

Université Louis Pasteur
Strasbourg I
2004

THESE

présentée à la
FACULTE DES SCIENCES DE LA VIE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR
Domaine : BIOLOGIE MOLECULAIRE ET CELLULAIRE

par

Didier LINK

Étude de la variabilité moléculaire et du pouvoir pathogène d'isolats naturels du virus
des nervures jaunes et nécrotiques de la betterave : caractérisation des protéines
codées par les ARN-3 et -5 de différents isolats.

Soutenue le 13 décembre 2004 devant la Commission d'Examen :

Jean-Luc SOUCIET	RAPPORTEUR INTERNE
Mireille JACQUEMOND	RAPPORTEUR EXTERNE
Michel RAVELONANDRO	RAPPORTEUR EXTERNE
Olivier LEMAIRE	EXAMINATEUR
David GILMER	DIRECTEUR DE THESE

Institut de Biologie Moléculaire des Plantes (IBMP-CNRS), Strasbourg

Liste des abréviations

Abréviations générales

ADN :	Acide désoxyribonucléique
ADNc :	ADN complémentaire
ARN :	Acide ribonucléique
ARNm :	ARN messenger
CP-RT :	protéine de capsid et domaine de trans-lecture
dNTP :	désoxyribonucléotides
ddNTP :	di-désoxyribonucléotides
GDP :	déoxy-guanosine-5'-diphosphate
GTP :	déoxy-guanosine 5'-triphosphate
HR :	hypersensitive response
NES :	séquence d'exportation nucléaire
NLS :	séquence d'importation nucléaire
ORF :	cadre ouvert de lecture
PCR :	réaction de polymérisation en chaîne
PTGS :	post-transcriptional gene silencing
rATP :	adénosine 5'-triphosphate
rCTP :	cytosine 5'-triphosphate
rGTP :	guanosine 5'-triphosphate
rUTP :	uridine 5'-triphosphate
RFLP :	Restriction Fragment Length Polymorphisms
RT-PCR :	transcription inverse suivie d'une PCR
SDS-PAGE :	gel de polyacrylamide en conditions dénaturante
SSCP :	Single-Strand Conformation Polymorphisms
UV :	ultra-violet
UTP :	uridine 5'-triphosphate

Abréviations des virus

BNYVV :	<i>Beet necrotic yellow vein virus</i>
BSBMV :	<i>Beet soil-borne mosaic virus</i>
BSBV :	<i>Beet soil-borne virus</i>
BVQ :	<i>Beet virus Q</i>
CMV :	<i>Cucumber mosaic virus</i>
E1A :	<i>Equine adenovirus 1</i>
HIV-1 :	<i>Human immunodeficiency virus type 1</i>
MVM :	<i>Minute virus of mice</i>
PVX :	<i>Potato virus X</i>
RSNV :	<i>Rice stripe necrosis virus</i>
SBWMV :	<i>Soilborne wheat mosaic virus</i>
SV40 :	<i>Simian virus 40</i>
TMGMV :	<i>Tobacco mild green mosaic virus</i>
TMV ou VMT :	<i>Tobacco mosaic virus</i>
TYMV :	<i>Turnip yellow mosaic virus</i>

Table des matières

La rhizomanie	1
Description de la rhizomanie	1
Importance économique	2
Description de l'agent viral : le BNYVV	2
Classification	2
Organisation génétique	3
Transmission naturelle de la maladie	6
Plantes hôtes	7
Propagation et contrôle de la maladie en milieu naturel	7
Variabilité moléculaire et pathogénicité des isolats viraux	10
Les différents isolats viraux	10
Distribution géographique	10
Importance des différents ARN sur la pathogénicité virale	10
Présentation du travail	11
1 Épidémiologie et Phylogénie du BNYVV	13
1.1 Introduction	13
1.1.1 Les différentes méthodes de reconstruction phylogénétique	14
1.1.2 Évaluation des dendrogrammes obtenus	16
1.1.3 Enracinement d'un dendrogramme	17
1.1.4 Choix de la meilleure méthode de reconstruction phylogénétique	17
1.2 Échantillonnage	18
1.3 Analyse des séquences	19

1.4	Structure des populations virales	20
1.4.1	Étude de la phylogénie de l'ARN-5	21
1.4.2	Étude de la phylogénie de l'ARN-2	22
1.4.3	Étude de la phylogénie de l'ARN-3	22
1.5	Interprétation globale de l'étude phylogénétique	23
1.6	Conclusions	25
2	Etude de l'effet de la séquence "tétrade" sur les propriétés biologiques de la protéine P25	29
2.1	Introduction	29
2.2	Effet des acides aminés 67 à 70 sur les symptômes foliaires de plantes hôtes	31
2.2.1	Effet sur <i>Chenopodium quinoa</i>	31
2.2.2	Effet sur <i>Beta vulgaris</i>	32
2.2.3	Effet sur <i>Tetragonia expansa</i> et <i>Beta macrocarpa</i>	33
2.3	Effet de la tétrade sur les propriétés d'activation de transcription de la protéine P25 fusionnée au domaine GAL4BD	35
2.4	Interaction entre protéines P25 issues de différents isolats et effet de la tétrade sur la dimérisation	37
2.5	Conclusions	40
3	Influence de l'ARN-5 et de l'expression de la protéine P26 sur les symptômes induits par le BNYVV	42
3.1	Les voies de transport nucléo-cytoplasmiques dans les cellules eucaryotes .	42
3.1.1	Le pore nucléaire	43
3.1.2	L'adressage nucléaire	45
3.1.3	Mécanismes de l'import nucléaire	46
3.1.4	Mécanismes de l'export nucléaire	48
3.1.5	Cycle de la protéine RanGTPase et orientation du transport nucléo-cytoplasmique	49
3.1.6	Régulation du transport nucléo-cytoplasmique	49

3.1.7	Le transport nucléo-cytoplasmique de protéines et acides nucléiques viraux : entre activation et inhibition des échanges	50
3.1.8	Inhibition du trafic nucléo-cytoplasmique par les virus qui ont un cycle répliatif strictement cytoplasmique : exemple du poliovirus .	51
3.1.9	Le transport nucléo-cytoplasmique chez le BNYVV	52
3.2	Utilisation de la EGFP	52
3.2.1	Propriétés de la "Green Fluorescent Protein" (GFP)	52
3.2.2	Améliorations et modifications des propriétés spectrales de la GFP .	53
3.2.3	Applications en virologie moléculaire	54
3.3	Etude de la protéine P26 codée par l'ARN-5	56
3.3.1	Effet sur la symptomatologie	56
3.3.2	Construction de clones et de virus recombinants exprimant la protéine P26	56
3.4	Étude du contexte d'initiation de la traduction de la protéine P26	58
3.5	Construction du clone d'ADNc complet et infectieux de l'ARN-5	60
3.6	Construction d'un nouveau vecteur d'expression viral dérivé de l'ARN-5 .	61
3.7	Localisation sub-cellulaire de la protéine P26	62
3.7.1	Expression de la protéine en contexte viral	62
3.7.2	Expression transitoire en cellules BY-2	65
3.7.3	Immunolocalisation de la protéine P26	68
3.8	Recherche de protéines interagissant avec la protéine P26 : utilisation du système double hybride de levure	69
3.8.1	Limites	69
3.8.2	Recherche d'interactions entre les protéines de capsid, la protéine P25 et la protéine P26	70
3.8.3	Caractérisation du domaine de la protéine P26 responsable de l'activation de la transcription	71
3.9	Caractérisation des propriétés biologiques de mutants de la protéine P26 . .	73
3.10	Conclusion	74

4	Conclusion générale et perspectives	78
5	Matériels et méthodes	81
5.1	ARN	81
5.1.1	Extraction à partir de tissus végétaux	81
5.1.2	Analyse des ARN par la technique de northern-blot	82
5.1.3	Transcription inverse	83
5.1.4	Obtention de clones ADNc complets d'ARN viraux par les techniques dites de 5' et 3' RACE	83
5.1.5	Transcription <i>in vitro</i>	84
5.1.6	Analyse de l'encapsidation des ARN viraux	85
5.2	ADN	85
5.2.1	Amplification d'un fragment d'ADN par PCR (Polymerase Chain Reaction)	85
5.2.2	Digestion par les enzymes de restriction	86
5.2.3	Déphosphorylation d'un vecteur	86
5.2.4	Analyse des fragments d'ADN par électrophorèse en gel d'agarose	86
5.2.5	Purification de fragments d'ADN.	87
5.2.6	Mutagenèse dirigée par "overlap extension" PCR	87
5.2.7	Mutagenèse dirigée par délétion	88
5.2.8	Ligation	88
5.2.9	Transformation d'organismes	88
5.2.10	Analyse des clones recombinants et purification de l'ADN plasmidique	90
5.3	Vecteurs de clonage et d'expression	91
5.3.1	Vecteurs de clonage	91
5.3.2	Vecteurs d'expression <i>in vivo</i>	92
5.3.3	Vecteur de transcription/traduction <i>in vitro</i>	93
5.4	Traitement des séquences	93
5.5	Organismes	94

5.5.1	Virus	94
5.5.2	Bactéries	94
5.5.3	Plantes hôtes	94
5.5.4	Levures (<i>Saccharomyces cerevisiae</i>)	95
5.6	Protéines	96
5.6.1	Analyse des protéines	96
5.6.2	Immunodétection (Western-blot)	97
5.6.3	Interaction protéine-protéine dans le système double-hybride de le- vure	98
5.6.4	Activation de la transcription dans le système simple-hybride de levure	101
5.6.5	Activité ONPG (O-NitroPhényl- β -D-Galactopyranoside)	101
5.6.6	Activité β -galactosidase dans l'essai "colony lift"	102
5.7	Transfection de cellules BY-2 de tabac par biolistique	102
5.7.1	Principe	102
5.7.2	Cultures de cellules BY-2 de tabac	102
5.7.3	Préparation des cellules pour le bombardement	103
5.7.4	Précipitation de l'ADN plasmidique sur les particules de tungstène et bombardement	103
5.7.5	Matériel	104
5.7.6	Application à l'analyse de protéines fusionnées à l'EGFP	104
5.8	Localisation <i>in situ</i> de protéines virales par microscopie	104
5.8.1	Préparation des échantillons et observation au microscope confocal	104
5.8.2	Préparation des échantillons et observations au microscope électro- nique	105
5.9	Milieux de culture	106
5.9.1	Milieu LB	106
5.9.2	Milieu YPD	106
5.9.3	Milieu SD	107

Appendices du chapitre 1	108
Appendice du chapitre 3	113



Figure 1 : Symptômes de la rhizomanie de la betterave sucrière. A) Essai variétal en champ : lignée de betteraves sensibles infectées, présentant une jaunisse des feuilles, entourées par des betteraves tolérantes à la maladie. B) Effet de la maladie sur la taille et l'aspect des pivots et la densité du chevelu racinaire, betteraves rhizomaniées à gauche et saine à droite. On remarque les anneaux brunâtres de nécrose sur les pivots atrophiés. C) Développement anormal du chevelu racinaire d'un pivot de betterave infectée. D) Détail du jaunissement des nervures d'une feuille de betterave infectée.

Introduction générale

La rhizomanie

Les virus forment un règne à la frontière du vivant et sont des parasites obligatoires des procaryotes et eucaryotes tant végétaux qu'animaux. En effet, la machinerie cellulaire de leur hôte leur est nécessaire pour se multiplier.

Les virus ont un patrimoine génétique de type ARN ou ADN, simple-brin (de polarité positive ou négative) ou double-brin.

Les virus de plantes, ou phytovirus, infectent des espèces végétales très variées, dont des plantes d'intérêts agronomiques, aussi bien sous des climats tropicaux que tempérés. Ils sont répartis en une quarantaine de familles et leur génome est dans la majorité des cas formé d'ARN simple brin de polarité positive.

La betterave sucrière n'est pas épargnée par ces parasites. L'une des maladies majeures, qui inquiète les professionnels du secteur, est la rhizomanie.

Description de la rhizomanie

En Italie, durant les années 1950, dans la plaine du Pô et la vallée de l'Adige ont été décrits les premiers symptômes d'une maladie de la betterave sucrière (Canova, 1959). Des îlots de betteraves qui présentent des chloroses jaunes, similaires à celles des plantes présentées sur la Figure 1 A, ont pu être observés. Après arrachage, les plantes symptomatiques possèdent un pivot¹ atrophié (Figure 1 C), une prolifération des radicelles et la présence d'anneaux de nécrose sur coupe transversale. Ces symptômes sont caractéristiques de la

¹ Racine principale s'enfonçant verticalement dans le sol

Table 1 : Classification des virus faisant partie de l'ancienne famille des *Furovirus*. **Virus** : *Soil-borne wheat mosaic virus* (SBWMV) ; *Oat golden stripe virus* (OGSV) ; *Sorghum chlorotic spot virus* (SCSV) ; *Potato mot-top virus* (PMTV) ; *Beet soil-borne virus* (BSBV) ; *Broad bean necrosis virus* (BBNV) ; *Peanut clump virus* (PCV) ; *Indian peanut clump virus* (PCV) ; *Beet necrotic yellow vein virus* (BNYVV) ; *Beet soil borne mosaic virus* (BSBMV) ; *Rice stripe necrosis virus* (RSNV) ; *Barley stripe mosaic virus* (BSMV). **Hôte naturel** : principales familles végétales infectées par le virus. **Vecteur** de transmission : Pb *Polymyxa betae* ; Pg *Polymyxa graminis* ; Ss *Spongospora subterranea*. **Transmis par graines** : transmission du virus par les graines. **Nombre d'ARN** composant le génome du virus. **Type de polymérase** : similaire au SBWMV (S) ou au BNYVV (B). **poly A** : présence d'une extrémité 3' polyadénylée. **TGB** : organisation des protéines de mouvement sous forme de Triple Gene Block. **CP-RT** : suppression du codon de terminaison de la protéine de capsidie donnant lieu à la production de la protéine de trans-lecture. **Genre actuel** : classification actualisée du virus.

Virus	Hôte naturel	Vecteur	Transmis par graines	Nombre d'ARN	Propriétés du génome				Genre actuel
					Type de polymérase	Queue Poly(A)	TGB	RT-CP	
SBWMV	<i>Graminae</i>	Pg		2	S	-	-	+	<i>Furovirus</i>
OGSV	<i>Graminae</i>	Pg		2	S	-	-	+	<i>Furovirus</i>
SCSV	<i>Graminae</i>			2	S	-	-	+	<i>Furovirus</i>
PMTV	<i>Solanaceae</i>	Ss		3	S	-	+	+	<i>Pomovirus</i>
BSBV	<i>Chenopodiaceae</i>	Pb		3		-	+	+	<i>Pomovirus</i>
BBNV	<i>Leguminosae</i>			3	S	-	+	+	<i>Pomovirus</i>
PCV	<i>Leguminosae Graminae</i>	Pg	+	2	S	-	+	-	<i>Pecluvirus</i>
IPC	<i>Leguminosae Graminae</i>	Pg	+	2	S	-	+	-	<i>Pecluvirus</i>
BNYVV	<i>Chenopodiaceae</i>	Pb		4/5	B	+	+	+	<i>Benyvirus</i>
BSBMV	<i>Chenopodiaceae</i>	Pb		4	B	+	+	+	<i>Benyvirus</i>
RSNV	<i>Graminae</i>	Pg		2		+			<i>Benyvirus</i>
BSMV	<i>Graminae</i>		+	3	S	+	+	-	<i>Hordeivirus</i>

rhizomanie (Figure 1 B). Cette prolifération racinaire est à l'origine du nom de la maladie, la rhizomanie (rhizo = racine, mania = démente). Dans de rares cas, les jeunes plantes sont rabougries, fanées et la plante peut finir par mourir. Sur des plantes infectées de façon systématique par le virus, on peut observer une coloration jaune claire des feuilles qui peut être suivie par une nécrose le long des nervures (Figure 1 D).

L'équipe de Tamada (1975) a démontré que cette maladie est due à un phytovirus, nommé virus des nervures jaunes et nécrotiques de la betterave (*Beet necrotic yellow vein virus*, ou BNYVV). Comme pour beaucoup de phytovirus, son nom est dû aux symptômes qu'il provoque.

Importance économique

La rhizomanie provoque de graves pertes de rendement en sucre, pouvant atteindre 50 % ou plus dans les cas d'infections sévères.

En raison du mode de transmission de la maladie (voir paragraphe , page 6), les surfaces de culture betteravières atteintes par la rhizomanie en Europe ne cessent de progresser. Entre 1990 et 2000, les surfaces infestées se sont étendues de 250000ha à 600000ha (source Institut Technique de la Betterave industrielle française).

Description de l'agent viral : le BNYVV

Classification

Ce virus multipartite, à ARN simple brin de polarité positive, a été initialement classé dans le genre *Furovirus* qui a été redéfini en cinq genres distincts : Furovirus, Pomovirus, Pecluvirus, Benyvirus et Hordeivirus (Torrance and Mayo, 1997; Shirako *et al.*, 2000). Le BNYVV fait désormais partie du genre isolé des Benyvirus. Ce genre comprend le BNYVV, qui en est l'espèce type, et le *Beet soil-borne mosaic virus* (BSBMV). Cette nouvelle classification a été effectuée sur la base des propriétés de chaque virus, ici le vecteur de transmission, le nombre d'ARN, le type de polymérase, la présence ou non d'une séquence poly-A en 3', la présence de trois ORFs constitutives du "triple gene block" (TGB voir

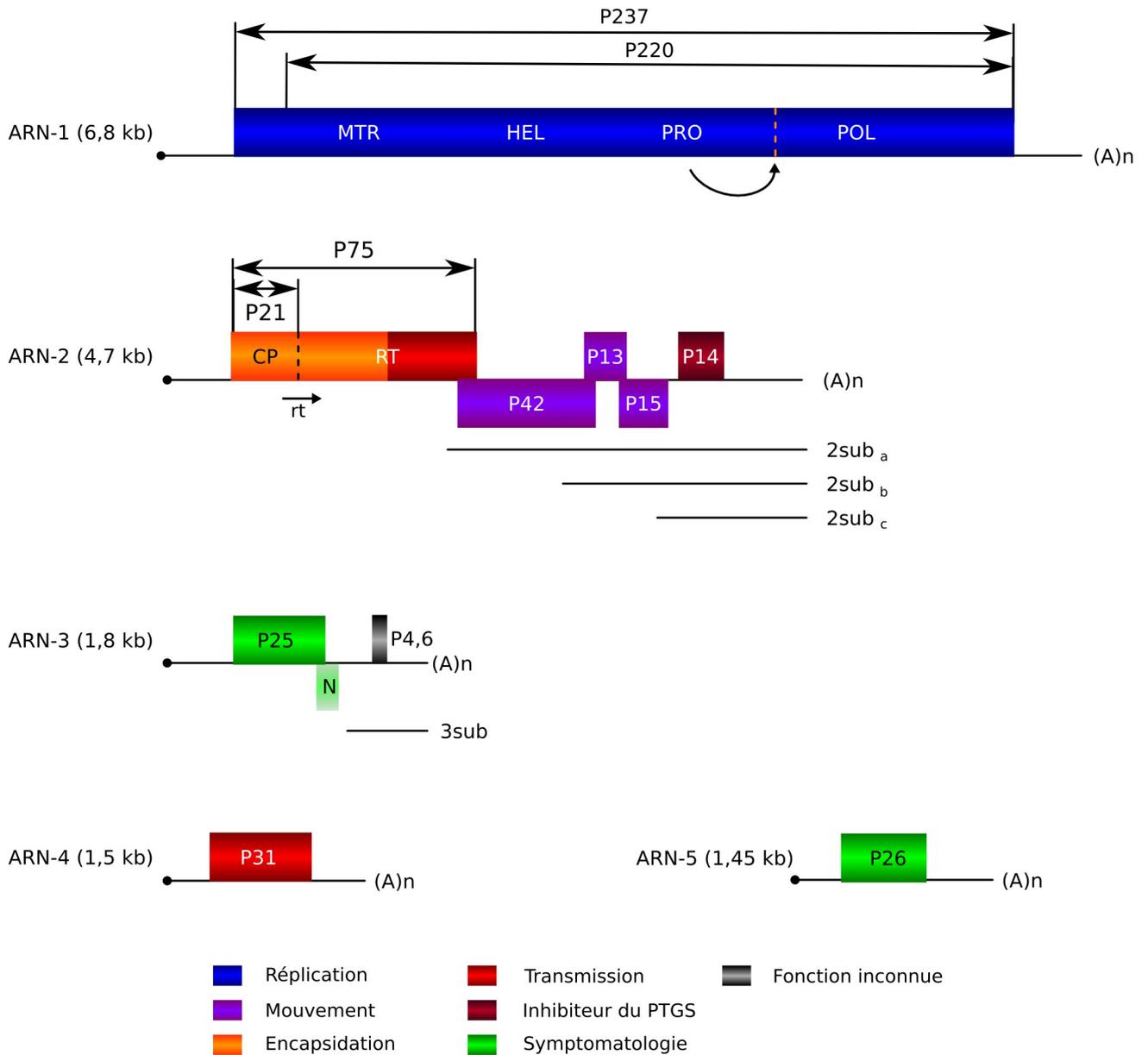


Figure 2 : Organisation génétique du BNYVV. Le nom et la taille de chaque ARN sont précisés. Le rond noir représente la coiffe (m7Gppp) présente à l'extrémité 5' et (A)n représente la séquence poly-A en 3'. L'initiation sur le premier et le second AUG de l'ARN-1 produisent respectivement les protéines P237 et P220 délimitées par des flèches. MTR (méthyl-transférase), HEL (hélicase), PRO (protéase) et POL (ARN-polymérase ARN dépendante), indiquent les positions relatives des différents domaines fonctionnels de la protéine codée par l'ARN-1. Le site d'auto-clivage de la polyprotéine est précisé par les pointillés oranges pointés par la flèche. La suppression du codon stop (rt) de la protéine mineure de capsid (P21 ou CP) conduit à la synthèse de la protéine P75. Les produits de traduction sont indiqués par leur masse moléculaire préfixée par la lettre P. Les positions des ARNs subgénomiques produits à partir des ARN-2 et -3 sont indiqués.

paragraphe , page 4) et la production d'une protéine de CP-RT (voir Table 1). Le *Rice stripe necrosis virus* (RSNV Morales *et al.*, 1999) et le *Burdock mottle virus* sont des espèces classées provisoirement dans ce genre.

Organisation génétique

Généralités

Les particules virales sont formées de bâtonnets à symétrie hélicoïdale contenant un canal central. Leur diamètre constant est proche de 20nm et leurs longueurs moyennes sont de 390, 265, 105, 90 et 80 nm en fonction du type d'ARN contenu (Tamada *et al.*, 1989). L'hélice de pas droit de 2,6 nm est formée de 49 sous-unités de protéine de capsid par répétition axiale de 4 tours. Chaque sous-unité protéique interagit avec 4 bases d'ARN (Steven *et al.*, 1981).

Cinq ARN distincts ont été décrits pour le BNYVV (Richards and Tamada, 1992). On les nomme ARN-1 à -5 dans l'ordre décroissant de leur taille. L'ARN-1 est composé de 6,8 kb, l'ARN-2 de 4,7 kb, l'ARN-3 de 1,8 kb, l'ARN-4 de 1,5 kb et l'ARN-5 de 1,45 kb. Leur organisation génétique est présentée sur la Figure 2. Les cinq ARN sont coiffés par une 7-méthylguanine à leur extrémité 5' et sont polyadénylés en 3'. On retrouve des similarités de séquences au niveau des 8-9 premiers nucléotides 5' de chaque ARN (Table 2) et des 70 derniers nucléotides précédant la séquence poly(A) (Table 3).

Ce virus multipartite présente une propriété intéressante : seuls les ARN-1 et -2 sont nécessaires à sa multiplication sur feuille de plante hôte. Les petits ARN (-3, -4 et -5) sont requis uniquement pour le processus d'infection naturelle : lors de la transmission par le vecteur (ARN-4), la prolifération racinaire et les symptômes de la maladie (ARN-3 et -5). La multiplication du virus par inoculation mécanique de feuilles peut engendrer des délétions internes voire la perte de ces petits ARN (Richards and Tamada, 1992).

Table 2 : Conservation des séquences 5' des ARN du BNYVV. Alignement des 20 premiers nucléotides de chaque ARN.
 Les étoiles (*) représentent une identité des 5 ARN, le point (.) représente une identité de 4 ARN.

ARN	Accession Genbank	5'
ARN-1	X05147	AAATTCGATTCTTCCATTC
ARN-2	X04197	AAATTCCTAACTATTATCTCC
ARN-3	M36894	AAATTCAAAATTTACCATTA
ARN-4	M36896	AAAT-CAAATCTCAAATATA
ARN-5	AY823407	AAATTCAAAGTACTTTCATA
Identité		****.**. .

Table 3 : Conservation des séquences 3' des ARN du BNYVV. Alignement des 75 derniers nucléotides de chaque ARN.
 Les étoiles (*) représentent une identité des 5 ARN, le point (.) représente une identité de 4 ARN.

ARN	Accession Genbank	3'
ARN-1	X05147	TTGTTGGGTTGTAATGTACTGACTGGGTGTGAATTGTACCAGTCATTAAGGGTTTACTATCAGTATATTGATAT
ARN-2	X04197	GATTAAGGTTGCAGTGTACTGACTGGGTGTGAATTGTACCAGTCCATGTAGGGTCTGTTTTCAGTATATTG----
ARN-3	M36894	AATTAAGGTTAAGATGTACTGACTGGGTGTGAAATGTACCAGTCCTTGTAGGGTTCTTTGTCAGTATATTGAC--
ARN-4	M36896	AAATATGGTTGAGATGTACTGACTGGGTGTGAAATGTACCAGTCTTTATAGGGTTCTCTGTCAGTATATTGAC--
ARN-5	AY823407	AAATAAGGTTGAGATGTACTGACTGGATGTGAAATGTGCCAGTCTTCATAGGGTTCTCTGTCAGTGTATTGACA-
Identité		. *. ****. .*****.***** ***.***** ..*****. * *****.*****.

L'ARN-1

Il contient une seule grande phase ouverte de lecture qui code pour une protéine d'environ 237 kilodaltons (kDa) appelée protéine P237 (Bouzoubaa *et al.*, 1987)².

La traduction *in vitro* de cet ARN peut s'initier à deux sites, le premier AUG se situe à la position 154 (produisant la protéine P237) et le deuxième en 496 (produisant la protéine P220), majoritairement utilisé *in vivo* (Jupin *et al.*, 1988; Hehn *et al.*, 1997).

Cette polyprotéine contient plusieurs domaines fonctionnels. On trouve en N-terminal un domaine impliqué dans une activité méthyl-transférase, suivi d'un domaine hélicase, un domaine protéase (de type papaïne, Koonin and Dolja, 1993) et finalement un domaine contenant l'activité ARN-polymérase-ARN-dépendante (Figure 2).

Le produit de 220kDa est clivé auto-catalytiquement, entre le domaine protéase et le domaine polymérase. C'est l'activité protéase de la protéine P220 qui produit une protéine de 150kDa et la polymérase de 66kDa (Hehn *et al.*, 1997), constitutives du complexe de réplication, encore appelé réplisome virale. Dans ce complexe, l'activité méthyl-transférase est nécessaire pour coiffer les ARN viraux au cours de la réplication cytoplasmique du virus. L'activité hélicase permet de séparer les ARN doubles brins et de dérouler les structures secondaires des brins d'ARN lors de la réplication.

L'ARN-2

La phase ouverte de lecture en 5' proximal de cet ARN code pour la protéine majeure de capsid du virus de 21 kDa (P21, voir Figure 2) (Bouzoubaa *et al.*, 1986). Cette séquence est séparée, par un codon stop ambre (UAG), d'une plus longue phase ouverte de lecture, en phase avec celle de la protéine P21. Dans environ 10% des événements de traduction de la protéine P21, la suppression du codon de terminaison ambre produit une protéine de fusion de 75kDa (P75 ou RT pour readthrough, Ziegler *et al.*, 1985; Niesbach-Klösgen *et al.*, 1990; Haeberlé *et al.*, 1994), la protéine mineure de capsid. Cette protéine P75 est retrouvée à une extrémité des particules virales (Haeberlé *et al.*, 1994) et participe à la transmission du virus par le champignon vecteur grâce à un motif KTER (Tamada *et al.*, 1996b).

² Les produits de gènes viraux sont désignés par leur masse moléculaire préfixée par la lettre P

Ce domaine RT est immédiatement suivi par trois phases ouvertes de lecture partiellement chevauchantes qui codent pour trois protéines de 42 kDa (P42), 13 kDa (P13) et 15 kDa (P15, voir Figure 2). Ces trois protéines sont essentielles au mouvement du virus de cellule à cellule (Gilmer *et al.*, 1992a). Cette organisation particulière que l'on retrouve entre autres chez les Potexvirus, les Pomovirus, les Pecluvirus et les Hordeivirus (Morozov *et al.*, 1989) est appelée "Triple Gene Block" ou TGB.

Enfin, la dernière phase ouverte de lecture code pour une protéine de 14 kDa (P14) qui est un suppresseur de Post Transcriptional Gene Silencing (PTGS) (Dunoyer *et al.*, 2002).

A l'exception des protéines de capsid directement exprimées à partir de l'ARN-2, ces séquences codantes 3' proximales sont traduites à partir de trois ARN sub-génomiques de 2,6 kb (2sub_a), 1,4 kb (2sub_b) et 0,7 kb (2sub_c) (Gilmer *et al.*, 1992a) codant respectivement pour la protéine P42, les protéines P13 et P15, et la protéine P14 (Figure 2).

L'ARN-3

Cet ARN contient 2 ORFs codant pour deux protéines, une de 25 kDa (P25) et une de 4,6 kDa (P4.6). Cette dernière est produite à partir d'un ARN sub-génomique (3sub, voir Figure 2) correspondant aux 550 derniers nucléotides de l'ARN-3 (Balmori *et al.*, 1993). Il existe des formes délétées naturelles de l'ARN-3 (Bouzoubaa *et al.*, 1991) qui permettent l'expression d'un gène cryptique³ N qui induit des symptômes sévères de nécrose (Jupin *et al.*, 1992). Ce gène cryptique peut ainsi être exprimé soit par le rapprochement de sa phase ouverte de lecture en 5' de l'ARN-3, soit par sa fusion avec celle de la protéine P25 (Jupin *et al.*, 1992).

L'ARN-3, en particulier la protéine P25, est essentiel à l'expression des symptômes de prolifération racinaire caractéristiques de la rhizomanie (Koenig *et al.*, 1991; Tamada *et al.*, 1999) ainsi que des symptômes chlorotiques jaunes observés sur feuilles (Jupin *et al.*, 1992).

L'ARN-4

Il contient une seule phase ouverte de lecture (Bouzoubaa *et al.*, 1985) codant pour une protéine de 31 kDa (P31), essentielle à la transmission du virus par *Polymyxa betae*

³ Qui est caché, qui n'est pas immédiatement compréhensible, déchiffrable ou identifiable

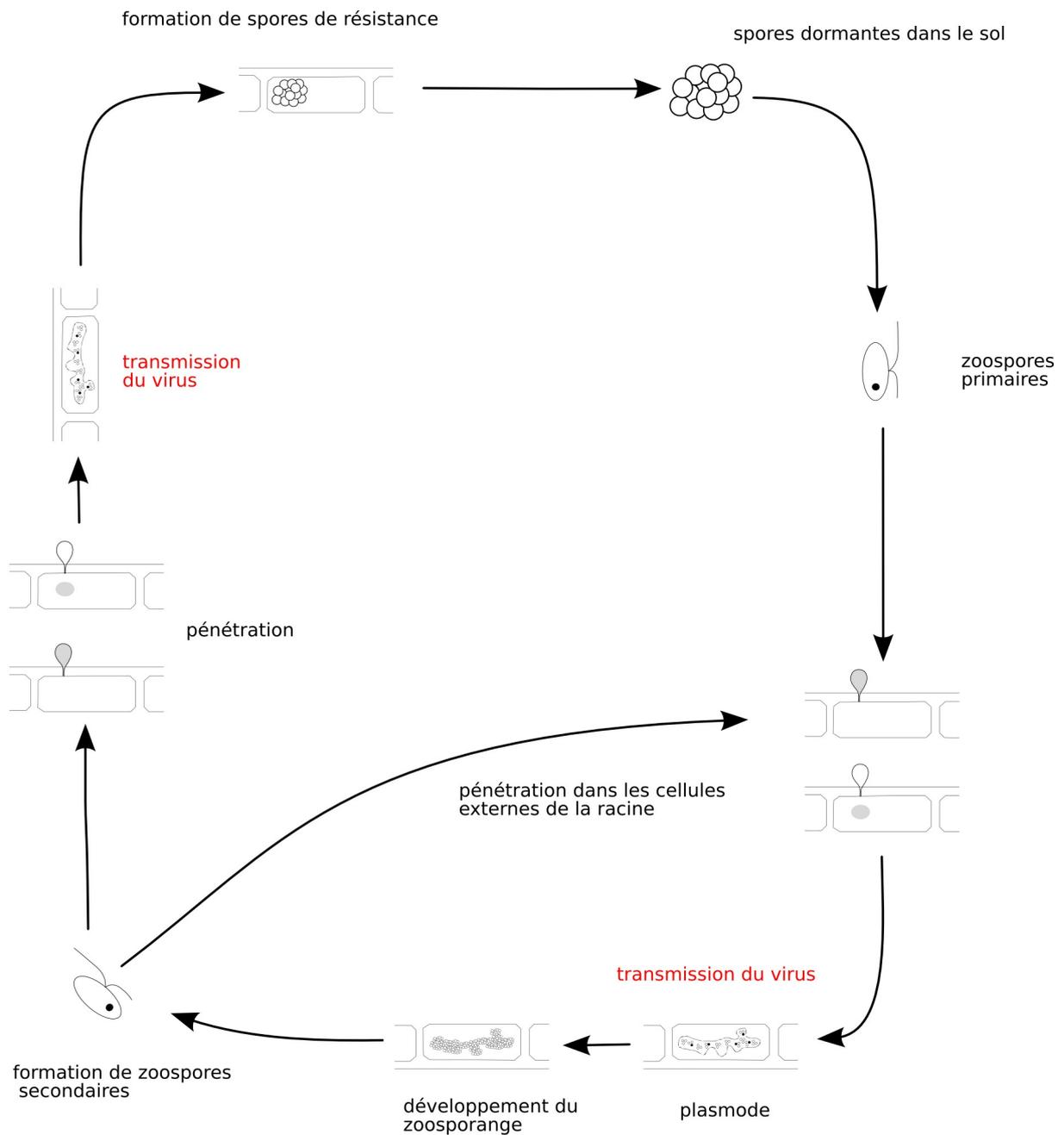


Figure 3 : Schéma du cycle de vie d'un plasmodiophoromycète (ici *Polymyxa graminis*, espèce très proche de *Polymyxa betae*). Le cycle est divisé en deux phases d'infection, en bas le développement du protozoaire lors de la saison de culture, en haut la phase de formation des formes de résistance. En rouge sont notés les phases du cycle où s'effectue probablement la transmission du virus du vecteur à la plante hôte ou l'acquisition du virus par le plasmodiophoride.

(Lemaire *et al.*, 1988; Tamada and Abe, 1989).

L'ARN-5

Cet ARN, découvert au Japon (Tamada *et al.*, 1989) puis en France dans la zone de Pithiviers (Koenig *et al.*, 1997), contient une seule phase ouverte de lecture codant pour une protéine de 26 kDa (P26) (Kiguchi *et al.*, 1996). Sa présence provoque une aggravation des symptômes en combinaison avec les ARN -3 et -4 (Tamada *et al.*, 1996a).

Transmission naturelle de la maladie

Le BNYVV est transmis par un organisme unicellulaire du sol de la famille des Plasmodiophoridés : *Polymyxa betae*. Il a longtemps été considéré comme un champignon (Keskin, 1964) mais semble plus proche des protozoaires (Kanyuka *et al.*, 2003). *P. betae* est un parasite obligatoire des racines et des tubercules (Scholten and Lange, 2000) et son cycle de vie est proche de celui de *P. graminis*, décrit dans la Figure 3 (Kanyuka *et al.*, 2003).

