dilution modifie le nombre de lésions locales (LL) de manière linéaire (voir le rapport de LL 100/50). Ensuite, les deux variants ont été inoculés en vis-à-vis, sur les deux moitiés d'une même feuille, et à la même concentration, pour évaluer leur éventuelle différence de pouvoir infectieux. Que ce soit à 50 ou 100 ng/µL on remarque que le VIT1 est en moyenne 1,65 fois plus infectieux que le VIT3, à la même concentration spectrophotométrique. Pour réaliser des inoculations à un ratio 1/1 il faudra donc utiliser une concentration 1,65 fois plus élevée pour VIT3 que pour VIT1.

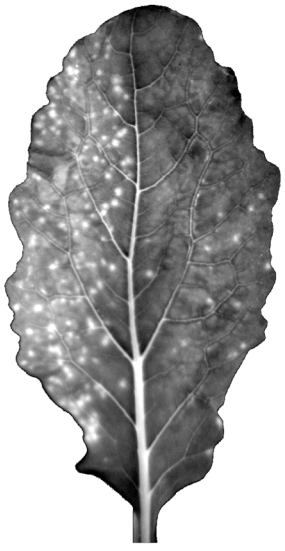


Figure 23 : Inoculation en vis-à-vis pour quantifier le pouvoir infectieux. Dans le but d'évaluer la qualité des purifications virales en terme de pouvoir infectieux, des inoculations sont réalisées de chaque coté de la nervure centrale, ici à 2 concentrations différentes pour un même clone (50 et 100 ng/µL). L'image correspond à une photo sous U.V. d'une feuille inoculée par VIT3 : 128 et 16 lésions locales sont respectivement visibles sur les parties gauche et droite de la feuille inoculée. Ces valeurs sont incluses dans le tableau 1.

pétri (une par feuille analysée), et les feuilles prétraitées sont dénervées et découpées afin d'y être déposées. Un soin particulier est consacré à ce dépôt afin d'éviter les bulles d'air et de maximiser les contacts entre la face inférieure du limbe et le tampon de lyse. Les boîtes de pétri sont alors délicatement parafilmées avant une incubation pendant une nuit à 25°C et à l'obscurité. Il est possible de n'incuber que 2 à 3 heures en doublant les concentrations d'enzymes.

Après cette macération, les boîtes sont légèrement agitées pour remettre les protoplastes en suspension, avant de les filtrer au travers de Miracloth, et ainsi éliminer les plus gros débris. Le filtrat est alors centrifugé (Centrifuge 5810 R, Eppendorf) à 60 × *g* pendant 5 minutes, en prenant soin de programmer une accélération lente (4/9) et une décélération nulle (sans frein). Enfin, le surnageant est éliminé à l'aide d'une pipette, avant de resuspendre précautionneusement le culot de protoplastes avec un volume (10 mL) de D-Mannitol à 0,5 M. Cette étape finale de centrifugation/rinçage est réitérée 2 fois afin de bien rincer avant de procéder à la fixation.

2.4.3 Fixation des protoplastes

A la fin de la protoplastation, le culot de cellule est délicatement repris dans 10 mL d'une solution fixatrice (4% paraformaldéhyde, 0,05% glutaraldéhyde, 10 mM NaH_2PO_4, 0,15 M D-Mannitol), et incubé durant 30 minutes à température ambiante sous agitation douce. Une alternative consiste à resuspendre le culot de protoplastes avec 5 mL de D-Mannitol 0,5 M et d'y ajouter 1% final de glutaraldéhyde avant d'incuber sous agitation durant 15 minutes. Le fixateur est ensuite éliminé par une étape de centrifugation/rinçage (60 × *g* 5 min., culot repris dans 10 mL) une première fois au D-mannitol 0,5 M, puis une seconde fois avec un mélange de D-Mannitol et de PBS (0,25 M et 0,5X, respectivement), pour terminer par un dernier rinçage au PBS 1X avant une resuspension finale des culots dans 10 mL de PBS 1X.

2.5 Décontamination et isolement des protoplastes

2.5.1 Décontamination du surnageant

Durant la protoplastation, un grand nombre de cellules sont cassées et libèrent potentiellement dans le milieu des virions, voire des génomes viraux, qui deviennent autant de contaminants dont il nous a absolument fallu nous affranchir. En effet, toute trace de contamination rendrait notre protocole caduc, et ce problème est probablement à l'origine du fait que la détection de virus par PCR sur protoplastes isolés n'ait jamais été développée avec succès auparavant.

Les cellules sont tout d'abord colorées par ajout de 3 µL de bleu de Coomassie (Brilliant Blue R) à une fraction de 250 µL d'une suspension de protoplastes préalablement fixés (environ 25 000), le tout laissé à incuber 10 minutes à température ambiante, avant de diluer jusqu'à un volume de 2 mL avec de l'eau distillée. La suspension colorée est alors centrifugée 2 minutes à 100 × *g,* puis le surnageant éliminé et le culot repris avec 1 mL d'eau distillée. À partir de cette dilution, 250 µL sont prélevés et dilués à nouveau avec 250 µL de trypsine à 0,2 % (2 mg dans 1mL de PBS 1X), 3 µL de DNAse (10 mg/mL) et 3 µL de MgCl2 (50 mM). Ce traitement de décontamination est mené durant 10 minutes à température ambiante puis l'activité de la trypsine est bloquée avec de l'aprotinine (2 µM final) en 5 minutes à température ambiante. La solution décontaminée est soumise à une dernière dilution avec 1 mL d'eau distillée pour obtenir environ 1 protoplaste par microlitre ; stockage possible quelques jours à 4°C.

2.5.2 Isolement des protoplastes

Sur une paillasse désinfectée, des gouttes de 2 µL de la solution de protoplastes décontaminée sont déposées sur des bandelettes de plastique (intercalaires de classeur) de 2 à 3 mm de large. Les cellules s'adsorbent alors sur le plastique durant 10 minutes, à la suite de quoi l'excédant de liquide est éliminé avec du papier absorbant. Après 10 minutes de séchage à température ambiante, les bandelettes sont rincées à l'aide d'une pissette d'eau distillée.

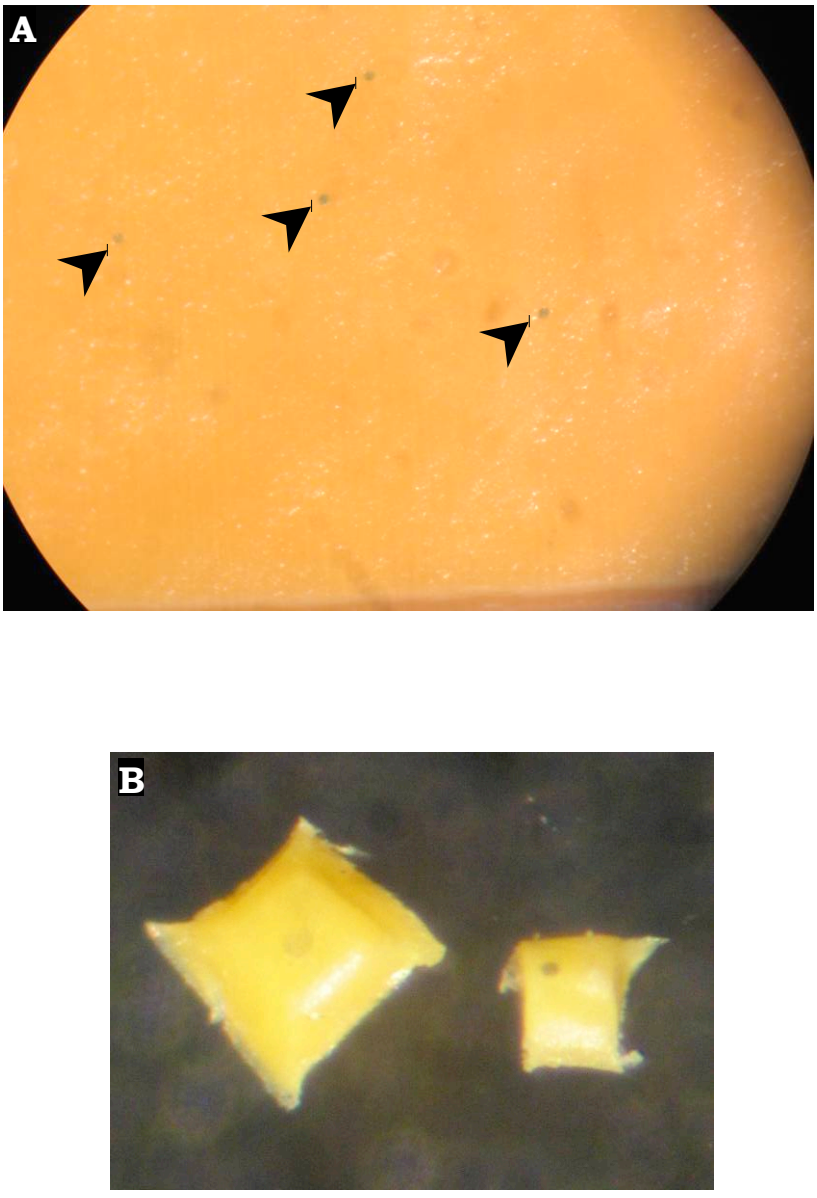


Figure 24: **Découpe des protoplastes fixés et déposés sur bandelette de plastique.** (A) Sous une loupe binoculaire, les bandelettes de plastiques sont observées afin d'identifier la présence de protoplastes colorés au bleu de Coomassie (voir les flèches). (B) Chaque protoplaste est alors isolé par découpe (environ 0,1 mm^2) à l'aide d'un scalpel et d'une pince fine stérilisés à chaque contact avec le plastique. La taille des protoplastes est de 35 et 16 µm respectivement à gauche et à droite.

Les protoplastes ainsi isolés et adsorbés sur les bandelettes de plastique, elles-mêmes déposées sur une lame de verre stérilisée à la flamme, sont visibles grâce à la coloration au bleu de Coomassie, et sont alors découpés individuellement sous une loupe binoculaire à l'aide d'un scalpel et d'une pince fine. Les outils sont repassés à la flamme entre chaque contact avec le plastique (figure 24). Au final, chaque découpe d'une surface approximative de 0,1 mm^2, contenant une seule cellule, est déposée dans un microtube de 0,2 mL afin de servir de matrice pour l'analyse de la diversité virale par Nested-PCR.

2.6 Détection des variants à partir de protoplaste fixé

2.6.1 Libération de l'ADN viral et Nested-PCR-1

Dans chaque puits d'une plaque de PCR contenant une découpe de plastique avec un protoplaste isolé, 14 µL d'une solution de « libération » de l'ADN sont ajoutés (tampon PCR 1,43X ; $MgCl_2$ 7,14 mM et Protéinase K à 143µg/mL). Après 30 minutes à 37°C, puis 30 minutes à 95°C pour inactiver la protéinase K, chaque puit est complété avec 6 µL du complément de mix PCR à savoir : les dNTP (125 µM final), les amorces externes de la Nested-PCR-1 (« Big One » forward et reverse à 0,25 µM chacune) et 1 unité de Taq polymérase. Le programme de la Nested-PCR-1 est le suivant : une première étape de dénaturation de 15 minutes à 95°C, suivie de 40 cycles d'amplification (94°C 45s ; 54°C 45s ; 72°C 60s), et d'une dernière étape d'élongation de 10 minutes à 72°C. Chaque produit de PCR ainsi obtenu est ensuite dilué au 30^{ième}, et un à cinq microlitres serviront de matrice lors de l'étape suivante en fonction de la méthode employée (Nested 2 classique ou HRM, voir ci-dessous).

2.6.2 Nested-PCR-2

2.6.2.1 Par PCR classique

A partir de 1 µL de matrice (Nested-PCR-1), la Nested-PCR-2 est effectuée de manière similaire mais avec deux couples d'amorces internes (ici : « Net 1 R&F » et « Net 3 R&F »), permettant la discrimination de nos deux variants, ainsi qu'une quantité de Taq polymérase divisée par deux (soit

0,5U). Le programme utilisé pour la Nested-PCR-2 est le suivant : une première étape de dénaturation de 5 minutes à 94°C, suivie de 35 cycles d'amplification (95°C 15s ; 56°C 15s ; 72°C 30s) et d'une dernière étape d'élongation de 10 minutes à 72°C. Les produits de PCR sont déposés sur gel d'agarose afin de déterminer la présence/absence de chacun des deux virus marqués dans chaque protoplaste isolé (figure 25).

2.6.2.2 Avec le LightCycler® 480 : kit HRM

Le kit HRM (High Resolution Melting master, Roche) permet d'effectuer des PCR en temps réel et de définir la composition qualitative des échantillons testés d'après le profil obtenus lors de l'étape finale de dénaturation (fusion) des produits PCR. L'avantage de cette technique réside dans l'utilisation d'un seul couple d'amorces pour la discrimination de plusieurs clones viraux, éliminant ainsi tout risque de fluctuation dans les efficacités de PCR qui dépendrait du couple d'amorces utilisé pour chaque variant. A partir de 2 µL de matrice (Nested-PCR-1), le kit est utilisé suivant les recommandations du fournisseur, à savoir que les réactions se font dans 10 µL final avec les concentrations suivantes : 1X de MasterMix, 3 mM $MgCl_2$ et 225 nM de chaque amorce (« Vlight » F&R). Le programme paramétré dans l'appareil est le suivant : 5 minutes de dénaturation à 95°C, puis 45 cycles (95°C 15s, 65°C 15s, 72°C 25s) avec une activité « touchdown » sur les 20 premiers cycles (65 à 60°C, -0,25°C par cycle). L'analyse des profils de fusion des produits PCR est effectuée avec le logiciel « LightCycler®480 Genotyping Software » – livré avec la machine – sous le mode « Gene Scanning » ou « Tm Calling ».

2.6.3 Vérification de l'efficacité la méthode de décontamination

Afin de contrôler si la méthode de décontamination mise au point était suffisamment efficace, nous avons protoplasté ensemble (dans la même boîte de pétri dès le stade initial de macération) des morceaux de feuilles infectées provenant de deux plantes (mono-)inoculées, l'une par VIT1 et l'autre par VIT3. Dans cette expérience, les cellules infectées soit par VIT1 soit par VIT3 sont en contact dès qu'elles sont libérées de leur paroi cellulosique et durant tout le processus ultérieur de préparation. A l'issu de cette co-macération,

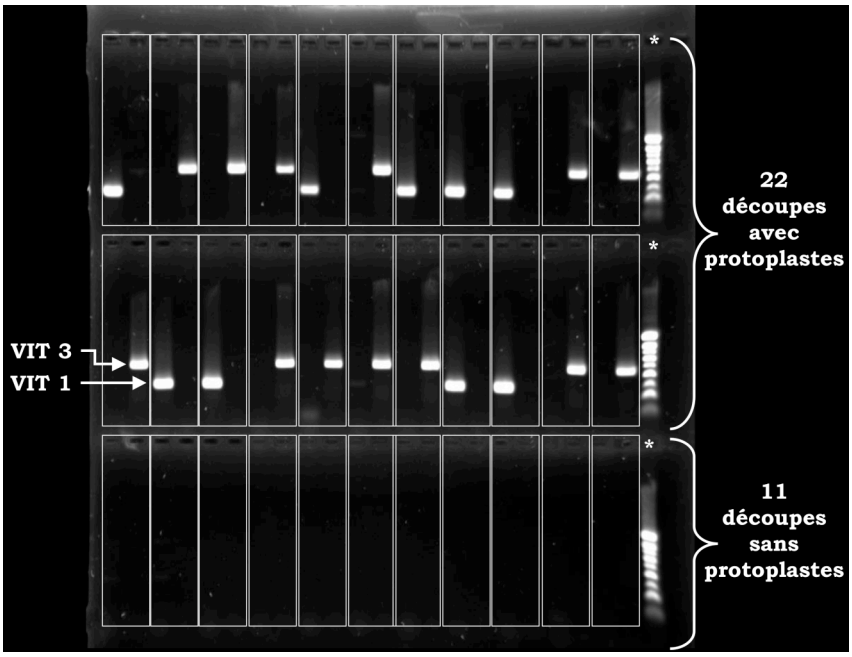


Figure 25 : Validation de la méthode de décontamination des protoplastes fixés. Après préparation et isolement des protoplastes (issus de la macération simultanée de feuilles infectées soit par VIT1 soit par VIT3, dans le même contenant), l'efficacité de la décontamination a été vérifiée à partir de découpes avec et sans protoplastes. La technique de Nested-PCR classique nécessite 2 couples d'amorces pour l'analyse de chaque découpe. Les rectangles blancs délimitent l'analyse des différentes découpes, comprenant chacune un puits pour la détection de VIT1 et un autre pour celle de VIT3. Le fragment d'amplification produit par les amorces spécifiques de VIT3 (F-Net3 et R-Net3) est plus long que celui généré par les amorces de VIT1 (F-Net1 et R-Net1) comme précisé sur la gauche de la figure. Un seul variant est significativement détecté dans chaque cellule. (*) désigne les puits dans lesquels ont été chargés le marqueur de taille 0,1-1kb.

	Découpe sur bandelette		**Profil détecté (en nombre (en fréquence))**			
Technique	Nature	Nombre	Négatif	VIT1	VIT3	VIT1 + VIT3
Nested-PCR Classique	avec protoplaste	77	7	26	42	2 (2,6%)
	sans protoplaste	25	24	0	0	1 (4%)
HRM	avec protoplaste	79	8	53	16	2 (2,5%)
	sans protoplaste	15	15	0	0	0 (0%)

Tableau 2 : Validation de la méthode de décontamination des protoplastes fixés. Synthèse de tous les résultats obtenus d'après les gels d'agarose réalisés (exemple avec la figure 25) et d'après les analyses en HRM. La très faible proportion de découpes détectées comme co-infectées, confirme l'efficacité presque parfaite de cette méthode de décontamination (voir Matériels et Méthodes).

les protoplastes ont été fixés, décontaminés, déposés sur bandelettes et enfin isolés par découpe dans l'objectif de vérifier la non-présence des 2 variants dans chaque cellule, par une Nested-PCR classique, ou par HRM. Les résultats démontrent la fiabilité de la méthode, puisque seule une fraction raisonnablement faible (2,6%) des protoplastes analysés a montré une détection des deux variants (tableau 2). Pour les 2 protoplastes faussement détectés comme co-infectés, il existe deux explications possibles : (i) sur cette découpe, 2 protoplastes se chevauchaient de façon indécelable à la loupe binoculaire ; et (ii) nous avons détecté une contamination résiduelle du protoplaste isolé, ou de la découpe de plastique (comme cela semble très rarement possible, *cf.* tableau 2).

Quoi qu'il en soit, la détection de tels faux positifs reste très faible, et nous est apparue comme totalement compatible avec la poursuite du projet sur la base de cette méthodologie.

2.7 Détermination des fréquences relatives de chaque variant dans une feuille

Pour déterminer les fréquences relatives des deux variants au sein de la population d'une feuille, 3000 à 3500 protoplastes fixés (en solution) sont amplifiés par Nested-PCR-1 avant d'effectuer une PCR quantitative en SYBR® Green sur les produits de la Nested-PCR-1. La mesure par SYBR® Green nécessite l'emploi d'un couple d'amorces spécifiques par puits et par variant à quantifier (par exemple : « Net 1 F&R » et « Tiny 3 F&R »). Afin de limiter les variations qui seraient imputables à de trop faibles quantités initiales de copies matrices, la quantité minimale de matrice (produits de la Nested-PCR-1) utilisée est de 5 µL pour chaque puits. Sur chaque plaque, conjointement aux échantillons, une gamme étalon correspondant à chacun des deux variants a systématiquement été déposée (plasmides de concentrations connues), afin d'en permettre la quantification. Le milieu réactionnel préparé selon le « qPCR Core kit for SYBR® green I No ROX » (Eurogentec, San Diego, CA) comprend, en plus de la matrice, 2,5 µL de tampon 10X, 250 mM de chaque amorce, 200 mM des dNTP, 3,5 mM de $MgCl_2$, 0,75 µL de SYBR® green I, 0,625 unité de Hot GoldStar DNA

polymerase, et de l'eau stérile pour compléter à 25 µL final. Le programme d'amplification PCR en temps réel est le suivant : après 10 minutes de dénaturation à 95 °C, les 40 cycles successifs (15 s à 95°C, 15 s à 56°C et 30 s à 72°C) sont clôturés par un cycle final de 60 s à 95°C, 30 s à 55°C et 30 s à 95°C. Les PCR et les analyses de courbes de dénaturation/fusion sont réalisées avec l'appareil Mx3005P™ QPCR System (Stratagene, La Jolla, CA) et le logiciel livré conjointement (Mx3005p software).

2.8 Pour le calcul de la MOI

Un programme (détails en Annexe 2) a été écrit sous le logiciel « R » par G. Thébaud (INRA, Montpellier) afin de calculer la MOI moyenne la plus vraisemblable à partir des fréquences mesurées des deux variants dans la population et des taux de co-infections observés. La méthode du maximum de vraisemblance calcule la valeur de la MOI pour laquelle la probabilité d'observer les données expérimentales (taux de co-infection, connaissant les fréquences relatives des variants dans la feuille) est maximale.

Le modèle décrivant l'expérience repose sur les hypothèses suivantes : (i) indépendance d'infection, c'est à dire qu'une première infection n'influence pas la probabilité d'une autre infection par quelque variant viral que ce soit, (ii) distribution identique (homogène) des infections des cellules hôtes, c'est à dire que les protoplastes ont tous la même probabilité d'être infectés, (iii) expériences indépendantes, de sorte que les résultats expérimentaux d'une plante n'influencent pas ceux d'une autre plante, et enfin (iv) distribution identique des caractéristiques du système expérimental (évolution de la MOI dans le temps, conditions expérimentales…) sur toutes les expériences (réplicats).

Le nombre de virions infectant une cellule (MOI) est modélisé par une loi de Poisson. Le nombre de protoplastes co-infectés est modélisé comme une loi binomiale dont le paramètre correspondant à la probabilité de succès (c'est-à-dire de co-infection) ne dépend que de la MOI et de la fréquence relative des deux variants. Comme cette fréquence relative a été mesurée (dans chacune des feuilles échantillonnées) il est possible d'estimer, par la méthode du maximum de vraisemblance, la MOI (moyenne) la plus probable

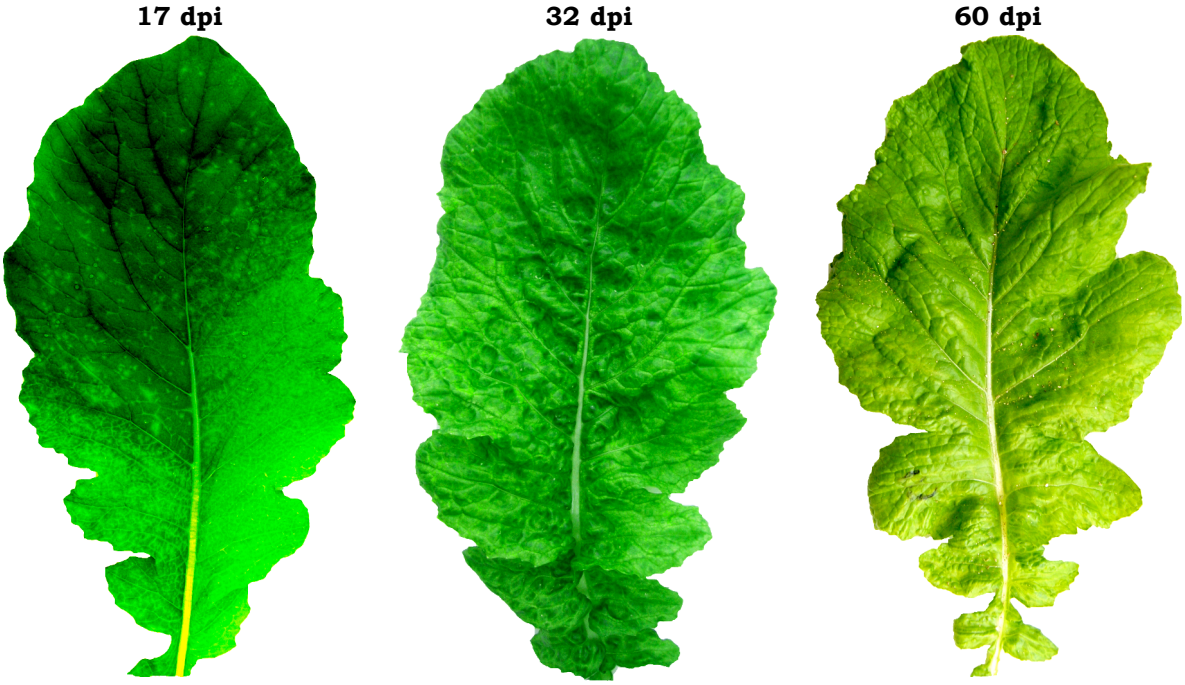


Figure 26 : Symptomatologie provoquée par le CaMV au cours du temps. Ces 3 photographies de feuilles présentent les symptômes typiquement observés à 17, 32 et 60 jours sur une plante co-inoculée par VIT1 et VIT3. A 17 dpi, la feuille n°7 arbore les premiers symptômes systémiques visibles avec une mosaïque basipétale (*cf.* Introduction générale, section 1.3.1.1. pour l'explication de cette répartition). A 32 dpi (la feuille n°13 par exemple) présente un gaufrage et une mosaïque généralisée. Ensuite à 60 dpi, les éclaircissements de nervure s'étendent et se rejoignent, d'où l'aspect jaunâtre uniforme de la feuille.

	17 dpi			**32 dpi**			**60 dpi**
Plante n°	***1***	***2***	***3***	***1***	***2***	***3***	***3***
Nombre de cellules infectées / testées	21/28	23/28	22/24	27/29	31/31	30/32	85/90
Nombre de co-infection	4	5	6	10	18	23	81
Taux de co-infection (%)	19	22	27	37	58	77	95

Tableau 3 : Evolution du taux de co-infection dans la plante au cours du temps. Sur 3 plantes inoculées par un mélange des deux variants du CaMV (ratio 1/1 de VIT1 et VIT3), des feuilles ont été prélevées à différents jours post-infection, protoplastées, puis analysées par Nested-PCR-HRM. A 17 dpi, c'est la feuille montrant les premiers symptômes systémiques – feuille n° 5 ou 6 selon la plante – qui a été prélevée. Ensuite, à 32 dpi, des feuilles systémiques supérieures – feuilles n°12 ou 13 selon la plante – ont été choisies. Pour chaque plante, il y a une augmentation du taux de co-infection entre le 17ième et 32ième jours après inoculation. La plante N°3 a été maintenue jusqu'à 60 dpi, et l'étude d'une de ses feuilles confirme que le taux de co-infection s'accroît continuellement et significativement (χ^2 ; P<0,01) au cours de l'infection de la plante.

au vu de l'ensemble des données expérimentales obtenues à un temps précis de l'infection.

3 Résultats

L'arrivée en post-doc de Serafin Gutiérrez a « dopé » les nombreuses réflexions quotidiennes et a considérablement enrichi l'interprétation de ces résultats.

3.1 Prologue exploratoire du taux de co-infection

Vu l'absence totale d'étude sur le sujet, et donc d'*a priori* sur la répartition au niveau cellulaire de deux variants génétiques neutres (ou quasi-neutre) dans une feuille, nous avons décidé de débuter par une première expérience exploratoire visant à mieux définir les protocoles ultérieurs.

Sur un lot de 3 plantes co-inoculées avec un mélange de VIT1 et VIT3 à un ratio 1/1, nous avons prélevé des feuilles à différents temps après inoculation, afin d'explorer le taux de co-infection des cellules par ces deux variants, et son éventuelle évolution, durant le développement de l'infection systémique. L'analyse par Nested-PCR-HRM des protoplastes isolés à partir des feuilles (figure 26) a mis en évidence plusieurs phénomènes très encourageants (tableau 3), et déterminants pour la suite de ce programme :

i) Dès le début de l'infection (17 dpi), le taux de co-infection des cellules par les deux variants est très élevé (20-30%), indiquant la faisabilité de notre procédure d'investigation.

ii) Il semble que le taux de co-infection ne soit pas une constante, mais augmente au fil de la progression de l'infection systémique dans les étages foliaires supérieurs. Entre le 17^{ième} et le 32^{ième} jour après inoculation ce taux passe en moyenne de 23 à 57 % de co-infection.

iii) L'unique feuille analysée à 60 jours indique que le taux de co-infection augmente encore avec le temps et qu'il pourrait même peut-être atteindre 100%. Ces observations nous ont engagés, dans les expérimentations ultérieures, à toujours considérer la mesure de la MOI dans le contexte d'une cinétique. Elles indiquent aussi

qu'au ratio 1/1 pour VIT1 et VIT3 le système pourrait être saturé à 100% de co-infection détectée, ce qui ne permettrait que l'évaluation d'une valeur minimale de la MOI. L'emploi de ratios différents, par exemple 1/10 où la co-infection serait alors 10 fois moins probable, repousserait la limite de saturation du système et serait plus approprié à la mesure de valeurs de MOI élevées.

3.2 Estimation de la MOI

3.2.1 Fréquence des variants et taux de co-infection

Connaître la fréquence relative des deux variants VIT1 et VIT3 dans la feuille analysée est essentiel pour le calcul de la MOI, car ces valeurs servent à définir les probabilités (et donc les valeurs théoriques) de co-infection, sous l'hypothèse où les deux variants ont la même capacité intrinsèque à infecter les cellules.

Considérer les fréquences au sein de l'inoculum initialement utilisé pour infecter la plante ne constitue pas une évaluation fiable de ce paramètre, car même si la taille efficace des populations de CaMV est grande, et si la fréquence relative des variants varie très peu durant la colonisation des feuilles (*cf* Article 2), il pourrait exister un biais non négligeable durant la phase de l'inoculation. Afin de contourner ce problème, nous avons choisi de mesurer le ratio des deux variants par Q-PCR, directement dans chacune des feuilles analysées. Les données présentées dans l'Article 2 démontrent une très faible variation des fréquences de VIT1 et VIT3 dans des plantes co-inoculées ; et ceci sur une période de 32 jours durant laquelle la population virale est passée de feuilles n à des feuilles de n+5 à n+12. En moyenne, les fréquences relatives de VIT1 et VIT3 avaient varié de +10% et - 4%, respectivement, sur cette longue période. Il nous paraît donc particulièrement raisonnable de considérer que les fréquences varient très peu sur une courte durée (colonisation d'une seule feuille), et que le ratio que nous mesurons dans chaque feuille reflète suffisamment fidèlement les fréquences des deux variants dans la population colonisatrice, et informe donc correctement sur leur probabilité respective de pénétrer et d'infecter chaque cellule.

Un lot de 6 plantes a été inoculé avec des ratios VIT1/VIT3 volontairement déséquilibrés, dans le but d'obtenir des plantes couvrant une large gamme dans leur fréquences relatives, et ainsi une bonne précision dans l'estimation ultérieure de la MOI.

Vingt et un, 42, 60 et 84 jours après l'inoculation, la feuille n-4 ou n-5 à partir du centre de la rosette (*cf.* Matériels et Méthodes) a été collectée sur chaque plante pour mesurer à la fois la fréquence des variants et le taux de co-infection des cellules. Pour chacune des feuilles prélevées, la fréquence relative des variants VIT1 et VIT3 a été estimée à partir d'un pool de 3000 à 3500 cellules (*cf.* Matériels et Méthodes), et le taux de co-infection a été évalué sur un échantillon de 30 protoplastes isolés.

Les courbes représentant l'évolution du taux de co-infection au cours du temps sont figurées dans la figure 27, toutefois l'ensemble des valeurs obtenues à 60 dpi est aussi présenté dans le tableau 4. Ce tableau dévoile que les fréquences des variants (VIT1/VIT3 en %) varient de 91/9 à 5/95, avec une seule valeur intermédiaire à 44/56, ce qui couvre une large gamme de fréquences relatives et procure l'assurance d'une précision suffisante pour le calcul final de la MOI, avec un taux de co-infection VIT1/VIT3 non saturé. En effet, si les taux de co-infection mesurés sont globalement élevés, comme cela été prévisible à 60 dpi, ils restent souvent très inférieur à 100% (tableau 4). Le taux moyen de co-infection sur ces 6 plantes passe de 23 à 73% entre 21 et 84 jours après inoculation. Le taux de co-infection a augmenté tout au long de l'expérimentation sur l'ensemble des plantes, à l'exception de celui de la plante n°3 qui a chuté après 60 dpi (figure 27). Ce changement brutal et unique à ce temps de prélèvement, provient potentiellement de la mise à fleur qui se produit plus ou moins précocement (au delà de 60 dpi) chez les plantes infectées par le CaMV. En conséquence cette plante au comportement atypique ne sera pas utilisée par la suite pour le calcul de la MOI.

Plante N°	*Fréquence des variants VIT1 (%) / VIT3 (%)*	*Effectif des protoplastes : co-infectés / mono-infectés*	*Taux de co-infection*
1	44 / 56	23 / 27	85,2 %
2	5 / 95	20 / 28	71,4 %
3	91 / 9	25 / 30	83,3 %
4	83 / 17	19 / 28	67,9 %
5	70 / 30	20 / 29	69,0 %
6	79 / 21	16 / 26	61,5 %

Tableau 4 : Evaluation de la fréquence des variants et du taux de co-infection cellulaire dans une feuille à 60 dpi. Six plantes ont été volontairement inoculées avec des ratios variables et déséquilibrés de VIT1 et VIT3, de manière à affiner le calcul final de la MOI. Soixante jours après inoculation, une feuille (la cinquième à partir du centre de la rosette) a été prélevée sur chacune des plantes afin d'être protoplastée. A partir de la suspension de protoplastes fixés, la fréquence des variants (VIT1 et VIT3) dans la feuille analysée est déterminée par Q-PCR-SybrGreen. En parallèle, le contenu viral de protoplastes décontaminés et isolés est analysé par Nested-PCR-HRM pour obtenir le taux de co-infection. La cinétique du taux de co-infection pour chaque plante – obtenue à partir de ces valeurs à 60 dpi plus celles à 21, 42 et 84 dpi – est représenté sous forme graphique dans la figure ci-dessous.

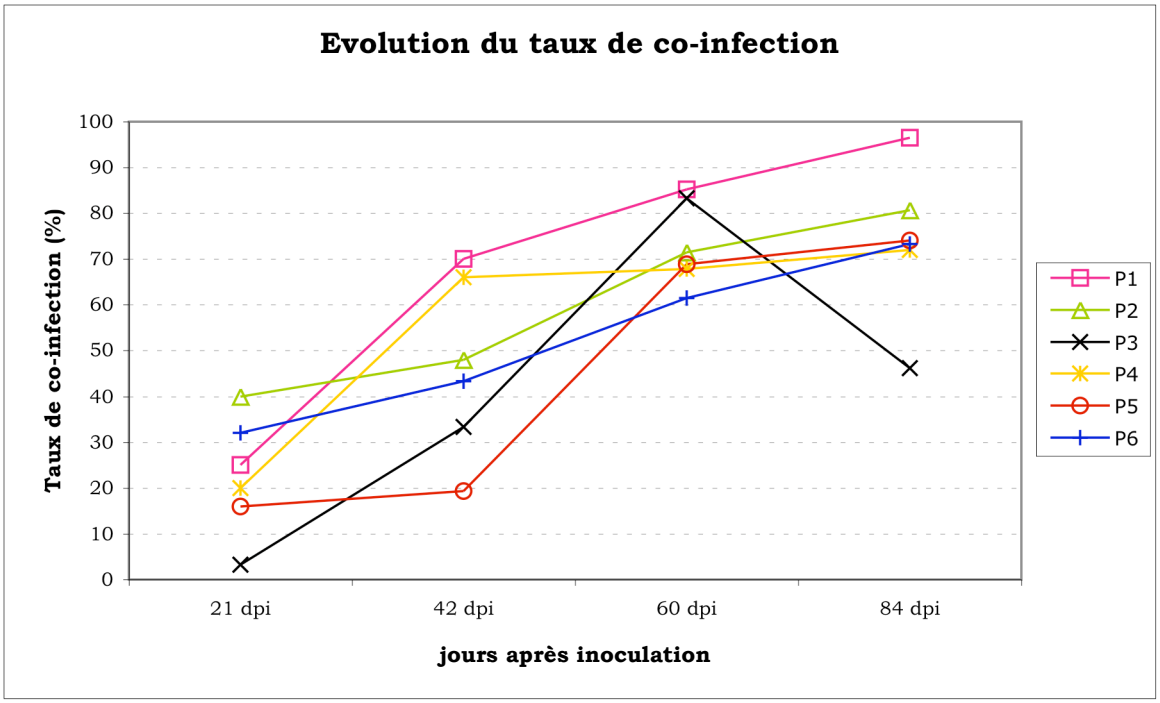


Figure 27: Evolution du taux de co-infection au cours du temps. Un lot de 6 plantes a été inoculé avec des ratios VIT1/VIT3 volontairement déséquilibrés (les valeurs mesurées à 60 dpi sont présentées dans le tableau 4), ce qui provoque logiquement des différences dans la fréquence de co-infection entre les différentes plantes. Le taux de co-infection a été analysé dans 4 feuilles par plante, prélevées à 4 temps successifs au fil du développement de l'infection systémique. Chaque courbe représente l'évolution du taux de co-infection pour une plante donnée. Le taux moyen de co-infection pour ces 6 plantes croît de 23 à 73 % entre 21 et 84 jours après inoculation.

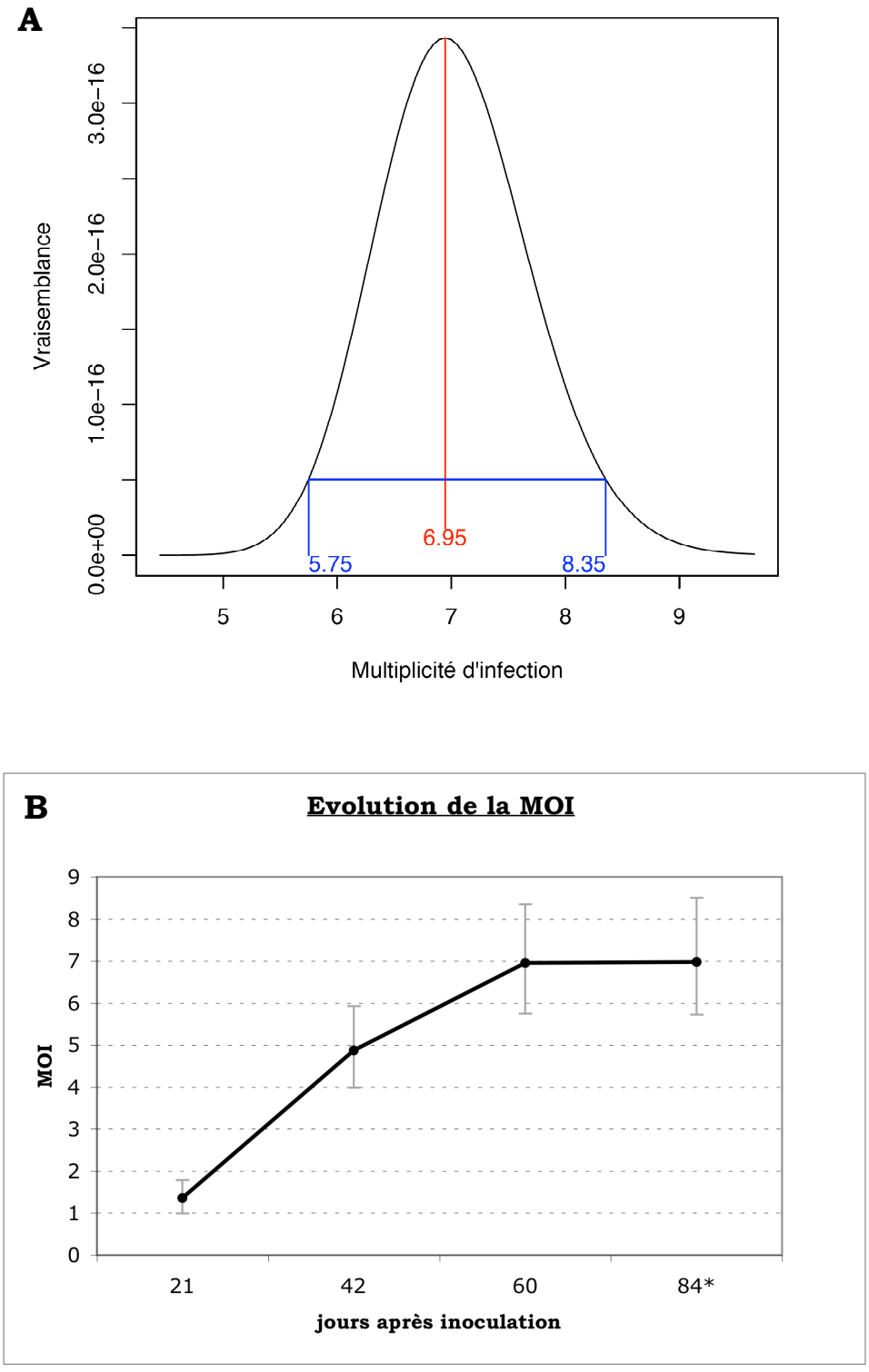


Figure 28: Représentation graphique de l'estimation de la MOI et de son évolution au cours du temps. (A) A partir des données recueillies sur 6 plantes à 60 dpi (voir tableau 4), la MOI est calculée par un programme statistique, régit par le logiciel « R », à l'aide de l'analyse par maximum de vraisemblance qui prédit une MOI moyenne proche de 7, avec un intervalle de confiance à 95% compris entre 5,75 et 8,35. Ce type de courbe a été établi de manière similaire pour tous les autres points de temps (21, 42 et 84 dpi). L'évolution de la MOI au cours du temps est figurée sur le graphique (B). Les barres verticales correspondent aux intervalles de confiance à 95%, donnés par le logiciel R comme montré en (A). * la MOI à 84 dpi est calculée sur 5 plantes, la plante n°3 ayant été retirée du fait de la chute atypique de son taux de co-infection.

3.2.2 Calcul de la MOI

Les données recueillies sur les 6 plantes co-infectées à différents ratios permettent d'estimer la MOI du CaMV par un calcul statistique, faisant intervenir le maximum de vraisemblance (voir Matériels et Méthodes).

Les valeurs obtenues pour la MOI sont de 1,36 à 21 dpi (avec un intervalle de confiance à 95%, $IC_{95\%}$ 0,99-1,79), 4,87 à 42 dpi ($IC_{95\%}$ 3,98-5,92), 6,95 à 60 dpi ($IC_{95\%}$ 5,75-8,35), et 6,98 à 84 dpi ($IC_{95\%}$ 5,72-8,5). Le suivi de ces plantes a clairement montré une augmentation de la MOI en fonction du temps, avec une phase de plateau dont le maximum semble atteint aux alentours de 60 dpi (figure 28).

4 Discussion

4.1 La MOI du CaMV *versus* autres virus d'hôtes pluricellulaires

Une MOI moyenne de l'ordre de 7 pour le CaMV est un résultat remarquable, puisque la valeur maximale disponible à ce jour a été déterminée à 4,3 chez les baculovirus (*cf.* Introduction, [180]). En ce qui concerne le HIV, bien que jusqu'à 8 provirus aient été observés dans le génome d'une cellule [181], la valeur moyenne de la MOI dans différents organes de l'hôte n'a jamais été établie, et ne peut donc pas être comparée. Chez les phytovirus, où plusieurs espèces de virus à ARN ne semblent pas ou peu capables de co-infecter les mêmes cellules d'une feuille [185,186,188], le cas du CaMV apparaît comme exceptionnel. Cependant, seules des études similaires sur un panel d'espèces virales très différentes infectant une large gamme d'hôtes végétaux permettra de conclure quant aux valeurs de MOI les plus fréquemment rencontrées, et si ces valeurs semblent similaires pour des virus phylogénétiquement proches.

4.2 Augmentation de la co-infection au cours du temps

Nos résultats mettent clairement en évidence un phénomène qui n'a jamais été étudié, ni même discuté auparavant ; il s'agit d'une augmentation de la MOI au fil du temps et du développement de l'infection systémique. Il est possible que les cellules soient surinfectées sans limite de temps (par une

circulation prolongée et libre des variants de cellule à cellule) et ainsi que la co-infection augmente avec la durée de l'expérience au sein d'une même feuille. Il est également possible que les feuilles soient assez massivement colonisées, dans un laps de temps relativement bref, et que les infections multiples cellulaires s'effectuent durant ce laps de temps seulement. Le taux de cellules co-infectées n'augmenterait alors pas sensiblement au sein d'une même feuille, mais plutôt d'une feuille n à une feuille n+1. Ce phénomène pourrait être lié à un accroissement de la charge virale dans les tissus vasculaires de la plante au fil du temps. L'expérience complémentaire décrite ci-dessous (trop préliminaire pour être mentionnée dans les résultats) visait à tester ces hypothèses afin de mieux comprendre la cinétique observée.

A partir d'une plante co-inoculée (VIT1/VIT3), deux prélèvements successifs ont été effectués sur une même feuille afin d'y suivre l'évolution du taux de co-infection. Sur les morceaux de limbe prélevés (en créneaux) de chaque côté de la feuille à 20 jours après inoculation, la méthode de Nested-PCR-HRM a montré que 52 % des protoplastes étaient co-infectés. Dix jours plus tard, avant que la feuille ne rentre en phase de sénescence, le limbe restant a été à son tour protoplasté et l'analyse a révélé un taux de co-infection très similaire de 56 % (tableau 5). Ces résultats suggèrent que, une fois l'infection établie de manière systémique, le taux de co-infection reste stable au sein d'une feuille, contrairement à ce qui est observé d'une feuille n à une feuille n+x. En outre, ces informations supplémentaires indiquent aussi que si des événements de surinfections contribuent effectivement au taux d'infection multiple, ils sont soit très rares, soit possibles durant un bref laps de temps (laps de temps qui se serait terminé avant le premier prélèvement sur cette feuille).

Ainsi, l'hypothèse la plus concevable pour soutenir l'accroissement de la MOI au cours du développement de l'infection systémique serait l'augmentation de la charge virale dans les tissus vasculaires de la plante, alimentant de plus en plus massivement et quasi-simultanément les cellules des feuilles néoformées. Or, de très récentes données montrent que l'accroissement de la charge virale du CaMV dans sa plante hôte est progressive jusqu'à atteindre une phase de plateau vers 35 dpi (communication personnelle [191]).

	20 dpi	30 dpi
Partie de feuille N°	*1*	*2*
Nombre de cellules positives	44/45	48/48
Nombre de co-infection	23	27
Taux de co-infection	52 %	56 %

Tableau 5 : Evolution du taux de co-infection au sein d'une même feuille au cours du temps. Suite à l'inoculation mécanique d'un plant de navet par deux variants du CaMV (ratio 1/1 VIT1 et VIT3), une feuille infectée de manière systémique a été choisie et prélevée de façon crénelée à 20 dpi (prélèvement de morceaux de feuilles en alternance de part et d'autre de la nervure centrale), puis analysée par Nested-PCR-HRM. A 30 dpi, le restant de la feuille a été à son tour analysé. La faible augmentation du taux de co-infection (+4%) après 10 jours suggère qu'il n'évolue pas significativement au sein d'une feuille et qu'il y serait donc fixé, très tôt après la colonisation.

D'après le détail des observations menées et de l'échantillonnage effectué, les feuilles récoltées à 60 dpi ont dû se former (et être colonisées) aux alentours de 40-45 dpi lorsque l'accumulation maximale du CaMV est déjà atteinte. Ceci expliquerait donc le décalage temporel entre la phase plateau de l'accumulation virale et celle plus tardivement observée de la MOI, ainsi que la légère inflexion constatée dans l'augmentation de la MOI entre 42 et 60 dpi puisque le plateau de la MOI pourrait être atteint dès 50-55 dpi.

Le fait que les surinfections ne semblent pas pouvoir se produire indéfiniment évoque la question de la protection croisée et un lien possible avec la MOI. En effet, la présence d'une infection au sein d'une cellule pouvant bloquer l'entrée et l'infection par un autre génome apparenté, le laps de temps durant lequel la surinfection reste possible pourrait déterminer le succès ou l'efficacité de la protection croisée. Bien que cette question spécifique soit intéressante, elle n'a pas été développée durant la thèse.

4.3 Inoculation des variants sur une même feuille ou sur feuilles séparées

Lors d'événements de mutation ou de recombinaison *in planta*, voire lors de l'inoculation de plusieurs particules virales par un même puceron (*cf.* chapitre suivant), chaque variant génétique ne se retrouve pas isolé dans une feuille mais co-existe dès le début (potentiellement dans la même cellule) avec d'autres variants plus ou moins apparentés. Ce sont ces arguments « biologiques » qui nous ont initialement confortés dans l'idée d'employer des solutions mixtes – contenant à la fois VIT1 et VIT3 – pour co-inoculer les feuilles d'une plante hôte. Ainsi nous observons le comportement de ces variants, en co-existence depuis les phases initiales de leur introduction (de leur « apparition ») dans la plante.