P. betae en lui-même ne provoque ni symptômes visibles, ni baisse de rendement sur la betterave sucrière. Les spores dormantes (voir Figure 3), en amas appelés cystosores, germent pour produire des zoospores bi-flagellées qui nagent jusqu'à atteindre une cellule d'épiderme de racine. Elles s'enkystent et pénètrent dans la cellule où elles forment des plasmodies (cytoplasme plurinucléé) qui se différencient entièrement en zoosporanges desquels de nouvelles zoospores seront libérées. La formation de spores dormantes n'est pas bien connue et ne se produit qu'à la fin de la saison de culture des plantes. Ces spores peuvent être virulifères ou avirulifères selon qu'elles contiennent ou non des particules virales. Les spores peuvent rester en dormance pendant de très longues périodes jusqu'à retrouver des conditions optimales. Ainsi, il a été observé que des spores virulifères dormantes pouvaient encore transmettre le BNYVV après plus de 15 ans (Abe and Tamada, 1986), expliquant notamment la persistance de la maladie dans les sols infestés.

Le BNYVV est véhiculé dans le cytoplasme des zoospores et les virions sont apparemment relâchés dans le cytoplasme des cellules infectées par pinocytose. Il n'existe aucune donnée concernant l'éventuelle multiplication du virus dans son vecteur. Cependant, les travaux récents de Driskel *et al.* (2004) tendent à montrer que le SBWMV est internalisé

dans les cystosores de *P. graminis* où sont également retrouvés l'ARN viral et la protéine de mouvement mais pas la protéine de capsid. La détection de la protéine de mouvement, exprimée par l'intermédiaire d'un ARN subgénomique, suggérerait que le virus est capable de se multiplier dans son vecteur. Une telle hypothèse devra cependant être vérifiée en recherchant les formes répliquatives des ARN viraux et les autres protéines virales.

Pour sa transmission, le BNYVV requiert l'expression de la protéine P31 codée par l'ARN-4 (Lemaire *et al.*, 1988; Tamada and Abe, 1989) mais aussi la protéine mineure de capsid qui comporte la séquence KTER essentielle à la transmission des particules virales (Tamada and Kusume, 1991; Haeberlé *et al.*, 1994; Tamada *et al.*, 1996b).

Plantes hôtes

Le spectre d'hôte du BNYVV est étroitement lié à celui de son vecteur, *P. betae* et reste restreint. Ce virus infecte entre autres la betterave sucrière (*Beta vulgaris* var. *saccharifera*), la betterave fourragère et l'épinard (*Spinacea oleracea* L) (Hugo *et al.*, 1996).

L'inoculation mécanique permet la transmission du virus à la plupart des espèces de la famille des *Chenopodiaceae* (famille incluse dans celle des *Amaranthaceae*) et à quelques espèces d'*Aizoaceae* (*Tetragonia expansa*) et de *Solanaceae* (*Nicotiana benthamiana*). Le virus tend à être restreint aux feuilles inoculées des plantes hôtes, sauf dans le cas de *Beta macrocarpa*, de l'épinard et de *Nicotiana benthamiana*. Des hôtes comme *Chenopodium quinoa*, *C. amaranticolor* et *Tetragonia expansa* répondent à l'infection virale par la formation de lésions locales.

Propagation et contrôle de la maladie en milieu naturel

Conditions de propagation de la maladie

De nombreux paramètres influencent l'incidence et la sévérité de la maladie au champ. L'apparition de la maladie se fait habituellement après l'introduction de terre infestée provenant de matériel de culture ou de résidus de terre collés aux pivots (appelé tarre-terre) provenant d'une précédente récolte et épandues dans un autre champ (Rush, 2003). Cette tarre-terre renferme une petite quantité de *P. betae* virulifère. La détection initiale de la ma-

maladie se fonde sur l'observation d'une tache de plantes à l'aspect jaunâtre (voir Figure 1 A) en plein champ durant la saison de culture de betteraves. Au fil des années, l'inoculum de base (quantité de zoospores virulifères) croît de manière exponentielle, jusqu'à 10 000 fois par saison de culture betteravière. La progression rapide de la maladie dans le champ a d'abord été attribuée à une migration facilitée des zoospores virulifères causée par le système d'irrigation. Il a été démontré que les opérations mécaniques de préparation du sol et de récolte sont les principaux responsables de la propagation de la maladie (Harveson *et al.*, 1996).

Après apparition de la population de *P. betae* virulifère, la densité d'inoculum est influencée par divers paramètres dont l'agressivité de l'isolat de *P. betae* introduit, la densité d'inoculum de la population de vecteur non virulifère en place et la culture d'un hôte sensible à *P. betae* et au virus.

Les avis divergent quant à la capacité plus ou moins élevée des populations virulifères à coloniser les racines des plantes hôtes (Rush, 2003). La pression d'inoculum est le paramètre le plus important pour déterminer quelle population de vecteur peut prendre l'ascendant. En effet, l'isolat qui infecte en premier le système racinaire de l'hôte aura un avantage conséquent lors du cycle infectieux. Si la population originelle de vecteur avirulifère est en faible densité, la population virulifère introduite aura la capacité d'accroître considérablement sa densité en une seule saison de culture.

La sévérité de la maladie et la perte de rendement sont maximales quand les conditions de multiplication du vecteur sont optimales (Rush, 2003) : des températures situées entre 25 et 30°C ainsi qu'une forte humidité du sol favorisent la prolifération de *P. betae*, qui est accentuée notamment dans les sols irrigués. Ces conditions augmentent également les possibilités d'infection des plantes par d'autres virus transmis par plasmodiophorides. *P. betae* peut aussi transmettre le BSBMV, le BSBV et le BVQ (Adams, 2002). Il a été démontré que la sévérité de la maladie était liée à la quantité de virus présente dans le sol et donc à la densité d'inoculum. Plus les conditions de multiplication du vecteur sont favorables, plus de générations de zoospores virulifères seront libérées dans le sol et plus la sévérité de la maladie et les pertes de rendement seront importantes (Rush, 2003).

Mesures de contrôle de la maladie

La mesure la plus efficace pour éradiquer la maladie est le traitement du sol infesté par fumigation avec du bromure de méthyl-1,3-dichloropropène. Cette mesure extrême est efficace pour réduire fortement la densité d'inoculum du champ mais a un coût et une toxicité prohibitifs pour être utilisée à grande échelle (Rush, 2003).

La meilleure stratégie pour réduire l'incidence de la rhizomanie reste les bonnes pratiques culturales. Ainsi, la plantation des graines dans des conditions de température basses (< 15°C) permet d'éviter une infection précoce et un développement rapide de la maladie. La réduction de l'irrigation et un meilleur drainage des parcelles permettent également de limiter l'amplification de *P. betae*.

Ces pratiques culturales sont à combiner avec l'utilisation de variétés résistantes au vecteur et au virus. La résistance à *P. betae* a été trouvée dans des espèces sauvages de betteraves (Scholten and Lange, 2000) et un travail important de sélection est maintenu afin d'obtenir des variétés de betteraves cultivées possédant cette résistance. La résistance au BNYVV a été introduite dans les variétés modernes qui possèdent le gène Holly (*Rz1*, Rush, 2003; Scholten and Lange, 2000). Les plantes possédant ce gène présentent une forte résistance au BNYVV qui n'est toutefois pas totale. De plus, les cultures de variétés résistantes dans des champs ayant une haute densité d'inoculum (vecteur virulifère) présentent une concentration en virus aussi élevée que des variétés sensibles (Scholten and Lange, 2000). Cet inoculum virulifère est donc maintenu dans le sol après formation des spores. D'autres sources de résistance ont été découvertes chez *Beta maritima* et sont en cours d'introduction par sélection classique dans les variétés modernes de betteraves sucrières.

L'utilisation d'un seul gène de résistance contre un virus à ARN est cependant risqué. En effet, le fort taux de mutations de ces virus combiné à une forte pression de sélection peut rapidement favoriser l'émergence de populations virales qui contournent cette résistance (Rush, 2003). L'utilisation de plantes doublement résistantes confèrent aux variétés une résistance accrue au BNYVV. Si ces variétés se comportent mieux en champs rhizomaniés, elles sont moins performantes dans des champs non infectés (M. Richard-Molard, ITB, communication personnelle, Scholten and Lange, 2000).

Isolats :

- A
- B
- P

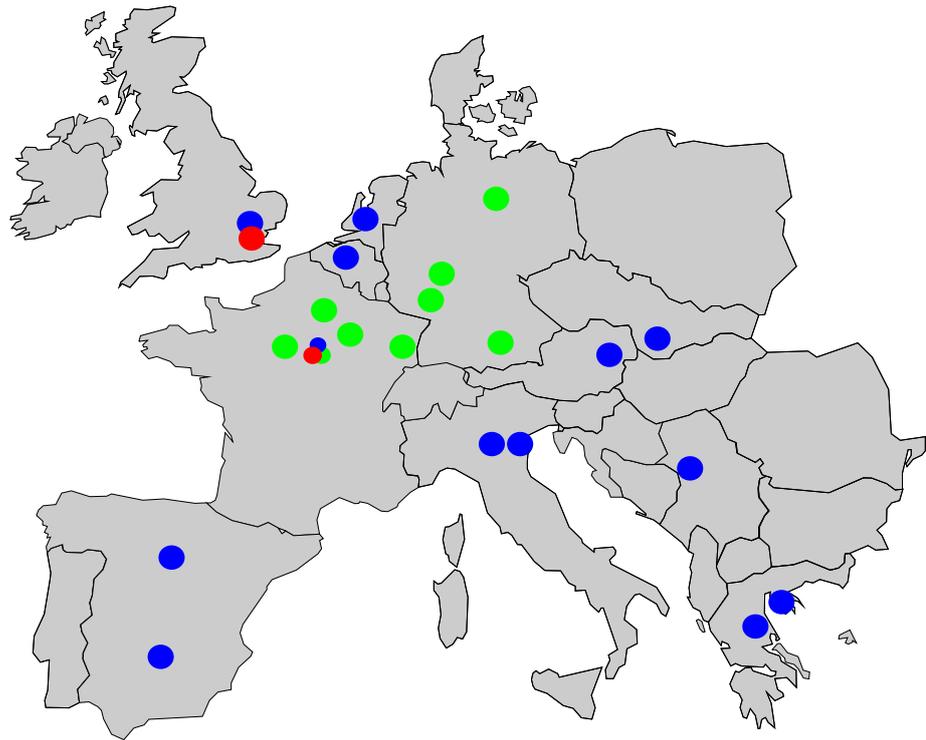


Figure 4 : Répartition européenne des différents isolats du BNYVV. Les disques bleus représentent une zone d'infection par l'isolat A du BNYVV, les disques verts par l'isolat B et les disques rouges par l'isolat P.

Variabilité moléculaire et pathogénicité des isolats

viraux

Les différents isolats viraux

Plusieurs groupes génotypiques de BNYVV existent et sont retrouvés dans des régions géographiques spécifiques et présentent des différences de pathogénicité. Trois isolats de BNYVV ont été identifiés. On les distingue par leur composition en ARN et par des analyses de "Restriction Fragment Length Polymorphisms" (RFLP) ou "Single-Strand Conformation Polymorphisms" (SSCP) de produits de RT-PCR spécifiques de la séquence CP d'isolats de BNYVV récoltés dans le monde entier (Kruse *et al.*, 1994; Koenig *et al.*, 1995). Ainsi, les isolats A et B contiennent uniquement les ARN-1 à -4 et sont identifiés par la séquence de leur protéine de capsid alors que l'isolat P est défini par la présence d'un cinquième ARN.

Distribution géographique

Le virus est probablement distribué partout dans le monde où se trouvent des industries sucrières betteravières : dans tous les pays d'Europe, au Moyen-Orient, en Chine, au Japon et aux Etats-Unis.

L'isolat A est très largement répandu dans le monde, notamment à travers l'Europe (voir Figure 4), les Etats-Unis, le Japon et la Chine. L'isolat B présente une distribution beaucoup plus restreinte. Il est retrouvé essentiellement en Allemagne et en France. L'isolat P, qui contient un cinquième ARN, est détecté uniquement en France dans le voisinage de Pithiviers, en Angleterre, au Kazakhstan (Koenig and Lennefors, 2000), et plus communément au Japon.

Importance des différents ARN sur la pathogénicité virale

Lors d'une infection naturelle on retrouve toujours les 4 premiers ARN et parfois, selon l'isolat, le cinquième ARN.

Les ARN-1 et -2 sont essentiels à la réplication et au mouvement du virus dans la plante. L'ARN-3 est responsable du développement des symptômes racinaires et foliaires caractéristiques de la rhizomanie. L'ARN-4 est important dans la transmission du virus par *Polymyxa betae*. L'ARN-5 est commun dans les isolats japonais du virus mais beaucoup plus rare en Europe de l'Ouest (Koenig and Lennefors, 2000). Il a été démontré que la quantité de virus dans le pivot de différentes variétés de betteraves, résistantes ou non, est bien plus élevée lors de l'infection par l'isolat P que par les isolats A et B (Heijbroek *et al.*, 1999), suggérant un rôle très important de l'ARN-5 dans l'efficacité d'infection naturelle de la betterave.

Présentation du travail

La présence aux environs de Pithiviers de souches de BNYVV très virulentes inquiète fortement l'Institut Technique de la Betterave industrielle française (ITB), d'où sa demande de recherches sur cet aspect. La maladie étant due à un phytovirus étudié dans le service de virologie moléculaire de l'IBMP, l'étude des particularités des isolats de type P a été initiée dans le laboratoire du Dr David Gilmer.

Des travaux effectués au laboratoire ont permis de mettre hors de cause la protéine P14 du BNYVV, inhibiteur de PTGS qui présente les mêmes caractéristiques d'expression et d'activité quel que soit l'isolat B ou P considéré (Sébastien Pfeffer, communication personnelle). Mon travail de thèse s'est donc articulé autour de trois approches ciblant deux autres déterminants de la pathogénicité connus et portés par les ARN-3 et -5. Ces trois approches sont décrites dans les trois chapitres suivants.

Le **chapitre 1** présente l'état des lieux de la structure de la population virale. La détection et la caractérisation d'isolats du BNYVV ont été réalisées dans l'ensemble des zones rhizomaniées européennes en collaboration avec l'Institut National de la Recherche Agronomique (INRA) de Colmar et de l'ITB qui a assuré la collecte des échantillons. Les résultats des analyses phylogénétiques effectuées ont permis d'estimer la stabilité dans le temps de séquences virales, de corrélérer les variations de séquences observées avec la pathogénicité des isolats et de proposer un modèle évolutif des populations de BNYVV. Les

résultats obtenus seront soumis pour publication dans la revue *Journal of General Virology* : "Extensive molecular epidemiology and phylogenetic analysis of worldwide collected beet necrotic yellow vein virus (BNYVV) isolates", en préparation.

Le **chapitre 2** présente l'analyse des propriétés moléculaires de la protéine P25 codée par différents isolats viraux. La caractérisation phénotypique d'une séquence hypervariable au sein de la protéine P25, codée par l'ARN-3, nous a permis de déterminer son importance dans l'induction de symptômes foliaires sur différentes plantes hôtes du virus. De plus, nous avons étudié l'importance de cette séquence hypervariable sur les propriétés connues de la protéine P25.

Le **chapitre 3** présente les résultats de la caractérisation des propriétés biologiques de l'ARN-5 et de la protéine P26. L'étude fonctionnelle par mutagenèse dirigée de la protéine P26 nous a permis de déterminer le codon d'initiation utilisé pour la traduction de cette protéine *in planta* à partir du clone complet infectieux de l'ARN-5 obtenu au laboratoire. La recherche de la localisation sub-cellulaire de la protéine P26 nous a montré qu'elle est activement importée dans le noyau des cellules infectées. Nous avons pu mettre en évidence l'importance du domaine N-terminal de la protéine dans cette localisation ainsi que l'importance de séquences pouvant soit induire l'export nucléaire soit la rétention cytoplasmique de la protéine. Enfin, la recherche des partenaires cellulaires et viraux de la protéine P26 en système double hybride, nous a permis de révéler sa forte capacité à activer la transcription dans la levure lorsqu'elle est fusionnée au domaine de fixation à l'ADN du facteur de transcription GAL4. Le domaine responsable de cette activation de transcription a pu être localisé puis caractérisé par mutagenèse dirigée. Les résultats obtenus seront complétés par une approche d'immunolocalisation avant de soumettre un article intitulé "Biological properties of the beet necrotic yellow vein virus RNA-5-encoded p26 protein" dans la revue *Journal of General Virology*". L'obtention du clone complet infectieux de l'ARN-5 nous a permis de construire un nouveau vecteur d'expression, le réplicon 5. Les propriétés de ce réplicon ont été analysées et comparées à celles du réplicon dérivé de l'ARN-3. Ces résultats ont fait l'objet d'un article accepté dans la revue "Journal of General Virology" intitulé "Use of a beet necrotic yellow vein virus RNA-5-derived replicon as a new tool for the expression of genes".

Chapitre 1

Épidémiologie et Phylogénie du BNYVV

1.1 Introduction

L'étude phylogénétique est le meilleur outil pour représenter la structure d'une population virale. La phylogénie (voir les tutoriels sur le site internet d'Infobiogen) est une reconstruction de l'histoire évolutive des êtres vivants. Le terme phylogénèse¹ a été introduit par Haeckel en 1860, qui l'a défini comme "l'histoire du développement paléontologique des organismes par analogie avec l'ontogénie ou histoire du développement individuel". La représentation de la phylogénèse d'un groupe de taxons² est l'arbre phylogénétique ou dendrogramme³.

On parle de reconstruction d'un dendrogramme car on cherche à découvrir les relations qui existent réellement entre les taxons. Il existe principalement deux types de reconstruction basés sur :

- les mesures de distances entre des séquences prises deux à deux, c'est à dire le nombre de substitutions de nucléotides ou d'acides aminés entre ces séquences,
- les caractères qui s'intéressent au nombre de mutations (substitutions, insertions ou délétions) qui affectent chacun des sites de la séquence.

¹ du grec phulon "race, tribu"

² Unité systématique de classification (famille, genre, espèce, etc.)

³ Vient du grec "arbre" et "dessin".

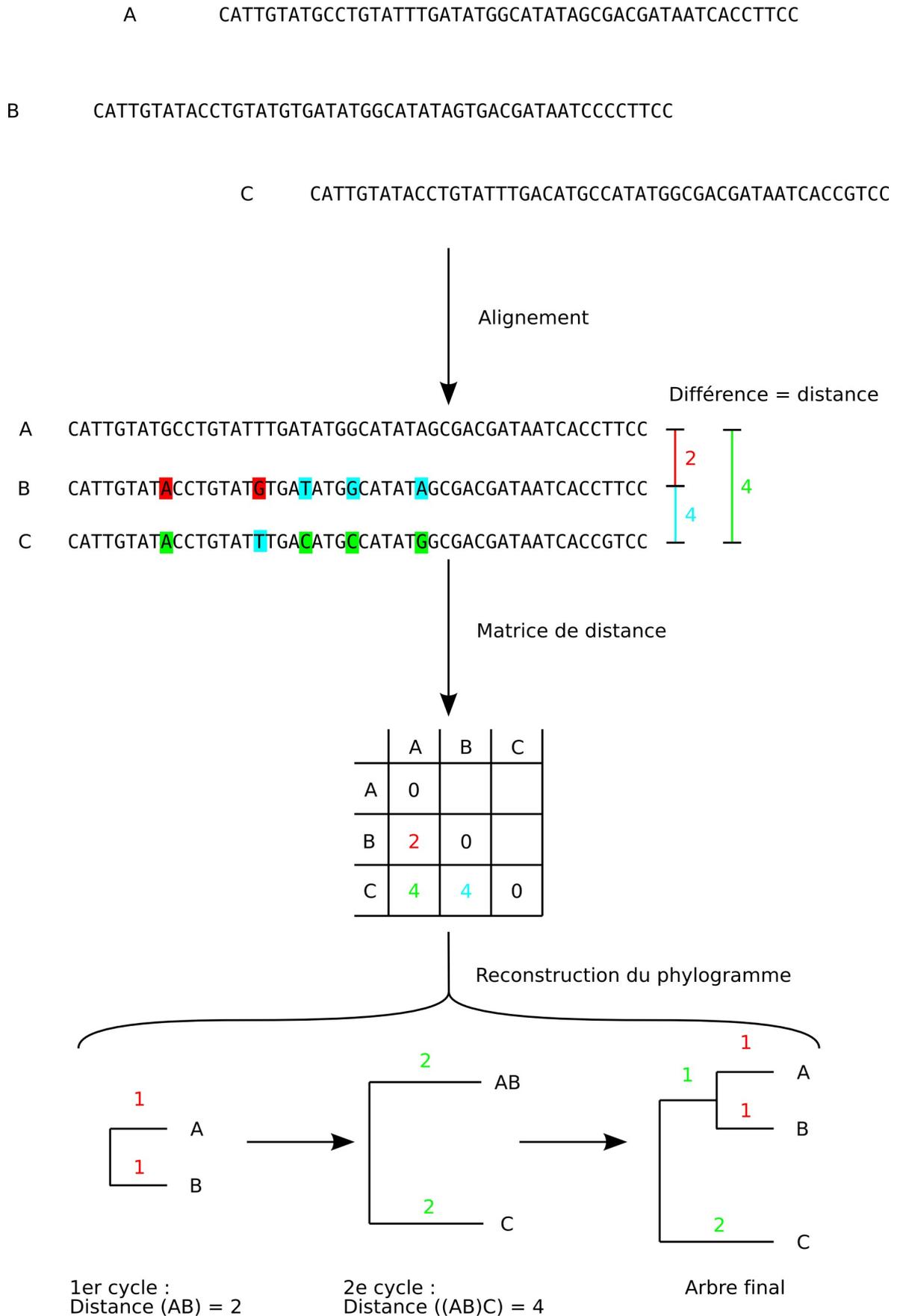


Figure 1.1 : Principe de reconstruction d'un arbre phylogénétique par la méthode des distances. Les 3 séquences initiales (A, B et C) sont tout d'abord alignées. Les différences nucléotidiques sont comptabilisées : en rouge entre les séquences A et B, en vert entre les séquences A et C, en bleu ciel entre les séquences B et C. La matrice de distance est produite et permet la formation de l'arbre phylogénétique par itérations, ici en trois étapes.

Ces reconstructions produisent trois types de dendrogrammes :

phénogramme un dendrogramme obtenu par mesure de distance où les relations entre taxons expriment des degrés de similitude globale sans polarisation des caractères⁴

cladogramme un dendrogramme exprimant les relations phylogénétiques entre taxons et construit à partir de l'analyse cladistique, c'est à dire la polarisation des caractères étudiés

phylogramme un dendrogramme des relations de parenté représenté par les branchements et du degré de divergence adaptative associé à chaque taxon, mesuré par la longueur des branches.

Ces dendrogrammes peuvent être enracinés, permettant ainsi de situer l'ancêtre commun des taxons : la racine. Un dendrogramme non enraciné représente des relations phylogénétiques intemporelles.

1.1.1 Les différentes méthodes de reconstruction phylogénétique

Mesure des distances

La reconstruction basée sur la mesure de distances produit des arbres non enracinés à partir de matrices issues des alignements de séquences. Ces reconstructions sont rapides et donnent de bons résultats pour les séquences ayant de fortes similarités. Deux méthodes sont utilisées :

Unweight Pair Group Method with Arithmetic mean (UPGMA) Cette méthode utilise un algorithme de clusterisation⁵ séquentiel dans lequel les relations sont identifiées dans l'ordre de leur similarité et la reconstruction de l'arbre se fait pas à pas grâce à cet ordre. Il y a d'abord identification des deux séquences les plus proches et ce groupe est ensuite traité comme un tout, puis on recherche la séquence la plus proche de ce groupe et ainsi de suite jusqu'à ce qu'il n'y ait plus que deux groupes. Cette méthode produit un phénogramme non enraciné. Un exemple est présenté dans la

⁴ Détermination de la séquence évolutive de la transformation d'un caractère depuis sa forme ancestrale vers sa forme actuelle

⁵ Formation de groupes de similarité

Figure 1.1. Son inconvénient majeur est sa sensibilité à des taux de mutations différents sur chaque branche, pouvant conduire à des erreurs de topologie⁶. En effet, deux séquences issues d'un ancêtre commun peuvent dériver indépendamment l'une de l'autre sans avoir le même taux de mutations, faussant les résultats de l'UPGMA basé sur le postulat d'une évolution parallèle avec des taux de mutation identiques.

Neighbor-Joining (NJ) Cette méthode corrige l'algorithme UPGMA (Saitou and Nei, 1987) afin de prendre en compte un taux de mutation différent sur les branches. Des corrections supplémentaires ont été apportées à l'algorithme afin de prendre en compte des mutations multiples à un locus donné selon les modèles de correction de Jukes-Cantor (Jukes and Cantor, 1969), de Kimura (Kimura, 1980) ou de Jin-Ney's (Jin and Nei, 1990). Elle produit des phylogrammes.

Analyse cladistique

La reconstruction basée sur les caractères est beaucoup plus précise, plus lente et permet d'obtenir des cladogrammes à partir de deux méthodes :

Parcimonie Elle postule qu'en minimisant le nombre total d'évènements évolutifs de l'arbre phylogénétique parmi toutes les combinaisons possibles, on obtient la phylogénie la plus vraisemblable pour un groupe d'espèces (Eck and Dayhoff, 1966; Kluge and Farris, 1969). Pour cela, la recherche exhaustive de toutes les topologies possibles permet de trouver un ou plusieurs arbres optimaux. Cette recherche systématique demande énormément de ressources en puissance de calcul et en temps, à un point tel qu'il devient difficile de trouver toutes les topologies au delà d'une limite de 10 séquences (plus de 34 millions d'arbres non enracinés et $2,13 \cdot 10^{15}$ arbres enracinés possibles). Toutefois, l'utilisation de techniques simplificatrices comme le "branch and bound" et les méthodes heuristiques permettent de réduire considérablement le temps nécessaire pour trouver les cladogrammes de longueur minimale. De ces trois algorithmes, seuls les deux premiers (recherche exhaustive et "branch and bound") permettent d'obtenir avec certitude l'arbre optimal. Ces trois méthodes permettent d'effectuer la reconstruction phylogénétique de séquences plus nombreuses.

⁶ Partie de la géométrie qui considère uniquement les relations de position

La parcimonie ne peut produire que des dendrogrammes sans informations sur la longueur des branches (mesure du degré de divergence adaptative).

Maximum de vraisemblance (ML) Cette méthode évalue en terme de probabilité l'ordre des branchements et la longueur des branches selon un modèle évolutif donné (Felsenstein, 1981, 1996). Les modèles utilisés sont de complexité variable en fonction des hypothèses formulées sur les données initiales. Les trois inconvénients de cette méthode sont :

- plus le nombre de paramètres introduits dans le modèle est grand, plus il y aura d'erreurs systématiques accumulées au long du calcul,
- les hypothèses évolutives de départ doivent s'appliquer pertinemment aux données pour obtenir l'arbre le plus juste,
- son utilisation est limitée à un faible nombre de séquences.

1.1.2 Évaluation des dendrogrammes obtenus

Après l'obtention d'un arbre, il faut évaluer la confiance que l'on peut avoir en celui-ci. Il existe plusieurs méthodes disponibles :

Le bootstrap Cette méthode, inventée par Efron en 1979, est la plus utilisée en phylogénie (Felsenstein, 1985). Elle part du postulat que les caractères évoluent de manière indépendante. Cette méthode est basée sur le mélange statistique des séquences initiales. Dans l'alignement de ces séquences, chaque colonne est mélangée par tirage au sort avec remise de chaque caractère. On obtient un nouvel alignement avec une duplication ou une perte de chaque caractère aléatoirement dans chaque colonne. Les données produites servent à établir un nouvel arbre. Cette procédure est itérée au moins 100 fois afin d'obtenir un ensemble de données statistiquement significatif permettant de produire de nouveaux arbres. Pour chaque noeud de l'arbre original, on dénombre les nouveaux arbres qui contiennent ce même branchement, permettant d'établir les valeurs de bootstrap. Par exemple, une valeur de 90 indique une confiance de 90% dans l'existence du branchement considéré.

Delete-half-Jackknifing C'est une méthode proche de celle du bootstrap, qui est plus ra-

pide et qui consiste à effectuer un ré-échantillonnage de la moitié des sites des séquences originales (Wu, 1986).

Permutation C'est une approche basée sur l'échange des colonnes de chaque caractère de l'alignement. Elle permet de tester si les données portent bien une information de taxonomie et ne sont pas le résultat d'un biais aléatoire des séquences initiales.

1.1.3 Enracinement d'un dendrogramme

Le plus souvent les méthodes de reconstruction vues ci-dessus ne produisent que des arbres non enracinés. Pour enraciner correctement ces arbres, on ajoute une séquence appelée "outgroup" apparentée et beaucoup plus ancienne que toutes les autres séquences. Cet "outgroup" doit être proche des séquences analysées sous peine d'obtenir des erreurs dans la topologie de l'arbre. En absence de séquence "outgroup", on peut employer la méthode dite de "mid-point rooting" qui consiste à positionner la racine de l'arbre à l'endroit où chaque séquence est à égale distance de toutes les autres.

1.1.4 Choix de la meilleure méthode de reconstruction phylogénétique

Parmi toutes ces méthodes d'analyse phylogénétiques, il est difficile de choisir l'algorithme idéal. Toutefois, il faut garder à l'esprit qu'une phylogénie fiable dépend plus de la qualité des séquences et de leur nombre que de la méthode employée. Cette affirmation a pu être démontrée par Tateno *et al.* (1994) grâce à la comparaison des méthodes de Neighbor-Joining, de parcimonie et de maximum de vraisemblance. Ces méthodes produisent des dendrogrammes de topologie identique si la divergence des séquences utilisées est inférieure à 5 % et leur longueur supérieure à 1000 nucléotides. Dans une telle situation, la méthode la plus simple sera avantageusement employée.

Ainsi, nous avons initié l'étude phylogénétique du BNYVV à la fois à partir des séquences nucléotidiques obtenues dans Genbank et collectées sur le terrain.

A

FC	AT GT CGAGT GAAGGT AGAT AT AT GACAT GGAAGGAT AT GT CACATA ATA G TTT AT GACCGAT CGAT GGGC CGT GTT T C	80
B2	AT GT CGAGT GAAGGT AGAT AT AT GACAT GGAAGGAT AT GT CACATA ATA A G TTT AT GACCGAT CGAT GGGC T CGT GTT T C	80
FC	GGACGT CGT GAGT GT TAT TAAACAAT CGCAT GCTAT GGA CTT GT CCAAGGCT GCGAAT CTAT CTATAA T TAAACT GCT T	160
B2	GGACGT CGT GAGT GT TAT TAAACAAT CGCAT GCTAT GGA CTT GT CCAAGGCT GCGAAT CTAT CTATAA T TAAACT GCT T	160
FC	T GGCAGGATT AGGCT CGGGT T G C A C T GACAATAAT CCT TTT GT GT C CCGAT GACCCGTT T CCACAGACACT AACT AT G	240
B2	T GGCAGGATT AGGCT CGGGT T G C T C T GACAATAAT CCT TTT GT GT C T CCGAT GACCCGTT T CCACAGACACT AACT AT G	240
FC	T A C GGT GCACTT GT GTT AT AT GTT AAT CT GT CT GACCCAGAAT T T GCGTT GATAAT GACT AAG T A A G T A C T T T A A C T G A	320
B2	T A T GGT GCACTT GT GTT AT AT GTT AAT CT GT CT GACCCAGAAT T T GCGTT GATAAT GACT AAG T A A T A C T T T A A C T G A	320
FC	T T CAGGGT TAGCAGATAA C GCAT CT GCTAAT GT CCGT AGAGAT GT GGT GT CCGGAAATAAAGC GAAT CAT CCGGT AAAA	400
B2	T T CAGGGT TAGCAGATAA T GCAT CT GCTAAT GT A CGT AGAGAT GT GGT GT C T G G A A A T A A A G C T G A A T C A T C C G G T A A A A	400
FC	CTGCTGG CACTAAT GAGAA T CTGCTTATACGTT A CCGT T AGT CTT GCT GGT T T G G C T C A A G C T C T T A G G C T T G A G G A A	480
B2	CTGCTGG TACTAAT GAGAA C T G C T T A T A C G C T T A C T G T C A G T C T T G C T G G T T T A G C T C A A G C T C T T A G G C T T G A A G A A	480
FC	T T AATGT GGACT C GGGATAA G TTT GAGGACCGTT C A A A T T A C C A T G G A C A C C T G T T C A A G G T A G A A C C A G T C C A C C C G G	560
B2	T T AATGT GGAC C G G G A T A A A T T T G A G G A C C G G T T C A A A T T A C C A T G G A C A C C T G T T C A A G G T A G A A C C A G T C C A C C C G G	560
FC	ACAA	564
B2	ACAA	564

96,45 % d'identité

B

FC	MSSEGRYMT WKDMSHN R F MT DRWARVSDVSVI KQSHAMDLSKAANLSI I	50
B2	MSSEGRYMT WKDMSHN R F MT DRWARVSDVSVI KQSHAMDLSKAANLSI I	50
FC	KTALAGL GSGW T DNNPFVSPMTRFPQTLT MYGALVLYVNLSDPEFALI MT	100
B2	KTALAGL GSGW S DNNPFVSPMTRFPQTLT MYGALVLYVNLSDPEFALI MT	100
FC	KI S T L T D S G L A D N A S A N V R R D V V S G N K A E S S G K T A G T N E N S A Y T L T V S L A	150
B2	KV N T L T D S G L A D N A S A N V R R D V V S G N K A E S S G K T A G T N E N S A Y T L T V S L A	150
FC	GLA Q A L R L E E L M W T R D K F E D R L K L P W T P V Q G R T S P P G Q	188
B2	GLA Q A L R L E E L M W T R D K F E D R F K L P W T P V Q G R T S P P G Q	188

97,87 % de similarité
97,34 % d'identité

Figure 1.2 : Exemple d'un alignement de séquences nucléotidiques (A) et protéiques (B) de la CP, codée par l'ARN-2 d'isolat A (FC) avec celle d'un isolat B (B2). La nature des acides aminés en position 62,103 et 172 permet de préciser l'appartenance de la séquence analysée à l'isolat A (T₆₂, S₁₀₃, L₁₇₂) ou B (S₆₂, N₁₀₃, F₁₇₂). Ces positions clés sont encadrées en bleu dans les deux alignements. A l'exception des changements de séquence codant pour les acides aminés en positions 17, 62, 102, 103 et 172, les changements nucléotidiques entre les deux séquences sont des mutations synonymes. Les pourcentages de similarité et d'identité sont précisés pour chaque alignement.

1.2 Échantillonnage

Pour réaliser l'étude phylogénétique des isolats viraux dans le temps et dans l'espace, deux approches complémentaires ont été adoptées. La première a consisté à reproduire, en serre, l'infection naturelle de betteraves par semis en caissettes dans du terreau contenant 10% de terre provenant de parcelles touchées par la rhizomanie. Nous avons pu ensuite caractériser les isolats présents dans ces terres après infection des plantes. La seconde a fait appel à une collecte de pivots de betteraves symptomatiques provenant de plusieurs régions de France, d'autres pays d'Europe et des Etats-Unis. Cette analyse a été effectuée pendant quatre années consécutives, de 2000 à 2004. Les échantillons de terres et de pivots ont été obtenus en collaboration avec les membres de l'Institut International de Recherche Betteravière (IIRB) et l'INRA de Colmar.

Des approches successives ont été mises en place pour caractériser les ARN viraux :

- L'analyse par RT-PCR multiplex ciblant spécifiquement les ARN-2 et -5, après extraction des ARN, nous a permis de confirmer la présence du virus dans les plantes infectées. La présence d'ARN-2 permet entre autre de confirmer que l'échantillon est infecté par le BNYVV. La présence d'ARN-5 permet de définir immédiatement l'infection de la plante par un isolat contenant les cinq ARN.
- L'analyse des séquences de chaque isolat par RT-PCR haute fidélité ciblant spécifiquement les ARN-2, -3 et -5. Ces trois ARN permettent de caractériser complètement les isolats viraux. Les ARN-1 et -4 étant obligatoirement présents et n'étant pas utilisés pour la caractérisation des isolats viraux, leur séquençage n'a pas été réalisé. Après séquençage de l'amplicon spécifique de la séquence codante pour la CP (ARN-2) et la protéine P25 (ARN-3), le type d'isolat A ou B est caractérisé.

Ainsi, l'ensemble des séquences obtenu par cette approche, combiné aux données de séquences d'isolats collectés dans Genbank, nous a permis d'obtenir les séquences de 57 gènes codant pour la protéine CP, 66 pour la protéine P25 et 42 pour la protéine P26. Les numéros d'accès des séquences sont présentés dans l'appendice du chapitre 1, page 108. Cet ensemble de séquences nous a permis d'affiner l'analyse phylogénétique du virus.

S113	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNVCRAI NLLTHDNRTTL	50
Pith	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTSL	50
S5	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
Ch	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
412USA	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
I12	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
S10	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
NL3	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
F2	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTSL	50
NM	MGDI LGAVYDLGHRPYLARRTVYEDRLI LSTHGNICRAI NLLTHDNRTTL	50
S113	VYHNNT KRI RFRGLLQAYRVPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
Pith	VYHNNT KRI RFRGLLQSYHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
S5	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
Ch	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
412USA	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
I12	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
S10	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
NL3	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
F2	VYHNNT KRI RFRGLLQVCHGYPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
NM	VYHNNT KRI RFRGLLQAYHGRPY CGFRALCRVMLCSLPRLCDI PINGSRDF	100
S113	VADPTRLDSSVNE LLVSNM LVTI HYDRVHNVPI HTDGFVVDFTT VFRGPG	150
Pith	VADPTRLDSSVNE LLVSNGLVI HYDRVHNVPI HTDGFVVDFTT VFRGPG	150
S5	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
Ch	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
412USA	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
I12	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
S10	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
NL3	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
F2	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
NM	VADPTRLDSSVNE LLVSNGLVI HYDRVHHVPLHT DGFVVDFTT VFRGPG	150
S113	NFLLPNATNFPRTSITTTDHVYMVCLVNTVNCVLRFESELA VVHSGLYAGD	200
Pith	NFLLPNATNFPRTSITTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYAGD	200
S5	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
Ch	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
412USA	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
I12	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
S10	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
NL3	NFLLPNATNFPRTTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
F2	NFLLPNATNFPRTTTDQVYMVCLVNTVDCVLRFESELT VVHSGLYTGD	200
NM	NFLLPNATNFPRTSITTTDQVYMVCLVNTVNCVLRFESELT VVHSGLYTGD	200
S113	VL D V D N N V I Q A P D G V D D D D	219
Pith	VL D V D N N V I Q A P D G V D D D D	219
S5	VL D V D N N V I Q A P D G V D D D D	219
Ch	VL D V D N N V I Q A P D G V D D D D	219
412USA	VL D V D N N V I Q A P D G V D D D H D	219
I12	VL D V D N N V I Q A P D G V D D D D	219
S10	VL D V D N N V I Q A P D G V D D D D	219
NL3	VL D V D N N V I Q A P D G V D D D D	219
F2	VL D V D N N V I Q A P D G V D D D D	219
NM	VL D V D N N V I Q A P D G V D D D D	219

Minimum 96,36 % d'identité

Figure 1.3 : Alignement des 219 acides aminés de 10 séquences représentatives de la protéine P25 codée par l'ARN-3 du BNYVV. Les acides aminés 67 à 70 qui constituent la tétrade sont encadrés en bleu. La séquence d'adressage nucléaire de la protéine (NLS) est encadrée sur fond vert. Le motif "zinc finger" de la protéine est encadré sur fond cyan. Le motif NES est encadré sur fond violet.

1.3 Analyse des séquences

Comme le montre la Figure 1.2 A, l'alignement de la séquence codante de la protéine de capsid d'un isolat viral de type A (FC) à celle d'un isolat de type B (B2) montre des variations de séquences nucléotidiques. Cependant, le pourcentage d'identité obtenu reste de 96,45%.

Le même alignement effectué avec les séquences protéiques de capsides de ces deux isolats permet d'obtenir 97,87% de similarité et 97,34% d'identité (Figure 1.2 B). On remarque un pourcentage d'identité plus fort dans le cas des alignements de séquences protéiques car les mutations synonymes sont masquées. Sur cette Figure, la nature des acides aminés 62, 103 et 172 permet de distinguer sans ambiguïté un isolat A (FC par exemple, avec la séquence TSL) d'un isolat B (B2 par exemple, avec la séquence SNF) selon Miyaniishi *et al.* (1999). Les séquences nucléotidiques codant pour la CP du BNYVV étant très proches (de 95,39 à 100% d'identité), nous avons utilisé la méthode de Neighbor-Joining pour la reconstruction phylogénétique à partir des 57 séquences disponibles.

La même approche a été effectuée sur les séquences d'ARN-3 codant pour la protéine P25. Des identités de 96,36 à 100% ont été obtenues à partir des alignements des séquences protéiques. L'analyse des alignements a permis de localiser des variations de séquences protéiques peu nombreuses sur toute l'étendue de la phase ouverte de lecture de 219 acides aminés (10 conservatives et 24 non conservatives sur la Figure 1.3). Ces variations sont régulièrement réparties sur toute la séquence de la protéine P25. Toutefois, nous avons pu observer qu'il existait au sein de la séquence de la protéine P25, une région constituée de 4 acides aminés hypervariables, en position 67 à 70, localisés en aval de la séquence d'adressage nucléaire de la protéine (Vetter *et al.*, 2004), et en amont d'un motif "zinc finger" (Jupin *et al.*, 1992) (Figure 1.3). Cette région hypervariable que nous avons nommé "tétrade" est située entre deux domaines très conservés (acides aminés 50 à 66 et 71 à 90 respectivement). Nous avons aussi identifié 11 "tétrades" distinctes constituées par les séquences AYRV, SYHG, VCHG, VLHG, ACHG, ALHG, AFHG, AFHR, AYHR, AYHG et AHHG. Nous avons ensuite tenté de corréler la nature de cette "tétrade" au type d'isolat A ou B défini par la séquence de la protéine de capsid. Ainsi, les "tétrades" AYHR et

A	S12	ATGGATATTGATCATTGTATGCCTGCATTTGATATGGCATATAGCGACGA	50
	FP	ATGGATATTGATCATTGTATGCCTGTATTTGATATGGCATATAGCGACGA	50
	S12	TGATCACCTCCCTTATTACATACTGAGATCGACTCATAATGTGGTTCGGG	100
	FP	TAAACACCTTCTTATTACATACAGATCGACTCATAATGTGGTTCGGG	100
	S12	ATGTTGATTATACTGGTTTTATATGTTATCCTTTCAGGTGATCTTAAT	150
	FP	ATGTTGATTATACTGGTTTTATATGTTATCCTTTCAGGTGATCTTAAT	150
	S12	GATAATGTTGAGGTCGGTGCTGATATTTATCACATGAAGATTAAGACTAT	200
	FP	GATAATGTTGAGGTCGGTGCTGATATTTATCACATGAAGATTAAGACTAT	200
	S12	GAGACTTAATGTTGACATTTATAACAACGATG---T GCCACAAATTTAC	247
	FP	GAGACTTAATGTCGATATTTATAACAATGATGATG T GCCACAAATTTTC	250
	S12	CAGGTTGGGTTAGGTTTATTGTGTTTTGCACACCTCCTGTGTCTTCTTGG	297
	FP	CAGGTTGGGTTAGGTTTATTGTGTTTTGTACACCTCCTGTGTCTTCTTGG	300
	S12	GTTAATGATGGGTGCTCTAGTCTGTTTTCACCTTTTGTAGGTGTGAATAG	347
FP	GTTAATGC CGGTGCTCTAGTCTGTTTTCACCTTTTGTAGGTGTGAATAG	350	
S12	TTTTATTGATCCTAAATTTGTCAAACGATGGTTCATGGGATTACTGTGT	397	
FP	TTTTATTGACCTAAATTTGTAAAACGATGGTTCATGGGATTACTGTGT	400	
S12	TACACGATGGTATTTATTGCTTGTGCTATCAAGAACATTTTACCGGAAGT	447	
FP	TACATGATGGTATTTATTGTTTGCCTATCAAGAAATTTTACCGGAAGT	450	
S12	TTTGAGTTTAACTTTCTGGACCTGTAATTATACCCTAACTAATGATGT	497	
FP	TTTGAGTTCAACTTTCTGGACCTGTAATTATACCCTAACTAATGATGT	500	
S12	TTGTTGGTCCCGGCTACTAATGTAGATTCTATATATGTGGCTTGTGTTG	547	
FP	TTGTTGGTCCCGGCTACTAGTGTGATTCTATATATGTGGCTTGTGTTG	550	
S12	CTAGTTGGCATGGTGATAGTGCTTTCATGTTGCAGTCCGATAGTGTGCT	597	
FP	CTAGTTGGTGTGGTGATAGTGCTTTATGTTGCAGTCCGATAGTGTGGCT	600	
S12	TGGGTACACAAGAGATTTTGGCAACGACCTGTGTTGGAACTTGGGCAATG	647	
FP	TGGGTACACAAGAGATTTTGGCAACGACCTGTGTTGGAAATTTGGGCAATG	650	
S12	GTTGGATGATCTACCAGATCATGATAAT-----SATTGTGGATGA	687	
FP	TTTAGATGATCTACCAGATCATGATAATAATAATGATTGTGGATGA	699	

93 % d'identité

B	S12	MDI DHCMPAFDMAYSDDHLPYYI LRSTHNVVRDVDYTGFI CYPLQVDLN	50
	FP	MDI DHCMPVFDMAYSDDNHL PYYI QRSTHNVVRDVDYTGFI CYPLQVDLN	50
	S12	DNVEVGADI YHMKI KT MRLNVDI YNNLDVATNLPGWRFI VFCTPPVSSW	99
	FP	DNVEVGADI YHMKI KT MRLNVDI YNNLDVATNFPGWRFI VFCTPPVSSW	100
	S12	VNDGCSSLFSPFVGVSFI DPKL SKRDGHGI TVLHDGI YCLCYQEHFTGS	149
FP	VNGGCSSLFSPFVGVSFI DPKLLKRDGHGI TVLHDGI YCLRYQEFFTGS	150	
S12	FEFNFRGPGNYTLTNDVCWSPAT NVDSI YVACVASWHGDSAFMLQSDSVS	199	
FP	FEFNFRGPGNYTLTNDVCWSPAT SVDSI YVACVASWCGDSAFMLQSDSVA	200	
S12	WVHKRFWQRPVLELQWLDLDPHDN---DCG	228	
FP	WVHKRFWQRPVLEFGQCLDDLDPHDN NNNDCCG	232	

98,86 % de similarité
90,15 % d'identité

Figure 1.4 : Exemple d'un alignement de séquences nucléotidiques (A) et protéiques (B) de protéines P26, codées par l'ARN-5 du BNYVV, provenant d'un isolat européen (FP) et d'un isolat japonais (S12). Deux sites de délétion caractéristiques sont encadrés en bleu. Les pourcentages de similarité et d'identité sont précisés pour chaque alignement.

AYHG sont toujours retrouvées dans des séquences de protéines P25 provenant d'isolats B. Les autres séquences sont toutes associées à des isolats A, nous révélant ainsi l'existence d'une grande variabilité de ces isolats viraux. Cette variabilité sera discutée au paragraphe 1.5.

Nous avons ensuite intégré les données de séquence de l'ARN-5 (Figure 1.4 A et B) lorsque celui-ci était présent dans l'échantillon. Les alignements de séquences protéiques de l'ARN-5 codant pour la protéine P26 présentent des pourcentages d'identité situés entre 90,15 et 100%. Toutefois, la valeur minimale est à nuancer. Il semble que l'on soit ici en présence de deux populations distinctes d'ARN-5, l'une européenne présentant des identités de séquences proches de 99%, l'autre asiatique possédant les mêmes caractéristiques. L'existence de deux délétions de séquences (Figure 1.4) influe grandement sur le calcul des pourcentages d'identité entre ces ARN-5. De plus, ces différences majeures entre ARN-5 européens et asiatiques permettent aisément de différencier les deux groupes.

Ces séquences étant très proches, comme dans le cas des protéines CP et P25, nous avons privilégié la méthode de NJ pour construire les arbres phylogénétiques.

L'absence de séquence nucléotidique de référence (outgroup) nous a conduit à produire des phylogrammes non enracinés. Notre étude est donc intemporelle et reste basée sur les distances entre séquences.

1.4 Structure des populations virales

Nous avons vu au paragraphe précédent que les isolats de BNYVV A et B peuvent se distinguer par la nature de la tétrade (acides aminés 67 à 70) présents sur la séquence de la protéine P25.

Dans le but de faciliter le travail d'analyse, nous avons incorporé cette séquence de tétrade au nom de chaque taxon pour que cette caractéristique apparaisse sur les arbres phylogénétiques.

Pour identifier aisément les clusters d'isolats A, B et P, un code couleur a été utilisé pour tracer les branches des dendrogrammes. Les clusters regroupant les isolats A apparaissent en bleu, les isolats B en vert et P en rouge.

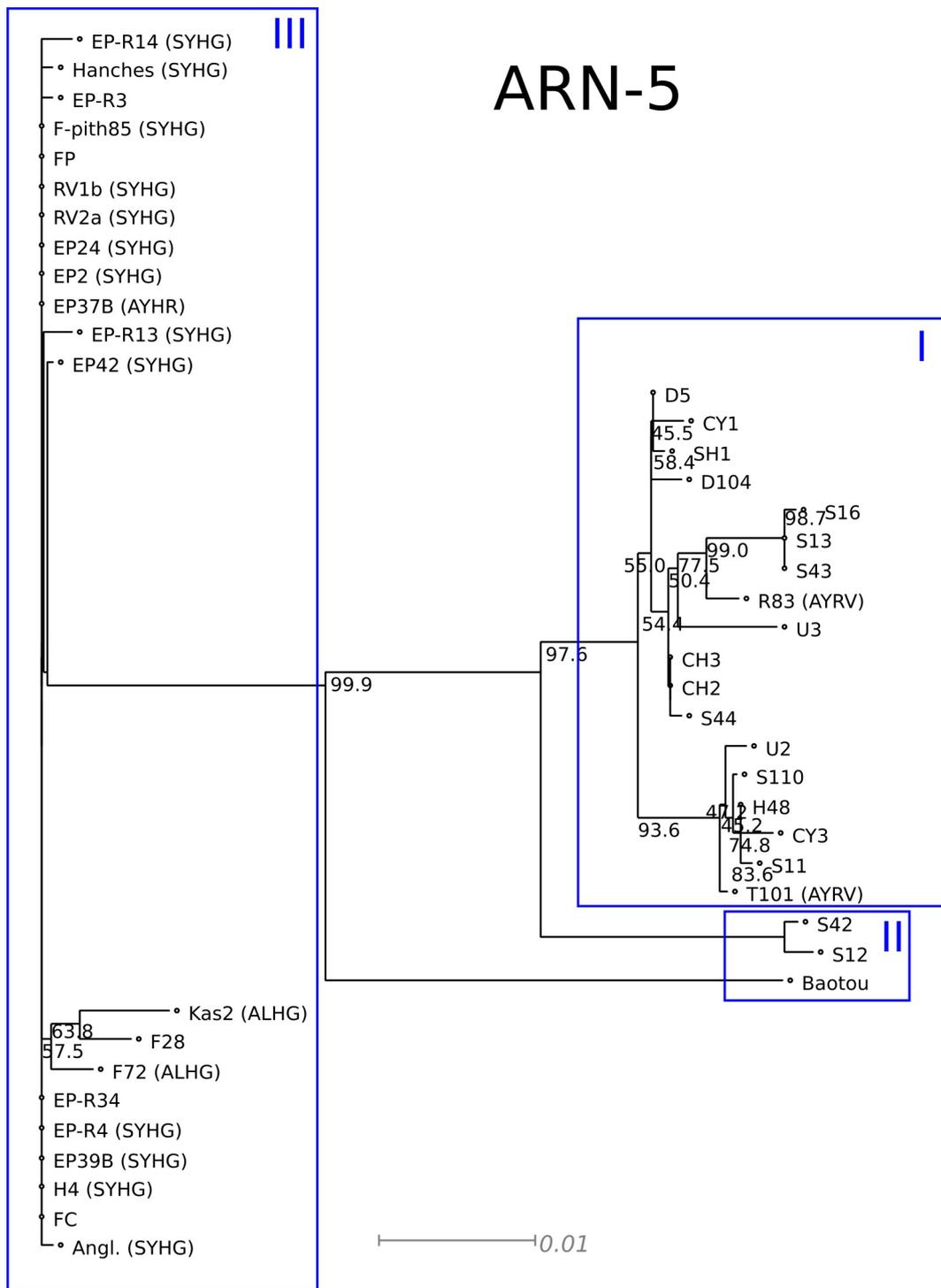


Figure 1.5 : Phylogramme représentant les distances génétiques des séquences nucléotidiques des ARN-5 du BNYVV. Les chiffres précisés sous les branches représentent le pourcentage de bootstrap effectué sur 1000 itérations. Seules les valeurs de bootstrap inférieures à 100% sont mentionnées. Entre parenthèses se trouve la séquence de la tétrade de la protéine P25 (si elle est connue). L'échelle en bas correspond à une substitution pour 100 nucléotides (1%). Les groupes similaires à ceux identifiés par Miyanishi *et al.* (1999) sont encadrés en bleus.

Les alignements des séquences nucléotidiques des phases ouvertes de lecture codant pour les protéines P26, CP et P25 ont été produits à l'aide de ClustalW (Thompson *et al.*, 1994). Les arbres phylogénétiques ont été produits à l'aide du logiciel SplitsTree4, en utilisant l'algorithme Hamming et en ignorant les gaps pour produire la matrice de distances initiale. Les dendrogrammes ont été obtenus avec la méthode NJ et en utilisant la longueur des branches pour représenter les distances entre séquences. Chaque arbre a été testé avec un bootstrap de 1000 répétitions.

1.4.1 Étude de la phylogénie de l'ARN-5

L'allure globale du phylogramme obtenu à partir de l'alignement des séquences nucléotidiques d'ARN-5 (Figure 1.5) permet de distinguer deux groupes de séquences, l'un issu des séquences d'origine asiatique (Figure 1.5, I et II), l'autre issu des séquences européennes (Figure 1.5, III). Cette ségrégation avait déjà été décrite par Miyanishi *et al.* (1999) à l'aide d'un échantillonnage moins conséquent et peut être observée par simple alignement de séquences (Figure 1.4) comme décrit précédemment.

Compte tenu de la nature de l'ARN-3 présent dans ces isolats (lorsqu'elle est connue), on remarque que tous les taxons proviendraient d'isolats A, à l'exception du taxon EP37B qui se distingue par la nature de type B de son ARN-3, déterminé par la séquence de tétrade AYHR.

On remarque également que les virus dont l'ARN-5 est d'origine anglaise (Angl.) et française (H4, RV1b, EP2, ...) proviennent tous de la même famille (motif SYHG) mais diffèrent dans trois cas par la nature du type d'ARN-3 rencontré dans chaque taxon et identifié par le motif tétrade (Kas2 et F72 associé à un motif ALHG, EP37B associé à AYHR).

Dans ce groupe III, nous retrouvons parmi les ARN-5 séquencés en 2004 (EP2, EP37B, ...) la séquence de l'ARN-5 obtenue en 1985 (F-pith85) qui est identique à la séquence EP39B, provenant du même endroit. Un tel résultat montre la très grande stabilité de la séquence de cet ARN au cours des vingt dernières années dans la région de Pithiviers. Une très forte pression de sélection existerait pour conserver cette séquence.

L'utilisation des séquences protéiques couplée à une correction de Jukes-Cantor pré-

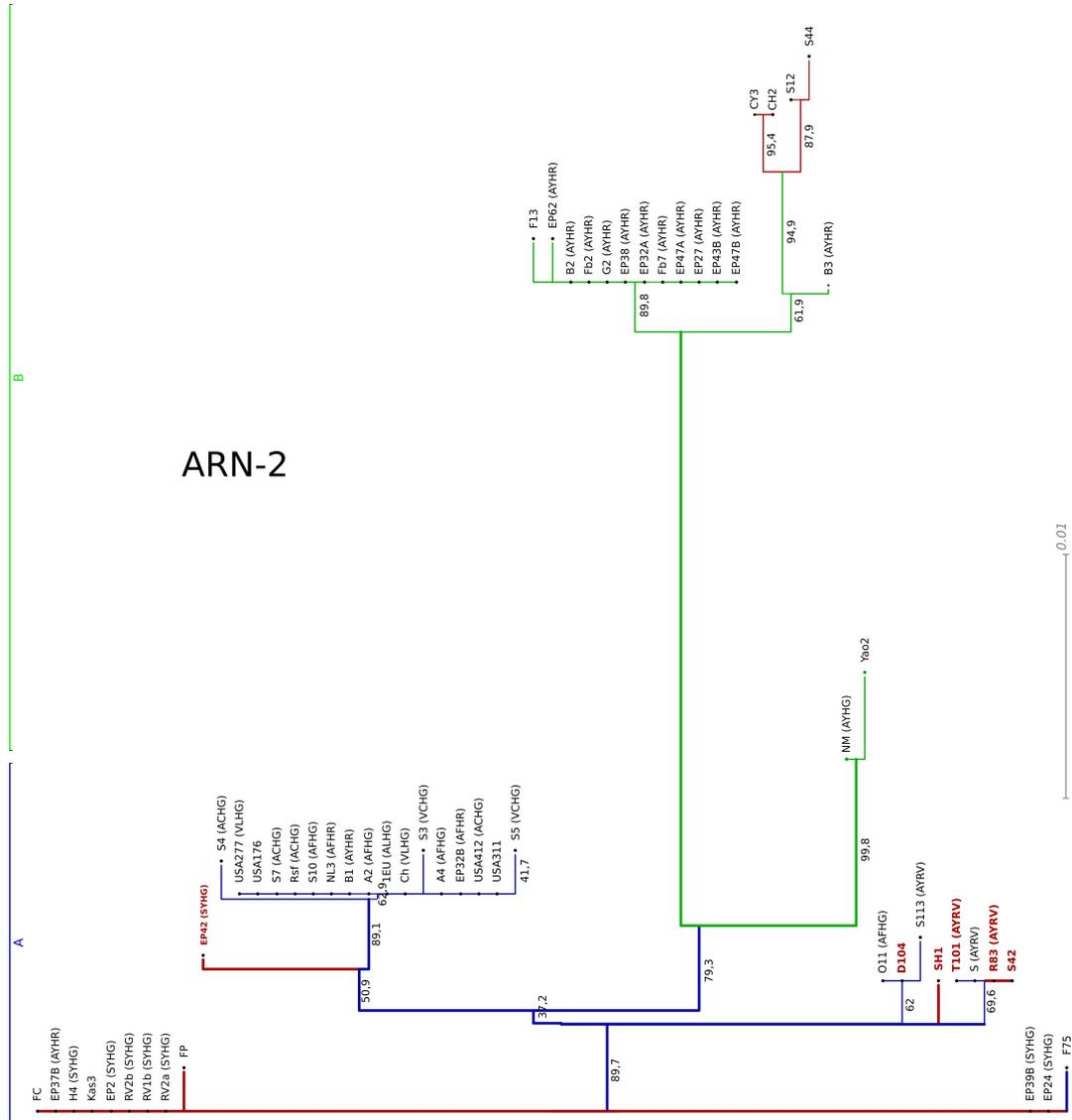


Figure 1.6 : Phylogramme représentant les distances génétiques des séquences nucléotidiques des ARN-2 du BNYVV. Les chiffres précisés sous les branches représentent le pourcentage de bootstrap effectué sur 1000 itérations. Seules les valeurs de bootstrap inférieures à 100% sont mentionnées. La couleur des branches représente l'isolat du virus correspondant : rouge = P ; bleu = A ; vert = B. Entre parenthèses se trouve la séquence de la tétrade de la protéine P25 si elle est connue. Les isolats A et B sont regroupés en deux groupes délimités par les barres bleues et vertes sur la gauche de l'arbre. L'échelle sur la droite correspond à une substitution pour 100 nucléotides (1%).

sente une répartition des taxons identiques au sein des groupes I, II et III, ce qui confirme nos observations (non montré).

Nous nous sommes ensuite intéressés à la distribution des isolats en analysant les séquences d'ARN-2 codant pour la protéine de capsid.

1.4.2 Étude de la phylogénie de l'ARN-2

L'allure globale du dendrogramme (Figure 1.6) laisse apparaître une nette ségrégation des taxons d'isolats A (bleu) et B (vert) qui se regroupent en deux clusters distincts. Au sein de chaque cluster, très peu de divergences sont observées.

En ce qui concerne les taxons possédant l'ARN-5, deux groupes se distinguent. Le premier est constitué des taxons d'origine européenne (FC, Kas3, FP, ...) très proches du cluster A. Le second est constitué des taxons d'origine asiatique (CY3, CH2, S12 et S44) retrouvés dans le cluster B. Par ailleurs, il est remarquable de trouver les isolats D104, SH1, T101, R83 et S42 asiatiques possédant l'ARN-5 dans un cluster de type A regroupant à notre connaissance trois souches asiatiques exemptes d'ARN-5 (O11, S113 et S).

La répartition du motif "tétrade" est bien corrélée avec l'origine A ou B des isolats viraux. Cependant, on retrouve encore le taxon EP37B qui possède un ARN-5, qui se localise dans le cluster A et présente une séquence hypervariable de la protéine P25 AYHR, caractéristique des isolats B.

La même approche sur les séquences protéiques en utilisant une correction de Jukes-Cantor révèle une même répartition des groupes de séquences au sein des clusters de taxons, à l'exception des séquences NM et Yao2 (de type B) qui se retrouvent branchées sur le cluster des CP asiatiques, sur le noeud CH2-CY3 plus précisément (données non montrées). Ceci révèle une petite incongruence à ce niveau des arbres.

1.4.3 Étude de la phylogénie de l'ARN-3

La reconstruction phylogénétique à partir des séquences nucléotidiques d'ARN-3 codant pour la protéine P25 est représentée Figure 1.7.

L'allure globale du phylogramme obtenu est proche de celui obtenu pour l'ARN-2, re-

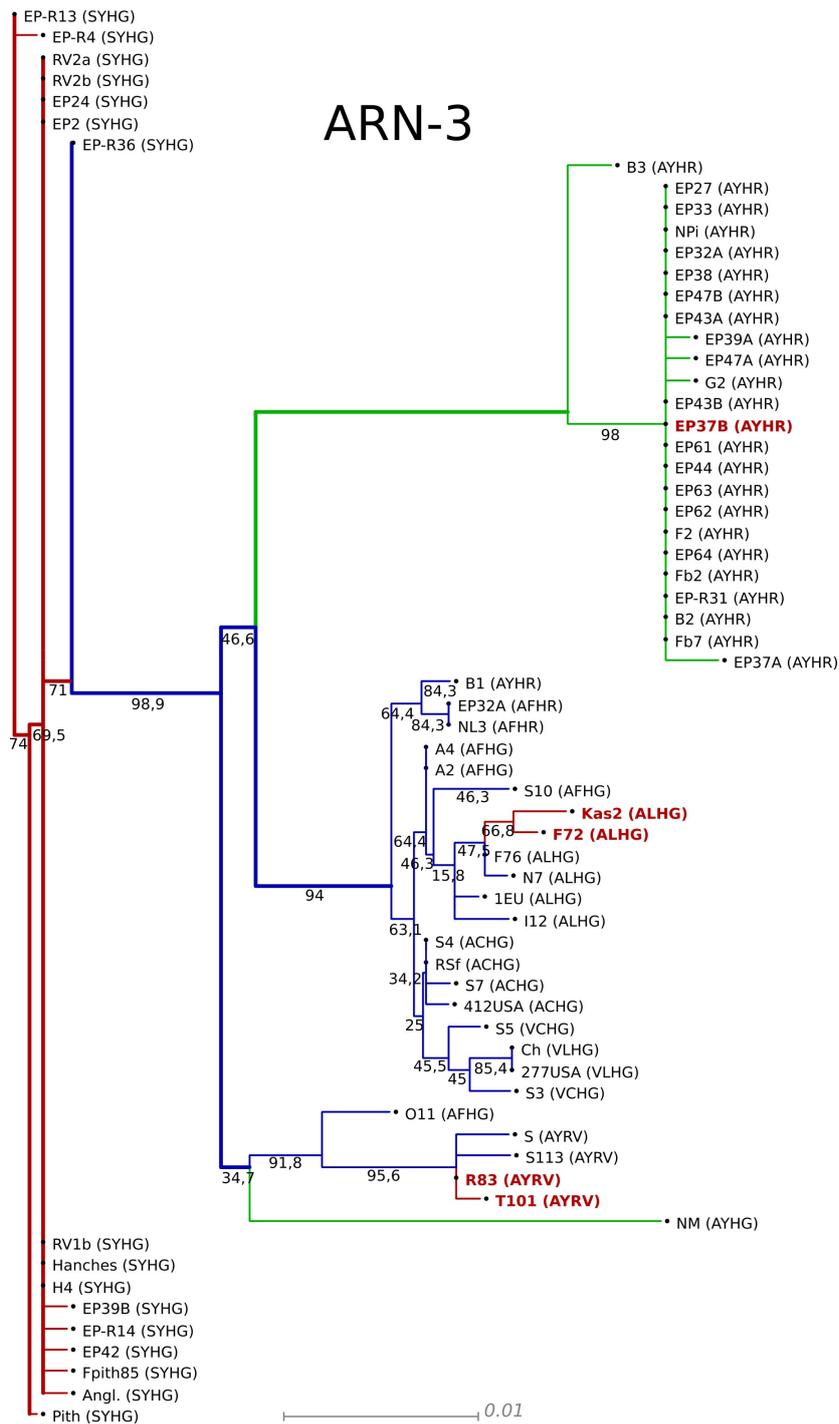


Figure 1.7 : Phylogramme représentant les distances génétiques des séquences nucléotidiques des ARN-3 du BNYVV. Les chiffres précisés sous les branches représentent le pourcentage de bootstrap effectué sur 1000 itérations. Seules les valeurs de bootstrap inférieures à 100% sont mentionnées. La couleur des branches représente l'isolat du virus correspondant : rouge = P ; bleu = A ; vert = B. Entre parenthèses se trouve la séquence de la tétrade de la protéine P25. L'échelle en bas correspond à une substitution pour 100 nucléotides (1%).

groupant les taxons de type A et B dans des clusters bien distincts. Cependant on remarque immédiatement une très grande diversité des ARN-3 au sein des taxons étudiés, en particulier pour les isolats A (bleu). Les séquences rattachées à cet isolat présentent un motif tétrade très varié. Dans le cas des isolats B (vert), les séquences nucléotidiques sont beaucoup moins variables. En effet, toutes les protéines P25 contiennent le motif tétrade AYHR, à l'exception notable de l'isolat chinois NM (AYHG) qui est classé parmi les isolats B sur la base de la séquence de la protéine CP (acides aminés S, N et F) mais que l'on retrouve proche du cluster de type A. Toutefois, les valeurs de bootstrap des noeuds séparant NM du cluster de type B sont faibles (34,7 et 46,6%) montrant que ce branchement n'est probablement qu'indicatif.

En ce qui concerne les isolats contenant l'ARN-5 (rouge), la variabilité des séquences d'ARN-3 est très faible et les protéines P25 codées par ces isolats sont toutes de type SYHG. Toutefois, on retrouve deux taxons contenant ce cinquième ARN qui sont contenus dans le groupe A (Kas2 et F72, séquence de la tétrade : ALHG) et toujours le taxon EP37B (AYHR), décrit précédemment, qui est retrouvé dans le cluster B.

L'alignement des séquences protéiques et la reconstruction phylogénétique par NJ en incorporant une correction de Jukes-Cantor produit un arbre identique (non montré).

1.5 Interprétation globale de l'étude phylogénétique

Avant d'effectuer une analyse détaillée des alignements et des arbres phylogénétiques produits pour les ARN-2, -3 et -5, il est nécessaire de rappeler les conditions expérimentales de collecte d'échantillons qui pourraient affecter l'interprétation des résultats. Nous ferons abstraction ici des possibilités d'erreur de recopie de matrice lors de la transcription inverse. Les erreurs au cours de l'étape de PCR ont été limitées par l'utilisation d'enzymes ayant une bonne activité correctrice. De plus les échantillons ont été analysés par plusieurs séquençages indépendants. Cependant, lors du prélèvement du matériel végétal infecté, la première étape incontournable consiste en un broyage des tissus afin d'isoler spécifiquement le matériel viral. Dans le cas d'une infection mixte par deux isolats différents, les particules virales se retrouvent ainsi mélangées et peuvent conduire à :

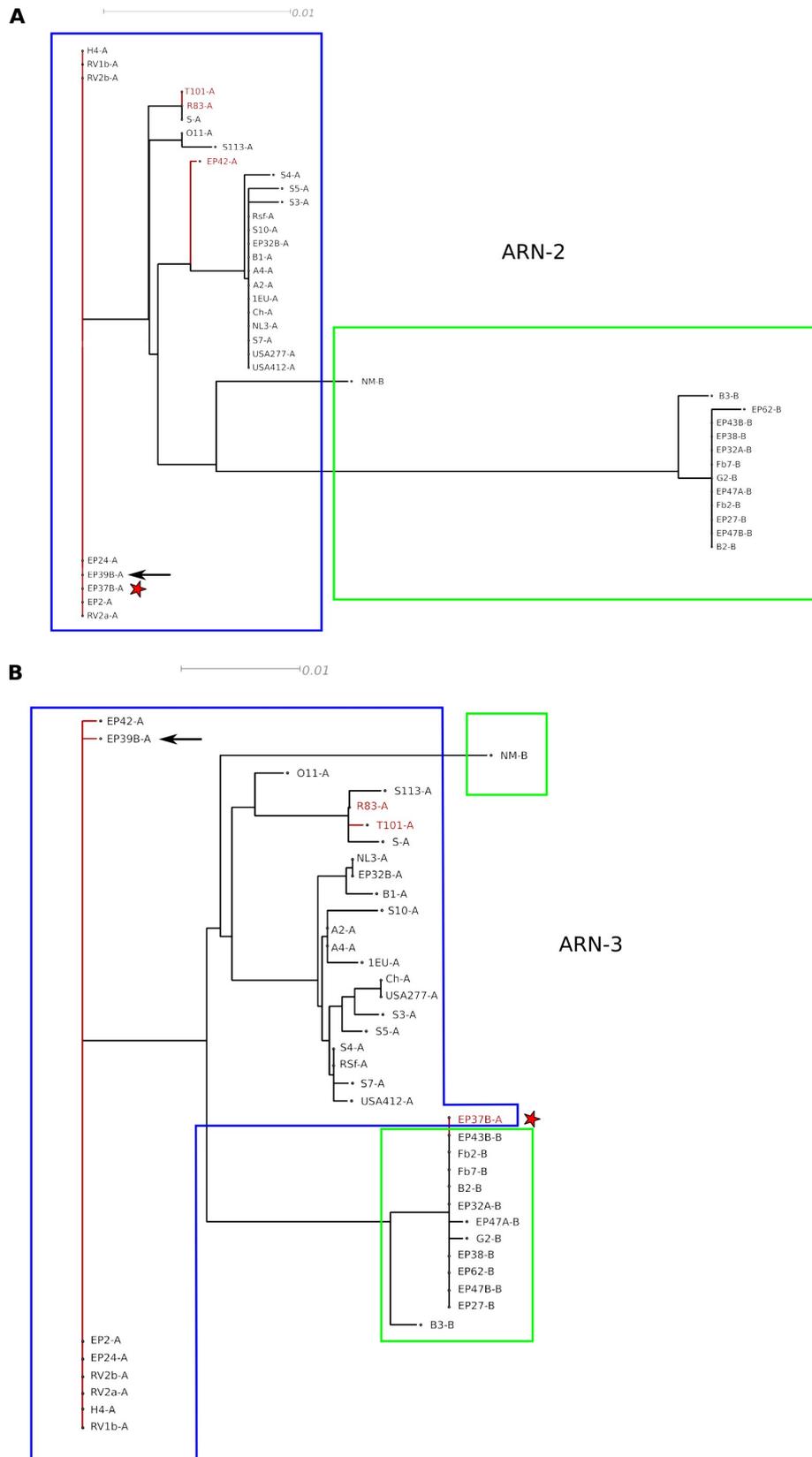


Figure 1.8 : Phylogrammes représentant les distances génétiques des séquences nucléotidiques des ARN-2 et -3 du BNYVV, chaque isolat ayant été séquençé pour au moins ces deux ARN. La couleur rouge des branches représente la présence de l'ARN-5 du virus. Le nom de chaque taxon est suivi du type (A ou B) défini par les séquences de l'ARN-2. L'échelle du haut de chaque dendrogramme correspond à une substitution pour 100 nucléotides (1%). Les cadres bleus désignent les clusters regroupant les isolats A, les cadres verts ceux de type B. L'étoile rouge et la flèche désignent respectivement les taxons EP37B et EP39B. A) Phylogramme des séquences nucléotidiques de la protéine CP. B) Phylogramme des séquences nucléotidiques de la protéine P25.