Toutefois, par souci d'exhaustivité, des inoculations séparées ont été réalisées sur des feuilles différentes de jeunes plants de navets (feuille 1 inoculée par VIT1 et feuille 2 inoculées VIT3, et inversement) . L'analyse par Nested-PCR-HRM des feuilles 5 à 8, prélevées 3 semaines après inoculation, révèle une répartition très hétérogène des variants (tableau 6). La

Plante N°	Feuille N°	Nombre de cellules		Mono-infection par		Co-infection
		Négatives	Infectées	VIT1	VIT3	(VIT1 + VIT3)
1	5	3	27	1	26	0 (0%)
	6	0	30	17	5	8 (27%)
	7	0	30	0	26	4 (13%)
	8	0	30	1	16	13 (43%)
2	5	0	30	26	0	4 (13%)
	6	1	29	17	5	7 (24%)
	7	7	23	21	0	2 (9%)
	8	1	29	23	0	6 (21%)

Tableau 6 : Influence de la séparation spatiale de l'inoculation sur la co-infection. La plante N°1 a été inoculée par VIT1 sur la première feuille, par VIT3 sur la suivante et inversement pour la plante N°2. A 21 jours après inoculation, les feuilles 5 à 8 de chaque plante ont été protoplastées et analysées par Nested-PCR. L'ensemble des feuilles présente un taux moyen de co-infection de 20%. Par contre, l'observation des feuilles individuelles révèle que le virus inoculé sur la feuille N°1 est peu disséminé dans les feuilles supérieures étudiées à l'exception de la feuille N°6. Le variant inoculé sur la feuille N°1 est rarement retrouvé en mono-infection, ce qui suggère que ce variant arrive dans les feuilles N°5, 7 et 8 après l'autre ou bien qu'il est en proportion bien plus faible.

distribution des deux variants dans ces 4 feuilles consécutives semble dépendante du choix des feuilles à l'inoculation. En effet, la feuille 1 du navet ne relargue apparemment que peu de virus vers les autres feuilles, exception faite de la feuille 6 à laquelle elle est directement liée via le réseau vasculaire (voir encadré 6). Ainsi la quasi totalité des feuilles 5, 7 et 8 comportent majoritairement le variant inoculé sur la feuille 2, infectant presque toutes les cellules, l'autre variant à plus faible occurrence ne pouvant donc que co-infecter.

Ce mode d'inoculation sur feuilles séparées permet effectivement d'observer des cellules co-infectées, mais le taux de co-infection est totalement biaisé par la distribution hétérogène des variants. Ces conditions expérimentales ne sont donc pas appropriées pour répondre à notre question, qui nécessite à l'évidence la présence des deux variants dans la même feuille. Le trajet des tissus vasculaires et les connexions entre feuilles étant une question complexe [83] (encadré 6), hors des objectifs de ce chapitre, nous avons exclu ce mode d'inoculation particulier.

4.4 Deux stratégies semblent s'être établies

Que ce soit directement par une détermination de la MOI, ou bien indirectement par une simple évaluation de la fréquence des événements de co-infection cellulaire, l'existence de deux situations bien distinctes selon les espèces virales semble se dessiner. En effet, il est possible d'anticiper soit (i) une très faible MOI (à peine supérieure à 1) dans les cas des cytomégalovirus et de plusieurs phytovirus à ARN, tels que les cucumovirus, les tobamovirus et les potyvirus, soit (ii) une forte MOI (≥3) comme dans les cas du phage ϕ6, des baculovirus, du HIV, et bien entendu du CaMV. L'origine ou les causes qui président à l'une ou l'autre de ces deux situations distinctes sont à ce stade totalement hypothétiques.

Il existe théoriquement des conséquences très importantes liées à la MOI, telles que la compétition entre génomes, la complémentation fonctionnelle, et le taux de recombinaison. Ces conséquences sont suffisamment importantes

pour penser que la MOI est un paramètre qui peut être contrôlé par les virus par le biais de mécanismes encore inconnus à ce jour.
Deux stratégies virales seraient alors envisageables, faible et forte MOI, chacune avec des avantages spécifiques discutés ci-dessous.

4.4.1 Faible MOI (≈1)

Cette stratégie se traduit par une majorité d'infections simples et peu de co-infection au niveau cellulaire. Cela tend à éliminer d'éventuelles compétitions entre variants [192], et offre ainsi la possibilité à chaque variant d'exploiter individuellement une cellule hôte à son rythme, sans risquer d'être éliminé par un compétiteur véloce. Par ailleurs, ce comportement d'exclusion de l'autre, permet d'éviter l'envahissement de la population par des particules virales défectives qui se répliquent généralement plus rapidement que le virus complet.
Une faible MOI peut aussi diminuer la taille efficace des populations, engendrant des goulots d'étranglement et une forte dérive génétique, à la fois durant l'infection de la plante et lors de la transmission par vecteur.

4.4.2 Forte MOI (≥3)

Cette stratégie se caractérise par un fort taux de co-infection cellulaire. Le fait que les variants puissent se retrouver aisément dans une même cellule favorise les événements de recombinaison et de *trans*-complémentation fonctionnelle. La compétition entre génomes peut aussi être augmentée, ce qui évoque un conflit possible entre coopération et compétition, et pose le problème de l'apparition de particules défectives interférentes.
Par ailleurs, une forte MOI permettrait : (i) de limiter les goulots d'étranglement lorsque l'infection de l'hôte passe par un cordon ou un amas de cellules restreint, ce qui pourrait augmenter globalement la taille efficace de la population virale durant l'infection de l'hôte (*cf. Ne* du CaMV au chapitre I), et (ii) de limiter les goulots d'étranglement intervenant potentiellement lors de la transmission par vecteur, en assurant un nombre élevé de virions disponibles dans chaque cellule échantillonnée par le vecteur (*cf.* chapitre suivant).

5 Perspectives

Les perspectives immédiates, découlant de ces résultats, sont nombreuses et brièvement évoquées ci dessous.

La durée durant laquelle la surinfection d'une cellule est possible pour le CaMV serait intéressante à évaluer ; cette information permettrait de mieux expliquer la cinétique de l'augmentation de la MOI au fil du temps. C'est à dire qu'il sera nécessaire de répéter l'étude de l'évolution du taux de co-infection et de la MOI au sein d'une seule feuille au cours du temps. Ainsi sur un lot de plantes co-infectées, en échantillonnant une partie de feuille (n°10 par exemple) à 21 dpi, puis en récoltant le restant de cette feuille, plus une feuille supérieure (par exemple n°20) à 42 dpi, nous pourrons quantifier la participation du niveau intra-feuille et celle du niveau inter-étage-foliaire à l'augmentation globale de la MOI. Si la MOI augmentait significativement aux deux niveaux, cela signifierait que la surinfection existe au sein des feuilles et nous envisagerions alors de réduire le délai entre les 2 prélèvements afin d'évaluer la durée durant laquelle cette surinfection peut avoir lieu.

En outre, la construction de clones viraux fluorescents du CaMV (ou la mise au point du FISH) permettrait une étude de la co-infection *in situ* afin de préciser la distribution de la MOI en fonction des différents tissus de la plante hôte. En effet, les suspensions cellulaires obtenues par la protoplastation sont constituées d'un mélange de cellules originaires de différents tissus, bien que la grande majorité proviennent du mésophylle.

Finalement, le dernier point qui nous intéresse, en aval de cette première estimation de la MOI chez le CaMV, serait de vérifier si tous les virus qui pénètrent une cellule s'y répliquent effectivement. En d'autres termes, est-ce que l'entrée de particules virales « inertes » provenant des cellules adjacentes (qui ne se répliqueraient pas, ne recombineraient pas, ne complémenteraient pas, et ne seraient disponibles que pour la transmission par vecteur) est un phénomène quantitativement important ? Si oui, notre méthode d'étude l'a probablement occulté puisque nous n'avons détecté que les variants fréquents dans chaque cellule (>1%), forcément issus de cycle(s) de

réplication. Malheureusement à ce jour, il n'existe pas de moyens susceptibles d'infirmer ou de valider ce phénomène.

Au delà de ces questions à court terme, plusieurs grands axes d'études peuvent être envisagés, dont certains ont été initiés et seront développés par Serafin Gutiérrez (Post-doc INRA 2008-2009) :

i) évaluer la MOI chez d'autres familles ou genres de virus de plantes (ex : *potyviridae*, bromoviridae, *tobamovirus, geminiviridae, nanoviridae*) afin de conforter l'hypothèse de l'existence de deux situations bien distinctes au sein des phytovirus (faible ou forte MOI)

ii) au moins pour deux espèces virales (une dans chaque situation), évaluer la MOI dans différents hôtes dans le but de montrer que la MOI est une propriété de l'espèce virale et non pas liée à la nature de l'hôte

iii) rechercher quels sont les mécanismes qui pourraient être à l'origine d'une régulation de la MOI ; soit, en d'autres termes, par quels moyens certains virus s'offrent-ils l'opportunité de rentrer à plusieurs dans les cellules de l'hôte, ou bien « ferment-t-ils la porte » pour se retrouver tout seul ?

La régulation de l'infection multiple pourrait s'effectuer au niveau des plasmodesmes, soit par une réduction de la taille d'exclusion limite empêchant l'entrée d'unités infectieuses supplémentaires, soit par un transport unidirectionnel préférentiel limitant ainsi les mélanges sauf à la frontière entre deux fronts de migration. Toutefois, il est aussi envisageable que la régulation puisse se faire à l'intérieur de la cellule – après le passage de la membrane plasmique – par l'intermédiaire de mécanisme tel que le « gene silencing » par exemple.

CHAPITRE III

Goulots d'étranglement imposés aux populations de CaMV lors de la transmission d'hôte à hôte par vecteur

Avant-propos du Chapitre III

Ce travail a été réalisé avec le puceron *Brevicoryne brassicae* car (i) c'est un parasite naturel du navet, (ii) c'est un des principaux vecteurs naturel du CaMV, et (iii) les individus adultes sont plus grands que ceux de *Myzus persicae,* ce qui est plus commode quand on réalise ses premiers « branchements » EPG.

Ce chapitre est construit sur la base structurale de l'article qui sera prochainement rédigé, c'est pourquoi certains travaux préliminaires réalisés en tout début de thèse n'y sont pas inclus et ont été placés en annexe 1.

La taille efficace des populations de CaMV est relativement large au cours de la colonisation successive des feuilles systémiques de leur hôte (voir au chapitre I, l'Article 2), et la multiplicité d'infection cellulaire est très élevée (voir chapitre II). Cependant, ces informations n'excluent en rien la possibilité d'un goulot d'étranglement génétique sévère, ailleurs au cours du cycle de vie du CaMV ; elles en facilitent même l'analyse.

1 Introduction

Il a été démontré expérimentalement pour des virus de bactéries [165,166], d'animaux [167-169] et plus récemment de plante [160], que les goulots d'étranglement génétiques répétés induisent une baisse significative de la valeur sélective moyenne (et/ou de la virulence) des populations virales, vraisemblablement liée au cliquet de Müller (*cf.* Introduction générale).

La transmission constitue une phase problématique majeure du cycle viral, car le virus endocellulaire doit quitter son hôte et passer par le milieu extérieur qui lui est défavorable, tout en restant infectieux afin de se maintenir dans l'environnement. Ces contraintes abiotiques, couplées au processus même de la transmission (contact mécanique, par vecteur…), vont non seulement potentiellement produire des goulots d'étranglement démographiques, pouvant influencer l'évolution de la population virale transmise [192-194], mais aussi imposer des contraintes sélectives, dans les

cas de transmissions par vecteurs qui nécessitent le maintien d'interactions moléculaires spécifiques [195]. Sans conteste, cette étape du cycle viral a été pressentie comme particulièrement critique, en termes démographique et génétique, pour les populations de pathogènes, par de très nombreux auteurs (dont [196]).

La quasi totalité des virus de plantes sont transmis par vecteur d'une plante hôte à une autre. Les aphides (pucerons), de l'ordre anciennement nommé des homoptères, constituent la famille la plus répandue et la plus vaste des vecteurs de virus de plantes. Les pucerons transmettent au moins 275 phytovirus, dont 75% appartiennent à la catégorie des virus non-circulants [122]. Cette catégorie (*cf.* Introduction générale) comporte des virus transmis selon une stratégie capside, où les virions sont retenus par une interaction directe entre leur protéine de capside et un récepteur hypothétique au niveau des stylets du vecteur, et des virus transmis selon une stratégie helper, pour laquelle l'interaction avec le vecteur est réalisée par une protéine non-structurale du virus (helper component, HC) qui constitue un « pont moléculaire » entre les deux.
Le mode de transmission non-circulant se caractérise par des périodes d'acquisition, de rétention et d'inoculation du virus qui sont très brèves (quelques minutes le plus souvent). Puisque, dans ce cas précis, le vecteur n'échantillonne que très succinctement et ponctuellement l'hôte infecté, il a toujours été intuitivement suspecté que ce mode de transmission devait imposer un goulot d'étranglement démographique très sévère aux populations virales. Pour cette raison, depuis plusieurs décennies déjà, des chercheurs ont tenté d'estimer le nombre de particules virales prélevées par les vecteurs de virus non-circulants.

Dès 1972, Walker et Pirone [197] ont mesuré qu'un minimum de 1000 particules virales du *Cucumber mosaic virus* (CMV ; stratégie capside) devaient être ingérées par le puceron, afin qu'il puisse ensuite infecter efficacement une plante sensible. En 1988, Pirone et Thornbury [198] ont montré que 80 % des pucerons infectants avaient initialement acquis moins de 500 virions du *Tobacco vein mottling virus* (TVMV), ou du *Tobacco etch*

virus (TEV), deux potyvirus ayant adopté la stratégie helper. Les mêmes auteurs ont aussi réussi à transmettre une fois le TVMV, avec un puceron n'ayant acquis que 15 particules virales. D'après ces résultats, il semble clair qu'un goulot d'étranglement existe réellement au niveau de la transmission non-circulante par vecteur, mais l'évaluation de sa taille reste toutefois largement imprécise. En effet, dans ces deux études, basées sur la mesure de radioactivité des pucerons nourris sur une solution contenant du virus radiomarqué, le nombre de virions estimé englobe toutes les particules virales ingérées par le puceron. Le nombre de particules à l'origine d'une infection suite à l'inoculation par un puceron devait être assez largement surestimé puisque:

- seule une partie des virions ingérés est retenue dans les stylets du puceron, le reste est digéré et perdu
- seule une fraction des virions retenus est « décrochée » et inoculée dans la nouvelle plante
- parmi ces dernières particules introduites dans les plantes, seule une fraction indéterminée participe effectivement à la nouvelle infection.

En 2006, une étude expérimentale a permis d'affirmer – par la perte de diversité génétique observée dans une population artificielle de CMV (12 mutants), suite à la transmission par deux vecteurs différents (*Aphis gossypii* et *Myzus persicae*) – que l'étape de vection d'hôte à hôte engendre un goulot d'étranglement génétique [199]. Toutefois, si ce goulot d'étranglement peut survenir à cause d'un très faible nombre de virions transmis, comme cela a été conclu par les auteurs, il est aussi possible qu'il résulte d'une répartition hétérogène – en « patch » – de chaque variant dans la plante source, l'ensemble n'étant donc pas pareillement accessible au puceron. Cette question n'a pas été tranchée, ni même clairement évoquée.
En 2007, une étude plus élaborée réalisée sur le *Potato virus Y* (PVY, potyvirus) a enfin confirmé la présence d'un goulot d'étranglement démographique (lié au nombre de virions transmis) lors de la transmission par *Myzus persicae*, grâce à la compétition *in vitro* entre un variant virulent et un autre avirulent pour la rétention au niveau des stylets [200]. Ces

travaux ont permis d'extrapoler qu'en moyenne un nombre extrêmement restreint de particules virales (0,5 à 3,2) sont transmises efficacement par un puceron-vecteur.

Afin d'examiner directement ce qui se passe *in vivo* lors de la transmission du CaMV par vecteur, nous avons utilisé des plantes, co-infectées par 6 clones viraux identifiables, comme source d'acquisition pour des pucerons sous contrôle EPG. Les acquisitions ont été menées selon 3 modalités différentes : 1 seule ponction intracellulaire (1 pd) par le puceron, 5 ponctions intracellulaire (5 pds), et 15 minutes d'ingestion de sève élaborée (phloème). Ce protocole particulier était destiné à répondre à la même question que celle posée par les études mentionnées ci-dessus sur le CMV et le PVY, tout en allant un peu plus loin et intégrer l'impact du comportement alimentaire du vecteur sur la taille du goulot d'étranglement engendré. En effet, il a antérieurement été démontré que le taux de transmission maximal du CaMV ne semble atteint que lorsque l'acquisition a eu lieu au sein du phloème [132].

2 Matériels et Méthodes

2.1 Plantes sources, pucerons et plantes test

Six variants transmissibles du CaMV (clones CaMV VIT1 à VIT6) ont été utilisés pour préparer indépendamment mais similairement deux inoculums mixtes (« MixVIT3 » et « Mix6VIT » ; voir Article 1, [190]). MixVIT3 et Mix6VIT ont chacun servi à l'inoculation de 2 plantes saines (stade 3 feuilles), à 15 jours d'intervalle ; plantes qui ont ensuite été utilisées de 14 à 33 dpi comme plantes sources pour les expérimentations d'acquisition par pucerons. L'ensemble des acquisitions sous contrôle EPG (voir ci-dessous) ayant duré 1 mois, les feuilles sources FS1 et FS2 provenant des plantes inoculées par MixVIT3 ont d'abord été utilisées durant deux semaines, jusqu'à apercevoir les premiers signes de sénescence, puis FS3 et FS4 provenant des plantes inoculées par Mix6VIT ont alors pris le relais. Toutes les plantes infectées ont été maintenues dans une chambre de culture

« insect-proof» à 26/20°C (jour/nuit), avec une photopériode de 16/8h (jour/nuit).

Le clone de *B. brassicae* est entretenu dans des conditions contrôlées [(22/18°C et 14/10h jour/nuit)] sur *Brassica rapa* cv. "Just Right". Ce clone a été isolé à partir d'une femelle virginipare, récoltée à Villa del Prado (Madrid, Espagne) par A. Fereres.

Des jeunes plants de *Brassica rapa* cv. « Just Right » au stade 2 feuilles servent de plantes tests. Après l'acquisition sous contrôle EPG (*cf.* ci-dessous), chaque puceron est déposé individuellement sur une unique plante test. Les plantes tests sont traitées au Confidor (Bayer Hispania Industria) après une période d'inoculation de 18-24h, puis transférées dans une serre insecte-proof. Toutes les plantes infectées par les pucerons sont récoltées 5 semaines après inoculation afin d'en purifier l'ADN viral (voir l'Article 1 pour la purification) [190].

2.2 Electro-Penetro-Graphie (EPG)

Les transmissions avec acquisition sous contrôle EPG ont été effectuées durant un séjour de 1 mois dans le laboratoire du Pr. Alberto Fereres (spécialiste européen de cette technique), financé par un projet bilatéral « Picasso ». Les plantes tests présentant des symptômes caractéristiques du CaMV après 5 semaines ont été collectées individuellement par Aranzazu Moreno à Madrid, congelées, puis envoyées sous carbo-glace à l'UMR BGPI.

2.2.1 Dispositif EPG (branchements)

Après avoir mis à jeûner quelques pucerons issus de l'élevage, la première étape avant l'EPG consiste à fixer, à l'aide d'une glue d'argent colloïde (Pelco Colloidal Silver no.16034, Ted Pella Inc., Reeding, CA, USA), un fil en or (3 cm de long, 20 µm de diamètre) sur le haut du thorax des pucerons (adultes aptères), préalablement immobilisés par aspiration sur un support à claires-voies, sous une loupe binoculaire. Ensuite, l'extrémité de ce fil d'or est collé avec cette même glue à un fil de cuivre (3 cm de long, 1 mm de diamètre), afin de permettre une connexion à une électrode de l'appareil de contrôle EPG. Une seconde électrode en cuivre (10 cm de long) est placée dans le pot, enfouie dans le substrat humide de la plante.

Les enregistrements d'EPG sont effectués à 100 Hz à l'aide d'un amplificateur à 4 canaux (Giga-99 DC-amplifier). Le signal analogique entrant est automatiquement converti en un signal numérique par l'appareil, ce qui permet l'enregistrement du graphe EPG sur le disque dur d'un PC, tout en visualisant sur l'écran les oscillations en temps réel. L'enregistrement et la visualisation directe sont effectués grâce au logiciel Stylet 3.0 (sur PC), tandis que la vérification ultérieure des graphes est réalisée (sur Macintosh) avec MacStylet v2.0 b10 [201] (encadré 8).

2.2.2 Modalités des périodes d'acquisition du CaMV

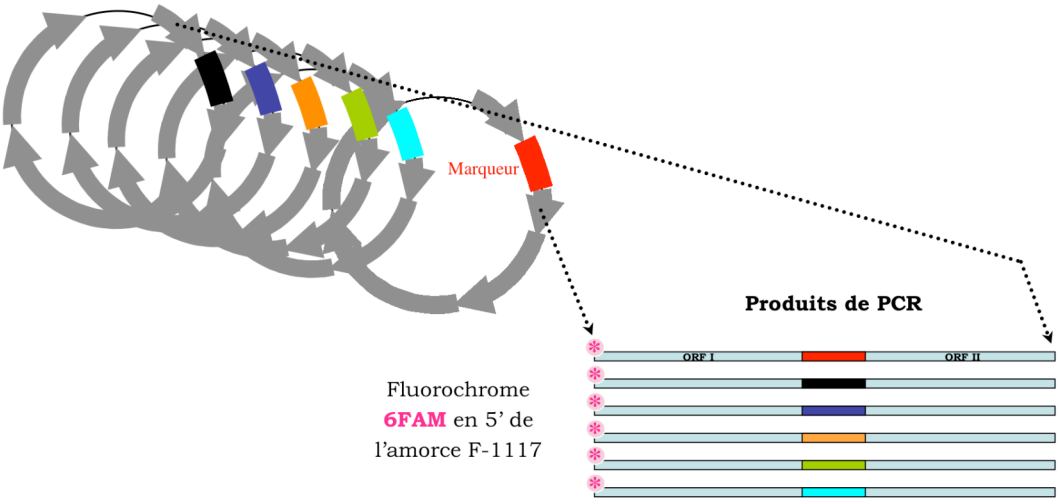
Afin d'étudier les possibles goulots d'étranglement, lors de la transmission et selon le comportement alimentaire du vecteur durant la phase d'acquisition, des pucerons ayant jeûné sont connectés à l'appareil d'EPG puis placés sur la feuille source. La phase d'acquisition est interrompue artificiellement en séparant délicatement le puceron de la feuille source (à l'aide d'un pinceau), au moment décidé par l'expérimentateur. Trois modalités d'acquisition ont été retenues, et leur identification est détaillée dans l'encadré 8 : (a) le puceron n'a effectué qu'une seule ponction intracellulaire (1 pd, potential drop) lors de la première piqûre ; (b) le puceron a (au moins) produit 5 pds ; (c) le puceron a effectué un nombre variable de ponctions intracellulaires, atteint le phloème et ingéré de la sève élaborée durant au moins 15 minutes (phase E2). Les pucerons n'ayant pas atteint le phloème après 4 heures d'expérience ont été éliminés.

2.3 Quantification des clones VIT dans les feuilles sources

La fréquence des différents clones VIT au sein de chaque feuille source a été déterminée par Quantitative Single-letter Sequencing (QSS), à partir de l'ADN viral purifié (technique QSS décrite dans l'Article 1).