- l'obtention de séquences d'ARN-2, -3 et -5 correspondant à un isolat viral inexistant dans la nature et qui posséderait par exemple un ARN-2 du premier isolat et des ARN-3 et -5 d'un second.
- l'obtention d'isolats ayant des caractéristiques de séquence chimérique comme par exemple une partie de la séquence de l'ARN-3 du premier isolat suivie de celle du second. Cette situation peut survenir lors de sauts de matrice au cours de l'étape de PCR (Koenig and Büttner, 2004).

Aucun moyen ne permet actuellement de contourner ce biais et nos résultats devront être interprétés en gardant à l'esprit cette limitation.

L'analyse globale de nos résultats montre que la variabilité moléculaire des isolats de BNYVV est relativement faible mais pourrait tout de même expliquer les changements d'agressivité du virus sur plante hôte.

Globalement, les isolats possédant l'ARN-5 se retrouvent dans les mêmes clusters, que ce soit dans la reconstruction phylogénétique basée sur la CP ou sur la protéine P25 (Figure 1.6 et 1.7).

Cependant, pour faciliter l'interprétation de nos résultats et éliminer les taxons moins informatifs, nous avons reproduit la même analyse sur les isolats dont au moins les séquences des ARN-2 et -3 ont été caractérisées. Pour cette analyse, la présence de l'ARN-5 est mise en évidence par la couleur rouge et le type de la protéine CP par la lettre A ou B ajoutée au nom du taxon. Par exemple, les isolats USA412 et EP37B seront renommés respectivement USA412-A et EP37B-A. Nous avons pu produire des alignements et des dendrogrammes pour 42 isolats (Figure 1.8 A et B).

Le phylogramme de l'ARN-2 (Figure 1.8 A) montre à nouveau la ségrégation nette entre les isolats A et B, sans ambiguïté possible. On remarque toujours la stabilité et le regroupement des isolats contenant un ARN-5. On peut noter que l'isolat EP42 est la seule exception et se retrouve dans un cluster de type A.

Pour le phylogramme de l'ARN-3 (Figure 1.8 B), on retrouve la même répartition entre isolats A et B. Les isolats A contenant l'ARN-5 sont regroupés à la base de l'arbre et les isolats B sont regroupés à l'extrémité opposée. Les autres isolats de type A forment un cluster très variable entre les deux. Sur ces phylogrammes, on remarque l'isolat EP37B (désigné

par une étoile rouge sur la Figure 1.8) qui possède un ARN-2 de type A, un ARN-3 de type B et contient un ARN-5. Une seconde analyse, récente, effectuée sur le même échantillon de betterave infectée (EP37B-2, numéro d'accèsion AY734502) a révélé la présence d'un ARN-2 de type A, d'ARN-5 et d'ARN-3 de type A. En effet, la protéine P25 exprimée dans ce cas présente la tétrade SYHG. Sur l'arbre, un tel isolat est localisé au même niveau que l'isolat EP39-B (désigné par une flèche sur la Figure 1.8). Ce résultat particulier met en lumière une situation de co-infection de la plante par deux isolats distincts, l'un de type A, l'autre de type B. Lors de l'analyse, ces deux populations se retrouvent mélangées, ce qui explique les résultats initiaux obtenus pour l'isolat EP37B. Nous ne pouvons cependant pas écarter la possibilité d'existence de pseudo-recombinants dans cet échantillon de betterave.

1.6 Conclusions

Les souches mondiales de BNYVV analysées dans cette étude montrent un maximum de 10% de différences de séquences. L'ARN-5 est très stable au cours du temps dans la région de Pithiviers, démontrant une très forte pression de sélection en faveur de la conservation de la séquence de cet ARN. Cette stabilité d'un phytovirus à ARN n'est pas un cas isolé. En effet, Fraile *et al.* (1997) ont montré que le TMGMV et le TYMV, géographiquement très éloignés, présentent des séquences très proches. De plus, les virus à ARN transmis par vecteurs, sont soumis à une forte pression de sélection pour la conservation de leur séquence CP et de tous les motifs protéiques exposés à l'interface vecteur/virus et cellule/virus (Chare and Holmes, 2004). Cette forte conservation pourrait suggérer un rôle très important de la protéine P26 dans le cycle viral. Ce rôle reste cependant à définir.

Nous avons également observé que l'ARN-5 était présent majoritairement dans des souches de type A (Kas3, FC, FP, ...) et rarement détectées dans des souches de type B (dans le cas des isolats asiatiques CY3, S44, CH2 et S12). Grâce à l'alignement des séquences nucléotidiques de l'ARN-5, nous avons pu définir deux groupes de séquence principaux dans la reconstruction phylogénétique. Le premier regroupant tous les ARN-5 européens, appelé type P, et le deuxième regroupant les ARN-5 japonais, que nous avons dénommé

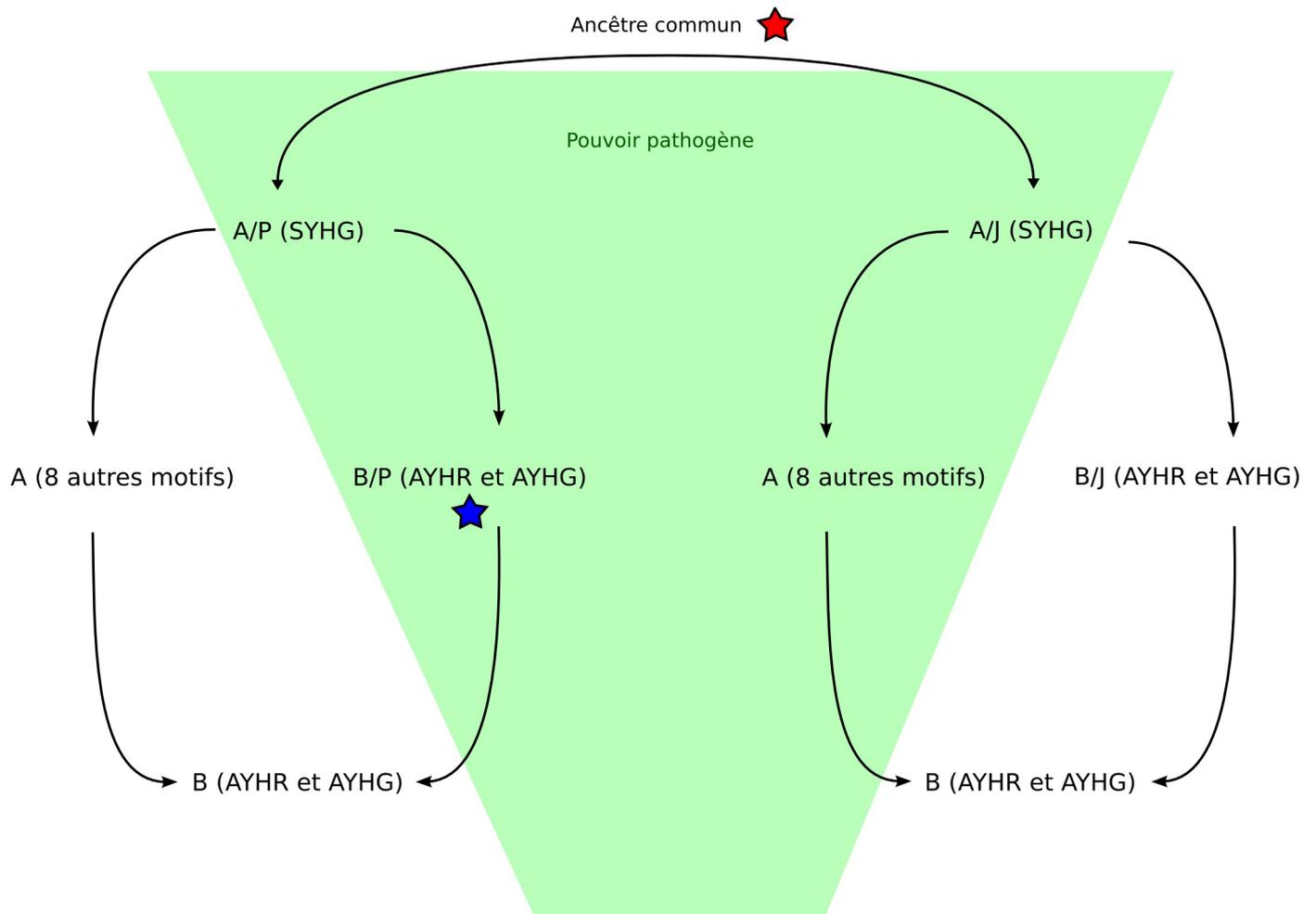


Figure 1.9 : Représentation schématique d'une hypothèse évolutive d'isolats de BNYVV. Le motif de tétrade correspondant au type de BNYVV est précisé entre parenthèses. L'étoile rouge montre l'emplacement d'un ancêtre commun hypothétique. L'étoile bleue montre la combinaison B/P qui reste à mettre en évidence. Le trapèze vert schématise une décroissance du pouvoir pathogène entre chaque type d'isolats, de A/P vers B.

type J. Ces dernières observations nous permettraient de penser que l'isolat P n'aurait pas une réelle existence en tant que sous-population.

Il existerait plutôt des souches de type A/P, A/J ou B/J. Avec cette nouvelle classification et nos observations, cinq isolats sont trouvés :

- A : cas des isolats T101, EP32B, USA412 ...
- B : NM, G2, EP47A, ...
- A/P : FC, H4, RV1b, ...
- A/J : D104, SH1, T101, R83 et S42
- B/J : CY3, CH2, S12 et S44

La combinaison B/P n'ayant pas été observée (Figure 1.9, étoile bleue). On constate une double dichotomie : A et B pour les ARN-2 et -3, P et J pour les ARN-5. Cette dernière montre clairement la séparation de deux sous-populations du virus entre l'Europe et l'Asie à partir d'un ancêtre commun, que notre étude ne permet malheureusement pas de situer (Figure 1.9, étoile rouge).

D'un point de vue évolutif, nous pouvons émettre une hypothèse visant à la stabilisation des isolats viraux et à la diminution de leur pouvoir pathogène, décroissant de l'isolat A/P à l'isolat B et qui est schématisée sur la Figure 1.9 par le tétraèdre vert. Cette perte de pouvoir pathogène rendrait les souches de virus plus adaptées à la survie de l'hôte et donc au maintien du virus dans les terres infestées. Dans cette hypothèse, les isolats de type A/P ou A/J, très pathogènes, pourraient perdre leur ARN-5. Leur ARN-3 subirait alors de fortes contraintes sélectives conduisant à l'apparition d'une forte variabilité de séquence. Le motif de tétrade tendrait à évoluer vers un motif de type B, AYHR, par mutation de la séquence de l'ARN-3 qui contrebalancerait la perte de fonction portée par l'ARN-5. Une autre voie évolutive aboutissant au même type d'isolat passerait par une variabilité de la séquence de l'ARN-3 aboutissant à un motif de tétrade de type B (AYHR ou AYHG) puis à la perte de l'ARN-5. Cette dernière voie permet d'expliquer la présence des isolats de type B/J mais pose la question du facteur déclenchant la variabilité initiale de l'ARN-3, facilement compréhensible dans la première voie.

Une telle hypothèse est bien sûr dépendante de l'absence de co-infection dans les échantillons analysés. L'existence des isolats A, B et A/P est bien établie par le nombre d'échan-

tillons analysés et la co-infection paraît un facteur assez marginal car seul l'isolat EP37B relève de ce cas de figure dans notre étude. Par contre, les séquences des isolats de type A/J et B/J proviennent uniquement d'accessions issues de GenBank. Aucune autre donnée expérimentale n'est disponible pour valider leur existence réelle.

L'existence des isolats caractérisés lors de cette étude pourrait également être discutée, si l'on s'en réfère aux travaux effectués sur le CaMV par Blanc *et al.* (2003) qui montrent la très forte diversité de ce virus après les premières étapes de réplication, cœur de la notion de "quasi-espèce". Seule une partie de la descendance est sélectionnée par ses propriétés biologiques. L'analyse par extraction effectuée sur betteraves infectées nous place dans le même contexte et ne nous permet pas de visualiser d'isolat réellement émergent.

Ainsi, seule l'utilisation d'approches expérimentales combinées pourra réellement confirmer notre modèle et lever certaines incertitudes. De telles approches pourraient consister en :

- l'échantillonnage des isolats de la zone asiatique pour confirmer la présence des types A/J et B/J
- l'infection de plantes pourrait être effectuée à l'aide de différentes combinaisons d'ARN et suivie au cours du temps
- le suivi de l'évolution d'un isolat prélevé à des endroits identiques en fonction du temps, si l'échelle de temps de l'évolution du virus est compatible avec la durée du test effectué
- la détection en conditions naturelles du type B/P en poursuivant l'échantillonnage de la zone Pithiviers. Un tel isolat a pu être produit au laboratoire (voir chapitre 3)
- la caractérisation individuelle d'ARN présents dans plusieurs cellules infectées d'une même plante hôte
- le mélange d'isolats A/P et B pour l'infection d'une même plante et l'analyse de l'isolat émergent

Pour ces deux derniers points, il faut savoir que Fraile *et al.* (1997) ont montré que la co-infection de *Nicotiana glauca* par le TMGMV et le TMV provoque la disparition progressive de la population de TMV. De plus, dans le cas du CMV (Cucumber mosaic virus), une analyse de la distribution spatiale d'une co-infection avec deux souches appar-

tenant à deux sous-groupes du virus induit leur exclusion mutuelle dans les tissus de l'hôte (Takeshita *et al.*, 2004). Il serait donc intéressant de savoir si les isolats du BNYVV sont capables de s'exclure mutuellement, malgré leur forte identité de séquence. Si tel est le cas, ces données auraient également une forte implication dans l'analyse des isolats lors de co-infection.

Cette analyse phylogénétique fait l'objet d'une publication intitulée "Extensive molecular epidemiology and phylogenetic analysis of worldwide collected beet necrotic yellow vein virus isolates" par Audrey SCHIRMER, Didier LINK, Benoît MOURY, Monique BEUVE, Alexandre MEUNIER, Claude BRAGARD, David GILMER et Olivier LE-MAIRE. Cette publication sera soumise à "Journal of General Virology".

Enfin, les analyses ont montré que le motif tétrade paraît être un excellent marqueur pour identifier les isolats de BNYVV. Le motif de tétrade est une séquence hypervariable située au milieu de deux îlots de stabilité. Ainsi, ce motif pourrait se situer dans une boucle joignant deux hélices α ou deux feuilletts β parfaitement conservés. Seules la cristallisation et la résolution de la structure tridimensionnelle de cette protéine pourra nous apporter des éléments de réponse à cette question. L'importance de la séquence "tétrade" dans le pouvoir pathogène d'isolats de BNYVV et dans la symptomatologie est abordée au chapitre suivant.

Chapitre 2

Etude de l'effet de la séquence "tétrade" sur les propriétés biologiques de la protéine P25

2.1 Introduction

Selon l'isolat viral et la plante hôte, les symptômes induits par l'infection peuvent fortement varier, allant d'une "simple" décoloration locale à une forte perturbation de la croissance et du développement, voire à la mort de la plante infectée. Les symptômes visibles d'une maladie sont induits par des processus complexes où des facteurs viraux et cellulaires modifient le métabolisme de la plante hôte. Les déterminants de la symptomatologie induite par le BNYVV sont portés par les ARN-3 et -5. Nous nous sommes donc intéressés aux propriétés biologiques de la protéine P25 codée par l'ARN-3 et aux propriétés biologiques de l'ARN-5 qui sont détaillées au chapitre suivant.

L'ARN-3 est responsable de l'apparition des symptômes caractéristiques de la rhizomanie et favorise l'amplification du virus dans les racines de betteraves (Koenig *et al.*, 1991; Richards and Tamada, 1992; Tamada *et al.*, 1999). La plus grande phase ouverte de lecture portée par cet ARN permet l'expression de la protéine P25 (Ziegler-Graff *et al.*, 1988). Cette protéine est responsable de l'aggravation des symptômes foliaires sur *Chenopodium quinoa*, *Tetragonia expansa* (Jupin *et al.*, 1992) et elle induit une prolifération racinaire sur

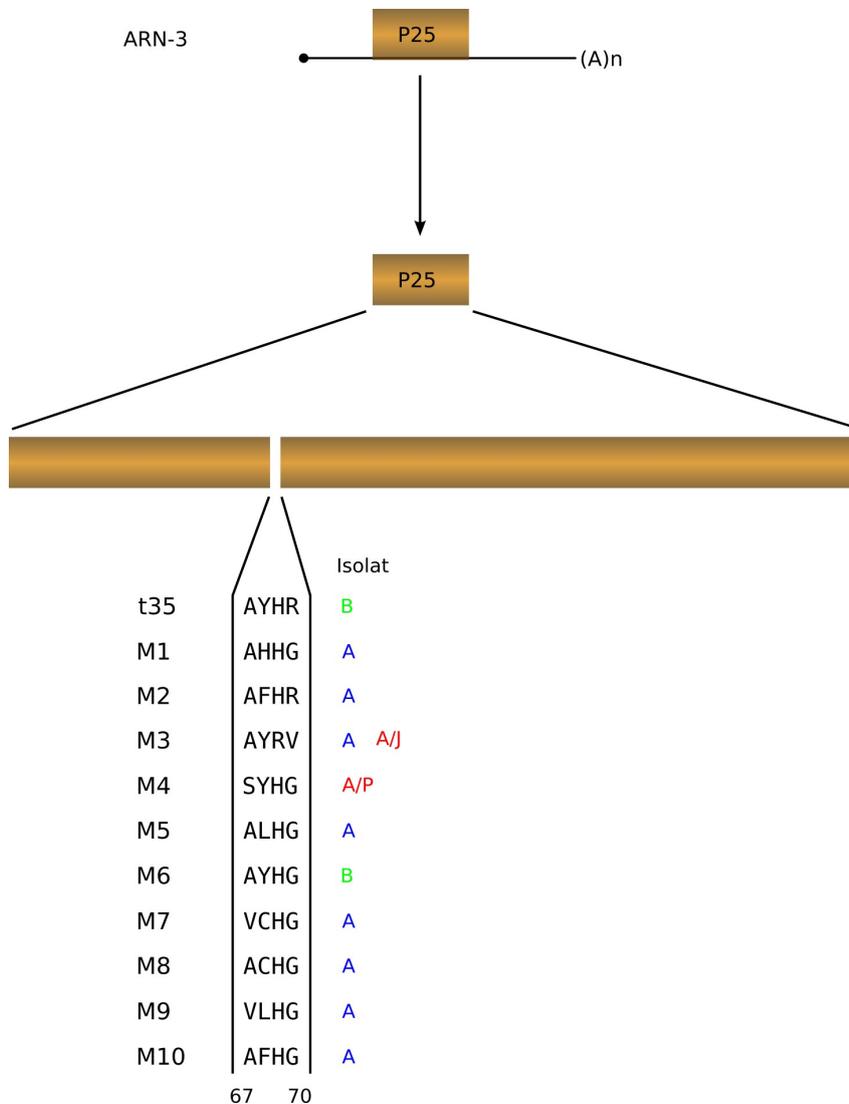


Figure 2.1 : Représentation simplifiée de l'ARN-3 et de la protéine P25 exprimée par sa phase ouverte de lecture 5' proximale. Les changements en aminoacides entre les positions 67 à 70 observés sur la séquence de protéines P25 issues d'isolats naturels caractérisés au cours de l'étude phylogénétique, ont été reproduits par mutagenèse dirigée du clone complet infectieux t35. La nature des acides aminés modifiés est précisée à côté de chaque mutant (M1 à M10) ainsi que l'isolat d'origine.

Beta macrocarpa (Tamada *et al.*, 1999). Quelques séquences et motifs particuliers ont déjà été caractérisés sur la protéine P25 (Figure 1.3 du chapitre 1) et certains de ces domaines sont importants dans l'établissement de la symptomatologie foliaire caractéristique, comme ce fût montré pour les motifs NLS et NES, essentiels à l'apparition de lésions locales de type chlorotique jaune sur *C. quinoa* (Vetter *et al.*, 2004) et le domaine C-terminal acide (Jupin *et al.*, 1992).

Comme nous l'avons vu au chapitre précédent, la variabilité de la séquence de la protéine P25 codée par l'ARN-3 est particulièrement marquée au niveau de la région directement en aval de la séquence NLS, c'est à dire au niveau des acides aminés 67 à 70. Nous avons donc cherché à caractériser l'effet biologique de protéines P25 mutées arborant de telles variabilités de séquences. Nous disposions au laboratoire des quatre clones ADNc permettant l'obtention de transcrits infectieux du BNYVV de type B. Nous avons donc réalisé des expériences de mutagenèse dirigée sur le clone complet de l'ARN-3 pour produire les mutants M1 à M10 (Figure 2.1). Ainsi, nous avons pu reproduire l'ensemble des 11 "tétrades" décrites et nous avons entrepris de caractériser les propriétés des protéines P25 ainsi modifiées.

L'approche utilisée a consisté à infecter des plantes hôtes et à analyser l'incidence de la séquence "tétrade" sur l'amplification des ARN viraux, sur l'accumulation des protéines P25 et surtout sur les symptômes foliaires. Plusieurs plantes hôtes ont été utilisées et inoculées avec différentes combinaisons d'ARN (1+2+3 ; 1+2+3+4 ; 1+2+3+4+5). Je ne présenterai ici que les résultats les plus significatifs obtenus sur *Chenopodium quinoa*, *Tetragonia expansa*, *Beta macrocarpa* et *Beta vulgaris*, inoculés en présence des ARN-1, -2 et -3.

Nous avons ensuite caractérisé la capacité de ces variants de la protéine P25 à activer la transcription et à dimériser. En effet, des travaux antérieurs réalisés au laboratoire par Elodie Klein (DEA, ULP 2003) et Guillaume Vetter (Thèse de doctorat, ULP 2003) avaient permis de montrer que la protéine P25 (isolat B) fusionnée au domaine de fixation aux acides nucléiques du facteur de transcription GAL4 (GAL4BD) était capable de créer un facteur de transcription chimérique fonctionnel. Ce facteur de transcription GAL4BD::P25 active (faiblement) la synthèse du gène rapporteur *His3* dans la levure, qui devient prototrophe pour cet acide aminé (voir chapitre Matériel et Méthodes). La protéine P25 est donc capable de

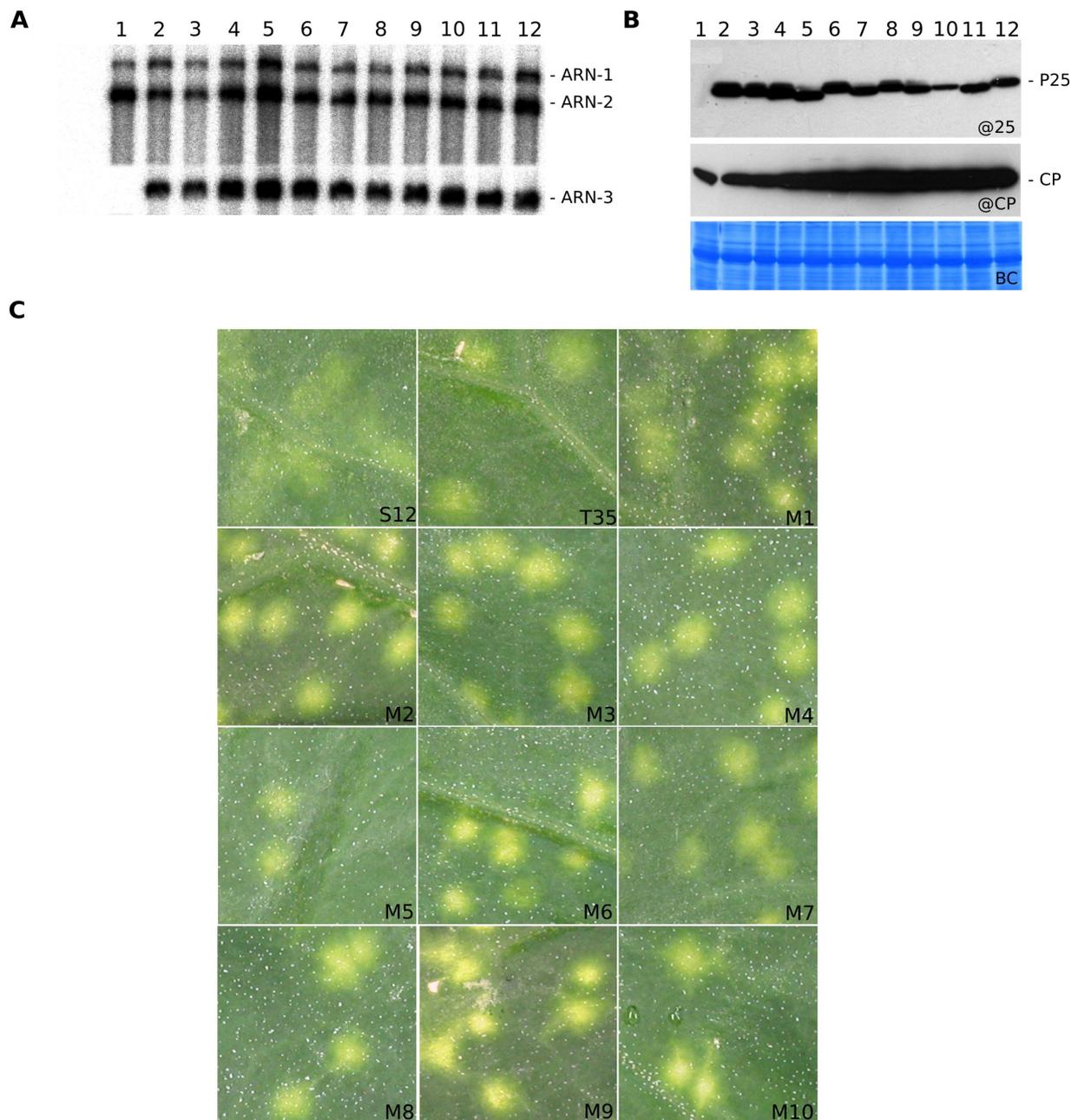


Figure 2.2 : Analyse de la réplication des ARN-3 recombinants (A) et de l'accumulation des protéines virales P25 et CP (B) dans les lésions locales de feuilles de *Chenopodium quinoa* inoculées par la souche helper Stras12 (1) supplémentée des transcrits t35 (2), M1 (3), M2 (4), M3 (5), M4 (6), M5 (7), M6 (8), M7 (9), M8 (10), M9 (11) et M10 (12), et de la symptomatologie foliaire 6 jours après infection (dpi). A) Analyse par northern-blot de l'accumulation des ARN viraux 6 dpi. Les quantités d'ARN totaux sont identiques. Sondes utilisées : ARN-1 et -2 (Lemaire *et al.*, 1988); ARN-3 (Lauber *et al.*, 1997). B) Analyse par western-blot de la présence des protéines P25 (@P25), de l'accumulation de la protéine de capside (@CP) et des protéines totales transférées sur membrane et colorées (BC). @P25 : immunodétection de la protéine P25; @CP immunodétection de la protéine CP; BC : coloration au Serva Blue. C) Symptômes observés 6 dpi après inoculation mécanique sur feuilles de *C. quinoa*. Les photographies ont toutes été prises dans les mêmes conditions.

remplacer le domaine d'activation de la transcription classiquement utilisé dans le système double hybride. Par mutagenèse dirigée, le domaine responsable de la transcription (AD1) avait pu être localisé entre les acides aminés 103 à 146 de la protéine P25. Ce domaine, fusionné à GAL4BD (GAL4BD::AD1) active fortement la transcription du gène *His3*. La particularité de ce système "simple hybride" réalisé avec la protéine P25 tient au fait que l'activation de la transcription observée dans les levures HF7C transformées par le vecteur pGBT9-P25 (GAL4::P25) disparaît lorsque les levures contiennent un vecteur pGAD424, même vide. Cette perte d'activation n'est pas observée si la protéine GAL4BD::AD1 est exprimée. Cette propriété particulière est mal comprise mais permet la recherche de partenaires cellulaires de la protéine P25 par la technique du double hybride (voir chapitre Matériel et Méthodes). Ainsi, Elodie Klein a pu démontrer au laboratoire que la protéine P25 interagissait avec elle-même. Ce résultat a pu être confirmé par Guillaume Vetter lors d'expériences de Far-western-blot. Nous avons donc entrepris de caractériser d'une part la capacité des variants de protéines P25 fusionnés à GAL4BD à activer la transcription en système simple hybride et d'autre part à rechercher si ces protéines pouvaient dimériser.

2.2 Effet des acides aminés 67 à 70 sur les symptômes foliaires de plantes hôtes

2.2.1 Effet sur *Chenopodium quinoa*

Le transcrit t35 du clone complet pB35 (Ziegler-Graff *et al.*, 1988) et des 10 variants M1 à M10 (Figure 2.1) ont été inoculés à des feuilles de *C. quinoa* en présence de la souche helper Stras12 qui contient les ARN-1 et -2, suffisants à produire une infection (Quillet *et al.*, 1989). Six jours après inoculation (6 dpi) les lésions locales ont été photographiées, découpées et leur contenu en ARN viral et en protéine CP et P25 analysé. Les résultats obtenus sont présentés dans la Figure 2.2. Ainsi, aucune différence entre la répllication de l'ARN-3 de type B (Figure 2.2 A, colonne 2) et les variants M1 à M10 (Figure 2.2 A, colonnes 3 à 12) n'a pu être constatée, suggérant que les modifications de séquences introduites dans l'ARN-3 ne modifient pas sa stabilité, ni sa capacité à être reconnu par la réplicase virale.

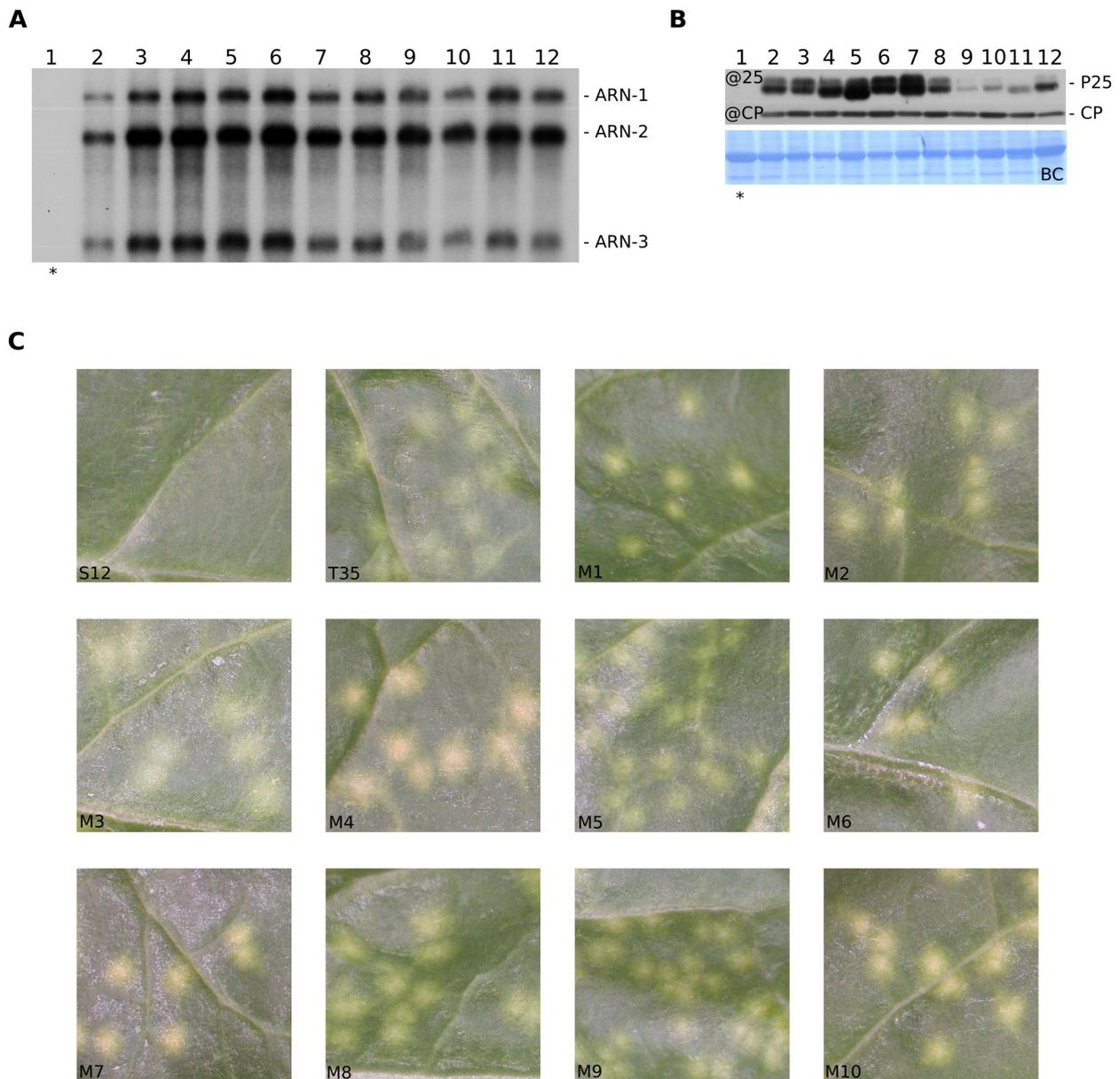


Figure 2.3 : Analyse de la réplication des ARN recombinants (A) et de l'accumulation des protéines (B) dans les lésions locales de feuilles de *Beta vulgaris* inoculées par la souche helper Stras12 (1) supplémentée des transcrits t35 (2), M1 (3), M2 (4), M3 (5), M4 (6), M5 (7), M6 (8), M7 (9), M8 (10), M9 (11) et M10 (12), et de la symptomatologie foliaire 6 jours après infection (dpi). A) Analyse par northern-blot de l'accumulation des ARN viraux 6 dpi. Sondes utilisées : ARN-1 et -2 (Lemaire *et al.*, 1988); ARN-3 (Lauber *et al.*, 1997). B) Analyse par western-blot de la présence des protéines P25 (@P25), de l'accumulation de la protéine de capsid (@CP) et des protéines totales transférées sur membrane et colorées (BC). @P25 : immunodétection de la protéine P25; @CP immunodétection de la protéine CP; BC : coloration au Serva Blue. C) Symptômes observés 6 dpi après inoculation mécanique sur feuilles de *Beta vulgaris*. Les photographies ont toutes été prises dans les mêmes conditions. * L'absence de détection de la protéine CP est expliquée par l'absence d'infection virale dans le prélèvement effectué sur la feuille inoculée asymptomatique.

De plus, l'expression des protéines P25 modifiées ne semble pas avoir d'incidence sur l'accumulation des ARN génomiques 1 et 2. L'analyse de la descendance par RT-PCR n'a pas révélé de modification de la séquence de chaque mutant. Nous avons également quantifié par western-blot l'accumulation relative de la protéine P25 (Figure 2.2 B, P25) par rapport à la protéine de capsid (Figure 2.2 B, CP), et ceci pour une quantité identique de protéines totales transférées sur la membrane (Figure 2.2 BC). Ainsi, nous avons constaté une légère décroissance de l'expression ou de la stabilité de certaines protéines P25 mutantes (M7 à M10) ainsi qu'une migration différentielle des protéines P25 des mutants M3, M5, M7 et M9 (Figure 2.2 C, colonnes 5, 7, 9 et 11) qui se déplacent plus rapidement dans le gel. La mobilité électrophorétique accrue de ces protéines mutantes peut refléter des changements de charges et/ou d'évènements de modifications post traductionnelles, liés directement ou indirectement à la nature de la tétrade introduite. L'infection de *C. quinoa* par les ARN-1 et -2 produit des lésions chlorotiques vertes, 6 jours après inoculation (Figure 2.2 C, S12). En présence d'ARN-3, l'expression de la protéine P25 induit l'apparition d'une chlorose jaune (Figure 2.2 C, t35) qui s'intensifie en fonction du temps (Jupin *et al.*, 1992). Parmi l'ensemble des variants testés, seules les constructions M5 et M7 produisent des symptômes comparables à ceux obtenus en présence de l'ARN t35. Les autres constructions induisent une chlorose jaune plus marquée (Figure 2.2 C) et qui apparaît plus précocement (non montré). Si ces différences de symptômes ont pu être reproduites, les faibles différences phénotypiques observées ne nous permettent pas d'établir un test biologique fiable. Nous avons donc choisi de tester d'autres plantes hôte.