2.4 Détection des clones VIT dans les plantes tests

Les 6 marqueurs VIT insérés entre les ORFs I et II du génome viral contiennent tous un site de restriction enzymatique, BglI, à des positions décalées. Ainsi, suite à la digestion enzymatique de produits PCR encadrant la région du marqueur, il est possible de déterminer la présence/absence de



Marqueur 1	CACGGCTAACAAGGTCCT**GCC**CCTTA**GGC**AATCAGT
Marqueur 2	ACGCAGCTTTCACACAATA**GCC**AATTT**GGC**TTTTGC
Marqueur 3	GCG**GCC**CGACT**GGC**ACTCAGCTGTTGGGTTCTGGTC
Marqueur 4	CGTTAAGGC**GCC**GCTAC**GGC**ACGAGTATAATCTTAT
Marqueur 5	CAGAGGCTCGGCGGCGACGTA**GCC**GGTTC**GGC**TACT
Marqueur 6	CGCGGGTTGGACCGG**GCC**TAAGT**GGC**CCAGAGATCC

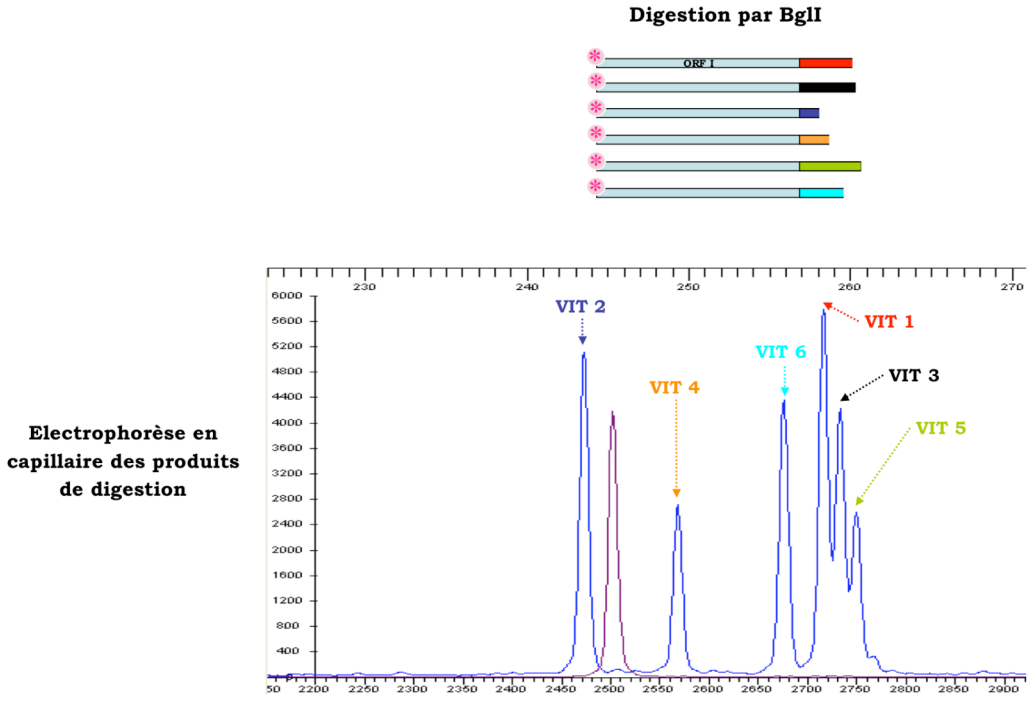


Figure 29: Détection des clones VIT dans les plantes tests par PCR-RFLP-CE. A partir d'ADN viral purifié (population de virus), la région contenant les marqueurs VIT (entre les ORFs I et II du CaMV) est amplifiée par PCR à l'aide de deux d'amorces dont l'une (amorce sens) est couplée à un fluorochrome (6FAM). Les produits PCR sont purifiés par phénol-chloroforme avant d'être digérés par une enzyme de restriction, BglI, dont le site de coupure GCCnnnnnGGC est présent à une position décalée dans chacun des 6 marqueurs VIT. Ainsi, les produits fluorescents issus de la digestion par BglI présentent tous des longueurs différentes. Suite à une nouvelle purification par phénol-chloroforme, les ADN viraux amplifiés et digérés sont dosés et soumis à une électrophorèse en capillaire. Sur l'électrophorégramme, les 6 pics bleus correspondent respectivement (de gauche à droite) aux VIT 2, 4, 6, 1, 3 et 5, tandis que le pic violet correspond au marqueur de taille (245 pb). L'analyse globale de la position des pics révèle une variation moyenne de ± 0,07pb, ce qui permet une discrimination sûre de chaque VIT au sein des populations d'ADN viral.

chacun des clones VIT en fonction du profil de restriction observé (pour ce produit de PCR) en électrophorèse capillaire. Cette technique de détection est ci-après dénommée PCR-RFLP-CE (PCR Restriction Fragment Length Polymorphism on Capillary Electrophoresis), et a été développé en collaboration avec a société Cogenics (ex Genome Express).

L'ADN viral extrait des plantes tests infectées (comme décrit précédemment : Article 1) est amplifié par PCR à l'aide des amorces (externes) suivantes : 6FAM-1117 5'-GGAAATGTCCAGCAATCTGAG-3' et Ris7 5'-GTTGGGTACCTAAGGCTTCTAATATCTC-3'. Les produits de PCR sont ensuite purifiés par phénol-chloroforme, digérés par BglI, puis re-purifiés et dosés afin de réaliser l'électrophorèse en capillaire ; cette dernière étape étant accomplie par Cogenics selon des conditions permettant de discriminer des fragments dont la longueur diffère d'un seul nucléotide. L'amorce 6FAM-1117 étant couplée en 5' à un fluorochrome (6FAM), les fragments 5' des produits de PCR, libérés par BglI, sont détectables et identifiables en sortie d'électrophorèse capillaire, comme illustré sur la figure 29. Ils signent la présence/absence du marqueur correspondant dans la population.

Les électrophorégrammes sont analysés sous le logiciel STRand (téléchargeable sur http://www.vgl.ucdavis.edu/informatics/strand.php), afin de visualiser les fragments d'ADN et d'identifier leur taille.

3 Résultats

3.1 Taux de transmission selon la modalité d'acquisition

Le CaMV étant plus efficacement transmis à partir du phloème, nous avons effectué une plus grande quantité de tests pour les modalités 1 et 5 ponction(s) intracellulaire(s), afin d'obtenir un nombre suffisant de plantes infectées pour les besoins de l'analyse. Après 5 semaines, les plantes tests infectées sont dénombrées, ce qui nous permet de calculer le taux de transmission pour chaque modalité d'acquisition (tableau 7). Les résultats obtenus confirment que la modalité d'acquisition la plus efficace, pour le succès ultérieur de la transmission, est effectivement l'accès au phloème avec plus de 80% de réussite. Toutefois, 5 pds suffisent pour atteindre environ 72% de transmission (non différent de l'acquisition au phloème, par

une analyse de χ^2 ; P=0,1412 avec 1 ddl), tandis qu'une seule ponction intracellulaire conduit seulement à un succès de transmission de 30 % (significatif P<0,05).

Cette efficacité de transmission, différente suivant le comportement alimentaire du puceron, était particulièrement intéressante pour notre projet car il devenait alors possible de tester l'existence d'une corrélation entre le taux de transmission et la taille du goulot d'étranglement produit par le vecteur.

3.2 Fréquence de transmission de chaque clone VIT

La composition de la population virale de chaque feuille source (FS1 à FS4) a été estimée par QSS (tableau 8). Les fréquences relatives obtenues pour chacun des clones VIT sont proches entre FS1 et FS2 et entre FS3 et FS4, et un peu plus éloignées entre ces deux couples de feuilles. Ce résultat est cohérent avec le fait que FS1-2 et FS3-4 proviennent de plantes inoculées avec MixVIT3 et Mix6VIT, respectivement (voir Matériels et Méthodes).

Les fréquences relatives « initiales » de chacun des clones CaMV VIT1-6 (au sein des FS) sont globalement similaires à leur fréquence de transmission observée sur un ensemble de 175 plantes tests, analysées par PCR-RFLP-CE (comparer la moyenne des fréquences dans les FS à la moyenne des fréquences de transmission, toutes modalités confondues, dans le tableau 8). En d'autres termes, sur l'ensemble du jeu de données, il semble que les chances de transmission de chacun des clones VIT soient très intimement corrélées à leur fréquence relative dans la population source. Ce résultat signifie que les différents marqueurs VIT1-6 semblent neutres, ou proches de la neutralité, dans les conditions expérimentales décrites. La nuance dans cette conclusion provient des légers écarts observés pour VIT2, et peut être VIT3, qui pourraient être significatifs et indiquer un faible coefficient de sélection associé à ces clones au moment de la transmission. De plus, en analysant les données du tableau 8 dans le détail, d'autres écarts mineurs sont observés pour certaines conditions: (i) une transmission accrue de VIT5 au dépend de VIT1 suite aux acquisitions avec 1 pd, (ii) une faible transmission de VIT1 suite aux acquisitions avec 5 pds. Ces variations

Modalité d'acquisition	Nombre total de plantes tests	Nombre de plantes infectées	Taux de transmission (%)
1 pd	245	74	30,2
5 pds	178	128	71,9
Phloème	73	59	80,8

Tableau 7 : Taux de transmission en fonction de la modalité d'acquisition. De nombreuses plantes tests ont été utilisées pour les modalités 1 et 5 ponction(s) intracellulaire(s), dans le but d'obtenir assez de plantes infectées, prévoyant que le taux de transmission dans ces conditions expérimentales seraient inférieur à celui obtenu après 15 minutes d'acquisition dans le phloème. Les taux de transmission mesurés après les acquisitions 5 pds (71,9%) et phloème (80,8%) ne sont pas significativement différents (χ^2 ; P=0,1412 pour 1 ddl).

Feuille Source (FS) n°	Fréquence des marqueurs VIT n°					
	1	**2**	**3**	**4**	**5**	**6**
1	22,3	8	38,8	16,9	11,3	2,7
2	18,8	12,7	47,7	9,9	10	0,9
3	29,3	7,6	42,7	12,2	4,6	3,6
4	29,1	7,4	38,6	16,7	4,6	3,6
Moyenne 1-4	**24,9**	**8,9**	**42,0**	**13,9**	**7,6**	**2,7**

Modalité	Acquisition sur FS n°	Fréquence de transmission des marqueurs VIT n°						Nombre de plantes tests
toutes (1 pd, 5 pds et phloème)	1	24,5	*14,5*	*25,5*	18,2	10,0	*7,3*	42
	2	18,2	*20,0*	40,0	9,1	11,8	0,9	50
	3	24,5	11,7	37,2	9,6	7,4	*9,6*	42
	4	22,0	*16,5*	37,4	9,9	*11,0*	3,3	41
	Moyenne FS 1-4	**22,2**	**15,8**	**34,8**	**11,9**	**10,1**	**5,2**	**175**
1 pd	1	16,0	*20,0*	20,0	16,0	20,0	8,0	13
	2	20,6	17,6	41,2	11,8	8,8	0,0	17
	3	20,0	10,0	33,3	6,7	*13,3*	*16,7*	15
	4	15,8	10,5	36,8	13,2	*21,1*	2,6	16
	Moyenne FS 1-4	**18,1**	**14,2**	**33,9**	**11,8**	**15,7**	**6,3**	**61**
5pds+Phl	**Moyenne FS 1-4**	**24,1**	**16,5**	**35,2**	**11,9**	**7,6**	**4,7**	**114**
5 pds	1	29,7	10,8	29,7	18,9	5,4	5,4	13
	2	18,4	21,1	42,1	5,3	10,5	2,6	18
	3	27,3	15,2	36,4	9,1	9,1	3,0	14
	4	25,9	*25,9*	40,7	*3,7*	3,7	0,0	16
	Moyenne FS 1-4	**25,2**	**17,8**	**37,0**	**9,6**	**7,4**	**3,0**	**61**
phloème (Phl)	1	25,0	14,6	25,0	18,8	8,3	*8,3*	16
	2	15,8	21,1	36,8	10,5	15,8	0,0	15
	3	25,8	9,7	41,9	12,9	0,0	9,7	13
	4	26,9	15,4	34,6	11,5	3,8	7,7	9
	Moyenne FS 1-4	**23,1**	**15,4**	**33,6**	**14,0**	**7,7**	**6,3**	**53**

Tableau 8 : Fréquence de transmission de chacun des marqueurs VIT. Les feuilles sources (FS) FS1-2 et FS3-4 proviennent respectivement de 2 lots de 2 plantes inoculées avec 2 inoculums mixtes indépendants, MixVIT3 et Mix6VIT (*c.f.* Matériels et Méthodes de ce chapitre). La fréquence relative de chacun des clones CaMV VIT1-6 au sein des quatre FS (tableau sur fond grisé) a été analysée par QSS (technique décrite dans l'article 1). Après transmission par pucerons sous contrôle EPG à partir des ces FS, 175 plantes tests infectées ont été analysées par PCR-RFLP-CE, et la fréquence de détection de chacun des clones VIT1-6 est notée en relation avec les modalités d'acquisition autorisées aux pucerons. Les fréquences indiquées en gris et italique sont significativement différentes de celles de la FS correspondante (χ^2 ou un test exact de Fisher lorsque le nombre de plante est inférieur à 5 ; P<0,05).

pourraient ne pas être significatives et une analyse statistique formelle et spécifiquement adaptée à notre jeu de données reste à développer.
En tout état de cause, ces données fournissent une appréciation indirecte pour suggérer que les clones viraux sont répartis de manière assez homogène au sein de la feuille source. Dans le cas contraire, en effet, la corrélation entre fréquences dans les FS et fréquences de transmission n'aurait pas été aussi évidente. Le fait que les fréquences moyennes de transmission de chaque VIT, avec acquisition à 1 pd, ne diffèrent pas énormément de celles obtenues avec acquisition dans le phloème est un élément en faveur de cette hypothèse.
Vu les résultats présentés au chapitre I (distribution homogène entre feuilles) et au chapitre II (MOI naturelle très élevée) ces observations étaient attendues et sont donc parfaitement cohérentes.

3.3 Diversité transmise selon la modalité d'acquisition

Le génotypage (PCR-RFLP-CE) a permis de dénombrer le nombre de clones VIT présents dans chacune des plantes tests infectées. Parmi le total des 175 plantes analysées, 145 contenaient jusqu'à 3 clones VIT, seulement 30 renfermaient plus de 4 clones VIT et aucune ne réunissaient l'ensemble des 6 clones VIT (tableau 9). En moyenne, de 2,1 à 2,7 clones VIT sont détectés dans les plantes tests, issues respectivement d'acquisitions à 1 pd et au phloème. Ces résultats soutiennent fermement l'existence, pour les populations virales, de goulots d'étranglement lors de leur transmission d'hôte à hôte par vecteur. Si l'on considère le fait que la répartition des variants génétiques du CaMV est homogène au sein de la plante, il est possible de conclure que ce goulot d'étranglement est clairement démographique (c'est le nombre de variants transmis qui est faible, et non pas un grand nombre de variants similaires).
Afin de se faire une meilleure idée du nombre de particules virales pouvant être transmises suite à une modalité d'acquisition spécifique, nous avons rapporté le nombre de marqueurs VIT transmis au nombre de plantes analysées pour chacune de ces modalités (figure 30). Ainsi, il est possible d'évaluer la fréquence des plantes tests contenant X marqueurs VIT suite à

Modalité d'acquisition	Nombre de clones VIT détectés dans les plantes tests							Nombre de plantes par modalité
	1	2	3	4	5	6	Moyenne	
1 pd	23	18	14	4	2	0	2,08	61
5 pds	20	19	13	7	2	0	2,21	61
Phloème	9	17	12	11	4	0	2,70	53
TOTAL	52	54	39	22	8	0	2,31	175

Tableau 9 : Nombre de marqueurs VIT transmis en fonction de la modalité d'acquisition. L'analyse par PCR-RFLP-CE a permis de déterminer la présence/absence de chaque clone VIT au sein des 175 plantes tests infectées. Le nombre de plantes tests contenant 1, 2, 3, 4, 5 ou 6 marqueurs est présenté en fonction de la modalité d'acquisition : (i) une seule ponction intracellulaire, 1 pd ; (ii) 5 ponctions intracellulaires, 5 pds ; (iii) 15 minutes d'ingestion de sève élaborée au sein du phloème.

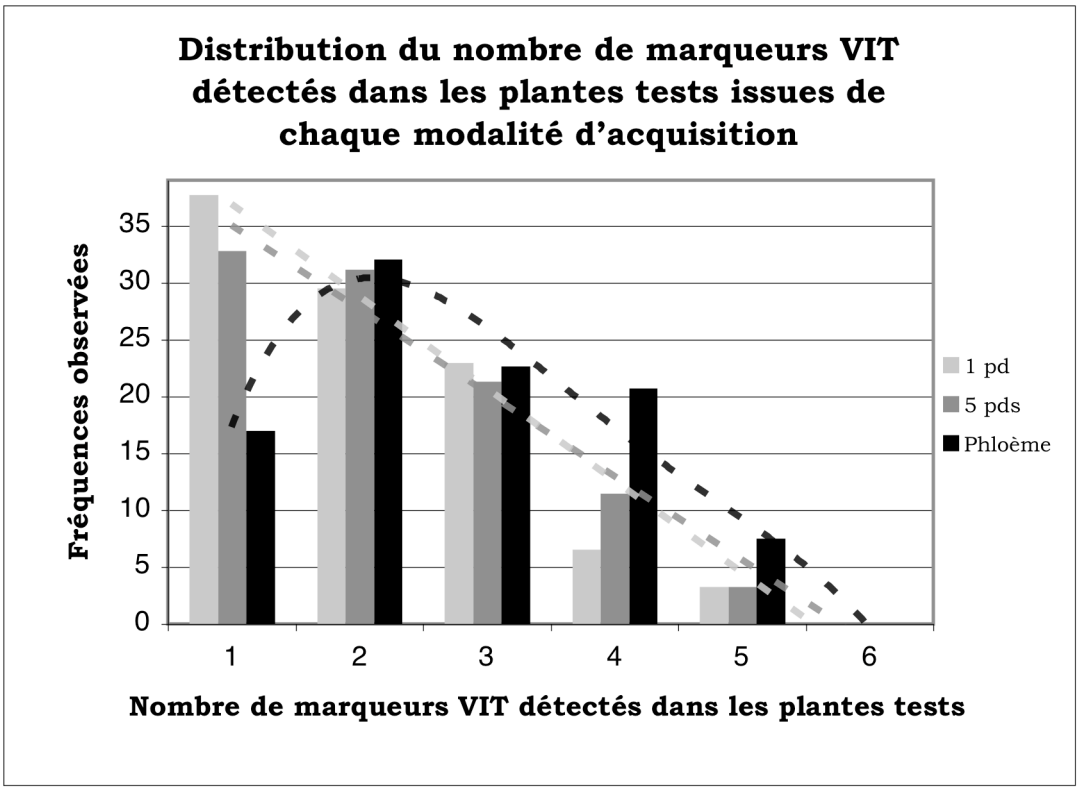


Figure 30: Distribution du nombre de marqueurs VIT détectés dans les plantes tests issues de chaque modalité d'acquisition. Le nombre de plantes analysées contenant « n » marqueurs VIT est rapporté au nombre total de plantes étudiées pour chaque modalité ; ceci dans l'optique de pouvoir comparer les modalités ayant des effectifs différents. Ainsi, lors de transmissions suite à des acquisitions avec 1 et 5 ponctions cellulaires, 33 à 38 % des plantes tests ne contiennent qu'un seul marqueur, et les cas de co-transmission de 2, 3, 4, voire 5 marqueurs diminuent régulièrement en fréquence (voir les droites discontinues grises). La tendance semble différente suite à l'acquisition au sein du phloème (voir la courbe discontinue noire), où le cas le plus fréquent (32%) semble correspondre à la co-transmission de 2 marqueurs, 1, 3 et 4 semblant un peu moins fréquents et 5 étant le cas le plus rare avec 8%.

	1 pd	**5 pds**	**Phloème**
1 pd		-	+
5 pds	0,9156 ± 0,0062		-
Phloème	0,0358 ± 0,0053	0,2593 ± 0,0148	

Tableau 10 : Test de différenciation par paires entre les modalités d'acquisition. Nous avons testé l'hypothèse selon laquelle la distribution du nombre de marqueurs VIT transmis est identique entre les différentes modalités d'acquisition. Pour ce faire nous avons utilisé le test exact de différenciation de population du logiciel « Arlequin » avec une chaîne de Markov comprenant 10 000 itérations. Dans la partie droite supérieure du tableau, les différences significatives ($P<0.05$) entre deux modalités d'acquisition sont signalées par un « + », tandis que les valeurs exactes des P-values sont indiquées dans partie inférieure gauche.

un mode d'acquisition donné. Il est intéressant de noter, que suite à 1 pd ou 5 pds une tendance assez homogène se dessine : les courbes du nombre de plantes tests contenant 1, 2, 3, 4 ou 5 marqueurs diminuent progressivement pour atteindre 0 plante avec 6 marqueurs co-transmis. Par contre, suite à une acquisition de 15 minutes minimum dans le phloème, la courbe adopte un profil « en cloche », le cas le plus fréquemment observé étant la co-transmission de 2 marqueurs, celle de 3 et/ou 4 marqueurs étant au moins aussi fréquente que la transmission d'un seul. La co-transmission de 5 marqueurs est ici aussi relativement faible et jamais les 6 n'ont été retrouvés dans une même plante test.

De récentes analyses, effectuées par Chiraz Jridi (Post-Doc, INRA) avec le logiciel Arlequin, mettent en évidence une différence significative entre les modalités d'acquisition 1 pd et phloème (voir tableau 10).

3.4 Evaluation de la taille du goulot d'étranglement : nombre de génomes viraux efficacement transmis en fonction de la modalité d'acquisition *(analyses futures)*

Nous n'avons malheureusement pas pu collecter et organiser l'ensemble du jeu de données présenté ici, suffisamment tôt pour pouvoir développer formellement une analyse statistique adaptée au test de la neutralité des marqueurs, et un programme mathématique pour l'évaluation de la taille du goulot d'étranglement. Ce travail se poursuit actuellement dans l'équipe d'accueil, en collaboration avec Yannis Michalakis (CNRS-IRD Montpellier), et ne pourra être finalisé et définitivement mis en forme qu'après la soutenance de ma thèse.

Il sera tout d'abord indispensable d'évaluer (et conclure sur) la neutralité des marqueurs, dans les conditions expérimentales décrites. S'il est possible de conclure positivement, alors, une analyse par maximum de vraisemblance sera développée, afin de déterminer précisément le nombre de particules virales transmises par le vecteur, qu'il acquière le virus dans les cellules de l'épiderme et du mésophylle, ou dans les tissus vasculaires. Ce n'est qu'alors que l'influence du comportement alimentaire des pucerons sur la taille du goulot d'étranglement engendré sera soit confirmée, soit infirmée.

4 Discussion

4.1 Neutralité des clones lors de la transmission

La neutralité des clones VIT au cours de la colonisation de la plante hôte a été illustrée dans l'Article 1. En ce qui concerne leur transmissibilité, elle n'a pas été formellement évaluée préalablement aux tests EPG. D'une part, les clones VIT sont isogéniques (à la seule exception du marqueur non codant) et codent tous pour les mêmes protéines dont P2, P3 et P4 (capside). Il n'y a donc aucune raison de penser à priori que des différences dans la nature de l'interaction virus-vecteur puissent exister, à moins d'imaginer des problèmes d'accumulation différentielle, soit quantitative, soit qualitative (spatiale par exemple), ce qui aurait dû apparaître lors du travail sur la *Ne* ou sur la MOI. Même si les taux de transmission avaient été mesurés indépendamment pour chacun des CaMV VIT1-6, ils n'auraient été que d'un intérêt limité. En effet, la seule évaluation qui soit réellement pertinente pour nos expérimentations correspond à leurs taux de transmission à partir de feuille sources co-infectées, information qui est spécifiquement contenue dans notre jeu de données.

Les résultats présentés dans le tableau 8 semblent très encourageants et suggèrent une neutralité ou une quasi-neutralité des marqueurs, puisque dans le cas contraire la fréquence de transmission de chaque clone après vection n'aurait pas été aussi proche de celle mesurée dans les feuilles sources. Les écarts (déjà mentionnés dans la partie « Résultats ») qui sont apparents dans le détail de l'analyse pourraient cependant être significatifs et un test statistique adapté reste à développer.

Nous avons pu noter que les taux de transmission observés dans l'ensemble de notre expérience sont significativement (χ^2 ; $P<0,05$ pour 1 ddl) plus élevés qu'antérieurement constatés dans le même laboratoire [132], où les taux de transmission mesurés étaient de 13%, 24% et 59%, pour des acquisitions dans 1 cellule de l'épiderme ou du mésophylle, plus de 5 cellules, et un accès au phloème, respectivement. Nous supposons que cette différence est imputable à un état physiologique dissemblable entre les

feuilles sources issues de plantes cultivées et infectées à Madrid [132] ou à Montpellier (nos résultats).

4.2 Cohérence et pertinence avec les études précédentes

En parfaite concordance avec l'étude similaire, menée sur les effets de la transmission par vecteurs sur les populations de CMV [199], nous avons constaté une perte aléatoire de la diversité génétique virale suite à la vection vers une nouvelle plante hôte. En d'autres termes, nous avons établi l'existence de fort goulots d'étranglement liés à la transmission par puceron pour le CaMV. L'ensemble des informations produites dans les chapitres I et II sont ici particulièrement enrichissantes, puisqu'elles nous permettent de proposer que cette perte de diversité génétique serait directement liée à un goulot d'étranglement démographique (voir aussi la discussion détaillée dans la section suivante). En effet, nous avons vu que les variants CaMV-VIT1-6 sont distribués de façon homogène dans les plantes hôtes, et même au sein des feuilles où chaque cellule individuelle semble en contenir un nombre relativement important. Dans cette hypothèse l'accessibilité des variants ne serait pas limitante, et c'est donc un faible nombre de génomes transmis qui réduirait la diversité des marqueurs dans les populations des plantes tests.
La taille moyenne pressentie pour ce goulot d'étranglement démographique, bien que cela n'ait pas encore été formellement établi par le calcul statistique, doit se situer autour de 2 ou 3, certainement entre 1 et 5. Cette valeur hypothétique est voisine ou similaire à celle qui a été publiée récemment pour le PVY, à savoir entre 0,5 et 3,2 particules virales transmises en moyenne [200] (aucun chiffre n'ayant été proposé pour le CMV).
Toutefois, le comportement alimentaire des pucerons peut être assez largement différent selon les morphes avec en particulier des durées d'acquisition au sein du phloème variables. Ainsi pour *Myzus persicae*, il a été montré en moyenne (suite à des enregistrements de 7h sous EPG) que les individus aptères passent 13,6% de leur temps en phase E2 (ingestion de sève élaborée) contre seulement 0,001 à 0,1% pour les morphes ailés [202]. Bien que nos résultats avec les 3 différentes modalités d'acquisition n'aient

été obtenus qu'à partir de pucerons aptères, il est probable que nos données soient extrapolables aux morphes ailés pour les modalités 1 pd et 5 pds.

4.3 Effet de la modalité d'acquisition ?

Si les taux de transmission observés entre 1 pd et 5 pds sont très divergents (30 et 72 %, respectivement), la diversité virale transmise est similaire avec respectivement 2,1 et 2,2 marqueurs VIT retrouvés en moyenne dans les plantes tests. A l'inverse, bien que les taux de transmission pour 5 pds et le phloème soient similaires, la diversité virale transmise s'accentue de manière perceptible lorsque l'acquisition se termine au sein du phloème. Au moins trois paramètres – potentiellement imbriqués – nous semblent devoir être pris en considération pour tenter d'expliquer ces observations surprenantes :

(i) le volume de matériel végétal prélevé par le vecteur est de plus en plus important pour les acquisitions à 1 pd, 5 pds et phloème.
(ii) la charge virale, même à volume égal, peut être différente suivant la nature du matériel végétal ingéré : cytoplasme pour les acquisitions à 1 et 5 pds, et sève élaborée pour les acquisitions dans le phloème.
(iii) l'accessibilité des variants génétiques peut être différente dans les différents tissus « visités » par le puceron.

Comment expliquer que le taux de transmission augmente entre 1 pd et les autres modalités d'acquisition ?

Il existe deux réponses possibles : i) soit le nombre de particules virales retenues augmente avec l'augmentation du volume ingéré, ii) soit le puceron est « saturé » dès qu'il a ponctionné une fraction virulifère mais la probabilité de cet événement augmente avec le nombre de piqûres « d'essais » opérées par le puceron (1 seule à 1 pd, 5 ou plus pour les autres modalités).

Pourquoi la diversité transmise n'augmente pas parallèlement au taux de transmission entre les modalités à 1 pd et 5 pds ?

Ceci semble indiquer que l'hypothèse i) (formulée ci-dessus) est fausse puisqu'elle impliquerait une accumulation de particules virales provenant de multiples cellules et devrait être corrélée à une augmentation de la diversité transmise. En conséquence, il semble plus probable que l'hypothèse ii) soit vraie, chaque nouveau contact avec des cellules virulifères ne pourrait alors (le puceron étant déjà saturé) que remplacer tout ou partie des particules déjà retenues. Si l'on suppose que le nombre de variants accessibles dans chaque cellule est souvent équivalent, alors la diversité génétique retenue n'augmente pas au fil des remplacements successifs intervenant lors de chaque ponction intracellulaire.

Pourquoi le taux de transmission n'augmente pas sensiblement entre des acquisitions à 5 pds et dans le phloème ?

Il est possible d'imaginer que lors des cinq premières ponctions intracellulaires, les chances de rencontrer une zone virulifère (corps clair) sont maximisées. C'est pourquoi aucune augmentation significative n'est observée par la suite en multipliant les ponctions, puisque le puceron est déjà saturé et que tout nouveau contact ne sera que remplacement : changement uniquement qualitatif au mieux.

Pourquoi la diversité génétique transmise est supérieure après une acquisition phloémienne ?

Il semble que la diversité accessible dans les tubes criblés du phloème soit un peu plus exhaustive que celle dans les cellules de l'épiderme ou du mésophylle. Le remplacement des particules virales déjà présentes dans les stylets se ferait alors par un échantillon plus homogène de l'ensemble des 6 variants VIT. En effet, il est aisé d'imaginer que lors de la phase de salivation intra-phloèmienne (E1) toutes les particules virales sont relarguées mais pas P2. Ensuite, lors

de la phase d'ingestion (E2) les P2 retiennent des complexes P3-virus présents dans la sève élaborée, dont les flux sont susceptibles d'optimiser l'homogénéisation de la population virale.

Si toute cette discussion a un sens, alors il nous faut admettre que le nombre de particules virales transmises est très faible car l'augmentation de diversité notée lors de l'acquisition phloèmienne reste très limitée et n'atteint qu'une moyenne légèrement inférieure à 3 marqueurs co-transmis.

5 Perspectives

5.1 A brève échéance

Finaliser les analyses statistiques pour déterminer la taille du goulot d'étranglement en terme de nombre de génomes (ou particules virales) efficacement transmis.

5.2 A plus long terme

5.2.1 Effets de deux comportements alimentaires distincts du vecteur sur l'évolution du CaMV

Si nous confirmons la différence de taille des goulots d'étranglement entre 1 pd et phloème, nous pourrons faire évoluer des lignées indépendantes sous ces deux modalités de transmission et ainsi produire des goulots d'étranglement répétés à moindre coût (*cf.* section 3.2.3. de l'introduction générale). Ensuite, nous pourrons tester l'effet de ces goulots liés à la transmission par pucerons sur l'évolution de la fitness – évolution qui n'a encore jamais été évaluée sous l'effet de transmissions par un quelconque vecteur – et tester si des comportements alimentaires distincts (mimant aptères et ailés par exemple) influencent différemment l'évolution du CaMV.

5.2.2 Quantification de la HC-transcomplémentation

Cette question (figurant initialement dans le projet de thèse) est théoriquement importante dans le contexte de la stratégie helper puisqu'un fort taux de HC-transcomplémentation légitimerait le maintien de cet élément « supplémentaire » pour la transmission. En effet, il est peu

concevable d'imaginer qu'un phytovirus puisse coder pour une protéine additionnelle – par opposition à la stratégie capside – si elle ne lui confère pas un certain avantage. Il est envisageable qu'un taux de HC-transcomplémentation élevé offre un potentiel d'adaptation accru dans un environnement changeant, avec par exemple le cas d'alternance de populations vectrices, qui dans une stratégie capside nécessite la modification plus contraignante et donc plus rare d'une protéine structurale. Ainsi, du fait de la stratégie helper et du phénomène de HC-transcomplémentation, il est possible de transmettre des génomes non-adaptés (ne codant pas pour un helper compatible avec le vecteur «du moment»), par le biais du helper compétant produit par un variant adapté. Il se développerait ainsi une coopération pour l'interaction avec le vecteur, plutôt qu'une compétition, entre variants d'une même quasi-espèce. Cette complémentation peut avoir lieu soit au sein d'une unique cellule (si la MOI est forte), soit lors de ponctions cellulaires successives (si la MOI est faible).
Pour l'étude de la HC-transcomplémentation chez le CaMV, nous pourrions utiliser une souche sauvage dite « self-transmissible » et un mutant non-self-transmissible (sans P2 ou codant pour une P2 non fonctionnelle, terminologie définie par Rémy Froissart). La détection du mutant non-self-transmissible dans les plantes tests infectées (par de la Q-PCR ou une autre méthode) nous permettrait ainsi d'estimer le taux de HC-transcomplémentation chez le CaMV.

5.2.3 Taille du goulot d'étranglement chez les virus circulants

Etant maintenant établi que les virus non-circulants (stratégie capside : CMV, et helper : PVY et CaMV) subissent un goulot d'étranglement lors de leur transmission par vecteur, il serait intéressant d'étudier ce qui se passe pour les virus circulants. En effet, même si une étude récente [203] a démontré que de très nombreuses particules virales infectieuses ($10^{3,4}$ à $10^{6,1}$ PFU selon le vecteur) du *West nile virus* (WNV) sont inoculées au nouvel hôte, rien n'est précisé concernant la diversité composant ce large inoculum. Dans le cas du *Venezuelan equine encephalitis virus* (VEEV), et pour un

vecteur donné, il a été montré que la moyenne géométrique était seulement de 11 PFU [204].

Pour une évaluation exhaustive du goulot d'étranglement dans les cas suscités, Il serait intéressant de quantifier précisément la diversité virale transmise d'un hôte à un autre. En effet, pour des virus circulants multipliants, il est possible qu'un grand nombre d'unités infectieuses sortant du vecteur constituent tout de même un fort goulot d'étranglement génétique, puisque le nombre entrant et infectant initialement le vecteur pourrait lui être très faible (moins de 15 ou une moyenne de 28 cellules intestinales du vecteur sont primo-infectées respectivement par le WNV ou le VEEV) [205,206].

Conclusion générale

Ces travaux de thèse ont permis tout d'abord d'ôter l'idée générale que tous les virus de plante sont soumis à de fort goulots d'étranglement durant de l'invasion systémique de leur hôte. En effet, pour le CaMV ce sont en moyenne 400 virions (plus d'une centaine à plusieurs milliers) qui initient l'infection de chaque feuille soit, 10 à 100 plus que pour les virus à ARN déjà étudiés.

Ensuite, il semblerait qu'une situation tranchée se dégage aussi pour la multiplicité d'infection cellulaire (MOI), avec une forte MOI (proche de 7 à 60 dpi) pour le CaMV, en opposition aux virus à ARN et tout particulièrement aux potyvirus, dont de nombreux indices indiquent une quasi totale exclusion territoriale, aussi bien à un niveau intra qu'inter-spécifique. Il serait intéressant de rechercher plus précisément les mécanismes impliqués lors de ces situations de faible MOI, dans le but de discerner si ces exclusions apparentes sont effectivement assimilables à un isolement des autres (individualisation), ou bien juste à une « valve anti-retour » permettant au virus de progresser de façon orientée.

La taille efficace et la multiplicité d'infection cellulaire pourraient être liées par le fait que la MOI puisse influencer directement le nombre de génomes viraux initiant l'infection d'une feuille néoformée, et ceci d'autant plus que le nombre de cellules de cette feuille en contact (relation) avec le reste de la plante est faible. Si tel est le cas, la taille efficace que nous avons calculée lors de la colonisation des feuilles (Article 2) serait éventuellement sous-estimée, puisque nos derniers prélèvement ont été effectués à 45 dpi, étape de l'infection où la MOI n'est pas encore maximale. Plus encore, il serait intéressant d'estimer la taille des populations fondatrices dans des feuilles de plus en plus hautes sur les plantes, prélevées suivant un protocole similaire à celui suivi pour la MOI, afin de tester un éventuel accroissement aussi de la *Ne* au fil du développement de l'infection systémique.

Finalement, des tests d'acquisition sous contrôle EPG ont démontré que la transmission du CaMV par vecteur induit un fort goulot d'étranglement (2 à 3 variants transmis en moyenne), directement imputable au vecteur et non pas à la distribution de la population virale au sein de la plante source. En ce qui concerne les acquisitions à 1 ou 5 pds, il est évident que la diversité transmise aux plantes tests dépend directement de la MOI des cellules ponctionnées, laquelle peut être estimée entre 1 et 3 en moyenne, de 14 à 33 dpi, moment où les feuilles sources ont été utilisées. Mais quelles sont les chances de prélever toute ou un maximum de la diversité en une seule ponction intracellulaire ?

Un projet indépendant de biologie cellulaire, développé dans l'équipe par Martin Drucker et Alexandre Martinière, a mis à jour un phénomène à la fois très intriguant et particulièrement bien intégré dans la problématique de ma thèse. Il démontre sans équivoque que tous les variants entrant dans une cellule vont se répliquer dans plusieurs corps d'inclusion (corps denses) dispersés dans le cytoplasme, mais que les composants du complexe transmissible (P2, P3 et virions) convergent ensuite dans un seul énorme corps d'inclusion clair, constituant ainsi une sorte de « point de ralliement unique» pour l'ensemble des variants de la cellule destinés à la rencontre avec le vecteur (figure 8). En l'état, il est impossible de conclure que la morphogenèse particulière du corps clair est une adaptation spécifiquement sélectionnée pour une meilleure accessibilité de la diversité virale pour le puceron vecteur. Toutefois, si l'on prend en compte le fait que le puceron ne prélève qu'un « aliquot » du contenu cytoplasmique, il est particulièrement troublant de découvrir que la synthèse du corps clair, unique dans chaque cellule, favorise l'accès simultané à l'ensemble des variants qui s'y sont répliqués.

Quelles sont les hypothèses qui supportent l'intérêt d'avoir un HC pour la transmission par vecteur ?

Hypothèse 1 : comme discuté dans la section 5.2.2. du chapitre III et comme initialement envisagé par Pirone et Blanc en 1996 [207], le HC favoriserait l'optimisation de l'échantillonnage de la quasi-espèce virale lors de l'acquisition par le vecteur, du fait de la HC-transcomplémentation, puisqu'elle induit une coopération à ce niveau plutôt qu'une compétition entre variants.

Hypothèse 2 : un projet indépendant, développé dans l'équipe par Marilyne Uzest et portant sur la caractérisation du récepteur chez l'insecte, semble montrer que P2 (du CaMV) et HC-Pro (des potyvirus) pourraient se lier directement ou indirectement à un même récepteur. Les co-infections de plantes par plusieurs espèces virales différentes, mais transmises par les mêmes insectes vecteurs, étant très fréquentes dans les champs, il est fort probable qu'une compétition sévère pour ce récepteur existe au moment de la transmission. Dans ce scénario, les molécules « helper » permettraient donc à une espèce virale dans la plante de convertir ces vecteurs « polyvalents » en vecteurs spécifiques, puisque seules les particules virales homologues pourraient s'y associer ultérieurement. Le HC, dans cette hypothèse, serait ainsi une arme peu coûteuse pour saturer le récepteur du vecteur.

Les résultats obtenus durant ma thèse sur le CaMV font pencher la balance en faveur de l'hypothèse 2. En effet, pour ce virus, étant donné la distribution homogène des variants dans les feuilles, et les cellules (forte MOI estimée), le rôle de la HC-transcomplémentation pour optimiser la diversité prélevée semble secondaire. En particulier nous n'avons pas pu détecter de différence entre la diversité transmise à partir d'une seule cellule et à partir de 5 cellules. Le maintien d'une molécule HC pour un virus comme le CaMV devient très difficile à justifier au travers de l'hypothèse 1.

Finalement, il semblerait donc, d'après les travaux expérimentaux menés, qu'au niveau des plantes hôtes et des vecteurs étudiés il existe des goulots d'étranglement successifs pour les virus à ARN, tandis que pour le CaMV seule l'étape de transmission par vecteur réduit fortement la taille des populations virales. Pour bien en évaluer les conséquences en terme d'évolution virale il devient alors nécessaire de disposer de données épidémiologiques complètes (diversité et dynamique des plantes hôtes et des vecteurs) car ces informations pourraient par exemple par une simple densité de vecteurs réduire le fort impact estimé.

.

Matériels & Méthodes

1 Matériels

1.1 Matériel végétal

La plante hôte du CaMV utilisée lors de ces travaux est *Brassica rapa*, cultivar « Just Right » (hybride F1) plus communément appelé navet. Les semences nous sont gracieusement fournies par TAKII Ltd. Japon.

1.2 Souches virales

Les clones viraux du CaMV utilisés dérivent tous de la souche sauvage Cabb-S et du clone infectieux correspondant pCa37, qui fut le premier isolat de CaMV séquencé en totalité [3]. Le détail des modifications de ce clone pour les besoins spécifiques des différents programmes de la thèse sont décrits dans les sections *ad hoc.*

1.3 Espèces de pucerons

Le puceron vert du pêcher, *Myzus persicae* Sulz. (clone fourni par G. Labonne, INRA Montpellier), a été systématiquement utilisé pour effectuer des tests de transmission classiques.

Le puceron cendré du chou, *Brevicoryne brassicae* (L.) (clone isolé à partir d'une femelle virginipare collectée à Villa del Prado, Espagne par A. Fereres, CSIC Madrid), a été employé pour effectuer les tests de transmission sous EPG. Les individus adultes sont plus grands que ceux de *Myzus persicae,* ce qui est fort pratique et utile quand on réalise ses premiers « branchements » EPG.

Ces deux espèces sont communément décrites comme les deux meilleurs vecteurs du CaMV (voir par exemple [55]).

1.4 Souches bactériennes

La souche XL1-Blue d' *Escherichia coli* a été utilisée pour le clonage et l'amplification de plasmides recombinants, car elle induit peu d'événements de recombinaison entre les molécules d'ADN. En effet, son génotype est : *recA1 endA1 gyrA96(nalR) thi-1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+).*

1.5 Plasmides

Le plasmide pBR322 est utilisé comme vecteur de clonage et d'amplification. C'est un vecteur à usage général d'une taille de 4361 paires de bases, présentant de 15 à 20 copies par cellule et portant les gènes de résistance à l'ampicilline et à la tétracycline.

2 Méthodes

2.1 Techniques relatives aux plantes

2.1.1 Culture des navets

Les plants de navet sont maintenus dans une serre insect-proof dans les conditions suivantes : 25/19°C jour/nuit selon une photopériode de 16/8h jour/nuit. L'arrosage est effectué à l'eau de ville, additionnée d'une solution nutritive.

2.1.2 Inoculation mécanique du CaMV

Au stade 2-3 feuilles, chaque plante préalablement saupoudrée avec un abrasif (Carborandum ou Célite AFA, Fluka) est frottée délicatement avec soit (i) un broyat de feuilles infectées dans de l'eau distillée, soit (ii) 20 µL de particules virales purifiées à une concentration d'au moins 1 µg/mL, soit (iii) 4 µg d'ADN plasmidique viral en solution (plasmide linéaire suite à la digestion par l'enzyme de restriction SalI).

2.2 Techniques relatives aux pucerons-vecteurs

2.2.1 Conditions d'élevage

Les pucerons sont élevés sur des plants de navets sains en cage d'isolement, avec une température maintenue à 20-23°C, une photopériode de 16h de jour pour 8h de nuit et une hygrométrie d'environ 80%.

2.2.2 Tests de transmission classiques

2.2.2.1 A travers membrane de Parafilm®

Avant l'acquisition de chacun des composants nécessaires à la transmission, les pucerons sont retirés de leur plante hôte d'élevage, et mis à jeûner durant au moins une heure dans un pot en verre. Ensuite, ils sont disposés

par groupes d'environ 500 dans des tubes en laitons dont l'une des extrémités est fermée par une membrane de Parafilm® étirée, sur laquelle est déposée la solution à tester (par exemple P2-P3-virus directement, ou bien seulement P2 qui est capable de s'attacher indépendamment des autres composants au récepteur chez le puceron [56]. Les pucerons sont autorisés à prélever cette solution sucrée (10 à 15% saccharose pour la palatabilité) en piquant à travers le Parafilm®. Cette étape, constituant un " repas ", peut être reconduite avec les mêmes pucerons pour l'acquisition d'une autre solution (par exemple : P3-virus). Après le dernier repas d'acquisition qui peut aussi s'effectuer sur une feuille infectée, les aphides sont déposés à l'unité ou par lot de 10 sur les plantes tests saines à l'aide d'un pinceau. Ils sont éliminés 16 heures plus tard par un traitement aphicide systémique (Pirimor G, Sopra ou Confidor, Bayer Cropscience). Dans les cas où la transmission est positive, des symptômes d'infection sont visualisables dès 10 jours (les plantes sont toutefois systématiquement surveillées durant 4 semaines après le test).

2.2.2.2 Transmission de plante à plante

Les transmissions de plante à plante par pucerons-vecteurs se font de manière similaire aux tests à travers membrane de Parafilm®, sauf que l'acquisition de virus s'effectue directement (exception faite des virus déficients en P2) sur la plante infectée avant le transfert des pucerons sur les plantes tests saines.

2.3 Techniques et stratégies utilisées pour le clonage

L'ensemble des techniques de biologie moléculaire évoquées ci-dessous sont des techniques courantes, décrites en détails dans les manuels de laboratoires tels que le Maniatis [208] ou Current Protocols in Molecular Biology [209].

2.3.1 Amplification par PCR (Polymerase Chain Reaction)

Il s'agit d'une amplification exponentielle de l'ADN réalisée par une série de cycles du type dénaturation-hybridation/amorçage-élongation avec deux amorces oligonucléotitiques spécifiques des extrémités du fragment à

amplifier. A titre d'exemple, la réaction peut-être effectuée avec : 0,5 à 3 µL de matrice (ADN viral purifié par exemple), des dNTPs (400 µM final), le ThermoPol Reaction Buffer (1X de chez New England Biolabs), les amorces (0,2 µM final chacune), du sulfate de magnésium (4 mM final), 2 unités enzymatiques de Vent$_R$® DNA Polymerase et de l'eau distillée stérile pour un volume réactionnel total de 100 µL. Le programme exécuté par le thermocycleur (Mastercycler gradient, Eppendorf) comprend (toujours à titre d'exemple) : 7 minutes de dénaturation à 95°C, puis 35 cycles avec 1 minute de dénaturation, 1 minute d'hybridation à 50°C et 2 minutes d'élongation à 75°C, avant de terminer par 10 minutes d'élongation et de maintenir à 20°C.

2.3.2 Digestion par des enzymes de restriction

Les digestions d'ADN par des enzymes de restriction sont réalisées classiquement selon les conditions préconisées par les fournisseurs.

2.3.3 Déphosphorylation du plasmide

Afin de prévenir la recircularisation du plasmide lors de la ligation, les vecteurs linéarisés par digestion enzymatique sont déphosphorylés par une phosphatase alcaline (issue de la crevette ou bien du veau) suivant les recommandations des fournisseurs. Le vecteur ainsi déphosphorylé est purifié sur gel d'agarose.

2.3.4 Purification de fragments d'ADN sur gel d'agarose

Les molécules d'ADN sont séparées selon leur taille par migration électrophorétique sur gel d'agarose (1 à 1,5% en général) préparé avec du tampon TAE 0,5X (20 mM Tris-acétate, 0,5 mM EDTA). Après migration, le gel est incubé pendant quelques minutes dans un bain contenant du bromure d'éthidium (BET) afin de révéler sous rayons ultra-violet (U.V.) les différents fragments d'ADN.

Le fragment d'intérêt est découpé et placé dans un microtube afin de dissoudre l'agarose dans un volume de phénol et d'eau distillée à 65°C. Après deux étapes d'extraction au chloroforme/alcool isoamylique (ratio 24/1), l'ADN est précipité à l'éthanol 70%/acétate de sodium, culotté par centrifugation, et repris dans un petit volume d'eau distillé.

2.3.5 Ligation

La ligation du plasmide linéarisé et déphosphorylé (vecteur) avec le fragment d'ADN d'intérêt (insert) a lieu sur la nuit à 16°C dans un volume réactionnel de 10 µL. Le milieu réactionnel contient :

Vecteur : 10 à 50 ng

Insert : rapport molaire insert/vecteur compris entre 2 et 3

Tampon de ligation 10X : 1 µL

ADN ligase T4 : 1 Unité

2.3.6 Transformation et sélection

100 µL de bactéries compétentes XL1 blue, fraîchement sorties du congélateur -80°C, sont ajoutés au produit de ligation. Après au moins vingt minutes d'incubation sur glace, les bactéries sont alors soumises à un choc thermique à 42°C durant 45 à 60 secondes, puis à nouveau immédiatement refroidies sur glace. L'ajout de 300 µL de milieu LB (Luria Bertani) permet d'initier la croissance bactérienne, et d'induire la résistance à l'antibiotique portée par le plasmide vecteur, durant une courte période d'incubation de 30 minutes sous agitation à 37°C. Enfin, 200 µL de cette pré-culture bactérienne sont étalés sur une boite de Pétri contenant du milieu LB-Agar et l'antibiotique de sélection. Les boites sont incubées durant une nuit à 37°C et seules les bactéries résistantes à l'antibiotique de sélection, contenant donc le plasmide, formeront des colonies.