2.2.2 Effet sur *Beta vulgaris*

La même expérience effectuée sur *Beta vulgaris* (variété Dorena, tolérante à la rhizomanie) nous a conduit au même type de résultats (Figure 2.3). Ainsi, aucune différence d'accumulation relative des ARN-3 n'a pu être observée (Figure 2.3 A). L'analyse de l'accumulation des protéines P25 et CP nous a révélé des différences d'expression et/ou de stabilité plus marquées que sur *C. quinoa*. Ainsi, pour des charges protéiques totales (BC) et virales (CP) comparables, les protéines P25 mutées sont retrouvées avec des intensités variables et sous forme de doublets qui apparaissent selon les expériences de façon plus ou

moins marquée. Ces doublets pourraient correspondre à des formes alternatives des protéines P25 ayant subi ou pas des événements de modifications post traductionnelles. On retrouve également la migration électrophorétique différentielle des mutants M3, M5, M7 et M9 précédemment observée dans les échantillons de *C. quinoa* infectés (Figure 2.2 B). Sur *B. vulgaris*, les protéines M7, M8 et M9 s'accumulent moins que les autres formes. Quoi qu'il en soit, nous n'avons pas pu directement corréler les taux d'accumulation des protéines P25 aux symptômes foliaires. En effet, sur *B. vulgaris*, l'infection par les ARN-1 et -2 seuls est asymptomatique (Figure 2.3 C, S12) et les symptômes foliaires deviennent visibles en présence d'ARN-3 (Figure 2.3 C, t35 à M10), les symptômes chlorotiques jaunes sont peu marqués pour les ARN-3 t35, M3, M5, M8, M9 et plus intenses dans le cas des variants M1, M2, M4, M6, M7 et M10. Cette faible variabilité phénotypique ne nous a pas non plus permis d'établir de corrélation directe entre le type de tétrades et le type de chlorose observée.

2.2.3 Effet sur *Tetragonia expansa* et *Beta macrocarpa*

T. expansa est une plante qui répond à l'infection de manière hypersensible, ce qui se traduit par l'apparition de lésions locales visibles en présence d'ARN-3.

Comme précédemment, nous avons analysé le comportement des variants d'ARN-3 exprimant les protéines P25 modifiées au niveau des acides aminés 67 à 70. L'analyse par northern-blot des ARN viraux présents dans les lésions locales ne nous a pas permis d'observer de différences de répllication des ARN viraux (Figure 2.4 A). L'accumulation des protéines P25 et CP, analysées sur le même nombre de lésions locales, nous a permis de constater que sur cette plante hôte, les protéines P25 mutées des constructions M3, M5, M7 et M9 présentent toujours les mêmes différences de migration électrophorétique que nous avons observées sur *C. quinoa* et *B. vulgaris* (Figure 2.4 B). Sur cette image, les doublets protéiques observés précédemment sont bien visibles mais moins intenses que dans les deux autres hôtes (*C. quinoa* et *B. vulgaris*). Sur la même membrane, à charge virale constante (CP), les protéines P25 des mutants M7 à M10 s'accumulent plus faiblement que les autres protéines. Ce résultat, parfaitement reproductible, traduit probablement plus une modification de la stabilité des protéines P25 de ces mutants M7, M8, M9 et M10 qu'une

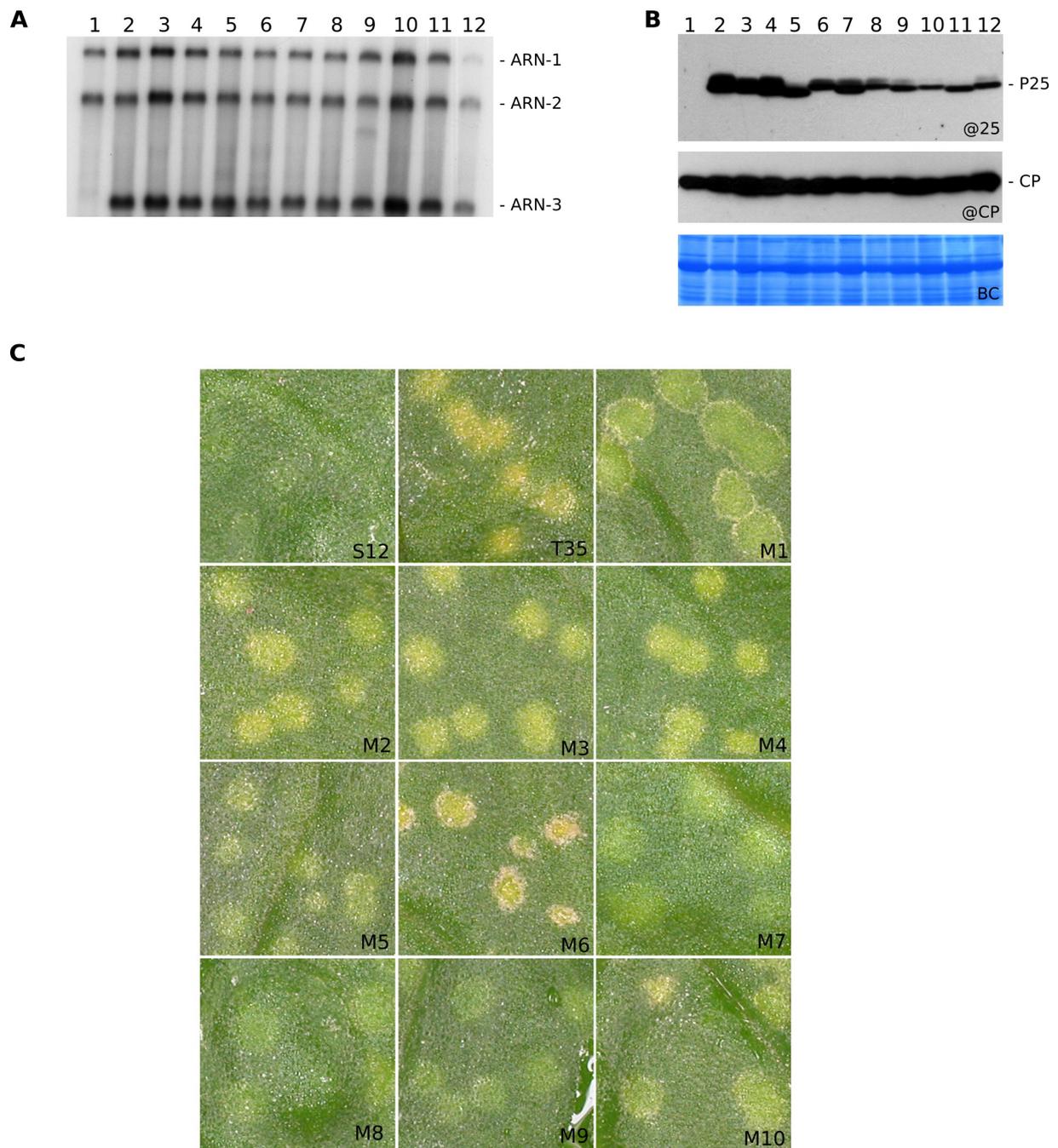


Figure 2.4 : Analyse de la réplication des ARN recombinants (A) et de l'accumulation des protéines (B) dans les lésions locales de feuilles de *Tetragonia expansa* inoculées par la souche helper Stras12 (1) supplémentée des transcrits t35 (2), M1 (3), M2 (4), M3 (5), M4 (6), M5 (7), M6 (8), M7 (9), M8 (10), M9 (11) et M10 (12), et de la symptomatologie foliaire 6 jours après infection (dpi). A) Analyse par northern-blot de l'accumulation des ARN viraux 6 dpi. B) Analyse par western-blot de la présence des protéines P25 (@P25), de l'accumulation de la protéine de capsid (@CP) et des protéines totales transférées sur membrane et colorées (BC). @P25 : immunodétection de la protéine P25 ; @CP immunodétection de la protéine CP ; BC : coloration au Serva Blue. C) Symptômes observés 6 dpi après inoculation mécanique sur feuilles de *Tetragonia expansa*. Les photographies ont toutes été prises dans les mêmes conditions.

modification de leur taux d'expression puisque l'accumulation des ARN-3 est comparable dans toutes les constructions testées. Le phénotype des lésions obtenues sur *T. expansa* (Figure 2.4 C) est plus contrasté et donc plus informatif que sur *C. quinoa* et *B. vulgaris*. En effet, si les lésions induites par la souche helper Stras12 sont quasiment invisibles (Figure 2.4 C, S12) et n'apparaissent que sous lumière UV (non montré), la présence d'ARN-3 et en particulier l'expression de la protéine P25 de type B (tétrade AYHR) induit l'apparition de petites lésions chlorotiques jaunes (Figure 2.4 C, t35). Ce phénotype jaune est accentué dans le cas du mutant M6 qui contient, lui aussi, une protéine P25 de type B (tétrade AYHG, Figure 2.4 C M6). Les lésions observées sont dans ce cas jaunes et nécrotiques. Trois autres phénotypes sont obtenus pour les mutants contenant des P25 de type A ; les mutants M2 (AFHR), M3 (AYRV) et M4 (SYHG), tous de type A, induisent des lésions chlorotiques vertes jaunissantes. D'autres mutants de type A : M7 (VCHG), M8 (ACHG) et M9 (VLHG) induisent des lésions chlorotiques vert pâle. Enfin, les variants M1 (AHHG), M5 (ALHG) et M10 (AFHG) se caractérisent par des lésions chlorotiques vertes entourées d'un anneau de nécrose blanc, très prononcé dans le cas de M1, moyennement pour M10 et discret pour M5.

Ces expériences démontrent donc l'implication du motif tétrade dans la réponse de la plante vis à vis de l'infection virale localisée aux feuilles infectées.

Nous avons analysé l'effet de la tétrade sur le mouvement à longue distance du virus, c'est à dire la capacité des ARN-3 recombinants à induire une infection systémique sur *Beta macrocarpa*. En effet, si les ARN-1 et -2 induisent une infection locale sur cette plante hôte, l'ARN-3, et en particulier l'expression de la protéine P25, conduit à l'infection généralisée de la plante (Lauber *et al.*, 1998; Tamada *et al.*, 1999, et observations personnelles). Tous les variants de l'ARN-3 se sont révélés capables d'infecter de manière systémique ce type de plante hôte (non montré). En conclusion, si la séquence de la tétrade est en partie impliquée dans la reconnaissance de l'infection virale par la plante, elle n'influence pas la fonction de la protéine P25 à induire la systémie.

La répartition nucléo-cytoplasmique des protéines P25 des mutants M1 à M5 a été comparée à celle de la protéine P25 (type B). Ces protéines ont été fusionnées à l'extrémité C-terminale de la protéine EGFP (voir chapitre 3) et exprimées en contexte viral dans *B.*

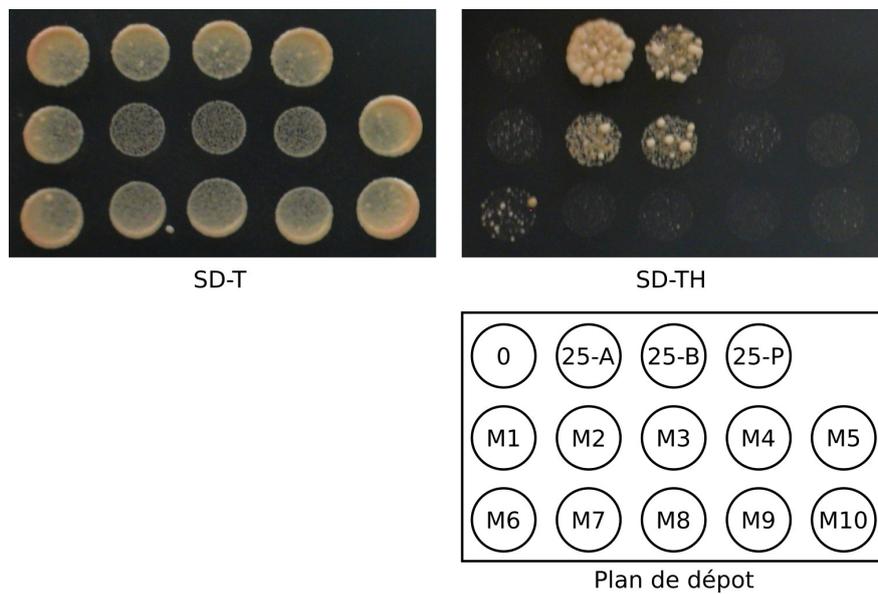


Figure 2.5 : Analyse en système dit "simple hybride" de l'effet de l'expression de facteurs de transcription chimériques GAL4BD::P25 sur l'activation de la transcription du gène rapporteur *His3*. Les levures HF7C exprimant la protéine GAL4BD fusionnée ou non ('0') à la protéine P25 provenant d'un isolat A, B ou P et des variants M1 à M10, sélectionnées sur milieu SD-T ne peuvent pousser sur un milieu dépourvu en histidine (SD-TH) que lorsque le gène rapporteur *His3* est exprimé. Le plan de dépôt des colonies est indiqué.

macrocarpa et *C. quinoa*. Aucune différence de répartition nucléo-cytoplasmique pouvant expliquer les symptômes différentiels obtenus n'a pu être observée (non montré). Nous avons donc poursuivi la caractérisation des protéines recombinantes en analysant leur capacité à activer la transcription dans la levure lorsqu'elles sont fusionnées au domaine de fixation à l'ADN du facteur de transcription GAL4 (GAL4BD).

2.3 Effet de la tétrade sur les propriétés d'activation de transcription de la protéine P25 fusionnée au domaine GAL4BD

Les séquences codantes des protéines P25 de type A (accession AJ505294), B (accession A44503) et A/P (accession AY734505, abrégé P par la suite) et celles des variants M1 à M10 ont été fusionnées en phase à la séquence 3' du domaine de fixation à l'ADN du facteur de transcription GAL4 (GAL4BD) dans le vecteur pGBT9. Les plasmides recombinants (pGBT9-P25A, B, P, M1 à M10) ont servi à transformer des levures HF7C pour sélectionner les clones exprimant ces protéines de fusion sur milieu minimum dépourvu de tryptophane (SD-T, voir chapitre Matériel et Méthodes). Les clones obtenus ont ensuite été repiqués à la fois sur milieu SD-T, pour visualiser leur croissance, et sur milieu SD-TH (dépourvu de tryptophane et d'histidine) pour visualiser l'activation de la transcription du gène rapporteur *His3*. En effet, ce gène est placé sous le contrôle du promoteur Gal1. Lorsqu'il est exprimé grâce à un facteur de transcription Gal4 recombinant fonctionnel, il confère une prototrophie des levures vis à vis de l'histidine. Cette prototrophie se caractérise par une croissance sur milieu SD-TH. Les résultats obtenus sont présentés sur la Figure 2.5. Comme on peut le constater, l'ensemble des levures recombinantes transformées avec le vecteur exprimant GAL4BD seul ("0") ou les vecteurs exprimant les protéines de fusion GAL4BD-P25, est capable de croître sur milieu SD-T (Figure 2.5, SD-T). Cependant, en absence d'histidine seules les levures exprimant les protéines de fusion P25 provenant de l'isolat A (25-A), B (25-B), M2, M3 et M6 sont capables de croître, démontrant l'activation de la transcription du gène rapporteur *His3* par le facteur de transcription chimérique. Cette

expérience, qui confirme les résultats obtenus au laboratoire par Elodie Klein et Guillaume Vetter apporte des indications supplémentaires quant à l'importance de la séquence ou de la structure de la protéine P25 dans sa capacité à recruter l'ARN polymérase II. Si le domaine responsable de l'activation de la transcription, localisé entre les acides aminés 103 à 146 (AD1), est suffisant, il n'en est pas moins sensible aux séquences environnantes. En effet, les domaines d'activation de la transcription AD1 isolés des protéines P25-A, -B et -P restent capables d'activer la transcription (non montré). Lorsqu'ils sont replacés dans leur contexte protéique, seuls les domaines AD1 des protéines P25 provenant d'isolats de type A (tétrade AFHR) et B (AYHR, clone pB35) sont toujours capables d'activer la transcription. La protéine P25 de type P n'est quant à elle pas fonctionnelle (Figure 2.5, 25-P). L'incidence de la nature de la tétrade est également observée dans nos expériences de simple hybride. Seules les protéines P25 de type B contenant une tétrade d'origine B (AYHR, 25-B ; AYHG, M6) et A (AFHR, M2 ; AYRV, M3) sont capables d'activer la transcription du gène rapporteur *His3* (Figure 2.5). Pour les constructions 25-B, M2, M3 et M6, à l'exception des quatre acides aminés variables, l'ensemble de la séquence protéique et en particulier le domaine AD1 est conservé. Ce résultat démontre donc qu'il existe bien une corrélation entre la séquence située en aval du NLS et la structure/fonction de la protéine P25. Cet effet est particulièrement visible dans le cas de la protéine P25A (AFHR) et de la protéine P25 M2 (AFHR, dans un contexte protéique P25 de type B) qui n'activent pas la transcription avec la même efficacité (Figure 2.5). Cette activité différente est illustrée ici par la croissance plus lente des levures exprimant GAL4BD::M2 par rapport à celles exprimant la protéine P25 de l'isolat A.

Cette activité de transcription, détectée dans la levure, dont l'existence *in planta* n'est pas encore démontrée, permet toutefois d'attribuer une caractéristique fonctionnelle supplémentaire à la protéine P25 exprimée dans la levure. Cependant, des limites liées aux contraintes structurales de la protéine P25 ont pu être observées. Aussi, nous devons isoler les séquences des protéines P25 complètes provenant des isolats mondiaux pour caractériser leur capacité à recruter l'ARN polymérase II lorsqu'elles sont fusionnées à GAL4BD.

Nous avons toutefois poursuivi notre étude en recherchant si ces protéines P25 de type A, P et les variants M1 à M10 étaient capables de dimériser, comme la protéine P25 de

type B, décrit dans l'introduction de ce chapitre.

2.4 Interaction entre protéines P25 issues de différents isolats et effet de la tétrade sur la dimérisation

La caractérisation des partenaires cellulaires et viraux qui interagissent avec la protéine P25 permettrait de préciser son rôle dans le cycle viral. Plusieurs approches sont envisageables. La co-immunoprécipitation de protéines permet de caractériser par microséquençage les protéines co-précipitées avec la protéine candidate. Cette identification peut s'avérer fastidieuse en absence d'informations de séquences suffisantes de l'hôte utilisé, ce qui est le cas pour la betterave et *C. quinoa*. La recherche d'interactions par le système double hybride permet de s'affranchir du microséquençage, la séquence de la protéine étant directement accessible à partir du vecteur "proie". Nous avons donc choisi cette approche pour caractériser des partenaires de la protéine P25.

La première étape de cette étude a consisté à vérifier que les protéines P25 fusionnées au domaine de fixation à l'ADN de Gal4 perdaient leur fonction d'activation de la transcription dans des levures exprimant l'autre vecteur navette pGAD424. La seconde étape a consisté à exprimer dans les mêmes cellules, à la fois la protéine P25 fusionnée à GAL4BD et les protéines P25 fusionnées au domaine d'activation de la transcription de GAL4 (GAL4AD). Si les deux protéines de fusion interagissent, elles restaurent un facteur de transcription fonctionnel qui active la transcription du gène rapporteur *His3* et qui rend les levures capables de croître sur milieu minimum dépourvu d'histidine. Nous avons donc inséré l'ensemble des séquences codant pour les protéines P25 décrites précédemment en fusion avec la séquence GAL4AD dans le vecteur d'expression pGAD424. La souche de levure Y187 a été transformée par ces vecteurs recombinants et les transformants sélectionnés sur milieu minimum dépourvu de leucine (voir chapitre Matériel et Méthodes).

Les levures HF7C de signe sexuel MAT α ont été transformées par le vecteur pGBT9 vide ou contenant les différentes séquences de la protéine P25. Les levures Y187 ont un signe sexuel MAT α et contiennent les vecteurs pGAD424 vide ou exprimant les différentes séquences de protéines P25. Ces deux ensembles de souches ont été mises en contact pour

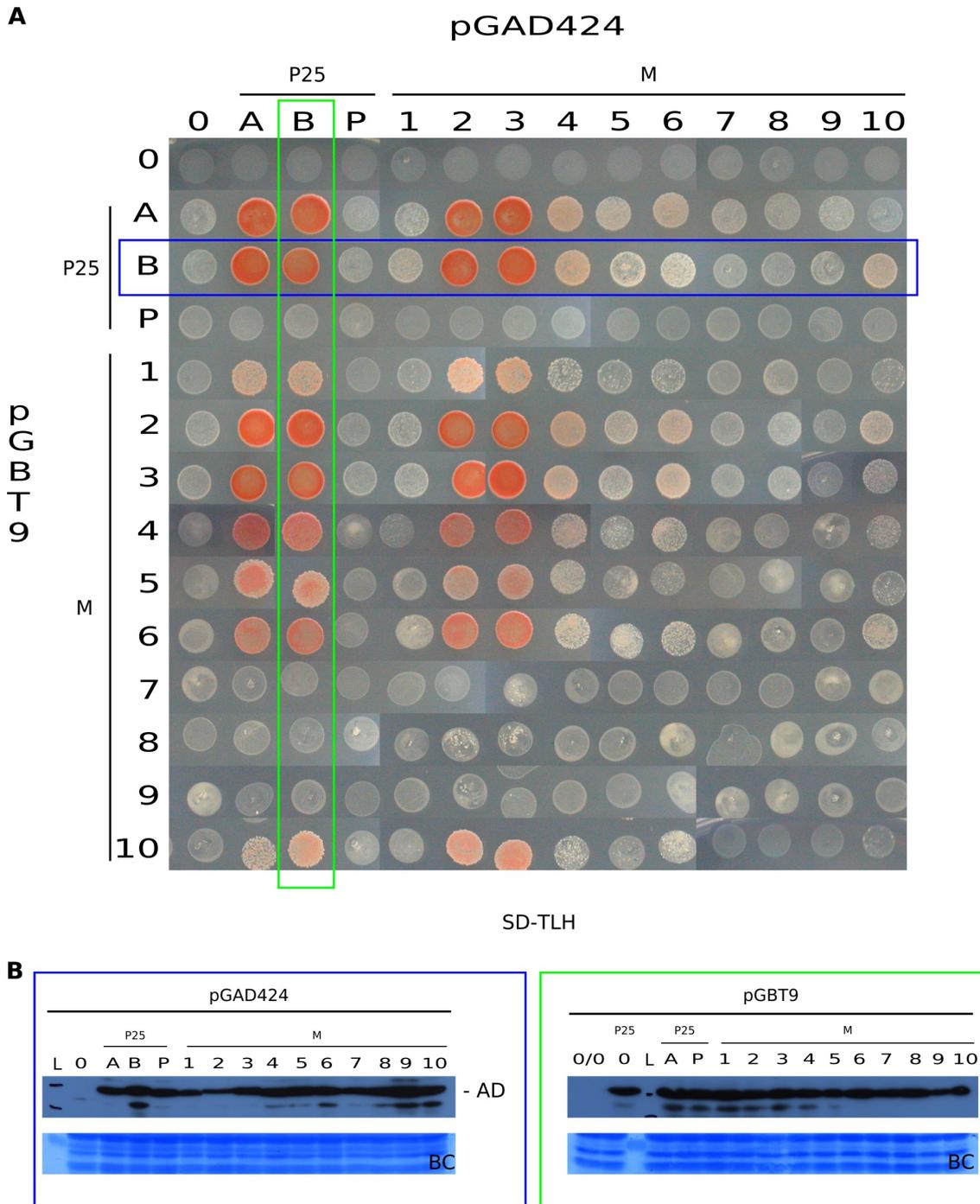


Figure 2.6 : Caractérisation de l'existence d'interaction entre les protéines P25 provenant de différents isolats par la technique de double hybride. A) Les levures diploïdes sont sélectionnées par leur croissance sur milieu dépourvu en histidine (SD-TLH). Les levures exprimant le facteur de transcription GAL4 reconstitué grâce à l'interaction des protéines P25 qui active la transcription du gène rapporteur *His3* et les rend prototrophes vis à vis de l'histidine. Les cadres bleu et vert représentent les levures diploïdes sélectionnées pour rechercher l'expression des protéines de fusion GAL4AD par western-blot (B). "0/0" correspond aux deux vecteurs vides. L : marqueur de masse moléculaire.

produire des levures diploïdes après "mating", sélectionnées sur milieu minimum dépourvu de tryptophane et de leucine (SD-TL) (voir chapitre Matériel et Méthodes). L'ensemble des 196 clones de levures diploïdes a ensuite été repiqué sur milieu SD-TL (non montré) et sur milieu minimum dépourvu de tryptophane, leucine et histidine (SD-TLH) pour rechercher leur prototrophie vis à vis de ce dernier acide aminé. Cette prototrophie traduit l'activation de la transcription du gène rapporteur *His3* par un facteur de transcription chimérique (voir paragraphe précédent) ou reconstitué par l'interaction des deux protéines de fusion (voir chapitre Matériel et Méthodes). L'expression des protéines de fusion a également été recherchée par western-blot (voir plus bas). Les résultats obtenus sont résumés sur la Figure 2.6. Cette expérience permet dans un premier temps de visualiser l'absence d'activation de la transcription du gène rapporteur *His3* dans les levures exprimant à la fois le vecteur pGBT9 exprimant les protéines P25 A, B, M2, M3 et M6, et le vecteur pGAD424 vide (Figure 2.6 A, colonne 0, lignes P25A, P25B, M2, M3 et M6). Ainsi, lorsqu'une croissance est observée, celle-ci traduit l'activité du facteur de transcription GAL4 reconstitué par l'interaction de deux protéines P25. Nous avons pu retrouver l'interaction entre les protéines P25 de type B décrit par Elodie Klein et Guillaume Vetter (Figure 2.6 A colonne P25B, ligne P25B). Nous nous sommes intéressés au niveau d'expression des protéines de fusion et nous les avons analysées en utilisant des anticorps monoclonaux reconnaissant spécifiquement les domaines GAL4AD et GAL4BD. Ainsi, nous avons extrait et analysé par western-blot les protéines totales des levures diploïdes contenant tous les types de plasmides recombinants : les diploïdes contenant le vecteur pGAD-P25B et l'ensemble des vecteurs pGBT9 (Figure 2.6 A, rectangle vert) et les diploïdes contenant le vecteur pGBT9-P25B et l'ensemble des vecteurs pGAD424 (Figure 2.6 A, rectangle bleu). Seules les immunorévélations effectuées à l'aide de l'anticorps dirigé contre le domaine GAL4AD sont présentés sur la Figure 2.6 B. Les mêmes détections effectuées à l'aide de l'anticorps dirigé contre GAL4BD ne nous ont pas permis de détecter de protéines. Ainsi, l'ensemble des protéines de fusion GAL4AD-P25 sont exprimées dans les levures diploïdes à la masse moléculaire attendue de 49kDa (Figure 2.6 B, cadre bleu). On peut noter ici que le niveau d'expression des protéines GAL4AD::M2 et M7 est plus faible que pour les autres protéines détectées. Le signal obtenu est spécifique puisqu'aucun signal de masse moléculaire identique n'est ob-

tenu pour les levures diploïdes exprimant la protéine GAL4AD non fusionnée (Figure 2.6 B, cadre bleu, colonne 0) et les protéines GAL4AD et GAL4BD non fusionnées (Figure 2.6 B, cadre vert, colonne 0/0). Pour les levures diploïdes exprimant la protéine GAL4AD::P25B et les protéines GAL4BD recombinantes, la protéine est immunorévélee dans l'ensemble des pistes (Figure 2.6 B, cadre vert) démontrant que l'expression des protéines GAL4BD, fusionnées ou pas, n'affecte pas l'accumulation de la protéine GAL4AD::P25B.

Pour poursuivre l'interprétation de nos résultats, nous avons dû estimer que les protéines GAL4BD::P25 étaient exprimées dans les levures avec la même efficacité et la même stabilité que les fusions GAL4AD::P25 en émettant toutefois une réserve dans le cas des protéines GAL4BD fusionnées à M7 et M9 pour lesquelles aucune interaction n'a pu être détectée (voir Figure 2.6 A, lignes M7 et M9).

Ainsi, nous avons constaté qu'une interaction P25/P25 existait pour l'isolat A (AFHR, AJ505294) (Figure 2.6 A, colonne A, ligne A). Cette situation n'est pas retrouvée dans le cas de la protéine P25 provenant de l'isolat P (SYHG, AY734505) (Figure 2.6 A, colonne P, ligne P). Une telle absence d'interaction n'est pas liée à la nature de la tétrade SYHG puisque la protéine M4 (SYHG) est capable de dimériser avec elle-même et d'autres protéines P25 (par exemple Figure 2.6 A, colonne B, ligne 4); ce qui n'est pas le cas de la protéine P25P (Figure 2.6 A, colonne B, ligne P). La situation est différente pour la protéine P25 provenant de l'isolat naturel A (AFHR) qui est capable d'interagir avec celle provenant de l'isolat B (AYHR) (Figure 2.6 A, colonnes A et B, lignes A et B). Le même comportement est observé dans le cas de la protéine issue de l'isolat B (comparer lignes et colonnes A et B). Cette analyse démontre bien l'importance de la séquence complète de la protéine P25 dans le maintien d'interactions P25/P25 et en particulier l'effet des acides aminés 67 à 70 dont la nature bouleverse les propriétés électrophorétiques de la protéine P25B et sûrement sa structure tridimensionnelle. En effet, les variants M2, M3 ne modifient pas les propriétés de dimérisation de la protéine P25B (comparer ligne 25B et lignes M2, M3 ou comparer colonne 25B aux colonnes M2, M3), atténuent son efficacité (comparer ligne et colonne 25B aux lignes et colonnes M1, M4, M5, M6 et M10 respectivement) ou la suppriment (comparer ligne et colonne 25B et ligne et colonne M8). Les mutants M7 et M9 peuvent être classés dans ce dernier cas mais compte tenu de la réserve décrite

précédemment ils ne sont pas commentés.

De telles interactions, ou absence d'interaction entre protéines P25 de type B, A ou P ne sont évidemment possibles que dans le cas d'infections mixtes conduisant deux isolats à coexister dans la même cellule infectée. Ainsi, l'analyse de la dimérisation des différentes protéines P25 doit être prioritairement lue sur la Figure 2.6 dans le sens de la diagonale, c'est à dire entre protéines P25 de même origine. C'est pourquoi seules les protéines P25 A (AFHR), B (AYHR), M2 (AFHR), M3 (AYRV) et plus faiblement M4 (SYHG) et M6 (ACHG) peuvent dimériser. Cependant, tout comme dans le cas des analyses de l'activation de transcription, ces tests d'interaction devront être réitérés à partir des séquences complètes des protéines P25 issues d'isolats naturels.

2.5 Conclusions

Pour analyser l'ensemble des caractéristiques de la protéine P25 abordées dans ce chapitre, j'ai résumé dans la table 2.1 les résultats phénotypiques dans *T. expansa*, la capacité des protéines P25 à activer la transcription et à dimériser et leur stabilité. Nous constatons qu'il existe une corrélation directe entre la capacité des protéines P25 à former des dimères et leur capacité à induire des lésions chlorotiques jaunes sur *T. expansa*. C'est le cas pour les tétrades AYHR (B, t35), AFHR (A, M2), AYRV (A, M3), SYHG (P, M4) et ACHG (B, M6). Les autres variants induisent quant à eux des lésions chlorotiques vertes : AHHG (A, M1), ALHG (A, M5), VCHG (A, M7) ACHG (A, M8), VLHG (A, M9) et AFHG (A, M10) et sont incapables de dimériser. De plus, à l'exception du mutant M4 (SYHG), les protéines P25 qui dimérisent sont également capables d'activer la transcription du gène rapporteur *His3*. La nature des acides aminés de la tétrade a une influence directe sur les propriétés de la protéine P25 de type B mais nous n'avons pas testé ici l'influence des modifications des autres acides aminés de la séquence naturelle de chaque isolat, qui pourraient compenser les modifications de charges ou de repliement induits par notre mutagenèse. Cependant, la comparaison des propriétés biologiques des protéines P25 provenant d'isolats naturels A (AFHR), B (AYHR) et A/P (SYHG) nous apporte des renseignements précieux sur les fonctions de ces protéines qui diffèrent en système levure. Ainsi, nous avons démontré

Table 2.1 : Synthèse des fonctions putatives d'activation de transcription, des propriétés de dimérisation de protéines P25 modifiées provenant d'isolats naturels, de leur stabilité *in vivo* et des symptômes induits 6 dpi sur feuilles de *T. expansa*. La nature du motif tétrade de la protéine P25 est précisé ainsi que l'origine de l'isolat. NT : non testé.

Variant	Tétrade	Isolat	Symptômes	Activation de la transcription	Dimérisation	Stabilité
t35	AYHR	B	Jaune	+	++	+
A	AFHR	A	NT	++	++	NT
M1	AHHG	A	Vert, nécrose périphérique	-	-	+
M2	AFHR	A	Vert jaune	+	++	+
M3	AYRV	A	Vert jaune	+	++	+
M4	SYHG	A/P	Vert jaune	-	+	+
M5	ALHG	A	Vert, nécrose périphérique	-	-	+
M6	AYHG	B	Jaunes nécrotiques	+	+	+
M7	VCHG	A	Vert pâle	-	-	±
M8	ACHG	A	Vert pâle	-	-	±
M9	VLHG	A	Vert pâle	-	-	±
M10	AFHG	A	Vert, nécrose périphérique	-	-	±

que la protéine P25 provenant d'isolats A et B peuvent dimériser et porter les fonctions d'activation de la transcription, mais que la protéine P25 d'un isolat A/P n'est pas capable d'effectuer ces fonctions. Notre hypothèse évolutive décrite au chapitre 1 et appliquée à ces résultats nous amène à suggérer un gain de fonction entre la protéine P25 d'un isolat A/P et celle d'un isolat B. Ce gain de fonction pourrait être lié à un bouleversement de la structure de la protéine, comme semble le démontrer notre analyse par mutagenèse dirigée du motif tétrade. Nous devons reproduire l'analyse de la symptomatologie liée à l'expression des protéines P25 provenant d'isolats naturels pour rechercher l'incidence de leur expression sur l'aggravation des symptômes. Ainsi, en testant l'effet sur la symptomatologie foliaire (sur *C. quinoa*, *B. vulgaris* et *T. expansa*), l'activation de la transcription et les capacités de dimérisation de ces protéines, nous pourrions conforter les résultats résumés dans la table 2.1. Nous pourrions peut être montrer que la protéine P25 est un facteur d'avirulence, reconnu par le produit d'un gène de résistance qui limite l'infection virale. En absence de reconnaissance, la réaction de défense n'est pas ou peu activée (réponse HR) et conduit à la réaction compatible et donc à l'invasion de la plante. En effet, il est intéressant de noter que les tétrades provenant des isolats américains 412USA (ACHG, M8) et 277USA (VLHG, M9) n'induisent pas de réponse sévère de la plante. De tels isolats sont décrits comme émergents et très agressifs sur betterave.

Il sera donc intéressant de rechercher les facteurs cellulaires associés à ces protéines P25 pour les comparer. Pour cela, nous envisageons de cribler la banque d'ADNc d'*Arabidopsis thaliana* disponible au laboratoire, en utilisant comme appât les protéines P25 de type B et A/P. J'ai déjà réalisé ce criblage (voir chapitre Matériel et Méthodes) mais la faible efficacité de transformation obtenue ne m'a pas permis d'obtenir de résultats concluants. Pour cribler une seconde fois cette banque, les séquences des protéines P25 de type B et A/P seront placées en fusion avec GAL4BD dans le vecteur pGBKT7 qui permet l'expression de protéines de fusion étiquetée myc et favorise leur détection. Cette approche permettra de résoudre le problème rencontré lors de la tentative de détection des protéines de fusion.

Chapitre 3

Influence de l'ARN-5 et de l'expression de la protéine P26 sur les symptômes induits par le BNYVV

Pour initier l'étude de l'ARN-5 et de la protéine P26, nous avons utilisé une protéine chimérique obtenue par fusion entre la protéine EGFP (Enhanced Green Fluorescent protein) et la protéine P26. L'utilisation de cette protéine recombinante nous a permis, entre autre, de déterminer la localisation sub-cellulaire de la protéine virale. Avant de présenter en détail les résultats que nous avons obtenus, je passerai en revue dans cette introduction le mécanisme de transport de protéines à travers l'enveloppe nucléaire puis je présenterai les propriétés de l'EGFP et les applications liées à son utilisation.