2.3.7 Criblage des colonies

2.3.7.1 Par digestion enzymatique

Les colonies ayant poussées sur la boite de Pétri sont piquées et mises en culture dans 2 mL de LB + antibiotique durant une nuit à 37°C. L'ADN plasmidique est purifié par la méthode de « boiling » ou de « lyse alcaline » (protocoles classiques décrits dans [208]). L'ADN plasmidique est alors digéré par une ou plusieurs enzymes de restriction afin de vérifier la nature des ADN recombinants présents au sein de chaque colonie bactérienne.

Lorsque l'expérience se révèle positive, une fraction (850 µL) de culture bactérienne saturée qui a préalablement été ensemencée à partir de ce clone, est stockée à -80°C dans un tube contenant 150 µL de glycérol stérile.

2.3.7.2 Par PCR sur des pools de colonies

Cette technique s'avère très pratique lorsque les colonies qui se sont développées sur la boite de Pétri sont très nombreuses et que le nombre de plasmides recombinants à l'issue de la ligation est faible.

Elle consiste à piquer 5 colonies, à l'aide de 5 cure-dents, et de les plonger dans un même tube de 100 µL d'eau distillée afin de créer un petit pool de bactéries qui servira de matrice pour la réaction de criblage par amplification PCR. Seuls 2,5 µL de ce pool sont utilisés comme matrice dans un volume réactionnel final de 25 µL. Les amorces spécifiques de la construction plasmidique recombinante sont employées à une concentration de 200 nM, en présence de 200 µM de dNTP, 1,5 mM de $MgCl_2$, 1X de tampon PCR sans magnésium et 1 unité de Taq polymérase.

2.3.8 Amplification et purification de plasmides recombinants

Pour amplifier les plasmides vecteurs ou recombinants obtenus, un erlenmeyer contenant 200 à 250 mL de LB et l'antibiotique de sélection est ensemencé avec le clone bactérien désiré et incubé sous agitation durant 16 heures à 37°C. La culture bactérienne saturée est alors traitée avec le kit Plasmid Midi Kit (QIAGEN) selon les recommandations du fabriquant afin de purifier l'ADN plasmidique.

2.4 Techniques relatives aux virus

2.4.1 Purification de virus

Les feuilles infectées sont récoltées, dénervées, puis broyées à l'aide d'un robot mixer dans du tampon KH_2PO_4 0,5M - Na_2SO_3 (sulfite de sodium) 0,75 % (avec un rapport 2 volumes/poids). Ce broyat, fin et homogène, est filtré au travers d'une mousseline. Le jus ainsi obtenu est additionné de Triton X100 à 2,5 % final et d'Urée à 1 M final. L'ensemble est mis à 4°C sous agitation pendant 20 à 27 heures.

La solution est clarifiée par une centrifugation de 10 minutes à 2 000 × *g*. Le surnageant est ultra centrifugé durant 2 heures à 4°C et 28 000 rpm [4]

[4] Pour les rotors Beckman la conversion en × *g* (ou RCF pour Relative Centrifugal Force) est estimable depuis le site web : http://www.beckmancoulter.com/resourcecenter/labresources/centrifuges/rotorcalc.asp

(rotor Type 45 Ti, Beckman). Les culots sont repris avec de l'eau distillée (1-1,5 mL par culot), et mis à agités toute la nuit à 4°C.

La suspension ainsi obtenue est alors transférée dans des microtubes de 1,5 mL pour une clarification de 5 minutes à 10 000 × *g*. Le surnageant est chargé sur gradients de saccharose 10-40 % pour une ultracentrifugation à 25 000 rpm, 3 heures à 4°C (rotor SW 41 Ti, Beckman). Par une illumination verticale du gradient, la bande virale – visible en opaque bleuté suite à la diffraction de la lumière – est récupérée à l'aide d'une seringue. Le volume de l'ensemble des bandes virales récupérées est ajusté à 24 mL avec de l'eau distillée et les virions sont culottés par une ultracentrifugation finale durant 3 heures à 4°C et 38 000 rpm (rotor SW 41 Ti).

Le surnageant est éliminé et chaque culot est repris dans 300 µL (pour 100 g de feuilles fraîches) d'une solution comprenant : 1 CMC (concentration micellaire critique) final de βOG et 20 mM de Tris pH 6,9 ou alors avec une solution d'Hepes à 10mM pH 7.2. La resuspension des culots se fait à 4°C.

D'après l'équation suivante : $\varepsilon^{0,1\%}_{260} = 7$ (c'est à dire que 7 unités de D.O. correspondent à 1 mg/mL de virus en solution), il est possible de calculer la concentration en virus de chaque échantillon.

2.5 Techniques relatives aux protéines

2.5.1 Production de protéines en systèmes hétérologues

2.5.1.1 Production de P2 et P3 en système baculovirus

Le baculovirus (AcMNPV : Autographa californica mono nuclear polyhedrosis virus) infecte les cellules de plus de trente espèces différentes de lépidoptères. Lors de la phase dite tardive-tardive (après 18 h d'infection) du cycle de réplication, ce virus produit deux protéines qui ne sont par ailleurs pas indispensables à la multiplication virale et à la diffusion de cellule à cellule (intra-hôte ou en culture cellulaire). Il est donc possible d'utiliser ce baculovirus comme un vecteur d'expression en remplaçant la séquence codante de l'une ces deux protéines tardives (dans notre cas la protéine P10) par une séquence d'intérêt. Pour ce faire, il faut tout d'abord insérer la séquence d'intérêt dans un vecteur plasmidique de transfert contenant le promoteur de P10 suivi d'un site multiple de clonage, puis du terminateur.

Une fois le gène d'intérêt inséré au site de clonage multiple, le vecteur plasmidique est co-transfecté, avec l'ADN viral du baculovirus, dans les cellules d'insectes, en vue d'obtenir un baculovirus recombinant.
La lignée cellulaire utilisée provient de cellules ovariennes de *Spodoptera frugiperda* (Lepidoptera) : Sf9 (American Type Culture Collection, Rockville M.D.). Ces cellules sont maintenues à 28°C dans un milieu de culture G3 (mis au point au laboratoire de Saint-Christol-lez-Alès) et elles doivent êtres repiquées tous les 3 à 5 jours suivant leur densité initiale.
Les baculovirus recombinants exprimant P2 et P3 étant déjà disponibles au laboratoire, nous n'avons eu qu'à infecter les cellules Sf9. Lorsque les cellules *Sf9* atteignent 75% de confluence dans des flacons de cultures de 75 cm^2 de surface, le surnageant est éliminé et la suspension de virus recombinant (milieu surnageant d'une culture infectée) est déposée au contact des cellules (0,5 mL pour 25 cm^2 de culture). Après une heure, du milieu frais G3 (11 mL) est ajouté à l'inoculum viral, et la culture est incubée à 28°C pendant 3 jours. Les cellules sont décollées de la surface plastique du flacon et culottées pendant 10 minutes, 150 × *g* (Multifuge 3S-R, Heraeus), à 4°C. Ensuite, le surnageant est éliminé (ou conservé comme suspension de virus stock), les culots cellulaires sont repris avec 1 mL de tampon SES (Standard Extract Solution : 200 mM Tris, 100 mM EGTA, 50 mM $MgCl_2$, pH 7,6), puis soumis à un cycle de congélation-décongélation, afin de casser les cellules d'insectes et de libérer les protéines avant le premier emploi. L'utilisation des protéines P2 et P3 pour les tests de transmission est directe (sans étapes supplémentaires de purification). Des contrôles négatifs ont permis de confirmer que l'ensemble des autres composants cellulaires et baculoviraux n'ont pas d'effet sur les tests de transmission [54].

2.5.1.2 Production de P3 en bactérie et purification

Le gène III a été cloné par l'équipe de P. Yot de l'IBMP-CNRS de Strasbourg dans le plasmide pET 3a, sous contrôle du promoteur et des signaux de traduction du bactériophage T7. La souche bactérienne BL21 possède une copie chromosomique de l'ARN polymérase T7 dont l'expression est

gouvernée par l'opéron lacUV5. La présence de 0,1 mM d'IPTG dans le milieu provoque la dérépression du gène, donc la production de l'ARN polymérase T7 et par conséquent l'expression du gène III.

Les cultures de bactéries BL21 pET3aIII sont incubées à 37°C jusqu'en fin de phase exponentielle de croissance ($D.O._{600nm}$ de 0,7 à 0,8), avant ajout de l'IPTG (1,5 mM final). Après 6 heures de production de la protéine P3, les bactéries sont culottées (4°C, 10 min, 6 000 × *g* dans une Sorval R26 Plus) et resuspendues dans du tampon P3 (20 mM Tris pH 8, 200 mM NaCl, 1 mM EDTA, 1X Complete). La protéine P3 est ensuite libérée et solubilisée dans le tampon P3 par sonication puis récupérée dans le surnageant après une centrifugation (10 000 × *g*, 15 min, 4°C) éliminant au culot tous les débris et protéines bactériennes insolubles. P3 étant thermostable, environ 90% des protéines bactériennes sont coagulées à 65°C durant 10 minutes, puis précipitées par une centrifugation supplémentaire identique. P3 partiellement purifiée dans le surnageant est alors stockée à –20°C.

2.5.2 Analyse des protéines et quantification

2.5.2.1 Électrophorèse SDS-PAGE & Western Blot

Les analyses des protéines des différents échantillons sont basées sur la technique classique d'électrophorèse en gel de poly-acrylamide (12 ou 15%; ratio acrylamide/bisacrylamide : 30/0,8) en conditions dénaturantes [210] qui ne sera pas détaillée ici.

Pour la révélation spécifique par détection immunologique (Western blot), les protéines contenues dans le gel SDS-PAGE après migration sont transférées sur membrane de nitrocellulose (Optitran BA-S 85, 0,45µm; Schleicher & Schuell). Le transfert est semi-sec (appareil Ancos) dans du tampon Tris-Glycine dépourvu de SDS (25 mM Tris, 192 mM Glycine, 20% méthanol) durant 1 à 2 h à 2 mA/cm^2. La membrane de nitrocellulose est alors colorée dans une solution de rouge Ponceau (Sigma ; préparée suivant les indications du fournisseur) afin de vérifier l'efficacité du transfert, puis rincée à l'eau distillée. La membrane de nitrocellulose est alors saturée et décolorée dans un tampon TS-Sat (50 mM Tris-HCl, pH 7.4, 200 mM NaCl, 5% lait écrémé en poudre), durant 30 minutes au moins sous agitation. Les

anticorps primaires, dirigés contre la capside virale (fourni par D. Leclerc, FMI, Bâle, Suisse) ou contre la protéine P2 sont préparés chez le lapin. Les anticorps secondaires, dirigés contre les IgG de lapin, sont couplés à la Phosphatase Alcaline (PA; Santa Cruz Biotechnologies). Ces anticorps sont dilués au millième dans du tampon TS-Sat, avant d'être mis en contact avec la membrane, sous agitation durant 2 à 4 heures à température ambiante, ou bien toute la nuit à 4°C. Trois rinçages de 10 minutes avec du TS sont effectués après chaque bain d'anticorps. Un rinçage supplémentaire de 5 minutes avec le tampon de révélation PA (100 mM Tris, 100 mM NaCl, 50 mM Chlorure de magnésium, pH 9,5) est effectué afin d'équilibrer la membrane. Enfin, la membrane est incubée avec la solution de révélation (10 mL tampon de révélation, 45 µL NBT à 30mg/mL, 35 µL BCIP) durant une période d'une durée variable, dépendante de l'intensité du rapport signal sur bruit. La réaction de révélation est alors stoppée par de multiples rinçages successifs à l'eau distillée, puis la membrane est séchée à plat entre deux feuilles de papier Whatman.

2.5.2.2 Quantification densitométrique des virus

Après avoir numérisé les membranes de nitrocellulose révélées spécifiquement par l'anticorps anti-capside, nous utilisons le logiciel NIH 1.62 Image pour procéder à l'analyse densitométrique de l'intensité des différentes bandes.

Références Bibliographiques

1. Qiu SG, Schoelz JE (1992) Three regions of cauliflower mosaic virus strain W260 are involved in systemic infection of solanaceous hosts. Virology 190: 773-782.

2. Jacquot E, Dautel S, Leh V, Geldreich A, Yot P, et al. (1997) Les pararétrovirus de plantes. Virologie 1: 111-120.

3. Franck A, Guilley H, Jonard J, Richards K, Hirth L (1980) Nucleotide sequence of cauliflower mosaic virus DNA. Cell 21: 285-294.

4. Cheng RH, Olson NH, Baker TS (1992) Cauliflower mosaic virus: a 420 subunit (T = 7), multilayer structure. Virology 186: 655-668.

5. van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, et al. (2000) Virus Taxonomy: The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. SanDiego: Academic Press. 1167 p.

6. Anonymous (1996) Guidelines on good plant protection practice - No. 6 Vegetable brassicas EPPO Bulletin 26: 311-347.

7. Karsies A, Merkle T, Szurek B, Bonas U, Hohn T, et al. (2002) Regulated nuclear targeting of cauliflower mosaic virus. J Gen Virol 83: 1783-1790.

8. Plant AL, Covey SN, Grierson D (1985) Detection of a subgenomic mRNA for gene V, the putative reverse transcriptase gene of cauliflower mosaic virus. Nucleic Acids Res 13: 8305-8321.

9. Hohn T, Fütterer J (1992) Transcriptional and translational control of genes expression in cauliflower mosaic virus. Curr Top Genet Dev 2: 90-96.

10. Kiss-Laszlo Z, Blanc S, Hohn T (1995) Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. Embo J 14: 3552-3562.

11. Bonneville JM, Sanfacon H, Futterer J, Hohn T (1989) Posttranscriptional trans-activation in cauliflower mosaic virus. Cell 59: 1135-1143.

12. Howarth AJ, Gardner RC, Messing J, Shepherd RJ (1981) Nucleotide sequence of naturally occurring deletion mutants of cauliflower mosaic virus. Virology 112: 678-685.

13. Wurch T, Kirchherr D, Mesnard JM, Lebeurier G (1990) The cauliflower mosaic virus open reading frame VII product can be expressed in Saccharomyces cerevisiae but is not detected in infected plants. J Virol 64: 2594-2598.

14. Schultze M, Jiricny J, Hohn T (1990) Open reading frame VIII is not required for viability of cauliflower mosaic virus. Virology 176: 662-664.

15. Kobayashi K, Tsuge S, Stavolone L, Hohn T (2002) The cauliflower mosaic virus virion-associated protein is dispensable for viral replication in single cells. J Virol 76: 9457-9464.

16. Kobayashi K, Nakayashiki H, Tsuge S, Mise K, Furusawa I (1998) Accumulation kinetics of viral gene products in cauliflower mosaic virus-infected turnip protoplasts. Microbiol Immunol 42: 65-69.

17. Li Y, Leisner SM (2002) Multiple domains within the Cauliflower mosaic virus gene VI product interact with the full-length protein. Mol Plant Microbe Interact 15: 1050-1057.

18. Cecchini E, Gong Z, Geri C, Covey SN, Milner JJ (1997) Transgenic Arabidopsis lines expressing gene VI from cauliflower mosaic virus variants exhibit a range of symptom-like phenotypes and accumulate inclusion bodies. Mol Plant Microbe Interact 10: 1094-1101.

19. Rothnie HM, Chapdelaine Y, Hohn T (1994) Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv Virus Res 44: 1-67.

20. Haas M, Bureau M, Geldreich A, Yot P, Keller M (2002) Cauliflower mosaic virus: still in the news. Molecular Plant Pathology 3: 419-429.

21. Ryabova LA, Hohn T (2000) Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems. Genes Dev 14: 817-829.

22. Ryabova LA, Pooggin MM, Hohn T (2006) Translation reinitiation and leaky scanning in plant viruses. Virus Res 119: 52-62.

23. Love AJ, Laird J, Holt J, Hamilton AJ, Sadanandom A, et al. (2007) Cauliflower mosaic virus protein P6 is a suppressor of RNA silencing. J Gen Virol 88: 3439-3444.

24. Haas G (2007) Etude de l'importation nucléaire de la protéine P6 du virus de la mosaïque du chou-fleur (CaMV) et de son rôle dans la suppression du RNA silencing antiviral. Strasbourg: Louis Pasteur. 140 p.

25. Thiebeauld O (2007) Réinitiation de la traduction dépendante de TAV au cours de l'infection des plantes par le virus de la mosaïque du chou-fleur : caractérisation fonctionnelle d'une nouvelle protéine cellulaire RISP. Strasbourg: Louis Pasteur. 120 p.

26. Torruella M, Gordon K, Hohn T (1989) Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. Embo J 8: 2819-2825.

27. Linstead PJ, Hills GJ, Plaskitt KA, Wilson IG, Harker CL, et al. (1988) The subcellular location of the gene 1 product of cauliflower mosaic virus is consistent with a function associated with virus spread. Journal of General Virology 69: 1809-1818.

28. Perbal MC, Thomas CL, Maule AJ (1993) Cauliflower mosaic virus gene I product (P1) forms tubular structures which extend from the surface of infected protoplasts. Virology 195: 281-285.

29. Huang Z, Han Y, Howell SH (2001) Effects of movement protein mutations on the formation of tubules in plant protoplasts expressing a fusion between the green fluorescent protein and Cauliflower mosaic virus movement protein. Mol Plant Microbe Interact 14: 1026-1031.

30. Thomas CL, Maule AJ (1999) Identification of inhibitory mutants of Cauliflower mosaic virus

movement protein function after expression in insect cells. J Virol 73: 7886-7890.

31. Stavolone L, Villani ME, Leclerc D, Hohn T (2005) A coiled-coil interaction mediates cauliflower mosaic virus cell-to-cell movement. Proc Natl Acad Sci U S A 102: 6219-6224.

32. Jacquot E, Geldreich A, Keller M, Yot P (1998) Mapping regions of the cauliflower mosaic virus ORF III product required for infectivity. Virology 242: 395-402.

33. Citovsky V, Knorr D, Zambryski P (1991) Gene I, a potential cell-to-cell movement locus of cauliflower mosaic virus, encodes an RNA-binding protein. Proc Natl Acad Sci U S A 88: 2476-2480.

34. Thomas CL, Maule AJ (1995) Identification of the cauliflower mosaic virus movement protein RNA- binding domain. Virology 206: 1145-1149.

35. Stussi-Garaud C, Mutterer J, Delecolle B, Rohfritsch O, Vantard M (1998) Routes intra- et intercellulaires utilisées par les virus de plantes pour leur mouvement de cellule à cellule. Virologie 2: 269-283.

36. Gordon K, Pfeiffer P, Futterer J, Hohn T (1988) In vitro expression of cauliflower mosaic virus genes. Embo J 7: 309-317.

37. Kruse J, Timmins P, Witz J (1987) The spherically averaged structure of a DNA isometric plant virus: Cauliflower mosaic virus. Virology 159: 166-168.

38. DuPlessis DH, Smith P (1981) Glycosylation of the Cauliflower Mosaic Virus Capsid Polypeptide. Virology 109: 403-408.

39. Chapdelaine Y, Kirk D, Karsies A, Hohn T, Leclerc D (2002) Mutation of capsid protein phosphorylation sites abolishes cauliflower mosaic virus infectivity. J Virol 76: 11748-11752.

40. Champagne J, Laliberte-Gagne ME, Leclerc D (2007) Phosphorylation of the termini of Cauliflower mosaic virus precapsid protein is important for productive infection. Mol Plant Microbe Interact 20: 648-658.

41. Leclerc D, Chapdelaine Y, Hohn T (1999) Nuclear targeting of the cauliflower mosaic virus coat protein. J Virol 73: 553-560.

42. Guerra-Peraza O, de Tapia M, Hohn T, Hemmings-Mieszczak M (2000) Interaction of the cauliflower mosaic virus coat protein with the pregenomic RNA leader. J Virol 74: 2067-2072.

43. Himmelbach A, Chapdelaine Y, Hohn T (1996) Interaction between cauliflower mosaic virus inclusion body protein and capsid protein: implications for viral assembly. Virology 217: 147-157.

44. Dautel S, Guidasci T, Pique M, Mougeot JL, Lebeurier G, et al. (1994) The full-length product of cauliflower mosaic virus open reading frame III is associated with the viral particle. Virology 202: 1043-1045.

45. Plisson C, Uzest M, Drucker M, Froissart R, Dumas C, et al. (2005) Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy. J Mol Biol 346: 267-277.

46. Leh V, Jacquot E, Geldreich A, Haas M, Blanc S, et al. (2001) Interaction between cauliflower mosaic virus ORFIII product and the coat protein is required for transmission of the virus by aphids. J Virol 75: 100-106.

47. Mougeot JL, Guidasci T, Wurch T, Lebeurier G, Mesnard JM (1993) Identification of C-terminal amino acid residues of cauliflower mosaic virus open reading frame III protein responsible for its DNA binding activity. Proc Natl Acad Sci USA 90: 1470-1473.

48. Leclerc D, Burri L, Kajava AV, Mougeot JL, Hess D, et al. (1998) The open reading frame III product of cauliflower mosaic virus forms a tetramer through a N-terminal coiled-coil. J Biol Chem 273: 29015-29021.

49. Tsuge S, Kobayashi K, Nakayashiki H, Mise K, Furusawa I (1999) Cauliflower mosaic virus ORF III product forms a tetramer in planta: its implication in viral DNA folding during encapsidation. Microbiol Immunol 43: 773-780.

50. Guidasci T, Mougeot JL, Lebeurier G, Mesnard JM (1992) Processing of the minor capsid protein of the cauliflower mosaic virus requires a cysteine proteinase. Res Virol 143: 361-370.

51. Woolston CJ, Covey SN, Penswick JR, Davies JW (1983) Aphid transmission and a polypeptide are specified by a defined region of the cauliflower mosaic virus genome. Gene 23: 15-23.

52. Modjtahedi N, Volovitch M, Mazzolini L, Yot P (1985) Comparison of the predicted secondary structure of aphid transmission factor for transmissible and non-transmissible cauliflower mosaic virus strains. FEBS Letter 181: 223-228.

53. Hébrard E, Drucker M, Leclerc D, Hohn T, Uzest M, et al. (2001) Biochemical Characterization of the Helper Component of Cauliflower Mosaic Virus. J Virol 75: 8538-8546.

54. Blanc S, Cerutti M, Usmany M, Vlak JM, Hull R (1993) Biological activity of cauliflower mosaic virus aphid transmission factor expressed in a heterologous system. Virology 192: 643-650.

55. Moreno A, Hebrard E, Uzest M, Blanc S, Fereres A (2005) A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species. J Virol 79: 13587-13593.

56. Uzest M, Gargani D, Drucker M, Hebrard E, Garzo E, et al. (2007) From the Cover: A protein key to plant virus transmission at the tip of the insect vector stylet. Proc Natl Acad Sci U S A 104: 17959-17964.

57. Leh V, Jacquot E, Geldreich A, Hermann T, Leclerc D, et al. (1999) Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. Embo J 18: 7077-7085.

58. Drucker M, Froissart R, Hebrard E, Uzest M, Ravallec M, et al. (2002) Intracellular distribution of viral gene products regulates a complex mechanism of cauliflower mosaic

virus acquisition by its aphid vector. Proc Natl Acad Sci U S A 99: 2422-2427.

59. Martinière A, Gargani D, Uzest M, Lautredou N, Blanc S, et al. Le Cauliflower mosaic virus utilise les microtubules pour permettre la formation d'un corps d'inclusion spécialisé dans la transmission; 2007 2007/01/28-2007/02/01; Aussois (France).

60. Espinoza AM, Medina V, Hull R, Markham PG (1991) Cauliflower mosaic virus gene II product forms distinct inclusion bodies in infected plant cells. Virology 185: 337-344.

61. Khelifa M, Journou S, Krishnan K, Gargani D, Esperandieu P, et al. (2007) Electron-lucent inclusion bodies are structures specialized for aphid transmission of cauliflower mosaic virus. J Gen Virol 88: 2872-2880.