3.1 Les voies de transport nucléo-cytoplasmiques dans les cellules eucaryotes

Grâce à l'expression de l'ensemble des protéines nécessaires à leur réplication, un grand nombre de virus à ARN (en particulier à ARN de polarité positive) se répliquent dans le cytoplasme des cellules infectées. On peut citer comme exemple le poliovirus et le BNYVV, qui utilisent la machinerie de traduction cellulaire pour la synthèse de l'ensemble des pro-

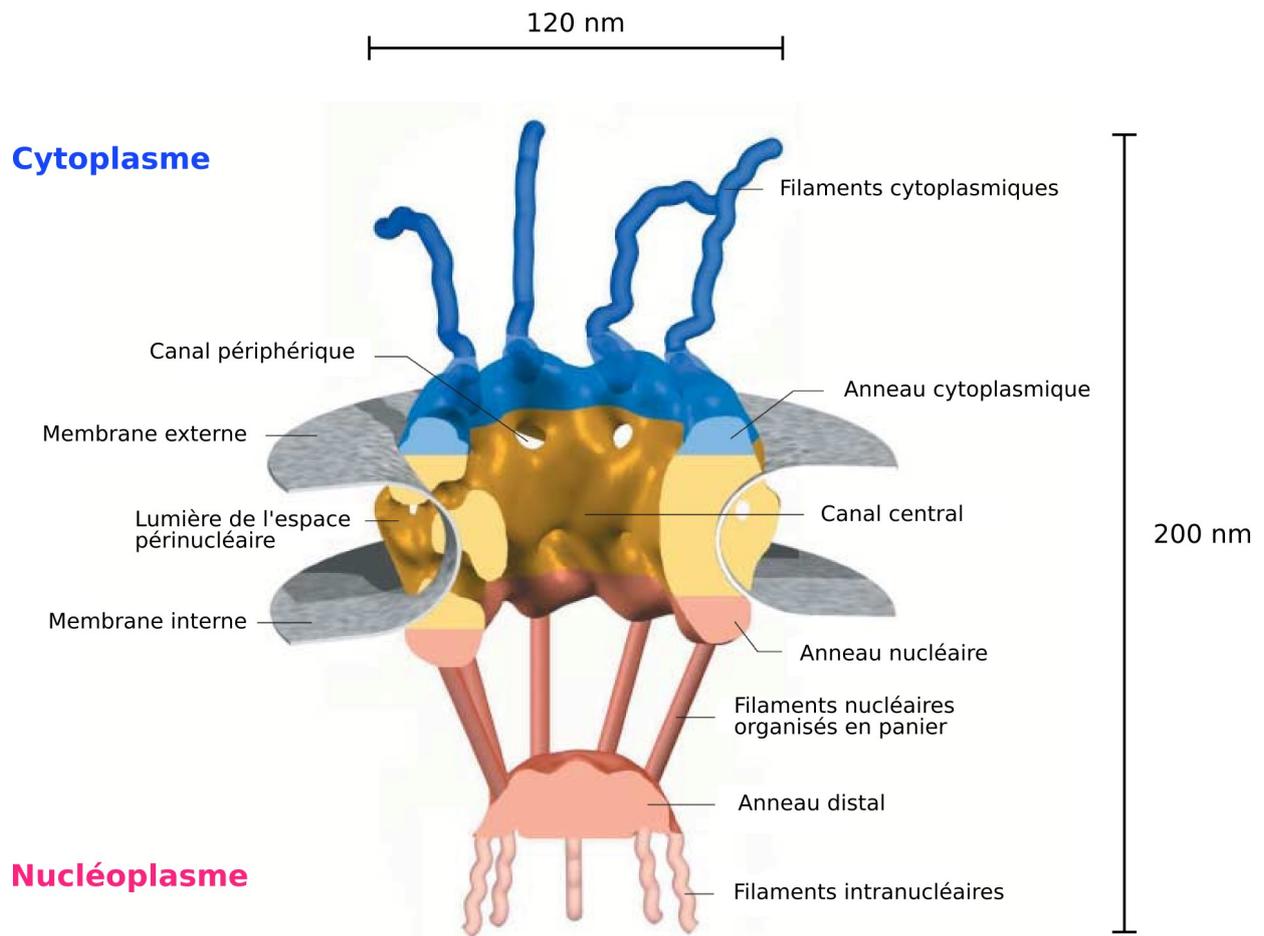


Figure 3.1 : Représentation schématique d'un pore nucléaire (Nuclear Pore Complex ou NPC) selon Fahrenkrog et Aebi (2003). La région centrale (jaune) contient un canal aqueux à travers lequel ont lieu les translocations. Les filaments cytoplasmiques (bleu) et intranucléaires (rose) sont impliqués dans les étapes initiales et finales de la translocation. La double membrane de l'enveloppe nucléaire est schématisée par la surface grise.

téines virales dont certaines, non structurales, seront directement impliquées dans le mécanisme de réplication du génome. Ces virus n'ont apparemment nullement besoin de localiser leur ARN génomique dans le noyau des cellules hôtes. Pourtant, certains sont capables d'exprimer des protéines qui possèdent des signaux de localisation nucléaire et qui vont être adressées dans le noyau des cellules infectées (CMV2b [Lucy *et al.*, 2000], poliovirus 3D [Weidman *et al.*, 2003], BNYVV P25 [Vetter *et al.*, 2004], VP-1 et VP-2 du MVM [Lombardo *et al.*, 2000]). Cette localisation nucléaire de protéines virales implique des interactions avec des facteurs cellulaires, tant dans le cytoplasme que dans le nucléoplasme, et des modifications de certaines fonctions cellulaires pouvant être corrélées à la pathogénicité.

Dans le cas de virus à ADN et des rétrovirus, l'expression et la réplication des génomes viraux implique obligatoirement une étape nucléaire, nécessaire entre autre à la transcription des ARN messagers viraux. Ainsi, certaines des protéines virales synthétisées dans le cytoplasme seront adressées au nucléoplasme pour y recruter ARN ou ADN génomique, éventuellement les exporter, et ainsi amener la formation de nouvelles particules infectieuses.

Ces protéines virales, tout comme certaines protéines cellulaires, possèdent des séquences particulières qui vont assurer leur adressage nucléaire puis éventuellement leur export du noyau vers le cytoplasme. De tels échanges nucléo-cytoplasmiques se déroulent à travers l'enveloppe nucléaire au niveau des pores nucléaires et sont soumis à des régulations fines.

3.1.1 Le pore nucléaire

Le noyau est séparé du cytosol par une enveloppe formée de deux membranes lipidiques. La membrane externe de l'enveloppe constitue un *continuum* avec le réticulum endoplasmique ; l'espace intermembranaire correspond à la lumière du réticulum. Le transport de molécules entre le cytosol et le nucléoplasme se produit à travers des pores appelés "Nuclear Pore Complex" (NPC), constitués de protéines (nommées nucléoporines) organisées en complexes (Figure 3.1). Les NPC peuvent être considérés comme les sites de fusion entre les membranes interne et externe de l'enveloppe nucléaire. Ils ont une masse molé-

culaire de 44MDa chez la levure et d'environ 60MDa chez les vertébrés (Cronshaw *et al.*, 2002). Ils se présentent sous la forme d'un cylindre de 120nm de diamètre et 200nm de long (Bednenko *et al.*, 2003). Chaque NPC est composé d'environ 50 à 100 nucléoporines chez la levure et chez les vertébrés. Plus de trente nucléoporines de levure ont été identifiées (Doye and Hurt, 1997; Stoffler *et al.*, 1999), une vingtaine chez les vertébrés (Stoffler *et al.*, 1999) et une dizaine chez les plantes (Rose and Meier, 2001). Une de leurs caractéristiques structurales est la présence de répétitions de motifs du type FXFG, GLFG ou FG (F correspond à la phénylalanine, G à la glycine, L à la leucine et X à un acide aminé quelconque). Chez les vertébrés et les plantes, certaines nucléoporines subissent une O-glycosylation avec des groupements N-acétylglucosamine (Heese-Peck and Raikhel, 1998). Le rôle de ces modifications post-traductionnelles n'est pas connu et n'est pas retrouvé chez la levure (Davis, 1995).

Le NPC apparaît comme une structure tripartite constituée de deux anneaux, l'un cytoplasmique et l'autre nucléaire, reliés par un cylindre central composé de huit unités formées de plusieurs nucléoporines (Figure 3.1). L'assemblage de ces unités forme un canal central d'une longueur d'environ 90nm et d'un diamètre de 45nm dans sa partie la plus étroite, permettant ainsi le passage de particules de 30 à 40nm (Fahrenkrog and Aebi, 2003). Le cylindre central est perforé par huit petits canaux d'un diamètre de 9nm, qui pourraient autoriser la diffusion passive d'ions et de petites molécules entre le noyau et le cytoplasme Hinshaw *et al.* (1992). Des données récentes semblent indiquer que ces canaux auraient plutôt un rôle de "zone tampon" lors des déformations du cylindre central occasionnées par le transport actif de molécules de grande taille (Fahrenkrog and Aebi, 2003). Huit fibrilles protéiques longues de 50nm, s'étendent vers le nucléoplasme à partir de l'anneau nucléaire et plongent dans le nucléoplasme pour se rejoindre au niveau d'un petit anneau et former ainsi une structure en panier.

Les pores nucléaires constituent ainsi une barrière à la fois sélective et perméable de l'enveloppe nucléaire. Les molécules inertes et les protéines d'une masse moléculaire inférieure à 40 – 45kDa ou ayant un diamètre inférieur à 9nm, pourraient diffuser passivement mais les molécules de taille supérieure nécessitent un transport actif. Le passage actif des macromolécules à travers les pores nucléaires est assuré par un système de transporteurs

Table 3.1 : Exemples de séquences impliquées dans le transport nucléo-cytoplasmique de protéines virales. Les motifs NLS et NES classiques sont imprimés en gras.

Protéine virale	Séquence	Référence
Séquences d'import nucléaire (NLS)		
SV 40 Antigène T	PKKKRKV	(Kalderon <i>et al.</i> , 1984a, 1984b)
HTLV-1 Rex	PKTRRRPRRSQRKRPTP	(Siomi <i>et al.</i> , 1988)
HIV-1 Rev	RQARRN RRRR WRERQR	(Kubota <i>et al.</i> , 1989)
HIV-1 Tat	GRKKRR QRRRAHQ	(Dang et Lee, 1989)
BDV N	PKRRLVDDA	(Boulikas, 1993)
Semliki NoLS-1	KPKKKK TTKPKTQP	(Favre <i>et al.</i> , 1994)
Semliki NoLS-2	RRRKRNRDAARRRRKQ	(Favre <i>et al.</i> , 1994)
MDV MEQ	RRRKRNRDAARRRRK	(Liu <i>et al.</i> , 1997)
Influenza-A NP	MASQGTKRSYEQM	(Wang <i>et al.</i> , 1997)
BDV P	P ₂₉ RPRKIPR ₃₆ et P ₁₈₁ PRIYQLPSAPT ₁₉₃	(Shoya <i>et al.</i> , 1998)
CMV2b	²² KRRRRR ²⁷	(Lucy <i>et al.</i> , 2000)
BDV P ₁₀	RTLLELVRR LN GN	(Wolff <i>et al.</i> , 2002)
PRRSV N	PGKKNKKKNPEKPHFPLATEDDVRHHFTPSER	(Rowland et Yoo, 2003)
BNYVV P25	⁵⁷ KRIRFR ⁶²	(Vetter <i>et al.</i> , 2004)
Signaux d'export nucléaire (NES)		
EIAV Rev	PLESDQWCRVLRQSLPEEKIP	(Fridell <i>et al.</i> , 1993)
FIV Rev	KKMMDLED RFRK LFGSPSKDEYT	(Mancuso <i>et al.</i> , 1994)
HIV Rev	LGLPPLERLTL	(Fischer <i>et al.</i> , 1995)
HSV ICP27	IDMLIDLGLDL	(Sandri-Goldin, 1998)
SqLCV BR1	VTKRVVSLEKDTLL IDLHGTTQL	(Ward et Lazarowitz, 1999)
BDV N	L TELEISSIFSCC	(Kobayashi <i>et al.</i> , 2001)
Influenza-A NEP	ILLRMSKMQL	(Paragas <i>et al.</i> , 2001)
Influenza-C NEP	L ₉₅ WLPMKSLSL ₁₀₅ et M ₁₂₂ KHQILTRLKL ₁₃₂	(Paragas <i>et al.</i> , 2001)
PRRSV N	L PTHHTVRLIRV	(Rowland et Yoo, 2003)
BNYVV P25	¹⁶⁹ VYMVCLVNTV ¹⁷⁸	(Vetter <i>et al.</i> , 2004)

protéiques mobiles, distincts des nucléoporines : les importines et les exportines. Ces transporteurs d'une masse moléculaire de 90 à 130 kDa constituent la famille des karyophérines (Rodriguez *et al.*, 2004; Ström and Weis, 2001).

3.1.2 L'adressage nucléaire

Le transport nucléo-cytoplasmique des protéines est régi par des séquences spécifiques localisées sur ces protéines. Ces séquences sont appelées respectivement séquence de localisation nucléaire (ou NLS pour "Nuclear Localisation Sequence") et signal d'export nucléaire (ou NES pour "Nuclear Export Signal"). La table 3.1 représente quelques NLS et NES caractérisés sur des protéines virales. En règle générale, ces signaux sont transférables d'une macromolécule à une autre, c'est à dire qu'ils peuvent conférer des compétences de transport nucléaire à des protéines qui sont restreintes au cytoplasme ou au noyau.

Les signaux d'import nucléaire

Le premier NLS à avoir été identifié est la courte séquence riche en acides aminés basiques (**PKKKRKV**) de l'antigène T du Virus Simien 40 (SV40) (Kalderon *et al.*, 1984a,b), une protéine à localisation nucléaire. Des séquences similaires, dénommées NLS classiques, ont été par la suite caractérisées chez de nombreuses protéines virales et cellulaires localisées dans le noyau. D'autres protéines comme la nucléoplasmine (Robbins *et al.*, 1991) et la plupart des protéines nucléaires de plantes, possèdent un NLS bipartite où deux motifs riches en lysines et arginines sont séparés par une séquence de 6 à 13 acides aminés quelconques (**KRPAATKKAGQAKKK** pour la nucléoplasmine par exemple). Des séquences d'adressage nucléaire différentes des NLS classiques et bipartites ont été identifiées sur les hnRNP ("Heterogeneous nuclear RNP"). Ainsi, les protéines hnRNP A1 et hnRNP K possèdent toutes deux des séquences d'adressage nucléaire distinctes constituées d'une quarantaine d'acides aminés appelés respectivement M9, riche en glycines (Siomi and Dreyfuss, 1995; Siomi *et al.*, 1997), et KNS (Michael *et al.*, 1997).

Les signaux d'export nucléaire

Les signaux d'export sont classiquement constitués de séquences riches en leucines ou, de manière plus générale, en résidus aliphatiques. Ils ont été identifiés sur de nombreuses protéines cellulaires et virales, de structures et de fonctions différentes (Fritz and Green, 1996). Les NES les mieux caractérisés sont ceux de la protéine Rev du HIV-1 (LPPLERLTL) (Bogerd *et al.*, 1995; Fischer *et al.*, 1995) et de la protéine kinase AMPc-dépendante cellulaire des eucaryotes (LALKLAGLDI) (Wen *et al.*, 1995). Un autre type de NES a été découvert sur les protéines hnRNP A1 et K. Ces séquences d'export nucléaire sont constituées des 40 acides aminés des NLS décrits précédemment (domaines M9 et KNS) qui interviennent également comme séquences d'export (Michael *et al.*, 1995, 1997).

3.1.3 Mécanismes de l'import nucléaire

Le transport nucléo-cytoplasmique est fortement orienté *in vivo*. En effet, les importines se lient dans le cytoplasme à leur substrats, les cargos, pour les transférer dans le noyau (Figure 3.2). A l'inverse, dans le cas des cargos réadressés dans le cytoplasme, les exportines interagissent avec leur cargos dans le noyau et les dirigent vers le cytoplasme. Généralement, la translocation des protéines à travers les pores nucléaires est réalisée exclusivement par les molécules de la famille des récepteurs de type importine β . L'importine β est constituée de deux domaines fonctionnels :

- un site de fixation du complexe RanGTP, localisé du côté N-terminal de la protéine
- un domaine d'interaction avec les importines α , situé à l'extrémité C-terminale (pour revue voir Nakielny and Dreyfuss (1999)).

Les importines α sont des adaptateurs additionnels qui jouent un rôle d'intermédiaire entre l'importine β et le cargo. Les importines α sont constituées de trois domaines fonctionnels :

- un site de liaison à l'importine β ou IBB ("Importin β Binding domain"), à caractère basique et localisé à l'extrémité N-terminale
- un domaine central hydrophobe appelé arm pour "Armadillo repeat" (par homologie

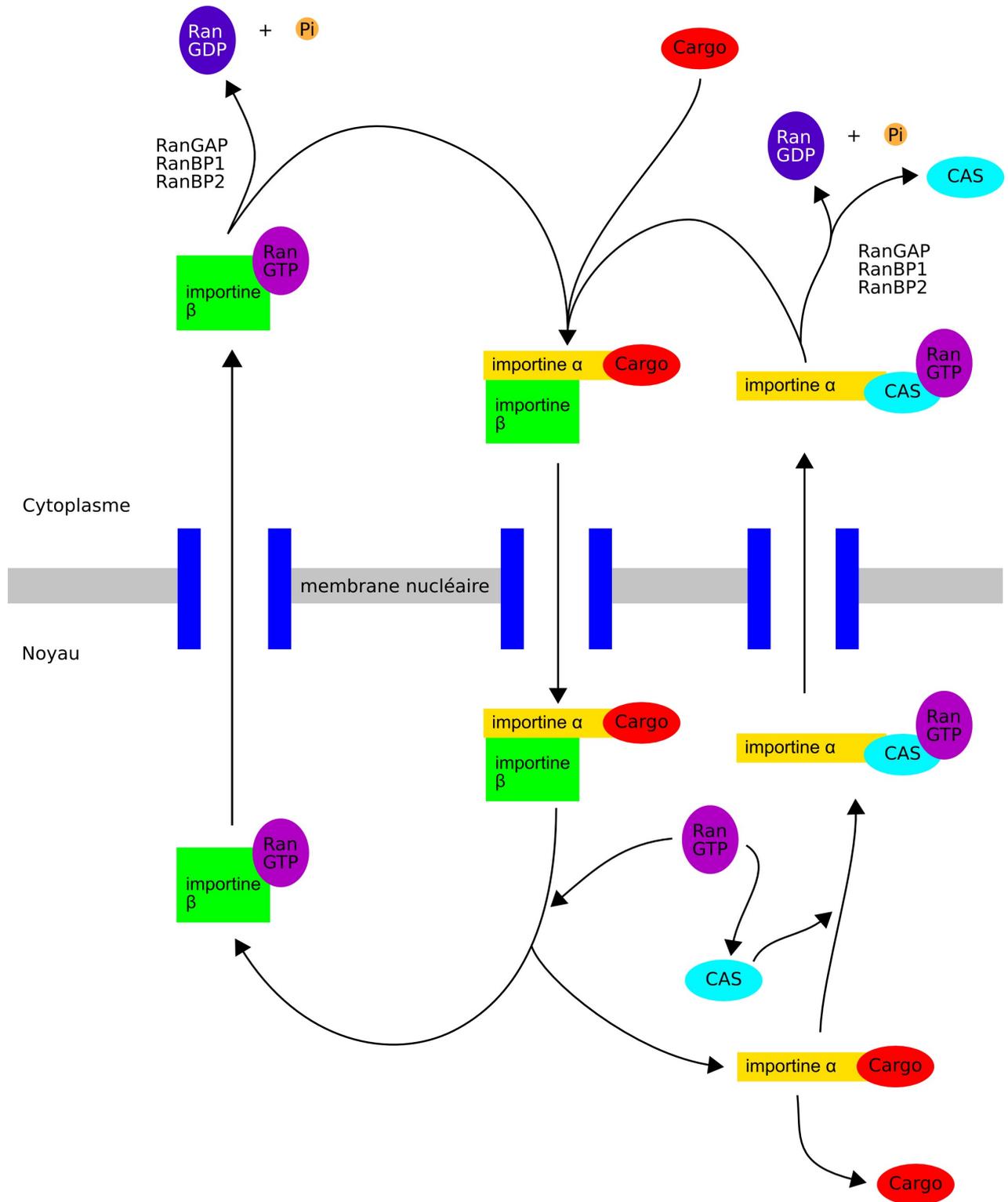


Figure 3.2 : mécanisme de l'import des protéines à travers le pore nucléaire. Le cargo est, dans un premier temps, reconnu dans le cytoplasme (en haut) par une importine α spécifique. Après prise en charge par l'importine β , le complexe ternaire formé est transloqué à travers le pore nucléaire (rectangles bleus) et libéré dans le noyau (en bas) car l'importine β va préférentiellement fixer la protéine Ran couplée à une molécule de GTP. L'importine α libère le cargo dans le nucléoplasme puis est ensuite exporté vers le cytoplasme après prise en charge par la protéine CAS couplée à la protéine RanGTP. Le complexe importine β -RanGTP est également exporté et recyclé au niveau du cytoplasme. L'hydrolyse du GTP pour les deux complexes, stimulée par le complexe RanGAP-RanBP1 et RanBP2, libère les importines α et β qui peuvent alors intervenir dans un nouveau cycle d'import.

avec la protéine Armadillo, un régulateur transcriptionnel de *Drosophila melanogaster*)

- une séquence acide située en C-terminal de la protéine et qui est impliquée dans l'exportation nucléaire de la protéine lors de son recyclage.

Il existe trois familles d'importine α qui possèdent environ 50 % d'identité de séquence. Ce pourcentage peut atteindre 85 % pour les membres d'une même famille. L'existence d'isoformes d'importines α implique qu'elles peuvent reconnaître un nombre distinct de protéines cibles portant des séquences NLS à caractère basique (pour revue voir Nakielny and Dreyfuss (1999)).

La voie d'import la plus étudiée est celle des protéines qui possèdent un NLS basique dit classique et est représentée sur la Figure 3.2. Cette voie est dirigée par l'importine β associée à son adaptateur, l'importine α . Cette dernière se lie dans le cytoplasme à la fois au NLS du cargo et à l'importine β . Dans le cytoplasme, après formation d'un complexe ternaire NLS/importine α /importine β , celui-ci va s'ancrer via l'importine β au pore nucléaire, puis être transloqué à travers le pore selon un processus encore inconnu. Toutefois, Bednenko *et al.* (2003) suggèrent que les motifs FG des nucléoporines sont activement impliqués dans cette translocation. Dans le noyau, le complexe se dissocie sous l'influence d'une RanGTPase qui a fixé une molécule de GTP (RanGTP, Moore and Blobel, 1993). Dans cet état, l'importine β possède une plus forte affinité pour RanGTP que pour l'importine α . La protéine RanGTP (Figure 3.2), intranucléaire (voir paragraphe 3.1.5, page 49), va entrer en compétition avec l'importine α et dissocier le complexe ternaire pour former de manière concomitante un complexe importine β -RanGTP. Le complexe binaire importine α /NLS est ensuite dissocié par l'interaction préférentielle de l'importine α avec la protéine CAS ("Cellular Apoptosis Susceptibility"), qui est l'exportine spécialisée dans le recyclage vers le cytoplasme de cette classe d'importine, laissant le cargo dans le noyau (pour revue voir Weis, 1998; Görlich and Kutay, 1999; Nakielny and Dreyfuss, 1999).

Certaines protéines sont importées dans le noyau en se fixant directement à l'importine β . Elles contiennent un motif de type IBB qui se caractérise par une séquence d'environ 40 acides aminés, riche en résidus basiques. Concernant les protéines nucléaires qui ne possèdent ni séquence NLS ni IBB (Görlich and Kutay, 1999), leur importation dans le

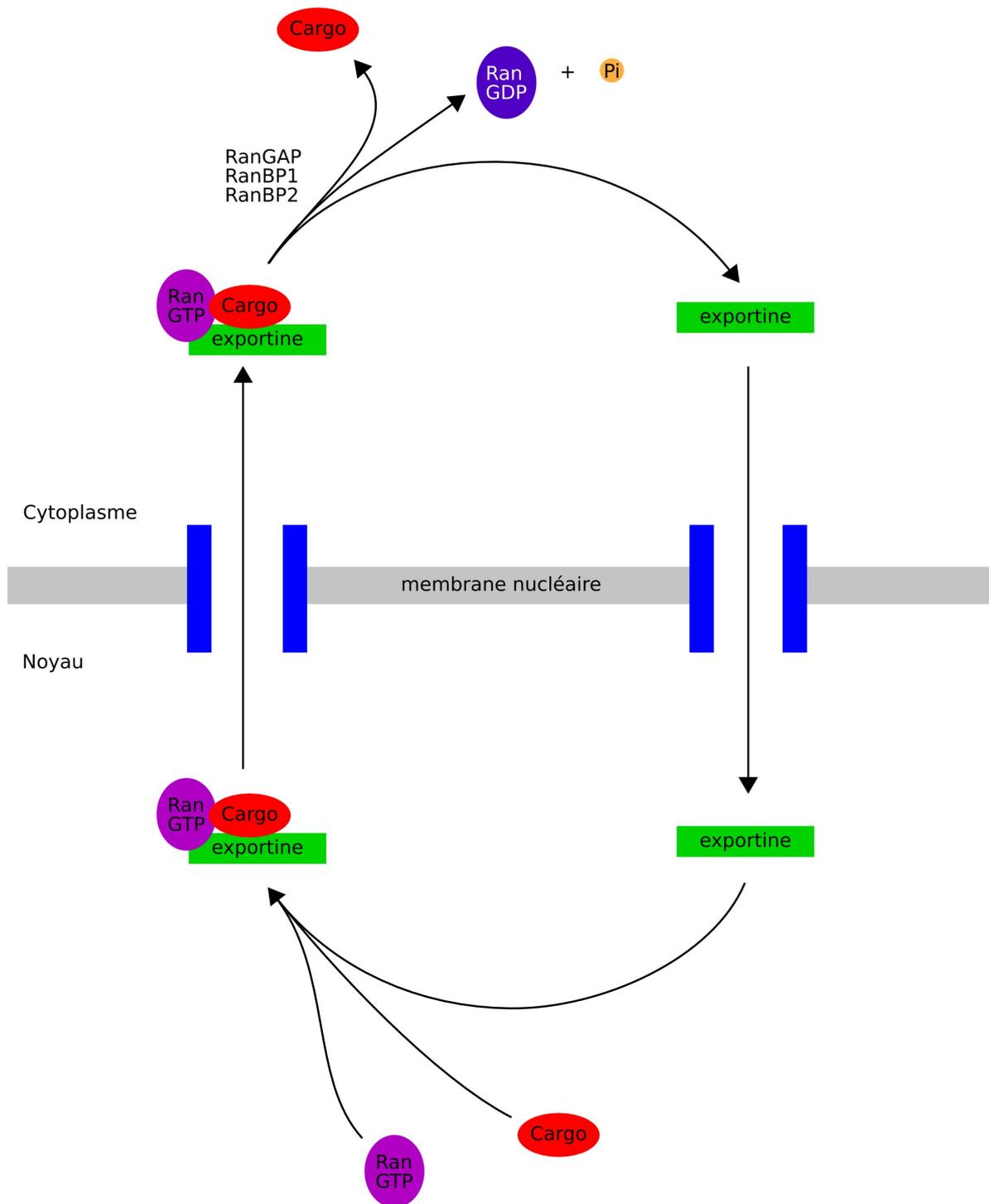


Figure 3.3 : mécanisme de l'export des protéines à travers le pore nucléaire. L'exportine fixe son substrat dans le noyau (en bas) pour former un complexe ternaire cargo-exportine-RanGTP qui est ensuite exporté dans le cytoplasme (en haut) où s'effectue l'hydrolyse du GTP. Cette réaction provoque la dissociation des trois constituants du complexe et le recyclage de l'exportine vers le noyau.

noyau dépend de récepteurs qui appartiennent dans la majorité des cas, à la famille des importines β comme par exemple la transportine 1 qui reconnaît le domaine M9 de la protéine hnRNP A1. Enfin, deux études récentes ont également mis en évidence chez les plantes une voie d'importation impliquant l'importine α indépendamment de l'importine β et de Ran (Hübner *et al.*, 1999; Miyamoto *et al.*, 2002).

3.1.4 Mécanismes de l'export nucléaire

L'export nucléaire, dirigé par les exportines, est régulé de manière inverse à celle de l'import (Weis, 1998; Görlich and Kutay, 1999; Nakielny and Dreyfuss, 1999). Le modèle le plus étudié est celui des protéines qui renferment un NES riche en leucines : un récepteur fixant spécifiquement ce type de NES est la protéine CRM1 ("Chromosome Region Maintenance protein 1") encore appelée exportine 1 (Fornerod *et al.*, 1997; Ossareh-Nazari *et al.*, 1997). Ce récepteur peut également reconnaître certaines protéines ayant des signaux d'export différents, comme dans le cas de la Snurportine 1 (Paraskeva *et al.*, 1999). L'export s'initie par la reconnaissance de la séquence NES du cargo par CRM1 qui s'associe à RanGTP. Le complexe ternaire cargo-CRM1-RanGTP est alors transféré dans le cytoplasme par un mécanisme impliquant la liaison de la protéine CRM1 au NPC. Dans le cytoplasme, l'hydrolyse en GDP du GTP associé à Ran est induite par le complexe Ran-GAP ("GTPase-Activated Protein"), RanBP1 et RanBP2 ("Ran Binding Protein" 1 et 2). Cette hydrolyse conduit à la dissociation de la protéine RanGDP du complexe, et à la dissociation du complexe CRM1/cargo. La protéine CRM1 est finalement importée dans le noyau dès sa dissociation, grâce à sa capacité à interagir avec des nucléoporines (Figure 3.3). D'autres récepteurs reconnaissent des séquences différentes des NES classiques (Görlich and Kutay, 1999) : on peut citer le cas de la protéine CAS qui est impliquée dans l'export des importines α vers le cytoplasme.

3.1.5 Cycle de la protéine RanGTPase et orientation du transport nucléo-cytoplasmique

Le cycle de la protéine RanGTPase joue un rôle clé en conférant une orientation aux évènements du transport nucléo-cytoplasmique (Izaurralde *et al.*, 1997; Nachury and Weis, 1999). Ainsi, en fonction du nucléotide qui lui est associé, la protéine Ran agit comme un marqueur à la fois de l'import et de l'export nucléaire. Selon le modèle proposé par Ström and Weis (2001) (Figures 3.2 et 3.3), chaque cycle d'import-export comprend l'hydrolyse d'une seule molécule de GTP. Un gradient RanGTP-RanGDP est généré par la compartimentalisation des régulateurs du cycle de Ran. Ainsi, la protéine RanGEF ("Guanine nucleotide Exchange Factor" ou encore RCC1) qui régénère RanGTP, est localisée dans le noyau où elle est majoritairement trouvée liée à la chromatine (Bischoff and Ponstingl, 1991). Par contre, la protéine RanGAP et ses cofacteurs, RanBP1 et RanBP2 qui stimulent l'hydrolyse du GTP par la protéine Ran, sont présents dans le cytoplasme (Bischoff *et al.*, 1994, 1995; Yokoyama *et al.*, 1995; Wu *et al.*, 1995). Cette distribution asymétrique suit celle des protéines RanGTP et RanGDP, respectivement présentes dans le noyau et le cytoplasme.

Le NPC contribuerait aussi à l'orientation du transport nucléo-cytoplasmique des cargos (Ben-Efraim and Gerace, 2001) car plusieurs nucléoporines qui se lient aux importines et exportines, présentent des localisations distinctes et asymétriques au sein de cette structure.

3.1.6 Régulation du transport nucléo-cytoplasmique

Plusieurs types de modifications post-traductionnelles interviennent dans la régulation du transport nucléo-cytoplasmique. Ainsi, la phosphorylation de l'antigène T du SV 40 module son importation (pour revue voir Jans and Hübner (1996)). A l'inverse, la phosphorylation des facteurs de transcription de la famille NFAT ("Nuclear Factor of Activated T cells"), chez les mammifères, (Chow *et al.*, 1997; Rao *et al.*, 1997) et du facteur de transcription Msn2p, chez la levure (Görner *et al.*, 1998) provoque leur rétention cytoplasmique. La phosphorylation des protéines du NPC par la protéine kinase C est aussi impliquée dans la régulation de l'export nucléaire des protéines (Beals *et al.*, 1997; Yang *et al.*, 1997) et des

ARN (Schröder *et al.*, 1988). La méthylation et la sumoylation contribuent aussi à la régulation du transport nucléo-cytoplasmique. Ainsi, de nombreuses hnRNPs sont diméthylées au niveau d'arginines par une méthyl-transférase avant leur exportation dans le cytoplasme (Liu and Dreyfuss, 1995; Henry and Silver, 1996). La sumoylation par le complexe SUMO ("Small Ubiquitin-related Modifier") de la protéine CtBP1 ("C-terminal binding protein of adenovirus E1A") et du répresseur transcriptionnel TEL est requise pour leur localisation dans le noyau (Lin *et al.*, 2003; Wood *et al.*, 2003).

3.1.7 Le transport nucléo-cytoplasmique de protéines et acides nucléiques viraux : entre activation et inhibition des échanges

Les virus ont largement contribué à la compréhension des mécanismes impliqués dans le transport nucléo-cytoplasmique. Comme nous l'avons précisé au début de cette introduction, pratiquement tous les virus à ADN effectuent leur réplication dans le noyau de la cellule hôte et certains virus à ARN exploitent la machinerie nucléaire pour réaliser une ou plusieurs étapes de leur cycle infectieux (pour revue, voir Whittaker and Helenius, 1998). Ces virus doivent donc franchir plusieurs barrières physiques et plus particulièrement, celle de l'enveloppe nucléaire pour acheminer leur génome dans le noyau. Ils utilisent le cytosquelette pour arriver jusqu'au noyau où ils délivrent leur ADN ou ARN génomique, en s'affranchissant des limites de taille imposées par la structure du NPC (pour revue, voir Sodeik, 2000; Campbell and Hope, 2003). L'import de l'ADN viral est particulièrement étonnant car il s'agit d'une macromolécule qui n'est généralement pas transportée de part et d'autre du noyau, et de ce fait, les virus à ADN doivent modifier les mécanismes de transport pré-existants.

Les virus qui ont un cycle de réplication nucléaire utilisent généralement les mécanismes classiques pour l'import et/ou l'export de leurs protéines et ARN, ceux-ci étant dotés de signaux reconnus par la machinerie cellulaire. L'export nucléaire des particules virales en fin de cycle infectieux est par contre plus problématique. Certains virus doivent

recourir à la mort cellulaire (Möller and Schmitz, 2003) ou au bourgeonnement à travers la membrane interne de l'enveloppe nucléaire pour s'échapper de la cellule hôte.

Des virus qui ont un cycle entièrement cytoplasmique peuvent inhiber le transport nucléo-cytoplasmique (voir paragraphe 3.1.8, page 51) afin d'avantager leur réplication. Ainsi, en empêchant l'export des ARN messagers cellulaires, ces virus favorisent la traduction de leurs propres messagers et donc, la production de protéines virales. L'inhibition du transport nucléo-cytoplasmique peut aussi empêcher le développement de la défense cellulaire contre les infections virales. Par exemple, dans la voie de l'interféron chez les animaux, l'activateur de transcription STAT, doit être importé dans le noyau pour permettre l'expression d'un ensemble de gènes de défense antivirale (pour revue : Stark *et al.*, 1998). De ce fait, l'inhibition de l'import de STAT ("Signal Transducers and Activators of Transcription") réduit l'efficacité de la réponse antivirale. Une étude récente a révélé un autre lien entre la réponse interféron et le NPC. En effet, Enninga *et al.* (2002) ont montré que l'expression de deux nucléoporines, Nup98 et Nup96, est induite après traitement des cellules à l'interféron- γ et que ces deux protéines lèvent le blocage de l'export des ARNm cellulaires dû à la protéine de matrice du virus de la stomatite vésiculaire, un Rhabdovirus.

3.1.8 Inhibition du trafic nucléo-cytoplasmique par les virus qui ont un cycle réplcatif strictement cytoplasmique : exemple du poliovirus

Un autre exemple d'inhibition du transport nucléo-cytoplasmique est celui observé lors de l'infection de cellules par le poliovirus. Weidman *et al.* (2003) ont montré que le pré-curseur 3CD est capable d'entrer dans le noyau des cellules infectées grâce à la présence d'une séquence NLS classique (KKKRD) localisée sur le domaine $3D^{pol}$ du polypeptide. Après clivage autoprotéolytique, la protéase $3C^{pro}$ induit le clivage de la protéine TBP ("TATAbox Binding Protein") qui inhiberait ainsi la transcription de gènes cellulaires (Gustin and Sarnow, 2001); la répartition nucléo-cytoplasmique de la protéine RanGTP et des importines reste inchangée (Gustin, 2003). Cependant, les nucléoporines Nup153 et p62 sont également protéolysées par la protéine $3C^{pro}$. Ces deux nucléoporines sont impliquées

dans l'import nucléaire de protéines à NLS classique (Moroianu *et al.*, 1997; Percipalle *et al.*, 1997; Shah *et al.*, 1998; Nakielny and Dreyfuss, 1999). Ainsi, les protéines La, PTB ("Polypyrimidine Tract Binding protein") et Sam68 ("Src-Associated in Mitosis protein") ne sont plus adressées dans le compartiment nucléaire et sont alors recrutées lors du cycle viral pour la traduction Cap indépendante du génome viral (La et PTB) et lors de la réplication de l'ARN (Sam68) (Gustin, 2003; Gustin and Sarnow, 2001; Belov *et al.*, 2000).

3.1.9 Le transport nucléo-cytoplasmique chez le BNYVV

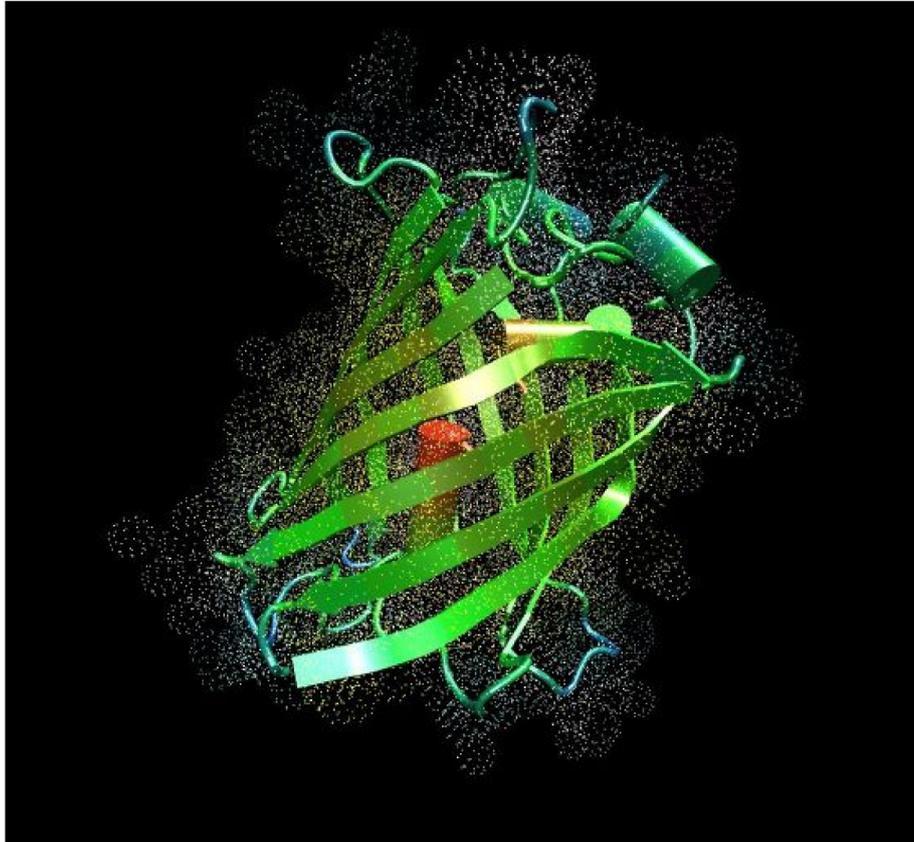
Si Haeblerlé *et al.* (1994) ont montré que la protéine P25 codée par l'ARN-3 était localisée dans le noyau et dans le cytoplasme des cellules infectées, il a fallu attendre 2004 (Vetter *et al.*, 2004) pour démontrer que cette localisation nucléaire était liée à la présence d'un NLS classique (**KRIRFR**). Une telle étude effectuée à l'aide de protéines fusionnées à l'EGFP ("Enhanced Green Fluorescent Protein") a également montré que la protéine P25 était activement réexportée dans le cytoplasme par la voie de l'exportine 1 grâce à un motif NES (**VYMVCLVNTV**). Le rôle exact de cette navette n'est pas encore connu mais pourrait expliquer le phénomène de démenche racinaire observé au champ. En effet, une telle prolifération racinaire n'est pas observée lors de l'infection de plantes par des isolats délétés d'une partie de la protéine P25 (Tamada *et al.*, 1999).

3.2 Utilisation de la EGFP

3.2.1 Propriétés de la "Green Fluorescent Protein" (GFP)

La GFP est une protéine qui a été isolée à partir d'*Aequoria victoria*, une méduse vivant dans l'océan Pacifique qui émet naturellement une fluorescence verte, due à l'expression de cette protéine. La fluorescence verte (509 nm) est stimulée par l'absorption de lumière bleue (470 nm) par le fluorochrome contenu dans la GFP. Cette lumière bleue, absorbée par la GFP, est émise *in vivo* lors d'un mécanisme calcium dépendant : l'activation de l'aequorine, une protéine de 21 kDa associée à l'oxygène et à la coelenterazine, conduit à l'oxydation de la coelenterazine en coenterazine, qui de son état excité revient ensuite

A)



B)

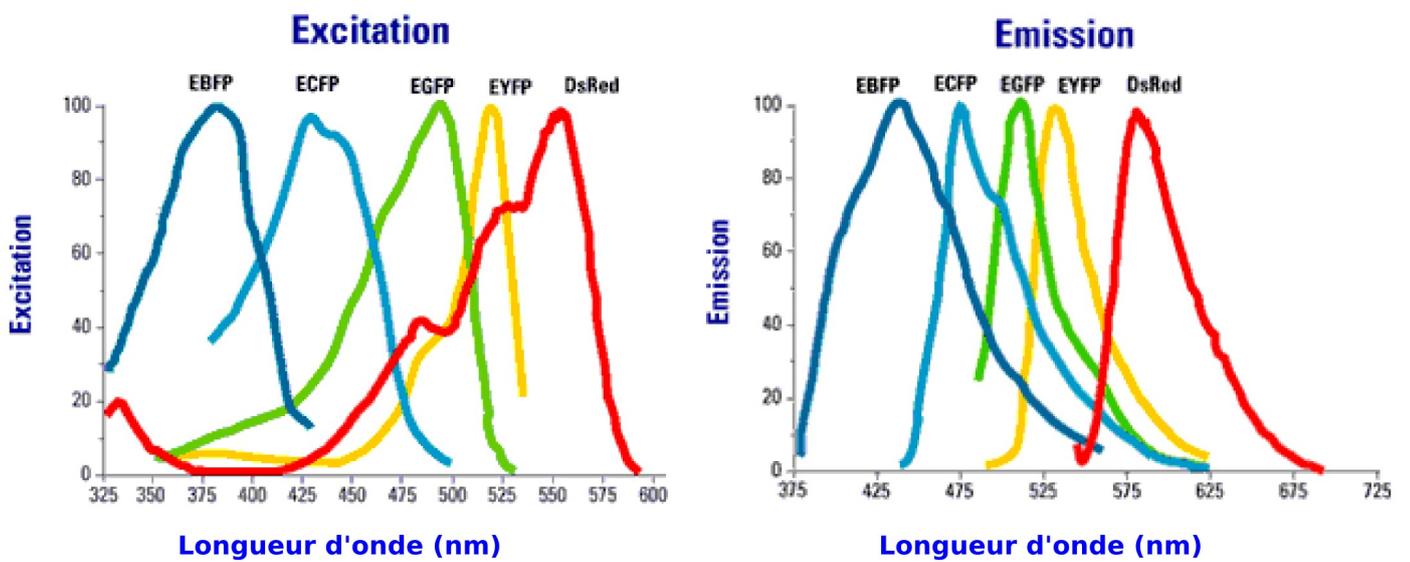


Figure 3.4 : Propriétés structurales et caractéristiques spectrales de protéines fluorescentes. A) Structure tridimensionnelle de la "Green Fluorescent Protein" (GFP). Les flèches vertes représentent les feuillets β et les cylindres sont des hélices α . B) Spectres d'excitation et d'émission de différentes protéines fluorescentes dérivée de la GFP (EBFP, ECFP, EGFP et EYFP) et de la DsRed (d'après : <http://www.clontech.com>).

à l'état stable en émettant une fluorescence à 470 nm (bleu). La GFP est une protéine de 27 kDa (Prasher *et al.*, 1992; Chalfie *et al.*, 1994), dont la structure tridimensionnelle a été résolue par diffraction aux rayons X (Ormö *et al.*, 1996; Youvan and Michel-Beyerle, 1996) et est présentée dans la Figure 3.4 A. Elle se présente sous la forme d'un tonneau très stable qui se structure sans l'intervention de chaperones. Elle comporte 11 feuillets β et 4 hélices α . Les acides aminés impliqués dans la fluorescence (SER65, TYR66 et GLY67) se cyclisent pour former le chromophore (Cubitt *et al.*, 1995; Cody *et al.*, 1993; Heim *et al.*, 1994). Son spectre d'absorption présente un pic majeur à 395 nm (bleu) suivi d'un pic mineur à 475 nm.

Cette protéine garde ses propriétés lorsqu'elle est fusionnée à d'autres séquences peptidiques, ce qui en fait un marqueur de choix pour des études de localisation *in vivo*.

3.2.2 Améliorations et modifications des propriétés spectrales de la GFP

La nécessité d'utiliser les UV comme source d'excitation du chromophore est létale pour les tissus irradiés et provoque des phénomènes de "bleaching" très rapides conduisant à la perte de fluorescence. La protéine a donc été modifiée génétiquement pour changer plusieurs de ses caractéristiques spectrales, sa brillance mais aussi pour adapter son expression aux cellules des eucaryotes supérieurs et pour éviter son agrégation à forte concentration (MET154THR et VAL164ALA). En effet, la détection *in vivo* de la GFP dépend de son seuil d'expression. Si une concentration cellulaire de 1 μ M de GFP est requise dans les cellules Hela, la forme mutée de GFP (SER65THR) (Cubitt *et al.*, 1995) est détectée à partir de 200 nM et l'Enhanced GFP (EGFP : PHE64LEU et SER65THR, codons optimisés) à partir de 100 nM.

Les spectres d'excitation et d'émission de cette protéine fluorescente ont pu être modifiés par mutagenèse dirigée. Ainsi, des dérivés bleus (EBFP et ECFP) et jaunes (EYFP) ont été obtenus (Figure 3.4 B). Plus récemment, une autre protéine fluorescente, DsRed, qui fluoresce en rouge a été isolée de l'anémone *Discosoma sp* (Figure 3.4 B). Cette dernière protéine forme naturellement des tétramères, pouvant provoquer son agrégation *in vivo*, ce

qui constitue son inconvénient majeur. Des modifications de cette protéine ont permis d'en produire un dérivé monomérique nommé mRFP (Campbell *et al.*, 2002).

Ce panel de protéines couvrant une gamme spectrale d'excitation et d'émission communes permet d'accéder à des expériences dites de FRET ("Fluorescent Resonance Energy Transfer") (Heim and Tsien, 1996; Matz *et al.*, 1999). Cette technique permet de mettre en évidence des interactions entre protéines grâce à l'utilisation de protéines de fusion (protéine fluorescente/protéine d'intérêt) puis par la mesure de transfert d'énergie associé à l'excitation d'un fluorochrome donné vers un autre fluorochrome. Les couples donneurs/accepteurs les plus utilisés actuellement sont les couples BFP/GFP et ECFP/EYFP.

3.2.3 Applications en virologie moléculaire

La GFP est devenue un outil incontournable pour les expériences de localisation *in situ*. Outre sa vocation de marqueur de protéines (Chalfie *et al.*, 1994) dont quelques exemples seront décrits plus bas, elle a permis de mettre en évidence *in vivo* des organites intracellulaires précis et, par la même occasion, a su en partie se substituer aux colorants vitaux ou aux immuno-marquages. En effet, la fusion de petites séquences d'adressage à la GFP comme, par exemple, la fusion du peptide KDEL, localise spécifiquement la GFP recombinante dans le réticulum endoplasmique rugueux qui fluoresce en vert. De même, la plupart des protéines du cytosquelette ont été couplées à la GFP, ce qui a permis d'accéder à l'imagerie dynamique des microfilaments, des microtubules et des filaments intermédiaires (Pang *et al.*, 1998). En biologie animale, la GFP est également utilisée comme marqueur de transformation d'organismes, outil qui permet de sélectionner rapidement les animaux porteurs de transgènes, ou peut être utilisé comme gène rapporteur dans l'étude de séquences promotrices et ainsi d'accéder à l'étude de leurs éléments régulateurs. En cancérologie, des cellules mutées marquées à la GFP et injectées à des souris permettent de suivre la formation de métastases. Enfin, cette protéine fluorescente est utilisée comme traceur d'infections. Elle permet, par exemple, de suivre l'évolution d'une infection bactérienne ou virale *in vivo*.

En virologie végétale, la GFP est un marqueur de choix pour suivre le déplacement de l'infection virale dans une plante. A l'inverse du marqueur GUS (β -glucuronidase), la GFP

est plus simple d'utilisation et n'est pas destructive pour la plante. L'équipe de Baulcombe *et al.* (1995) fut la première à utiliser la GFP comme marqueur d'infection. Pour cela, le gène de la GFP est inséré dans le génome du virus qui est ainsi utilisé comme vecteur d'expression transitoire. Le gène de la GFP peut être exprimé sous le contrôle d'un promoteur viral dupliqué, comme dans le cas du PVX·GFP. Le développement de l'infection virale dans la plante peut être très simplement suivi par illumination à la lumière UV et observation de la fluorescence verte de la protéine GFP. Par microscopie à fluorescence, il a été démontré que la GFP exprimée à partir du vecteur PVX était localisée à la fois dans le cytoplasme et dans le noyau des cellules infectées (Baulcombe *et al.*, 1995; Oparka *et al.*, 1996; Cruz *et al.*, 1996). Par la suite, des expériences similaires ont été réalisées avec le TMV (Epel *et al.*, 1996). Des informations sur la fonction des protéines virales ont également été obtenues grâce à l'utilisation de la GFP comme marqueur. C'est ainsi que l'importance de la protéine CP dans le mouvement de cellule à cellule du PVX (Baulcombe *et al.*, 1995) a été démontrée.

Enfin, la fusion de la GFP aux protéines virales a permis de déterminer la localisation sub-cellulaire d'un grand nombre de ces protéines. L'adressage aux plasmodesmes et l'interaction de la protéine de mouvement P30 du VMT a pu être ainsi confirmée à l'aide de la protéine de fusion GFP::P30. Plus tard, des expériences de colocalisation ont montré que cette protéine de mouvement suivait le réseau du cytosquelette (Heinlein *et al.*, 1995; Heinlein and Epel, 2004). L'expression hétérologue de la protéine de fusion GFP::P30 dans le vecteur PVX a permis de démontrer que l'adressage aux plasmodesmes et l'affinité de la protéine P30 avec le cytosquelette était une propriété intrinsèque à la protéine P30 (Oparka *et al.*, 1996). Outre les protéines de mouvement, les protéines de capsides peuvent également être marquées par la GFP. Dans le cas du PVX, la fusion de la GFP à l'extrémité N-terminale de la CP a permis de suivre le déplacement *in situ* des particules virales (Cruz *et al.*, 1996). Un tel système est utilisé dans le cas du BNYVV où la protéine mineure de capside est fusionnée à l'EGFP (Erhardt *et al.*, 2001). Dans ce même système viral, Vetter *et al.* (2004) ont pu montrer, grâce à des fusions à l'EGFP, que la protéine P25 du BNYVV était capable d'effectuer une navette entre le noyau et le cytoplasme de cellule infectées et transfectées.

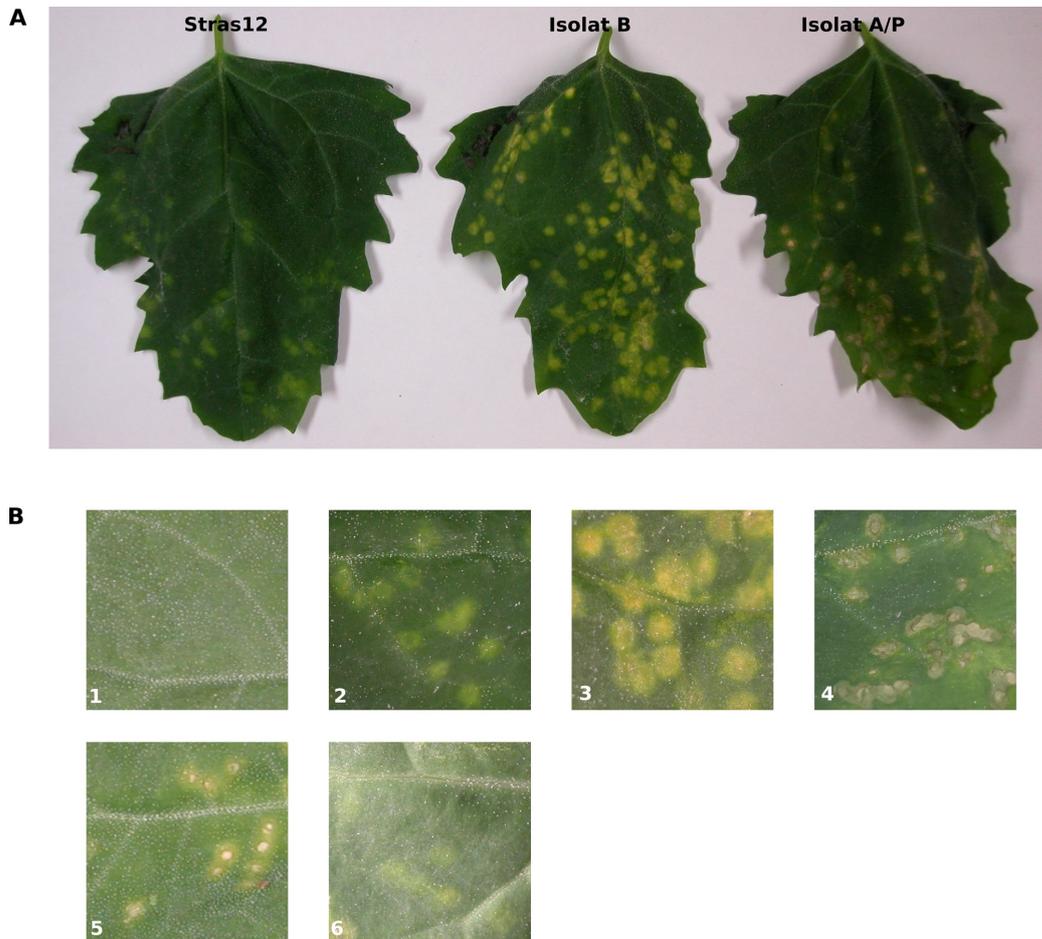


Figure 3.5 : Symptômes foliaires induits par inoculation mécanique d'ARN infectieux du BNYVV sur *Chenopodium quinoa*. A) Photographie de trois feuilles inoculées par des ARN de la souche helper Stras12 contenant les ARN-1 et -2 seuls (gauche), de l'isolat B (centre) et de l'isolat A/P (droite) contenant respectivement 4 et 5 ARN viraux. B) Détails des lésions induites par ces trois mêmes souches (2 à 4) ou d'une souche Stras12 exprimant la protéine P26 sauvage (5) ou mutée (6) par l'intermédiaire d'un ARN-3 recombinant, 1) Feuille saine ; 2) Stras12 ; 3) Isolat B ; 4) Isolat A/P ; 5) Stras12 + rep-P26 ; 6) Stras12 + rep-P26FS.

3.3 Etude de la protéine P26 codée par l'ARN-5

3.3.1 Effet sur la symptomatologie

La Figure 3.5 A illustre l'importance des petits ARN sur l'expression des symptômes sur feuilles de *Chenopodium quinoa* 6 jours après inoculation (dpi) mécanique d'ARN viraux infectieux. Les lésions chlorotiques vertes provoquées par l'infection avec les ARN-1 et -2 (Figure 3.5 A et B-2) de l'isolat B deviennent chlorotiques jaunes et nécrotiques en présence des ARN-3 et -4 du même isolat (Figure 3.5 A et B-3). En présence d'ARN-5, Figure 3.5 A (isolat A/P), les lésions se nécrosent fortement (Figure 3.5 B-4) et apparaissent plus précocement (5 dpi). Cette nécrose est liée à l'expression de la protéine P26 codée par l'ARN-5. En effet, son expression par l'intermédiaire d'un ARN-3 recombinant, appelé réplicon (voir paragraphe 3.3.2) permet de reproduire la nécrose en présence des ARN-1 et -2 (Figure 3.5 B-5). Cette nécrose est bien liée à l'expression de la protéine P26 puisqu'un mutant de décalage du cadre de lecture de cette protéine abolit la formation des nécroses (Figure 3.5 B-6) et conduit à des symptômes comparables à ceux obtenus avec les ARN-1 et -2 de la souche helper (comparer Figure 3.5 B, 5, 6 et 2).

3.3.2 Construction de clones et de virus recombinants exprimant la protéine P26

Pour initier l'étude du rôle de l'ARN-5 dans la pathogénicité virale, il était nécessaire de disposer de la séquence codante de la protéine P26. Pour cela, l'isolat A/P a été obtenu à partir de betteraves infectées naturellement. Les plantes infectées proviennent de cultures de betteraves dans du terreau contenant 10% de terre renfermant le vecteur virulifère issue de champs de Yèvre-La-Ville (proche de Pithiviers). Après croissance des betteraves durant un mois et extraction des ARN totaux de racines, une RT-PCR haute fidélité effectuée à l'aide d'amorces flanquant les séquences du gène codant pour la protéine P26 (BN5F : GTTTTTCCGCTCGCACAAGCG et BN5R : CGAGCCCGTAAACACCG-CATA) nous a permis d'obtenir un fragment d'ADN contenant la séquence de la protéine P26. L'amplicon obtenu a été placé dans un vecteur pGEM®-T et séquencé.

La séquence obtenue est parfaitement identique à celles de protéines P26 issues de GenBank et qui proviennent de la même région. Cette séquence contient trois codons initiateurs possibles, correspondant à des méthionines, en positions 439, 457 et 472 (positions issues de la séquence du clone ADNc de l'ARN-5 complet déterminée par la suite lors de ce travail, voir Figure 3.9).

La protéine produite à partir du premier codon initiateur possède une masse moléculaire calculée de 26,45 kDa, un pH iso-électrique de 4,43. Ce pHi très bas est expliqué par la présence de 29 résidus acides (12,5 % des résidus totaux) et 22 résidus basiques (9,48 %) répartis régulièrement sur toute la phase ouverte de lecture. Les outils bioinformatiques ne nous ont pas permis de trouver des motifs particuliers au sein de la séquence de la protéine P26. Ainsi, la protéine ne présente pas de motif "helix-turn-helix", caractéristique d'une protéine se fixant aux acides nucléiques, ni de structure "coiled-coil" pouvant suggérer une possibilité de dimérisation ou d'interaction protéine/protéine. Enfin, les recherches par alignement de séquences en interrogeant les banques de données n'ont révélé aucune homologie avec des protéines connues, ne donnant pas d'indication de fonctions potentielles.

Le clonage de la phase ouverte de lecture codant pour la protéine P26 dans un vecteur d'expression viral dérivé de l'ARN-3 disponible au laboratoire, le réplicon (Hehn *et al.*, 1994), a permis de tester l'influence de son expression sur la symptomatologie induite lors de l'inoculation mécanique de transcrits infectieux en présence des ARN-1 et -2. Une telle construction Rep-P26 est capable d'induire une nécrose comparable à celle observée en présence d'isolat A/P (Figure 3.5 B-5). Pour vérifier que c'est bien la protéine et non pas la séquence de l'ARN introduite dans le réplicon qui est responsable de ces symptômes, nous avons introduit une mutation de décalage de cadre de lecture (frameshift) dans la séquence codante de la protéine P26. Cette mutation a été produite par remplissage des extrémités cohésives du site *Bsr*GI et religation. Les transcrits du réplicon exprimant cette protéine mutée, inoculés en présence des ARN-1 et -2 n'est plus capable d'induire de nécrose, démontrant ainsi l'implication de l'expression de la protéine P26 dans la mort cellulaire (Figure 3.5 B 6).

En absence d'anticorps dirigés contre la protéine P26, nous avons fusionné la séquence de l'épitope HA (Hémagglutinine du virus de la grippe, peptide de séquence YPYDVPDYA)

à l'extrémité C-terminale de la protéine pour maintenir le contexte d'initiation et l'étudier. La construction obtenue, Rep-P26HA, est capable d'induire les mêmes symptômes nécrotiques que la protéine P26 non étiquetée.

3.4 Étude du contexte d'initiation de la traduction de la protéine P26

La séquence nucléotidique d'un ARN au voisinage d'un codon d'initiation influence sa reconnaissance par la machinerie traductionnelle. Ainsi, la présence d'une purine (R) en position -3 et d'une guanine (G) en +4 (gccRccAUGG) est optimale pour la traduction (Kozak, 2002). La sous-unité 40S du ribosome se fixe sur la partie 5' de l'ARN et se déplace vers l'extrémité 3' jusqu'à atteindre un codon d'initiation en contexte favorable. La traduction débute et peu de complexes d'initiation atteignent les AUG plus éloignés, qu'ils soient ou pas en contexte favorable. Des mécanismes d'initiation internes permettent la traduction à partir des AUG éloignés de l'extrémité 5' de l'ARN messager, comme dans le cas de l'initiation de la traduction de l'ARN de poliovirus (Michel *et al.*, 2001) ou de l'ARN 35S du CaMV (Ryabova *et al.*, 2004).

Concernant l'ARN-5, la phase ouverte de lecture la plus grande permettant l'expression de la protéine P26 possède trois codons d'initiation dont deux (le premier et le troisième) sont en contexte favorable. Ces deux codons possèdent une guanine en -3 et en +4 (Figure 3.6). Il est ainsi fort probable que le codon d'initiation utilisé pour la synthèse de la protéine P26 soit le premier AUG. Pour le vérifier, nous avons conduit des expériences de mutagenèse dirigée pour forcer l'initiation sur le premier, le second ou le troisième codon initiateur potentiel de la protéine étiquetée HA, exprimée par l'intermédiaire du réplicon dérivé de l'ARN-3. Dans ce réplicon, le premier codon d'initiation a vu son contexte d'initiation nucléotidique modifié mais celui-ci respecte la présence d'une purine en -3. Les constructions P26AUG2 et P26AUG3 forcent l'initiation de la traduction sur le deuxième et le troisième AUG de la protéine, respectivement. Les mutants P26HA-B et P26HA-C ont été conçus pour éliminer la possibilité d'initiation au troisième et aux deuxième et troisième AUG, respectivement, en les remplaçant par un codon sérine (UCU). Les mutants P26HA-

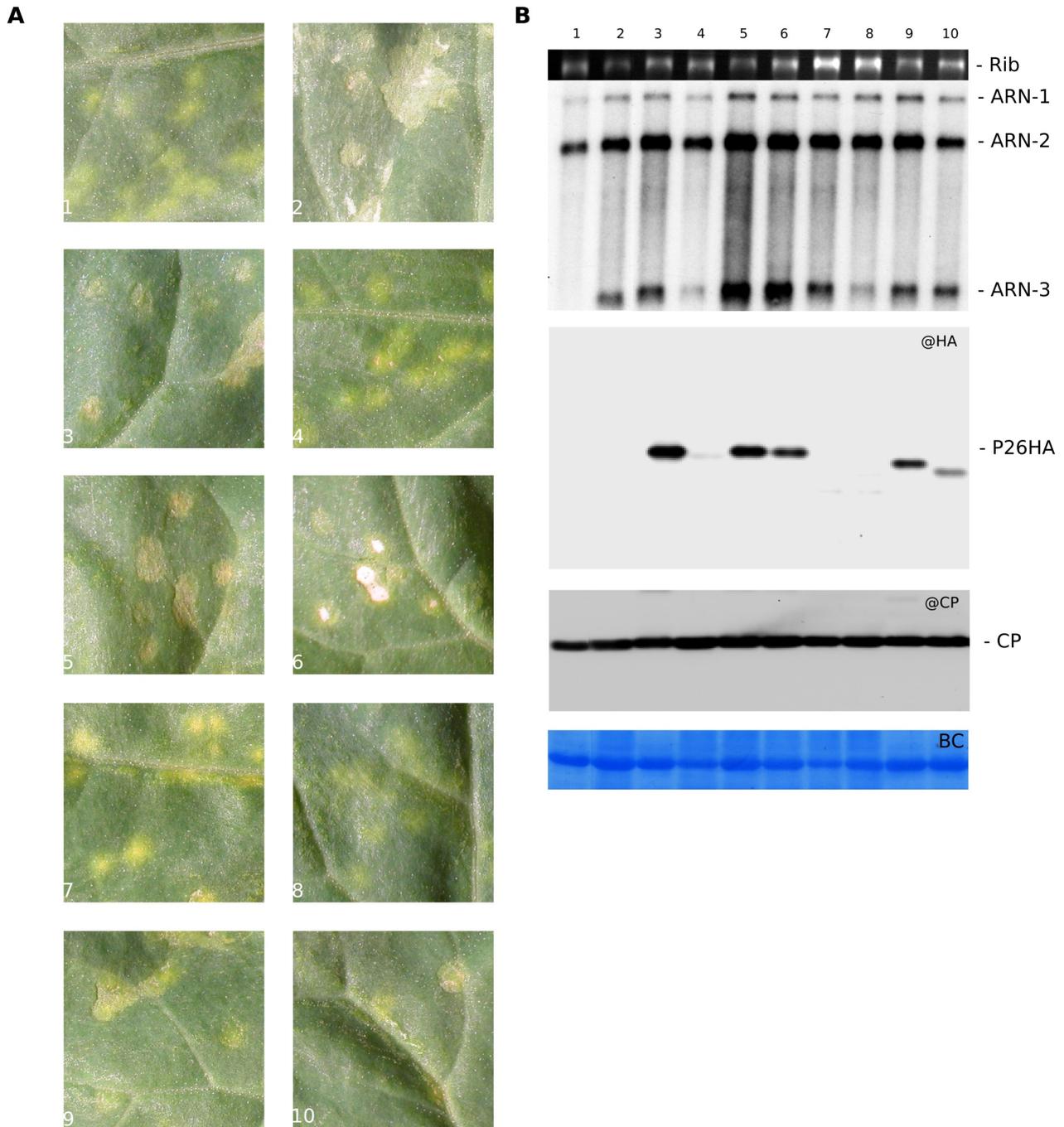


Figure 3.7 : Etude de l'influence du contexte d'initiation de la protéine P26 sur la symptomatologie foliaire et sur l'expression de la protéine. A) Symptômes observés 6 jours après inoculation mécanique sur feuilles de *C. quinoa*. 1 : Stras12 ; 2 : Stras12 + rep-P26 ; 3 : Stras12 + rep-P26HA ; 4 : Stras12 + rep-P26HA-A ; 5 : Stras12 + rep-P26HA-B ; 6 : Stras12 + rep-P26HA-C ; 7 : Stras12 + rep-P26HA-D ; 8 : Stras12 + rep-P26HA-E ; 9 : Stras12 + rep-P26AUG2 ; 10 : Stras12 + rep-P26AUG3. B) Analyse par northern-blot de l'accumulation des ARN viraux et détection par western-blot de l'accumulation de la protéine P26 étiquetée HA. Rib : ARN ribosomiques ; @HA : immunodétection de l'épitope HA ; @CP immunodétection de la protéine CP ; BC : coloration au Serva Blue.

A, P26HA-D et P26HA-E contiennent une mutation frameshift indiquée sur la Figure 3.6 par une flèche noire. Ces derniers mutants permettent de visualiser d'éventuelles initiations internes de la traduction sur le deuxième ou le troisième codon d'initiation. Ainsi, le mutant P26HA-A permet de visualiser une possible initiation interne à partir du troisième codon d'initiation, le mutant P26HA-E à partir du deuxième codon d'initiation et le mutant P26HA-D sur ces deux codons à la fois.

Les transcrits correspondant à l'ensemble de ces constructions ont été inoculés, en présence de la souche helper Stras12, à des feuilles de *C. quinoa*. Les lésions obtenues (6 dpi) ont été caractérisées phénotypiquement, les ARN viraux et les protéines virales recherchées. Les résultats obtenus sont représentés sur la Figure 3.7. Comme nous l'avons décrit précédemment, l'absence de synthèse de la protéine P26 conduit à la perte de la formation de lésions nécrotiques. Ainsi, au vu des symptômes obtenus, nous avons constaté que les mutants P26HA-B (Figure 3.7 A 5), P26HA-C (Figure 3.7 A 6), P26AUG2 (Figure 3.7 A 9) et P26AUG-3 (Figure 3.7 A 10) induisent toujours des lésions nécrotiques comparables à celles induites par l'expression de la protéine P26 (Figure 3.7 A 2) et P26HA (Figure 3.7 A 3). Les mutants P26HA-A (Figure 3.7 A 4), P26HA-D (Figure 3.7 A 7) et P26HA-E (Figure 3.7 A 8) n'induisent plus de nécrose mais des lésions chlorotiques vertes comparables à celles induites par la souche helper Stras12 (Figure 3.7 A 1). Ces différences de symptômes ne sont pas liées à l'absence de répllication des réplicons utilisés puisque la descendance virale de l'ensemble des constructions utilisées est détectée par northern-blot en utilisant des sondes spécifiques des ARN-1, -2 et -3. La recherche de l'expression des protéines P26 dans les lésions a été effectuée par western-blot à l'aide d'un anticorps monoclonal, couplé à la peroxydase, dirigé contre l'épitope HA. Sur la même membrane, nous avons recherché la protéine P26 et la protéine CP pour comparer le niveau d'accumulation du virus dans chaque échantillon, puis nous avons estimé la charge des protéines totales en colorant la membrane au Serva Blue. Cette approche permet une analyse comparative des différents échantillons entre eux. En effet, comme on peut le constater, la quantité de protéines totales déposées et transférées est similaire, tout comme l'accumulation de la protéine CP (Figure 3.7 B, CP et BC). L'utilisation de l'anticorps anti HA nous a permis de déceler l'expression de la protéine P26HA dans les lésions infectées par la construc-

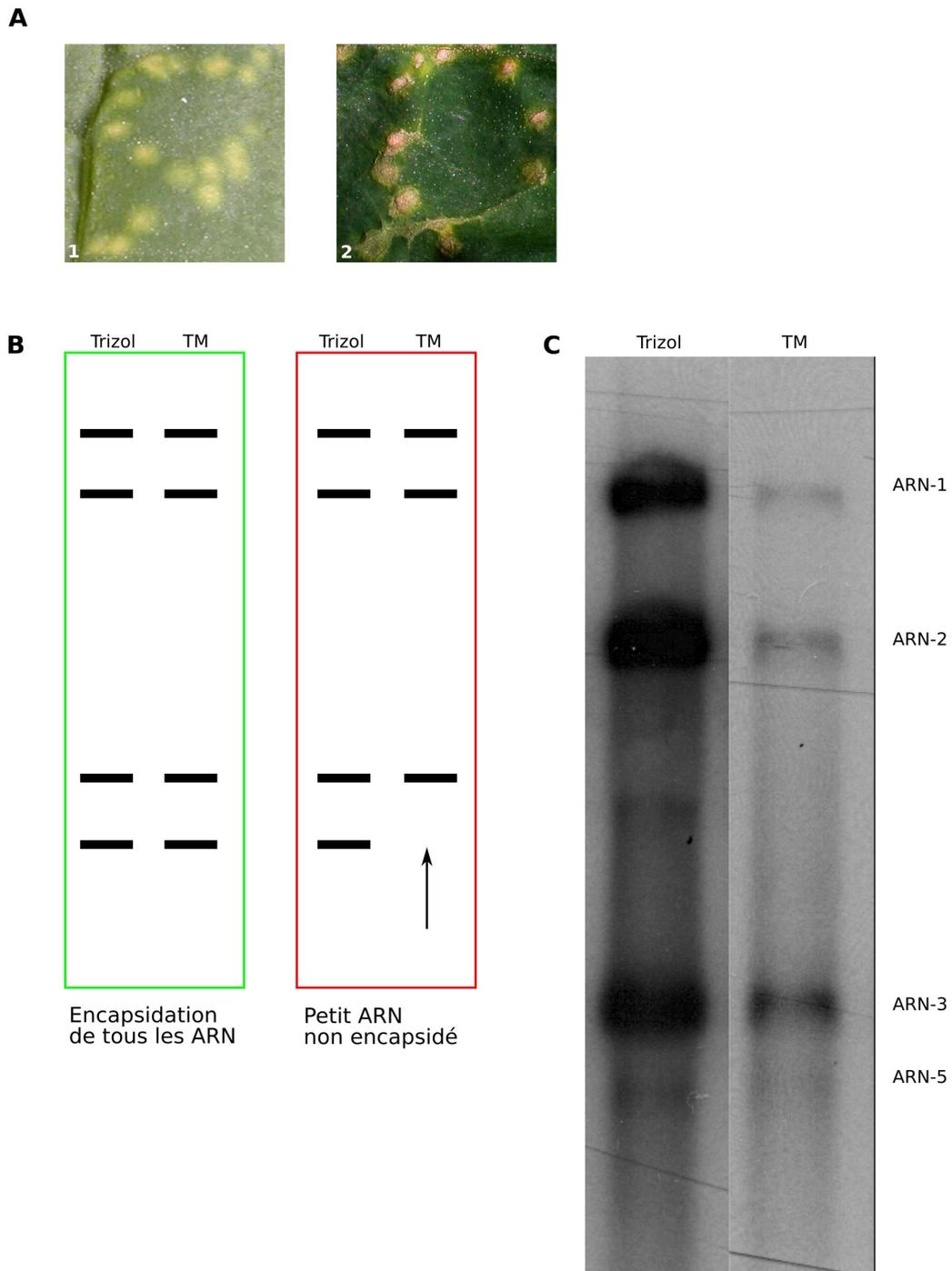


Figure 3.8 : Caractérisation des propriétés biologiques du clone ADNc de l'ARN-5. A) Symptômes obtenus sur des feuilles de *C. quinoa* après inoculation mécanique de Stras12 supplémenté d'ARN-3 (1) ou d'ARN-3 et -5 (2). B) Représentation schématique du test d'encapsidation utilisant deux extractions différentielles. La flèche noire désigne la position d'un ARN non encapsidé. C) Analyse par northern-blot de la présence des ARN viraux 1, 2, 3 et 5 à l'aide de sondes spécifiques. Les ARN viraux nu et encapsidés (Trizol) sont comparés aux ARN viraux encapsidés (TM).