62. Münch E (1930) Die Stoffbewegungen in der Pflanze: Jena: Verlag von Gustav Fischer, Germany. 234 p.

63. Mezitt LA, Lucas WJ (1996) Plasmodesmal cell-to-cell transport of proteins and nucleic acids. Plant Mol Biol 32: 251-273.

64. Hull R (1989) The movement of viruses in plants. Annu Rev Phytopathol 27: 213-240.

65. Citovsky V, Zambryski P (1993) Transport of nucleic acid through membrane channels : snaking through small holes. Annu Rev Microbiol 47: 167-197.

66. Epel BL (1994) Plasmodesmata: composition, structure and trafficking. Plant Mol Biol 26: 1343-1356.

67. Lucas WJ, Gilbertson RL (1994) Plasmodesmata in relation to viral movement within leaf tissues. Annu Rev Phytopath 32: 387-411.

68. Waigmann E, Curin M, Heinlein M (2007) Tobacco Mosaic Virus - a Model for Macromolecular Cell-to-Cell Spread. In: Waigmann E, Heinlein M, editors. Viral Transport in Plants: Springer-Verlag Berlin Heidelberg. pp. 29-62.

69. Sanchez-Navarro JA, Carmen Herranz M, Pallas V (2006) Cell-to-cell movement of Alfalfa mosaic virus can be mediated by the movement proteins of Ilar-, bromo-, cucumo-, tobamo- and comoviruses and does not require virion formation. Virology 346: 66-73.

70. Ritzenthaler C, Hofmann C (2007) Tubule-Guided Movement of Plant Viruses. In: Waigmann E, Heinlein M, editors. Viral Transport in Plants: Springer-Verlag Berlin Heidelberg. pp. 63-83.

71. Kasteel D, Wellink J, Verver J, van Lent J, Goldbach R, et al. (1993) The involvement of cowpea mosaic virus M RNA-encoded proteins in tubule formation. J Gen Virol 74 (Pt 8): 1721-1724.

72. Kasteel DT, Perbal MC, Boyer JC, Wellink J, Goldbach RW, et al. (1996) The movement proteins of cowpea mosaic virus and cauliflower mosaic virus induce tubular structures in plant and insect cells. J Gen Virol 77: 2857-2864.

73. Godefroy-Colburn T, Schoumacher F, Erny C, Berna A, Moser O, et al. (1989) The movement protein of some plant viruses. In: Springer-Verlag, editor. Parallels in cell to cell junctions in plants and animals. pp. 208-231.

74. Rojas MR, Zerbini FM, Allison RF, Gilbertson RL, Lucas WJ (1997) Capsid protein and helper component-proteinase function as potyvirus cell-to-cell movement proteins. Virology 237: 283-295.

75. Carrington JC, Jensen PE, Schaad MC (1998) Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. Plant J 14: 393-400.

76. Gomez de Cedron M, Osaba L, Lopez L, Garcia JA (2006) Genetic analysis of the function of the plum pox virus CI RNA helicase in virus movement. Virus Res 116: 136-145.

77. Roberts IM, Wang D, Findlay K, Maule AJ (1998) Ultrastructural and temporal observations of the potyvirus cylindrical inclusions (Cls) show that the Cl protein acts transiently in aiding virus movement. Virology 245: 173-181.

78. Huang Z, Andrianov VM, Han Y, Howell SH (2001) Identification of arabidopsis proteins that interact with the cauliflower mosaic virus (CaMV) movement protein. Plant Mol Biol 47: 663-675.

79. Huang Z, Han Y, Howell SH (2000) Formation of surface tubules and fluorescent foci in Arabidopsis thaliana protoplasts expressing a fusion between the green fluorescent protein and the cauliflower mosaic virus movement protein. Virology 271: 58-64.

80. Roberts IM, Boevink P, Roberts AG, Sauer N, Reichel C, et al. (2001) Dynamic changes in the frequency and architecture of plasmodesmata during the sink-source transition in tobacco leaves. Protoplasma 218: 31-44.

81. Lucas WJ, Ding B, van der Schoot C (1993) Plasmodesmata and the supracellular nature of plants. New Phytol 125: 435-476.

82. Esau K, Thorsch J (1985) Sieve plate pores and plasmodesmata, the communication channels of the symplast: ultrastructural aspects and developmental relations. Am J Bot 72: 1641–1653.

83. Leisner SM, Turgeon R, Howell SH (1992) Long distance movement of cauliflower mosaic virus in infected turnip plants. Molec Plant-microbe interactions 5: 41-47.

84. Urban LA, Ramsdell DC, Klomparens KL, Lynch T, Hancock JF (1989) Detection of Blueberry shoestring virus in xylem and phloem tissues of highbush blueberry. Phypathology 79: 488-493.

85. Verchot J, Driskel BA, Zhu Y, Hunger RM, Littlefield LJ (2001) Evidence that soilborne wheat mosaic virus moves long distance through the xylem in wheat. Protoplasma 218: 57-66.

86. Ueki S, Citovsky V (2007) Spread Throughout the Plant: Systemic Transport of Viruses. In: Waigmann E, Heinlein M, editors. Viral Transport in Plants: Springer-Verlag Berlin Heidelberg. pp. 85-118.

87. Roberts K, Love AJ, Laval V, Laird J, Tomos AD, et al. (2007) Long-distance movement of Cauliflower mosaic virus and host defence responses in Arabidopsis follow a predictable pattern that is determined by the leaf orthostichy. New Phytol 175: 707-717.

88. Kang J, Tang J, Donnelly P, Dengler N (2003) Primary vascular pattern and expression of ATHB-8 in shoots of Arabidopsis New Phytol 158: 443-454.

89. Melcher U (1989) Symptoms of Cauliflower Mosaic Virus infection in *Arabidopsis thaliana* and turnip. Bot Gaz 150: 139-147.

90. Wintermantel WM, Anderson EJ, Schoelz JE (1993) Identificatioon of domains within gene VI of cauliflower mosaic virus that influence systemic infection of *Nicotiana bigelovii* in a light dependant manner. Virology 196: 789-798.

91. Love AJ, Laval V, Geri C, Laird J, Tomos AD, et al. (2007) Components of Arabidopsis defense- and ethylene-signaling pathways regulate susceptibility to Cauliflower mosaic virus by restricting long-distance movement. Mol Plant Microbe Interact 20: 659-670.

92. Al-Kaff N, Covey SN (1996) Unusual Accumulations of Cauliflower Mosaic Virus in Local Lesions, Dark Green Leaf Tissue, and Roots of Infected Plants. Mol Plant Microbe Interact 9: 357-363.

93. Pique M, Mougeot JL, Geldreich A, Guidasci T, Mesnard JM, et al. (1995) Sequence of a cauliflower mosaic virus strain infecting solanaceous plants. Gene 155: 305-306.

94. N U (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Japanese Journal of Botany 7: 389-452.

95. Daubert S, Routh G (1990) Point mutations in cauliflower mosaic virus gene VI confer host-specific symptom changes. Mol Plant Microbe Interact 3: 341-345.

96. Schoelz J, Shepherd RJ, Daubert S (1986) Region VI of cauliflower mosaic virus encodes a host range determinant. Mol Cell Biol 6: 2632-2637.

97. Schoelz JE, Shepherd RJ (1988) Host range control of cauliflower mosaic virus. Virology 162: 30-37.

98. Daubert SD, Schoelz J, Debao L, Sheperd RJ (1984) Expression of disease symptoms in caulifloqer mosaic virus genomic hybrids. J Mol Appl Genet 2: 537-547.

99. Zijlstra C, Schärer-Hernandez N, Gal S, Hohn T (1996) *Arabidopsis thaliana* expressing the cauliflower mosaic virus ORF VI transgene has a late flowering phenotype. Virus Genes 13: 5-17.

100. Geri C, Cecchini E, Giannakou ME, Covey SN, Milner JJ (1999) Altered patterns of gene expression in Arabidopsis elicited by cauliflower mosaic virus (CaMV) infection and by a CaMV gene VI transgene. Mol Plant Microbe Interact 12: 377-384.

101. Palanichelvam K, Cole A, Shababi M, Schoelz J (2000) Agroinfection of *Cauliflower mosaic virus* gene VI elicits hypersensitive response in *Nicotiana* species. Mol Plant Microbe Interact 13: 1275-1279.

102. Love AJ, Yun BW, Laval V, Loake GJ, Milner JJ (2005) Cauliflower mosaic virus, a compatible pathogen of Arabidopsis, engages three distinct defense-signaling pathways and activates rapid systemic generation of reactive oxygen species. Plant Physiol 139: 935-948.

103. Anderson EJ, Qui SG, Schoelz JE (1991) Genetic analysis of determinants of disease severity and virus concentration in cauliflower mosaic virus. Virology 181: 647-655.

104. Qiu SG, Wintermantel WM, Sha Y, Schoelz JE (1997) Light-dependent systemic infection of solanaceous species by cauliflower mosaic virus can Be conditioned by a viral gene encoding an aphid transmission factor. Virology 227: 180-188.

105. Covey SN, Al-Kaff NS, Langara A, Turner DS (1997) Plants combat infection by gene silencing. Nature 385: 781-782.

106. Al-Kaff N, Covey S, Kreike M, Page A, Pinder R, et al. (1998) Transcriptional and posttranscriptional plant gene silencing in response to a pathogen. Science 279: 2113-2115.

107. Moissiard G, Voinnet O (2006) RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. Proc Natl Acad Sci U S A 103: 19593-19598.

108. Hohn T, Akbergenov R, Pooggin MM (2007) Production and Transport of the Silencing Signal in Transgenic and Virus-Infected Plant Systems. In: Waigmann E, Heinlein M, editors. Viral Transport in Plants: Springer-Verlag Berlin Heidelberg. pp. 127-157.

109. Dunoyer P, Voinnet O (2008) Mixing and matching: the essence of plant systemic silencing? Trends Genet 24: 151-154.

110. Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, et al. (2006) Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res 34: 6233-6246.

111. Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, et al. (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science 313: 68-71.

112. Moissiard G (2007) Induction, Suppression, Amplification of RNA Silencing during Viral Infection. Strasbourg: Louis Pasteur. 126 p.

113. Tang W, Leisner S (1998) Methylation of nonintegrated multiple copy DNA in plants. Biochem Biophys Res Commun 245: 403-406.

114. Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol 4: e363.

115. Blanc S (2007) Virus Transmission-Getting Out and In. In: Waigmann E, Heinlein M, editors.

Viral Transport in Plants: Springer-Verlag Berlin Heidelberg. pp. 1-28.

116. Fereres A, Collar JL (2001) Analysis of non-circulative transmission by electrical penetration graphs. In: Harris K, Smith OP, Duffus JE, editors. Virus-Insect-Plant interactions. San Diego: Academic Press. pp. 87-109.

117. Powell G, Tosh CR, Hardie J (2006) Host plant selection by aphids: behavioral, evolutionary, and applied perspectives. Annu Rev Entomol 51: 309-330.

118. de Vos M, Kim JH, Jander G (2007) Biochemistry and molecular biology of Arabidopsis-aphid interactions. Bioessays 29: 871-883.

119. Tjallingii WF (2006) Salivary secretions by aphids interacting with proteins of phloem wound responses. J Exp Bot 57: 739-745.

120. Will T, Tjallingii WF, Thonnessen A, van Bel AJ (2007) Molecular sabotage of plant defense by aphid saliva. Proc Natl Acad Sci U S A 104: 10536-10541.

121. Hébrard E, Froissart R, Louis C, Blanc S (1999) Les modes de transmission des virus phytopathogènes par vecteurs. Virologie 3: 35-48.

122. Powell G (2005) Intracellular salivation is the aphid activity associated with inoculation of non-persistently transmitted viruses. J Gen Virol 86: 469-472.

123. Froissart R, Michalakis Y, Blanc S (2002) Helper component-transcomplementation in the vector transmission of plant viruses. Phytopathology 92: 576-579.

124. Walker GP (2000) A beginner's guide to electronic monitoring of Homopteran probing behavior. In: Walker GP, A. BE, editors. Principles and Applications of Electronic Monitoring and other Techniques in the Study of Homopteran Feeding Behavior: Entomological Society of America. pp. 14-40.

125. Tjallingii WF, Prado E (2001) Analysis of Circulative Transmission by Electrical Penetation Graphs. In: Harris KF, Smith OP, Duffus JE, editors. Virus-Insect-Plant Interactions. San Diego: Academic Press. pp. 69-85.

126. Tjallingii WF, Hogen Esch T (1993) Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. Physiol Entomol 18: 317-328.

127. Powell G, Pirone T, Hardie J (1995) Aphid stylet activities during potyvirus acquisition from plants and an in vitro system that correlate with subsequent transmission. European Journal of Plant Pathology 104: 411-420.

128. Martin B, Collar JL, Tjallingii WF, Fereres A (1997) Intracellular ingestion and salivation by aphids may cause the acquisition and inoculation of non-persistently transmitted plant viruses. J Gen Virol 78: 2701-2705.

129. Kennedy JS, Day MF, Eastop VF (1962) A conspectus of aphids as vectors of plant viruses. London: Commonwealth Institue of Entomology. 114 p.

130. Broadbent L (1957) Investigation of virus diseases of brassica crops; Broadbent L, editor. Cambridge, UK: Cambridge University Press. 99 p.

131. Blackman RL, Eastop VF (1984) Aphids on the World's Crops: an Identification Guide. New York: John Wiley & Sons. 466 p.

132. Palacios I, Drucker M, Blanc S, Leite S, Moreno A, et al. (2002) Cauliflower mosaic virus is preferentially acquired from the phloem by its aphid vectors. J Gen Virol 83: 3163-3171.

133. Ng JC, Falk BW (2006) Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. Annu Rev Phytopathol 44: 183-212.

134. Bouchery Y, Givord L, Monestiez P (1990) Comparison of short- and long-feed transmission of the cauliflower mosaic virus Cabb-S strain and S delta II hybrid by two species of aphid: Myzus persicae (Sulzer) and Brevicoryne brassicae (L.). Res Virol 141: 677-683.

135. Markham PG, Pinner MS, Raccah B, Hull R (1987) The acquisition of a caulimovirus by different aphid species : comparison with a potyvirus. Ann Appl Biol 111: 571-587.

136. Moreno A, Palacios I, Blanc S, Fereres A (2005) Intracellular salivation is the mechanism involved in the inoculation of cauliflower mosaic vius by its major vectors *Brevicoryne brassicae* and *Myzys persicae*. Annals of the Entomological Society of America 98: 763-769.

137. Fereres A (2005) Communication personnelle.

138. Bradley RHE, Ganong RY (1955) Evidence that potato virus Y is carried near the tip of the stylet of the aphid vector Myzus persicae (Sulz.). Can J Microbiol 1: 775-782.

139. Bradley RHE, Ganong RY (1955) Some effects of formaldehyde on potato virus Y *in vitro*, and the ability of aphids to transmit the virus when their stylets are treated with formaldehyde. Can J Microbiol 1: 783-793.

140. Roossinck MJ (1997) Mechanisms of plant virus evolution. Annu Rev Phytopathol 35: 191-209.

141. Drake JW, Holland JJ (1999) Mutation rates among RNA viruses. Proc Natl Acad Sci U S A 96: 13910-13913.

142. Domingo E, Sabo D, Taniguchi T, Weissmann C (1978) Nucleotide sequence heterogeneity of an RNA phage population. Cell 13: 735-744.

143. Domingo E (2000) Viruses at the edge of adaptation. Virology 270: 251-253.

144. Domingo E, Holland J (1994) In: Morse SS, editor. Mutations and Rapid Evolution of RNA Viruses. New York: Raven Press. pp. 161-184.

145. Domingo E, Holland J, Ahlquist P (1998) Rna Genetics. Boca Raton: CRC Press. 272 p.

146. Eigen M (1993) Viral quasispecies. Sci Am 269: 42-49.

147. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, et al. (1982) Rapid evolution of RNA genomes. Science 215: 1577-1585.

148. Holland JJ, De La Torre JC, Steinhauer DA (1992) RNA virus populations as quasispecies. Curr Top Microbiol Immunol 176: 1-20.

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

150. Bull JJ, Meyers LA, Lachmann M (2005) Quasispecies made simple. PLoS Comput Biol 1: e61.

151. Wilke CO (2005) Quasispecies theory in the context of population genetics. BMC Evol Biol 5: 44.

152. Biebricher CK, Eigen M (2006) What is a quasispecies? Curr Top Microbiol Immunol 299: 1-31.

153. Coffin JM (1995) HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science 267: 483-489.

154. Domingo E, Escarmis C, Sevilla N, Moya A, Elena SF, et al. (1996) Basic concepts in RNA virus evolution. Faseb J 10: 859-864.

155. Eigen M, Biebricher C (1988) Sequence space and quasispecies distribution. In: Domingo E, Holland JJ, Ahlquist P, editors. RNA Genetics: Variability of RNA Genomes. Boca Raton: CRC Press. pp. 211-245.

156. Anderson JP, Daifuku R, Loeb LA (2004) Viral error catastrophe by mutagenic nucleosides. Annu Rev Microbiol 58: 183-205.

157. Pariente N, Airaksinen A, Domingo E (2003) Mutagenesis versus inhibition in the efficiency of extinction of foot-and-mouth disease virus. J Virol 77: 7131-7138.

158. Pariente N, Sierra S, Lowenstein PR, Domingo E (2001) Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. J Virol 75: 9723-9730.

159. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344-348.

160. de la Iglesia F, Elena SF (2007) Fitness Declines in Tobacco Etch Virus upon Serial Bottleneck Transfers. J Virol.

161. Chao L (1992) Evolution of sex in RNA viruses. Trends Ecol Evol 7(5): 147-151.

162. Muller HJ (1964) The Relation of Recombination to Mutational Advance. Mutat Res 106: 2-9.

163. Chao L (1988) Evolution of sex in RNA viruses. J Theor Biol 133: 99-112.

164. Lynch M, Burger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. Journal of Heredity 84: 339-344.

165. Chao L (1990) Fitness of RNA virus decreased by Muller's ratchet. Nature 348: 454-455.

166. de la Pena M, Elena SF, Moya A (2000) Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. Evolution Int J Org Evolution 54: 686-691.

167. Duarte E, Clarke D, Moya A, Domingo E, Holland J (1992) Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc Natl Acad Sci U S A 89: 6015-6019.

168. Escarmis C, Davila M, Charpentier N, Bracho A, Moya A, et al. (1996) Genetic lesions associated with Muller's ratchet in a RNA virus. J Mol Biol 264(2): 255-267.

169. Yuste E, Sanchez-Palomino S, Casado C, Domingo E, Lopez-Galindez C (1999) Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J Virol 73: 2745-2751.

170. Grimsley N, Hohn T, Hohn B (1986) Recombination in a plant virus : template-switching in cauliflower mosaic virus. EMBO J 5: 641-646.

171. Al-Kaff N, Covey SN (1994) Variation in biological properties of cauliflower mosaic virus clones. J Gen Virol 75: 3137-3145.

172. Pennington RE, Melcher U (1993) In planta deletion of DNA inserts from the large intergenic region of cauliflower mosaic virus DNA. Virology 192: 188-196.

173. Kiss-Laszlo Z, Blanc S, Hohn T (1995) Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J 14: 3552-3562.

174. Scholthof HB, Wu FC, Richins RD, Shepherd RJ (1991) A naturally occurring deletion mutant of figwort mosaic virus (caulimovirus) is generated by RNA splicing. Virology 184: 290-298.

175. Chenault KD, Melcher U (1994) Patterns of nucleotide sequence variation among cauliflower mosaic virus isolates. Biochimie 76: 363-375.

176. Blanc S, Froissart R, Martinière A, Uzest M, Gratraud P, et al. Genetic and biological diversity generated within a population of Cauliflower mosaic virus during a single host infection; 2004 October; Armore, Oklahoma, USA.

177. Froissart R (2002) Etude expérimentale de l'évolution du mode de transmission du Cauliflower mosaic virus. Paris VI: Pierre et Marie Curie. 141 p.

178. Olkkonen VM, Bamford DH (1989) Quantitation of the adsorption and penetration stages of bacteriophage phi 6 infection. Virology 171: 229-238.

179. Büchen-Osmond C ICTVdB - The Universal Virus Database, version 4.: International Committee on Taxonomy of Viruses.

180. Bull JC, Godfray HC, O'Reilly DR (2001) Persistence of an occlusion-negative recombinant nucleopolyhedrovirus in Trichoplusia ni indicates high multiplicity of cellular infection. Appl Environ Microbiol 67: 5204-5209.

181. Jung A, Maier R, Vartanian JP, Bocharov G, Jung V, et al. (2002) Multiply infected spleen cells in HIV patients. Nature 418: 144.

182. Meyerhans A, Jung A, Maier R, Vartanian JP, Bocharov G, et al. (2003) The non-clonal and

transitory nature of HIV in vivo. Swiss Med Wkly 133: 451-454.

183. Dang Q, Chen J, Unutmaz D, Coffin JM, Pathak VK, et al. (2004) Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc Natl Acad Sci U S A 101: 632-637.

184. Cicin-Sain L, Podlech J, Messerle M, Reddehase MJ, Koszinowski UH (2005) Frequent coinfection of cells explains functional in vivo complementation between cytomegalovirus variants in the multiply infected host. J Virol 79: 9492-9502.

185. Dietrich C, Maiss E (2003) Fluorescent labelling reveals spatial separation of potyvirus populations in mixed infected Nicotiana benthamiana plants. J Gen Virol 84: 2871-2876.

186. Takahashi T, Sugawara t, Yamatsuta T, Isogai M, Natsuaki T, et al. (2007) Analysis of the Spatial Distribution of Identical and Two Distinct Virus Populations Differently Labeled with Cyan and Yellow Fluorescent Proteins in Coinfected Plants. Phytopathology 97 1200-1206.

187. Jridi C, Martin JF, Marie-Jeanne V, Labonne G, Blanc S (2006) Distinct viral populations differentiate and evolve independently in a single perennial host plant. J Virol 80: 2349-2357.

188. Giritch A, Marillonnet S, Engler C, van Eldik G, Botterman J, et al. (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. Proc Natl Acad Sci U S A 103: 14701-14706.

189. Froissart R, Roze D, Uzest M, Galibert L, Blanc S, et al. (2005) Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. PLoS Biol 3: e89.

190. Monsion B, Duborjal H, Blanc S (2008) Quantitative Single-letter Sequencing: a method for simultaneously monitoring numerous known allelic variants in single DNA samples. BMC Genomics 9: 85.

191. Jridi C (2008) Accumulation of CaMV in turnip (communication personnelle).

192. Turner PE, Chao L (1999) Prisoner's dilemma in an RNA virus. Nature 398: 441-443.

193. Bergstrom CT, McElhany P, Real LA (1999) Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc Natl Acad Sci U S A 96: 5095-5100.

194. Elena SF, Sanjuan R, Borderia AV, Turner PE (2001) Transmission bottlenecks and the evolution of fitness in rapidly evolving RNA viruses. Infect Genet Evol 1: 41-48.

195. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3: 336-340.

196. Manrubia SC, Lazaro E (2006) Viral evolution. Physics of Life Reviews 3: 65-92.

197. Walker HL, Pirone TP (1972) Particle numbers associated with mechanical and aphid transmission of some plant viruses. Phytopathology 62: 1283-1288.

198. Pirone TP, Thornbury DW (1988) Quantity of virus required for aphid transmission of a potyvirus. Phytopathology 78: 104-107.

199. Ali A, Li H, Schneider WL, Sherman DJ, Gray S, et al. (2006) Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. J Virol 80: 8345-8350.

200. Moury B, Fabre F, Senoussi R (2007) Estimation of the number of virus particles transmitted by an insect vector. Proc Natl Acad Sci U S A 104: 17891-17896.

201. Febvay G, Rahbe Y, Van Helden M (1996) MacStylet, software to analyse electrical penetration graph data on the Macintosh. Entomologia Experimentalis et Applicata 80: 105-108.

202. Margaritopoulos JT, Souri A, Tsitsipis JA (2004) Feeding studies on gynoparae, males and apterous virginoparae of Myzus persicae on potato by electrical penetration graph. Journal of Applied Entomology 128: 522-527.

203. Styer LM, Kent KA, Albright RG, Bennett CJ, Kramer LD, et al. (2007) Mosquitoes inoculate high doses of West Nile virus as they probe and feed on live hosts. PLoS Pathog 3: 1262-1270.

204. Smith DR, Aguilar PV, Coffey LL, Gromowski GD, Wang E, et al. (2006) Venezuelan equine encephalitis virus transmission and effect on pathogenesis. Emerg Infect Dis 12: 1190-1196.

205. Scholle F, Girard YA, Zhao Q, Higgs S, Mason PW (2004) trans-Packaged West Nile virus-like particles: infectious properties in vitro and in infected mosquito vectors. J Virol 78: 11605-11614.

206. Smith DR, Adams AP, Kenney JL, Wang E, Weaver SC (2008) Venezuelan equine encephalitis virus in the mosquito vector Aedes taeniorhynchus: infection initiated by a small number of susceptible epithelial cells and a population bottleneck. Virology 372: 176-186.

207. Pirone TP, Blanc S (1996) Helper-dependent vector transmission of plant viruses. Annu Rev Phytopathol 34: 227-247.

208. Maniatis T, Fritsh EF, Sambrook J (1982) Molecular cloning : a laboratory manual. New-York: Cold Spring Harbor Laboratory, Cold Spring Harbor.

209. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, et al. (2003) Current Protocols in Molecular Biology; Chanda VB, editor: John Wiley & Sons Inc. approx. 5,300 p.

210. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 277: 680-684.

211. Monsion B (2003) Détermination de la taille du goulot d'étranglement induit par un insecte vecteur sur une population virale [DEA]. Saint-Christol-les-Alès: Université Montpellier II.

Annexes

1 Tests préliminaires de transmission par puceron

Ces résultats ont été acquis en début de thèse avec les clones viraux « Div » du CaMV qui ne contiennent pas le gène codant pour P2 contrairement au VIT construits ensuite durant la thèse. Les Div comportent donc le même type de marqueurs que les VIT mais insérés dans un génome Cabb-S (pCa37) préalablement délété du gène II. L'infectivité, la stabilité et la détection de ces clones Div ont été démontrées lors de mon Diplôme d'Etudes Approfondies (DEA) [211].

1.1 Evaluation de la neutralité des clones Div

A partir d'une solution mixte (nommée SM3) contenant 9 clones Div, 3 plants de navet ont été infectés mécaniquement. Trois semaines après infection, l'ensemble des feuilles des trois plantes infectées de manière systémique est collecté, poolé et broyé, afin d'en concentrer les particules virales avant d'effectuer (i) l'inoculation mécanique de 3 nouvelles plantules, et (ii) la purification de l'ADN viral. Après trois transferts similaires et successifs, la présence des différents clones Div à chacun des trois passages est évaluée par hybridation moléculaire (figure A1). Au fil des 3 passages, le clone Div 8 est toujours représenté très faiblement au sein des populations ; les clones Div 7, 9 et 17 semblent progressivement disparaître et donc être contre sélectionnés ; seuls les CaMV marqués Div 2, 3, 4, 10 et 11 donnent un signal constant au fil des passages et ont donc été considérés comme neutres ou quasi-neutres dans nos conditions expérimentales.

1.2 Préparation d'une solution mixte : « Mix5Div »

Les cinq clones Div (2, 3, 4, 10 et 11) présentant une neutralité apparente lors d'infections successives ont été choisis, afin d'infecter indépendamment des plantes et de purifier leurs particules virales (voir Matériels et Méthodes section 2.4.1). Ces solutions virales ont été dosées par un spectrophotomètre afin de préparer un mélange équimolaire nommé « Mix5Div » qui a servi à réaliser toute une batterie de tests de transmission (voir ci-dessous).

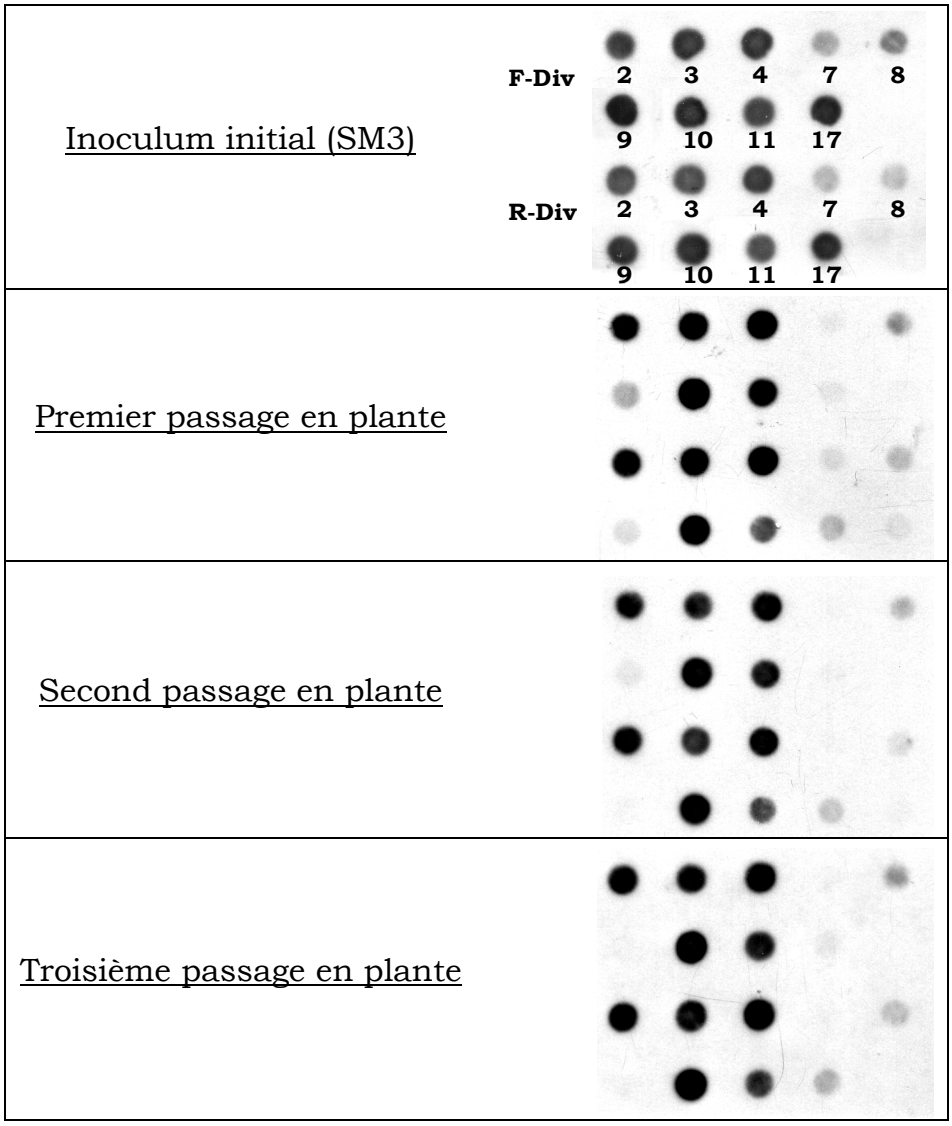


Figure A1 : Evaluation de la neutralité des clones Div. L'hybridation moléculaire des populations d'ADN viral, sur des membrane de nylon comportant les oligonucléotides forward et reverse correspondant aux différents marqueurs Div, met en évidence une neutralité apparente pour les clones 2, 3, 4, 10 et 11. Les autres clones Div disparaissent au fil des trois passages ou présentent une faible intensité tout au long de l'expérience.

1.3 Tests de transmissions avec *Myzus persicae*

1.3.1 Acquisition sur feuille infectée

Trois semaines après l'inoculation mécanique de plantules avec le Mix5Div, des pucerons nourris au préalable avec une solution contenant P2 au travers d'une membrane de Parafilm®, ont été déposés sur ces plantes infectées afin d'acquérir le virus. La durée d'acquisition s'étend de 2 à 27 minutes[5] ; les 25 minutes ayant été nécessaires pour disposer un seul puceron sur chacune des 360 plantes tests. Quatre semaines plus tard, 36,7% des plantes (132/360) sont infectées de manière systémique. L'ADN viral de 27 plantes entières (choisies de façon aléatoire) a été purifié par ultracentrifugation et analysé par génotypage (voir Matériels et Méthodes du Chapitre III). Les résultats illustrés par la figure A2 mettent en évidence le faible nombre de clones transmis : de 1 à 3 au maximum avec une moyenne de 1,63.

Toutefois, ne connaissant pas précisément la répartition des différents clones viraux au sein d'une feuille, nous pensions que les résultats observés pouvaient être sous-estimés du fait de l'éventuelle présence de plages de tissu foliaire infectées par un seul clone.

1.3.2 Acquisition sur membrane

Afin d'éliminer tout problème de répartition hétérogène du virus au sein des tissus végétaux, nous avons entrepris une large série de tests de transmission avec des acquisitions à travers membranes de Parafilm® (tableau A1). Certaines séries de tests ont offert la possibilité d'identifier les paramètres déterminants quant au taux de transmission finalement observé 4 semaines après inoculation. Par exemple, les tests du 18.XII.2003 ont validé le choix de ne faire qu'un seul repas d'acquisition avec P2-P3-virus par la suite. Malgré les tentatives d'optimisation, les taux de transmission obtenus n'ont pas été élevés. De ce fait, par deux fois (les 3.III.2004 et 18.III.2004) nous avons réalisé des tests contrôles avec d'autres purifications de protéines et de virus. Ceux-ci mettent en évidence la faible

[5] Il est fort peu probable que le puceron puisse accéder au phloème durant une période aussi brève.

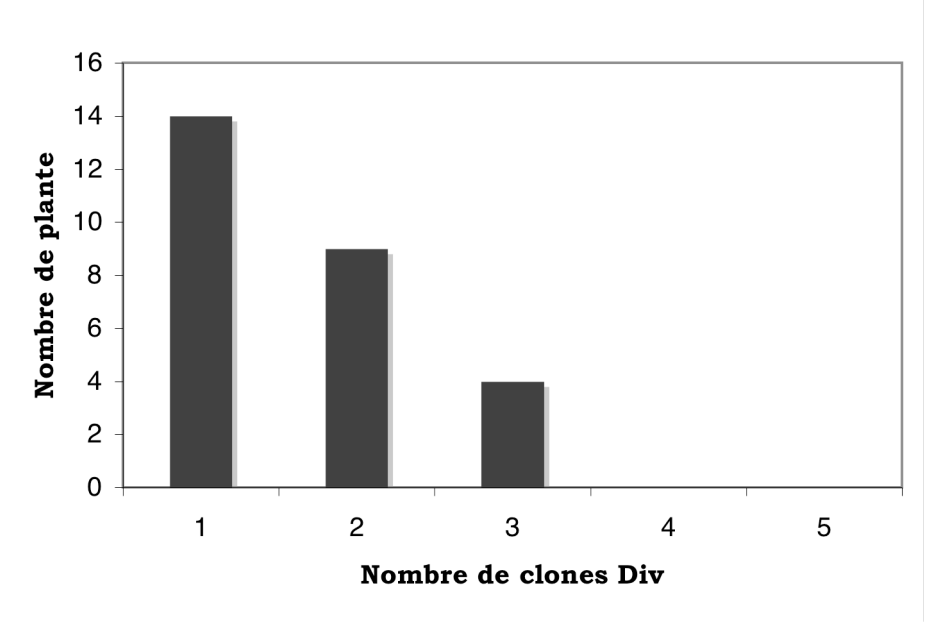


Figure A2 : Distribution du nombre de clones Div transmis par un puceron de plante à plante. Suite a un premier repas « P2 », chaque puceron a pu acquérir du virus sur une plante infectée par le Mix5Div (5 clones Div). L'analyse de 27 plantes infectées met en évidence le faible nombre de clone transmis : de 1 à 3 au maximum.



Figure A3 : Distribution du nombre de clones Div transmis par un puceron (acquisition sur membrane). Suite a un repas artificiel « P2-P3-virus » où virus correspond au Mix5Div (5 clones Div), les pucerons sont déposés à l'unité sur des plantes tests. L'analyse des 73 plantes infectées met en évidence le faible nombre de clones transmis : de 1 à 4 au maximum.

transmissibilité du Mix5Div par le puceron-vecteur en comparaison d'une souche sauvage du CaMV (Bji).

Néanmoins, sur l'ensemble des 24 dates où des tests de transmission ont été réalisés à 1 puceron par plante (tableau A1), 84 plantes sur 3120 (soit 2,7%) ont été infectées. L'analyse de 73 d'entre-elles par génotypage, révèle la présence de 1 à 4 clones Div (figure A3) avec une moyenne à 1,89 ce qui est légèrement supérieur aux résultats obtenus lors des tests de transmission à partir de feuilles infectées, mais qui préfigure toujours l'existence d'un fort goulot d'étranglement lié à la transmission par vecteur.

Ces résultats issus de tests avec un faible succès de transmission semblaient indiquer que les conditions expérimentales n'étaient pas optimales et que donc, nous risquions de sous-estimer le nombre de particules virales transmises par un puceron. L'utilisation de clone de CaMV ne produisant pas P2 nous interdisait de plus d'opérer à des transmissions dans des conditions « naturelles ».

Ces remarques expliquent que nous ayons abandonné ces séries d'expériences pour tout reprendre à zéro avec des clones de CaMV marqués exprimant P2, donc naturellement transmissibles de plante à plante. Ces nouveaux clones sont nommés VIT et l'ensemble des analyses qu'ils ont permis de développer constitue le corps des 3 chapitres de cette thèse.

Date	Variable(s) testée(s)	Composition des solutions d'acquisition en µL						Plantes infectées		Infection en %		N° du test à la récolte
		P2	P3	Virus	Sucrose	SES	Autre	10 puc/plte	1 puc/plte	10 puc/plte	1 puc/plte	
18.XII.2003	1 seul repas	100	5 d	10	40 (30%)	-	-	19/78	-	24,4	-	-
	1 repas P2 puis 1 repas P3-virus	100	-	-	30 (30%)	-	-	6/74	-	8,1	-	-
		-	7 d	13	40 (30%)	-	70 Tp				-	
15.I.2004	P2-P3 co-produitent + SES	150 #		39	50 §	-	-	20/40	14/120	50	11,7	1
16.I.2004	1 seul repas avec + de P3 et de virus	80	20 d	20	40	-	-	0/40	0/120			-
22.I.2004	Quantité de SES	175 #		25	50	-	-	11/40	6/120	27,5	5	2
		175 #		25	25 + 25 §	-	-	14/40	13/120	35	10,9	3
28.I.2004	Ratio des protéines : 1/3 P2 et 2/3 P3	100 £		25	50	-	-	9/40	8/120	22,5	6,7	4
	Ratio des protéines : 2/3 P2 et 1/3 P3	100 $		25	50	-	-	17/40	17/120	42,5	14,2	5
4.II.2004	P3 en SES ou PBS en présence de beaucoup de virus	50	25 p	50	50	25	-	0/40	0/120	-	-	-
		50	25	50	50	25	-	0/40	4/120	-	3,3	6
11.II.2004	P3 en SES en présence de peu de virus	50	25	5	50	25	45 ED	0/40	0/120	-	-	-
		50	25	5	50	25	45 ED	0/40	0/120	-	-	-
18.II.2004		50	25	5	50	25	45 ED	0/40	0/120	-	-	-
		50	25	5	50	25	45 ED	0/40	0/120	-	-	-
25.II.2004	Effet de la quantité de P2	180	10 d	30	80	-	-	7/40	2/120	17,5	1,7	7
		100	10 d	30	80	80	-	12/40	5/120	30	4,2	8
		40	10 d	30	80	140	-	1/40	1/120	2,5	0,8	9
3.III.2004	Témoin avec CaMV et protéines de Martin	50 h	20 p	10 ‡	40	80	-	14/40	-	35		-
	Test témoin avec Mix5Div et quantité variable de de P3	50 h	20 p	30	40	60	-	0/40	1/160	-	0,6	10
		50 h	10 p	30	40	70	-	3/40	0/200	-		11
10.III.2004	Acquisition 2-9 min. sur plante	100	-	-	30	-	-	-	45/120	-	37,5	2'
	Acquisition 10-17 min. sur plante	100	-	-	30	-	-	-	49/120	-	40,8	10'
	Acquisition 20-27 min. sur plante	100	-	-	30	-	-	-	38/120	-	31,7	20'
18.III.2004	Témoin avec CaMV et protéines de Martin	50 *	20 p	10 ‡	40	80	-	10/40	-	25	-	-
	Test témoin avec Mix5Div et quantité variable de de P3	50 *	20 p	30	40	60	-	1/40	0/160	2,5	-	12
		50 *	20 p	30	40	70	-	1/40	0/200	2,5	-	13
25.III.2004	Quantité P2-P3 avec incubation O.N. du mélange P2-P3-virus	120 $		45	40 (80%)	-	-	9/40	7/160	22,5	4,4	14
		60 $		40	40 (80%)	60	-	0/40	1/120	-	0,8	15
		20 $		40	40 (80%)	100	-	0/40	0/80	-		-
1.IV.2004	Ratio P2/P3 avec incubation O.N. du mélange P2-P3-virus	150 #		45	35 (80%)	-	-	6/40	1/120	15	0,8	16
		80 #		45	35 (80%)	-	50 P2	3/40	4/120	7,5	3,3	17

Tableau A1 : Synthèse des tests de transmission réalisés à partir du Mix5Div. A partir du Mix5Div de nombreux tests de transmission – sur membrane principalement – ont été entrepris. Les protéines P2 et P3 sont produites en cellules Sf9, le virus servant à l'acquisition correspond au Mix5Div et le sucrose est à 60% dans de l'eau distillée sauf indications contraires signalées : **#**, P2 et P3 co-produites avec une inoculation à 50/50 des baculovirus recombinants correspondants dans des cellules Sf9 ; **§**, sucrose 60 % en SES ; **£**, P2 et P3 co-produites en cellules Sf9 avec 1/3 et 2/3 comme ratios respectifs des baculovirus recombinants lors de l'inoculation ; **$**, P2 et P3 co-produites en cellules Sf9 avec 2/3 et 1/3 comme ratios respectifs lors de l'inoculation ; **‡**, particules virales purifiées de la souche B-JI du CaMV ; *****, P2-94 avec un histidine-tag en N-terminal exprimée en Sf9 ; **d**, P3 produite en bactéries, puis purifiée ; **h**, P2 avec un histidine-tag en N-terminal exprimée en cellules Sf9 ; **p**, P3 produite en cellules Sf9 mais resuspendue avec du PBS 1X ; **ED**, eau distillée ; **SES**, Standard Extract Solution ; **Tp**, tampon 10mM Tris pH7 avec 0,25mM $MgCl_2$; **O.N.** overnight.

2 Programme « R » pour calculer la MOI

Ci-dessous, figurent les détails bruts du modèle que nous avons utilisé pour calculer la MOI moyenne. Ce programme a été écrit en totalité par G. Thébaud (INRA, Montpellier), suite à de multiples discussions interactives ayant permis de s'assurer de l'adéquation de nos expérimentations biologiques avec les hypothèses sous-jacentes au modèle (voir chapitre II section 2.8.).

```
# Chargement à partir d'un fichier csv #

LitDon<-function(path)
  {data<-read.csv(path,sep=";",dec=",",as.is=T,header=T)
   jDon<<-as.matrix(unname(data))
  }



# Estimation et IC pour la proportion d'individus co-infectés par les 2 génomes
# et test d'adéquation (par bootstrap paramétrique utilisant la statistique du
# Chi-2). La 1ère difficulté est la gestion des 0 ; la 2ème difficulté est que
# le paramètre a été estimé d'après les données (il faut donc réestimer p pour
# chaque jeu de données simulé.

MaxVraisIC<-function(x,Binf,Bsup,conf.level=.95)
  {    # Equations pour Pcoi et dPcoi
   Pcoi<-function(l,x,i) {(1-exp(-l*x[i,1]))*(1-exp(l*(x[i,1]-1)))/(1-exp(-l))}
   dPcoi<-function(l,x,i)
     {x[i,1]*(exp(-l*x[i,1])-exp(l*(x[i,1]-1)))/(1-exp(-l))+
      (exp(l*x[i,1])+exp(-l*x[i,1])-2)/((1-exp(-l))*(exp(l)-1))
     }
   Pco<-function(l,vect) {(1-exp(-l*vect))*(1-exp(l*(vect-1)))/(1-exp(-l))}

       # Calcul de la vraisemblance
   vrais<-function(l,vbor)
     {v<-1
      for (i in 1:nrow(x))
        v<-v*choose(x[i,2],x[i,3])*((Pcoi(l,x,i))^x[i,3])*
           ((1-Pcoi(l,x,i))^(x[i,2]-x[i,3]))
      return(v-vbor)
     }

       # Estimation de lambda par maximum de vraisemblance
   Estimation<-function(obs,x)
     {f<-function(l)
        {somme<-0
         for (i in 1:nrow(x))
           somme<-somme+((obs[i]/Pcoi(l,x,i))-x[i,2])*dPcoi(l,x,i)/(1-Pcoi(l,x,i))
         return(somme)
        }
      estim<-uniroot(f,lower=Binf,upper=Bsup,tol=1e-8)$root
      return(estim)
     }
   estim<-Estimation(x[,3],x[,1:2])
```

```
      # Intervalle de confiance asymptotique
   Vmax<-vrais(estim,vbor=0)
   Vbornes<-Vmax*exp(-.5*qchisq(conf.level,1))
   B1<-uniroot(vrais,c(Binf,estim),vbor=Vbornes,tol=1e-8)$root
   B2<-uniroot(vrais,c(estim,Bsup),vbor=Vbornes,tol=1e-8)$root

      # Sortie graphique
   W1<-(3*B1-B2)/2
   W2<-(3*B2-B1)/2
   x11()
   curve(vrais(x,0),W1,W2,xlab="Multiplicité d'infection",ylab="Vraisemblance")
   segments(estim,Vmax/20,estim,Vmax,col="red")
   text(estim,Vmax/30,signif(estim,3),col="red")
   segments(B1,vrais(B1,0),B2,vrais(B2,0),col="blue")
   segments(B1,0,B1,vrais(B1,0),col="blue")
   segments(B2,0,B2,vrais(B2,0),col="blue")
   text(B1,0,signif(B1,3),adj=c(0,1),col="blue")
   text(B2,0,signif(B2,3),adj=c(1,1),col="blue")

      # Sortie texte
   cat("\n")
   cat("Matrice des effectifs observés : \n")
   colnames(jDon)<-c("Proportion de V4","Nbre total infectés",
                     "Nbre co-infectés")
   print(jDon)
   cat("\n \n")
   cat("Estimation de la multiplicité d'infection :",signif(estim,3),"\n")
   cat("Valeur maximale de la vraisemblance :",signif(Vmax,4),"\n \n")
   cat("        Calcul de l'intervalle de confiance basé \n",
       "          sur le rapport des vraisemblances \n")
   cat("Intervalle de confiance asymptotique à",conf.level*100,"% : [",
       signif(B1,3),",",signif(B2,3),"] \n \n \n")
  }



# Lance l'analyse #

AnalyseMOI<-function(chemin,Binf,Bsup,conf.level=.95)
  {malheur<-0
   LitDon(chemin)
   MaxVraisIC(jDon,Binf,Bsup,conf.level)
  }



# Pour appeler la fonction qui lance le modèle final intégrant
# la censure sur les cellules non-infectées
# A copier-coller dans R

rm(list=ls())
path0<-"C:/Documents/Goulot de thèse/"
source(paste(path0,"MOI/Estim_MOI.R",sep=""))
path<-paste(path0,"MOI/Cinétique/InterFeuilles/60dpi.csv",sep="")

AnalyseMOI(chemin=path,saisie=F,Binf=1e-6,Bsup=50,conf.level=.95)
```

3 Détails sur l'épigraphe

« Que nous soient donnés la pluie, la santé et l'enrichissement spirituel »
I:42

Est une bénédiction que prononce Brown le frontalier dans le tome I du cycle de *La Tour Sombre* écrit par Stephen KING

Voici en supplément quelques autres proverbes et dictons de l'Entre-Deux-Mondes délicatement sélectionnés parmi les 4 premiers volumes : I, Le Pistolero ; II, Les Trois Cartes ; III, Terres Perdues ; IV, Magie et Cristal.

« Il va falloir aller très vite ou on va se retrouver dans le four à houblon »
II:138

« La faute est toujours au même endroit, mes bébés : chez celui qui est assez faible pour ne pas l'assumer »
II:174

« Seul un fou s'imagine qu'il rêve avant de se réveiller »
III:318

« Le *ka* était une roue, dont le dessein était de tourner, et elle finissait toujours par revenir au point de départ »
III:488

« La roue du *ka* tourne, et le monde change »
III:499

« Si c'est le *ka*, il viendra comme le vent »
IV:215

« Les imbéciles sont les seuls sur terre à pouvoir absolument compter récolter ce qu'il méritent »
IV:219

« Ne renâcle point devant ta donne aux cartes ou dans la vie »
IV:233

« Que vos journées soient longues et vos nuits plaisantes »
IV:266