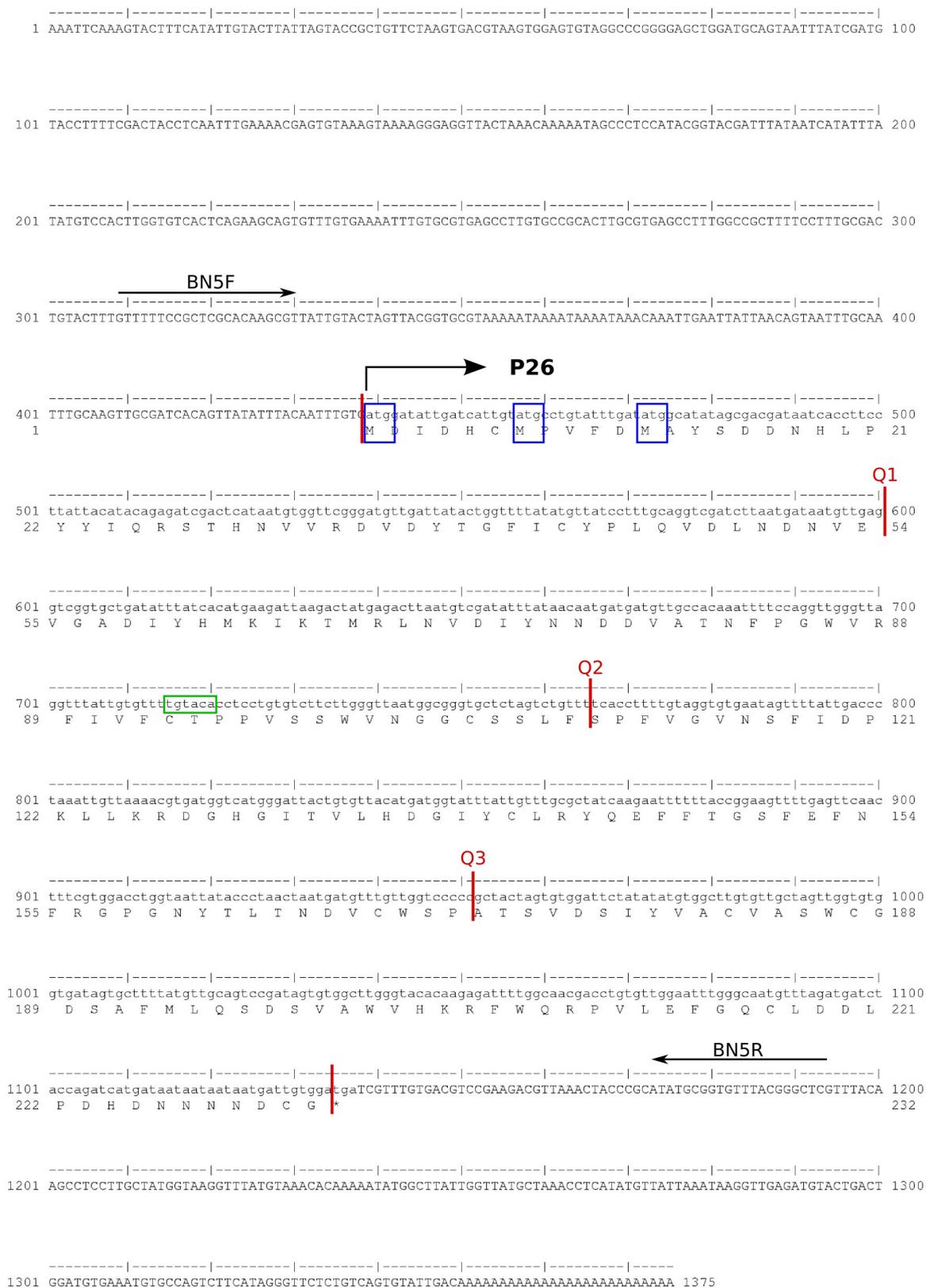


Figure 3.9 : Séquence annotée du clone ADNc de l'ARN-5 (numéro d'accèsion AY823407) provenant de la région de Pithiviers. La séquence de la protéine P26 est précisée en dessous de la séquence nucléotidique. Les trois sites potentiels d'initiation de la traduction sont encadrés en bleu. L'initiation de la traduction de la protéine P26 s'effectue au niveau de la flèche (voir paragraphe 3.4). Le site de restriction unique *Bsr*GI est encadré en vert. Les barres rouges sont placées aux positions divisant la protéine en quatre domaines qui seront étudiés par la suite. Les mentions Q1, Q2 et Q3, également en rouge, nomment les positions représentant les premier, deuxième et troisième quarts de la protéine en partant de la première méthionine. Les positions des amorces BN5F et BN5R sont indiquées par des flèches.

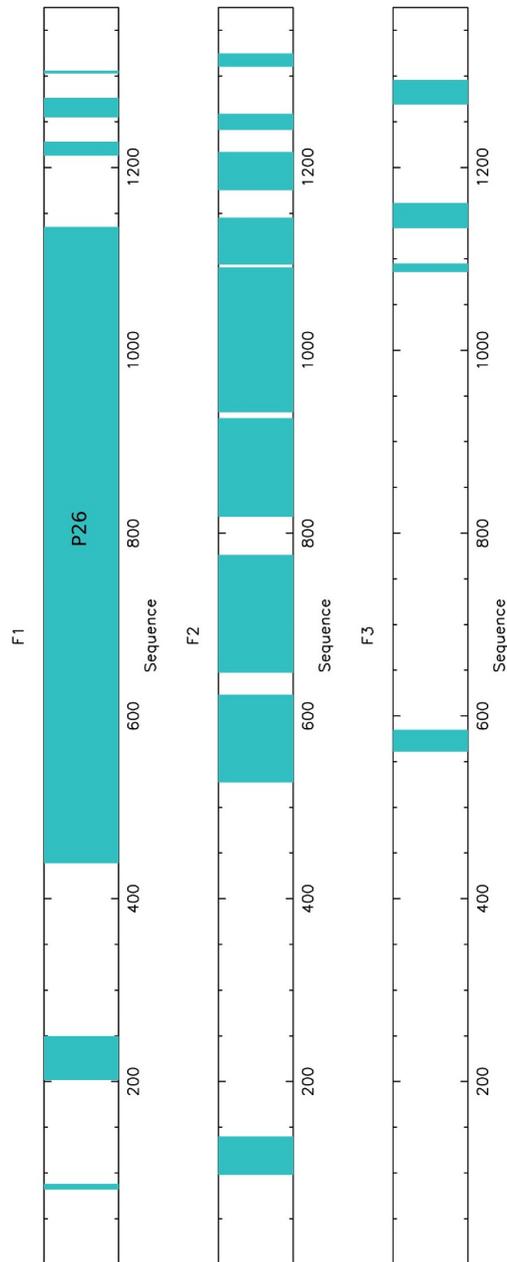


Figure 3.10 : Représentation schématique des phases ouvertes de lecture présentes dans les trois cadres de lecture du clone complet ADNc de l'ARN-5 du BNYVV. Chacune des trois boîtes noires représente la séquence complète de l'ARN-5 dans un cadre de lecture possible (sens direct). Chaque phase ouverte de lecture est schématisée par un rectangle bleu. La séquence codant pour la protéine P26 est retrouvée dans la phase de lecture F1 (position 439 à 1134).

tion Rep-P26HA, Rep-P26HA-B, -C et Rep-P26AUG2 et 3 (Figure 3.7 B, colonnes 3, 5, 6, 9 et 10 respectivement). La distance de migration des protéines P26HA, P26HA-B et -C sont identiques. Quant à celle des protéines P26AUG2 et 3, elles reflètent des tailles de protéines plus courtes. Grâce à cette première observation, nous pouvons d'ores et déjà conclure qu'uniquement le premier codon d'initiation est utilisé pour la synthèse de la protéine P26HA puisqu'aucune protéine de taille comparable à celle des protéines P26AUG2 et 3 n'est détectée (Figure 3.7 B, colonnes 9 et 10). Enfin, pour nous assurer de l'absence de mutants de réversion, nous avons effectué une analyse de la descendance des réplicons utilisés par RT-PCR et séquençage de l'amplicon. Aucun révertant ni pseudo révertant n'a pu être détecté au voisinage des 300 premiers nucléotides de la séquence codante.

3.5 Construction du clone d'ADNc complet et infectieux de l'ARN-5

L'obtention d'un clone ADNc complet et infectieux de l'ARN-5 s'est révélé primordial dans l'étude de l'expression de la protéine P26 lors de l'infection virale.

L'obtention de ce clone ADNc complet a été réalisée à partir d'ARN totaux de betteraves infectées par la méthode décrite précédemment (paragraphe 3.3.2, page 56). Cet ADNc a été obtenu par les techniques de 5' et 3' RACE décrites dans le chapitre Matériel et Méthodes (voir page 83). Les parties 5' et 3' des ADNc correspondant à l'ARN-5 ont été rassemblées dans un vecteur de clonage de type pUC sous le contrôle du promoteur du bactériophage T₇. Un alignement avec les séquences connues dans les bases de données nous a montré que le clone obtenu correspond bien à la séquence d'un ARN-5 de type A/P. Ce clone possédait une séquence polyadénylée courte et des nucléotides extra viraux provenant de la séquence adaptatrice incorporée durant l'étape de 3' RACE. Les transcrits correspondant à ce clone n'étaient pas infectieux en présence de la souche helper Stras12. Nous avons éliminé les nucléotides extra viraux et allongé la taille de la séquence polyadénylée par mutagenèse dirigée. Le clone obtenu nous a permis d'obtenir des transcrits infectieux. En effet, comme le montre la Figure 3.8 A, les transcrits inoculés en présence des ARN-1, -2 et -3 du BNYVV, de souche B, sont capables d'induire une nécrose caractéristique des lésions

locales sur *C. quinoa* (5 dpi) comme décrit précédemment. Nous avons purifié les ARN totaux (extraction Trizol) et les ARN encapsidés (extraction TM) de ces lésions locales pour les analyser par northern-blot. Ces extractions permettent de discriminer entre ARN viraux nus et encapsidés (Trizol) et ARN encapsidés (TM). Les ARN encapsidés sont protégés par la capsid virale de l'action des nucléases lors de l'incubation du broyat pendant 30 min à 37°C (Jupin *et al.*, 1990; Gilmer *et al.*, 1992b). Le principe de cette analyse différentielle est représenté sur la Figure 3.8 B. Nous avons ainsi montré que les transcrits de l'ARN-5 issus du clone infectieux pB55 étaient répliqués (Figure 3.8 C, Trizol) et encapsidés (Figure 3.8 C, TM) car ils ne sont pas dégradés lors de l'extraction TM. Une autre conclusion peut être tirée de cette analyse. La souche helper utilisée dans cette expérience provient de l'isolat B et prend en charge l'amplification et l'encapsidation de l'ARN-5. Ainsi, nous démontrons que des isolats B/P peuvent exister (voir Figure 1 de l'article en annexe page 113).

L'étape suivante a consisté à rechercher les phases de lecture ouvertes exprimées à partir de l'ARN-5 complet (numéro d'accession AY823407) dont la séquence est présentée Figure 3.9. Les régions codantes potentielles de l'ARN-5 sont présentées Figure 3.10. La séquence codante de la protéine P26 semble être la seule protéine pouvant être produite par cet ARN. La présence de 3 petites phases ouvertes de lecture dans la partie 5' non codante de l'ARN-5 pourraient correspondre à des éléments régulateurs visant à réduire le niveau d'expression de la séquence codante principale située en aval (Kozak, 2002). L'importance de ces trois séquences n'a pas été directement abordée dans cette étude.

3.6 Construction d'un nouveau vecteur d'expression viral dérivé de l'ARN-5

Afin de compléter les outils moléculaires nécessaires à l'étude de l'ARN-5, j'ai construit un vecteur d'expression viral à partir du clone complet pB55. Ce vecteur a été produit grâce à une mutagénèse dirigée par délétion (voir chapitre Matériel et Méthodes, paragraphe 5.2.7, page 88). Ainsi, la séquence codante pour la protéine P26 a été remplacée par une cassette de clonage contenant un site de restriction *NcoI* apportant le codon d'initiation

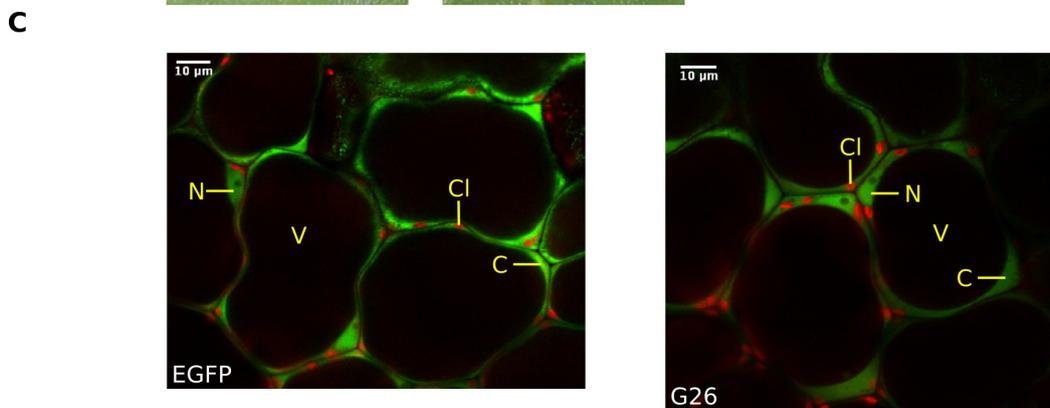
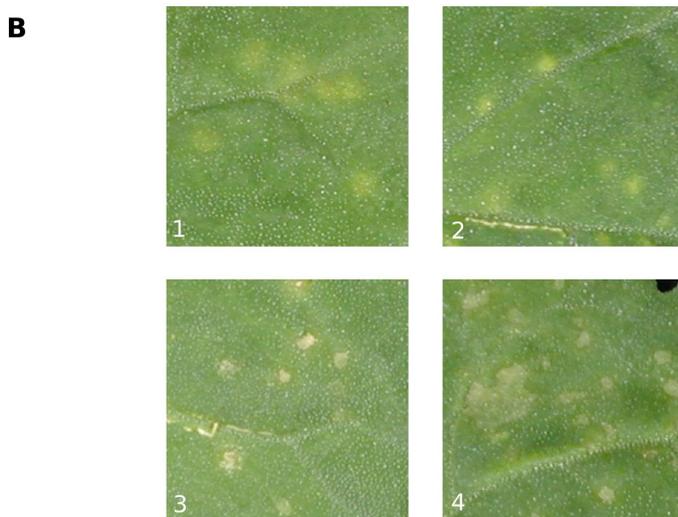
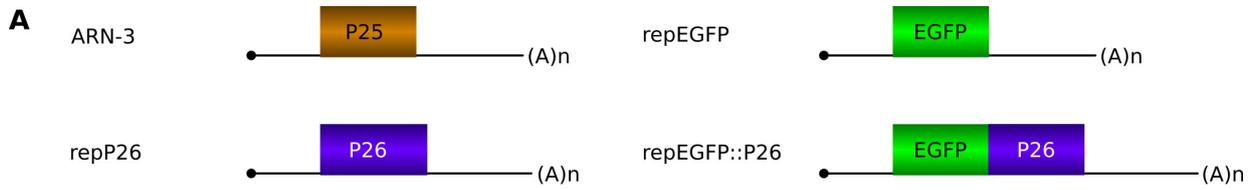


Figure 3.11 : Caractérisation des propriétés biologiques de la protéine de fusion EGFP::P26 exprimée en contexte viral. A) Représentation schématique de l'ARN-3 et des réplicons dérivés de cet ARN permettant l'expression de la protéine EGFP, P26 et EGFP::P26. B) Symptômes induits 6 jours après infection de feuilles de *C. quinoa* par une souche Stras12 (1) supplémentée de transcrits rep-EGFP (2), rep-P26 (3) et rep-EGFP::P26 (4). C) Comparaison de la localisation sub-cellulaire de la protéine EGFP::P26 à celle de la protéine EGFP dans des cellules de *C. quinoa* infectées. Les annotations jaunes désignent les compartiments sub-cellulaires visibles sur ces images : N : noyau ; C : cytoplasme ; V : vacuole ; Cl : chloroplaste. La barre blanche représente 10 μm.

AUG, trois codons glycine suivis des sites de restriction *EcoRV*, *BglIII* et *XbaI* puis un codon stop, en phase avec le codon d'initiation.

L'obtention et la caractérisation de ce vecteur d'expression fait l'objet d'une publication acceptée dans "Journal of General Virology" (voir annexes au chapitre 3, page 113). Cet article décrit la construction du réplicon appelé Rep-5 et son utilisation pour exprimer des protéines hétérologues (l'EGFP par exemple) à un taux environ cinq fois plus faible que celui du réplicon dérivé de l'ARN-3 (Figure 1 de la publication). Cet outil Rep-5 permet également de produire des protéines d'intérêt dans *B. macrocarpa*, une plante systémique, et ceci en présence de l'ARN-3 (Figure 2 A à D de la publication). Finalement, l'utilisation à la fois du réplicon dérivé de l'ARN-3 et de l'ARN-5 permet d'exprimer deux protéines distinctes dans la même cellule (Figure 2 E de la publication).

3.7 Localisation sub-cellulaire de la protéine P26

3.7.1 Expression de la protéine en contexte viral

La connaissance de la localisation sub-cellulaire d'une protéine peut apporter certains éléments de réponse quant à sa fonction dans la cellule. Ainsi, nous avons entrepris de caractériser celle de la protéine P26 en utilisant deux approches, l'une basée sur la construction d'une protéine de fusion EGFP::P26, l'autre basée sur l'utilisation de la protéine P26 étiquetée HA. La première approche, faisant appel à la protéine de fusion EGFP::P26, a été réalisée à la fois en contexte viral et dans des expériences d'expression transitoire. La seconde approche n'a été abordée qu'en contexte viral.

Expression de la protéine EGFP::P26 en contexte viral et caractérisation de mutants de délétion

Nous avons réalisé la fusion de la protéine P26 à l'EGFP dans le réplicon dérivé de l'ARN-3 pour obtenir le réplicon Rep-EGFP::P26 (dénommé G26 par la suite) représenté Figure 3.11 A. Les transcrits infectieux issus de ce réplicon ont été inoculés en présence de la souche helper Stras12 à des feuilles de *C. quinoa* et les symptômes comparés à ceux

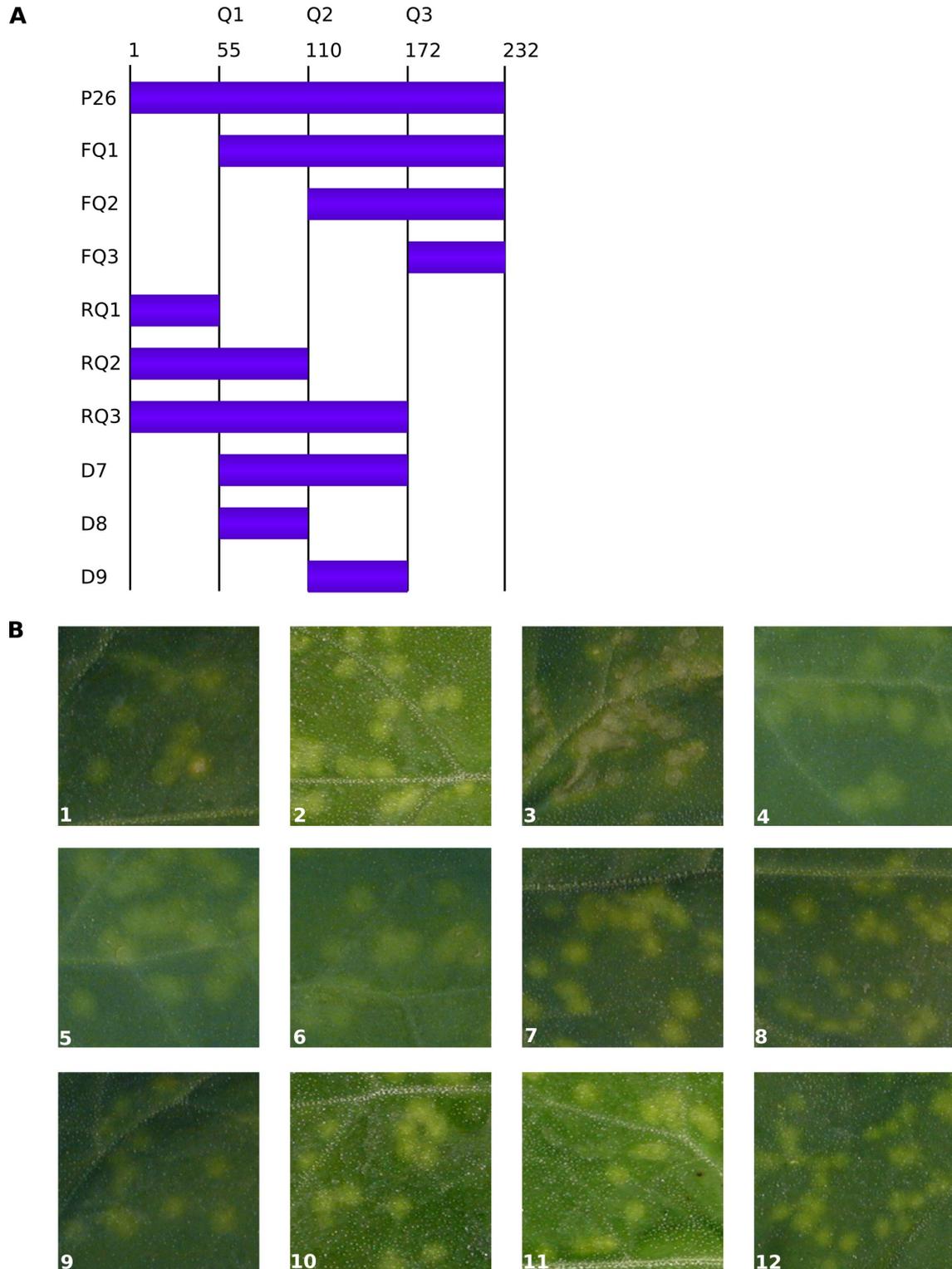


Figure 3.12 : Propriétés biologiques de la protéine P26 et de mutants de délétion sur feuilles de *C. quinoa*. A) Représentation schématique des mutants de délétion de la protéine P26. Chaque fragment peptidique (rectangle violet) est inséré dans le vecteur d'expression approprié, en fusion ou non avec la EGFP. B) Symptomatologie induite par inoculation de Stras12 seul (1), Stras12 + rep-EGFP (2), Stras12 + rep-EGFP::P26 (3), Stras12 + rep-EGFP::FQ1 (4), Stras12 + rep-EGFP::FQ2 (5), Stras12 + rep-EGFP::FQ3 (6), Stras12 + rep-EGFP::RQ1 (7), Stras12 + rep-EGFP::RQ2 (8), Stras12 + rep-EGFP::RQ3 (9), Stras12 + rep-EGFP::D7 (10), Stras12 + rep-EGFP::D8 (11) et Stras12 + rep-EGFP::D9 (12).

induits par un réplicon exprimant la protéine EGFP seule ou la protéine P26 seule (Figure 3.11 B). Comme le montre la Figure 3.11 B 4, la protéine EGFP::P26 reste capable d'induire des nécroses comparables à celles obtenues lors de l'expression du gène codant pour la protéine P26 non fusionnée (Figure 3.11 B 3). Cette nécrose n'est pas due à l'expression de la protéine EGFP qui n'induit pas de profonds changements de symptômes (comparer Figure 3.11 B 2 et 1). L'observation par microscopie confocale de cellules infectées au sein de ces lésions a permis de localiser la protéine G26 à la fois dans le noyau et dans le cytoplasme des cellules infectées, situation également trouvée avec la protéine EGFP non fusionnée (comparer les Figures 3.11 C, EGFP et G26). Si la localisation nucléaire de la protéine EGFP est due à une diffusion passive de la protéine dans le noyau (Baulcombe *et al.*, 1995; Oparka *et al.*, 1996; Cruz *et al.*, 1996), la localisation nucléaire de la protéine de fusion G26, d'un poids moléculaire de 53 kDa, tend à exclure une diffusion passive de la protéine à travers le NPC. Cette protéine de fusion G26 pourrait donc être recrutée comme cargo dans le cytoplasme, son lieu de synthèse, pour être acheminée dans le nucléoplasme par des facteurs d'import. L'absence de séquence NLS caractéristique sur la protéine P26 nous a conduit à effectuer l'étude de mutants de délétion pour rechercher le domaine de la protéine impliqué dans son adressage nucléaire.

Caractérisation des propriétés biologiques de mutants de délétion de la protéine P26

Nous avons construit neuf mutants de délétion dérivés de la séquence de la protéine P26, exprimés dans le Rep-3. Les bornes des délétions sont schématisées sur la Figure 3.9 (Q1 à Q3). Trois délétions récurrentes ont été effectuées à partir de l'extrémité N-terminale (FQ1 à FQ3), de l'extrémité C-terminale (RQ1 à RQ3) et des domaines internes D7 à D9 situées entre les bornes Q1-Q3, Q1-Q2 et Q2-Q3 respectivement (Figures 3.12 A et Figure 3.9). L'ensemble de ces formes délétées de la protéine P26 a également été placé en fusion avec l'EGFP pour produire les constructions GFQ1 à GD9. Les transcrits correspondant à ces mutants de délétion fusionnés ou non à l'EGFP, ont été inoculés sur feuille de *C. quinoa* en présence de la souche helper Stras12 et les symptômes observés 5 dpi. La Figure 3.12 B représente les résultats obtenus après infection et expression des protéines de fusion

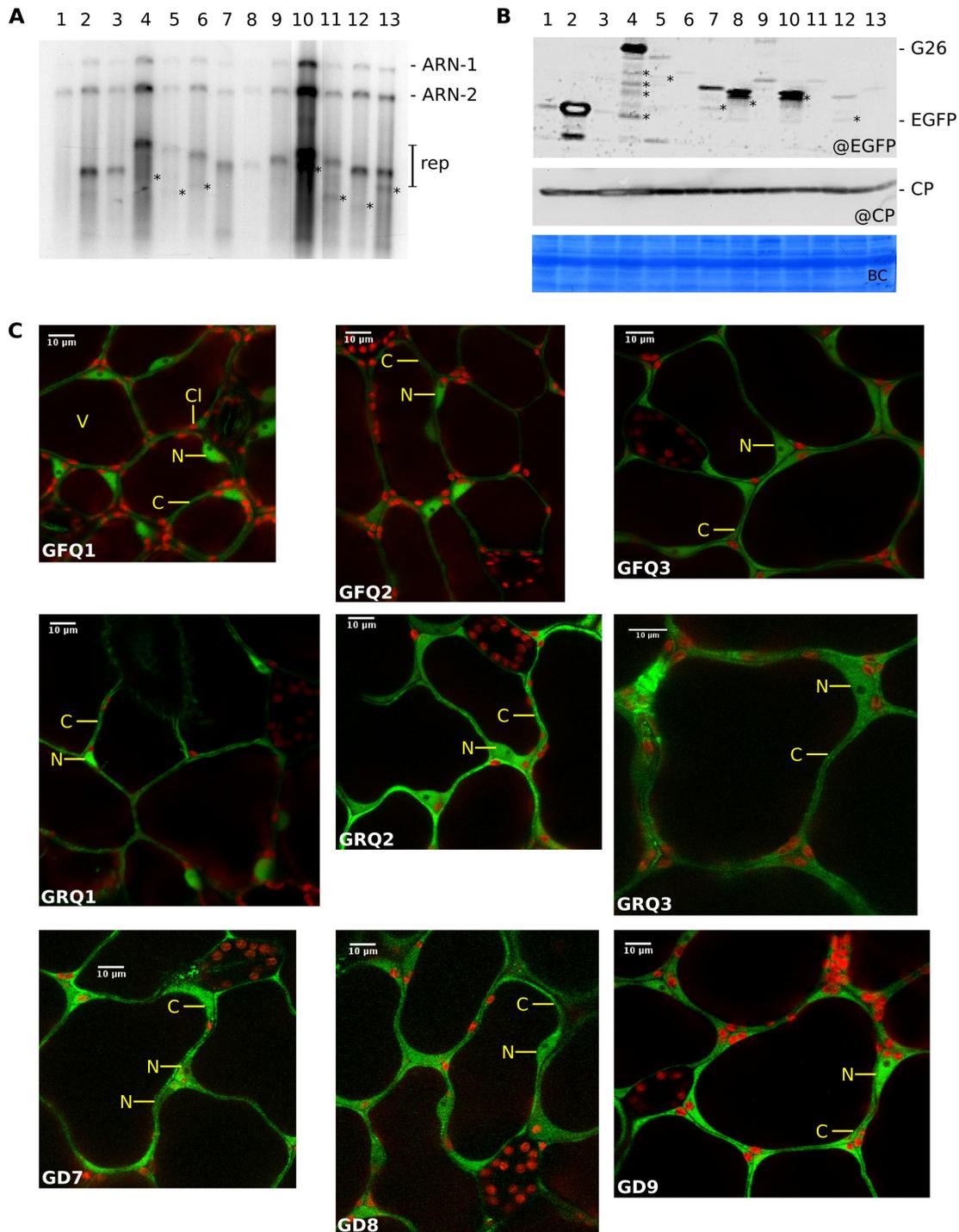


Figure 3.13 : Analyse de la replication des ARN recombinants exprimant les protéines de fusion EGFP::P26 tronquées (A), de l'accumulation des protéines de fusion (B) dans les lésions de feuilles de *C. quinoa* inocuées par la souche helper Stras12 (1) supplémentée des transcrits rep-EGFP (2), rep-P26 (3), rep-EGFP::P26 (4), rep-EGFP::FQ1 (5), rep-EGFP::FQ2 (6), rep-EGFP::FQ3 (7), rep-EGFP::RQ1 (8), rep-EGFP::RQ2 (9), rep-EGFP::RQ3 (10), rep-EGFP::D7 (11), rep-EGFP::D8 (12) et rep-EGFP::D9 (13), et de la localisation sub-cellulaire des protéines recombinantes (C). A) Analyse par northern-blot de l'accumulation des ARN viraux 5 dpi de feuilles de *C. quinoa* inocuées par la souche helper Stras12 supplémentée des transcrits des réplicons exprimant les protéines EGFP::P26 recombinantes. Les étoiles indiquent les positions des formes délétées des réplicons. B) Analyse par western-blot de la présence des protéines de fusion EGFP::P26 tronquées (haut), de l'accumulation de la protéine de capside (CP) et des protéines totales transférées sur membrane (BC). Les protéines EGFP::P26 tronquées apparaissent au PM attendu, mais on constate la présence de formes délétées, plus courtes, pouvant interférer sur l'interprétation de la localisation sub-cellulaire des protéines fluorescentes en contexte viral. @HA : immunodétection de l'épitope HA ; @CP immunodétection de la protéine CP ; BC : coloration au Serva Blue. C) Recherche par microscopie confocale de la localisation sub-cellulaire des protéines EGFP::P26 tronquées exprimées en contexte viral. Sur ces images, la EGFP fluoresce en vert et les chloroplastes en rouge. Les annotations jaunes désignent les compartiments sub-cellulaires visibles sur ces images : N : noyau ; C : cytoplasme ; V : vacuole ; Cl : chloroplaste. La barre blanche représente 10µm.

EGFP::P26 et ses dérivés. Si la protéine G26 est capable de produire une nécrose des lésions locales (Figure 3.12 B 3), aucun des mutants de délétion ne reproduit de phénotype nécrogène (Figure 3.12 B 4 à 12). Ces mutants induisent l'apparition de lésions chlorotiques comparables à celles obtenues en présence d'ARN-1 et -2, seuls ou supplémentés de Rep-EGFP (Figure 3.12 B, 1 et 2, respectivement). La même expérience effectuée avec les protéines non fusionnées à la EGFP nous a permis d'observer des symptômes foliaires identiques, c'est à dire une perte du pouvoir nécrogène des protéines tronquées (non montré). Ces résultats indiquent donc que l'ensemble de la protéine est requis pour l'induction des nécroses et tend à impliquer la structure tridimensionnelle de la protéine plutôt qu'un domaine protéique particulier.

Nous avons recherché la localisation sub-cellulaire des protéines de fusion EGFP::P26 tronquées, exprimées en contexte viral par microscopie confocale mais aussi caractérisé l'accumulation des réplicons par northern-blotting et le niveau d'expression des protéines de fusion par western-blotting, à l'aide d'un anticorps dirigé contre la protéine EGFP. L'ensemble de ces résultats est présenté dans la Figure 3.13. Ainsi, l'analyse de la réplication montre que l'accumulation des réplicons exprimant les protéines EGFP::P26 tronquées est très variable selon les transcrits inoculés en présence de souche Stras12. De plus, il apparaît au sein de la descendance des réplicons dérivés de l'ARN-3, des formes délétées (Figure 3.13 A, étoiles) et décrites précédemment par Bouzoubaa *et al.* (1991). L'analyse des protéines de fusion EGFP::P26 tronquées à l'aide d'un anticorps dirigé contre la protéine EGFP révèle la même disparité dans les taux d'expression. En effet, à charge de protéines constante (Figure 3.13 B, panneau BC) et infection virale comparable (Figure 3.13 B, panneau CP), l'accumulation des protéines EGFP, fusionnées ou pas aux différents domaines de la protéine P26, est détectée de façon très variable. Si l'on retrouve les protéines de fusion aux masses moléculaires attendues, on constate également la présence de formes tronquées de ces protéines, réagissant toujours avec l'anticorps dirigé contre la EGFP et qui pourraient correspondre à des protéines fluorescentes. De telles formes tronquées ne nous permettent pas d'interpréter avec confiance les résultats de localisation sub-cellulaires présentés sur la Figure 3.13 C et nous avons choisi d'adopter une approche biolistique pour exprimer transitoirement la protéine de fusion G26 et ses dérivés par l'intermédiaire du promoteur fort

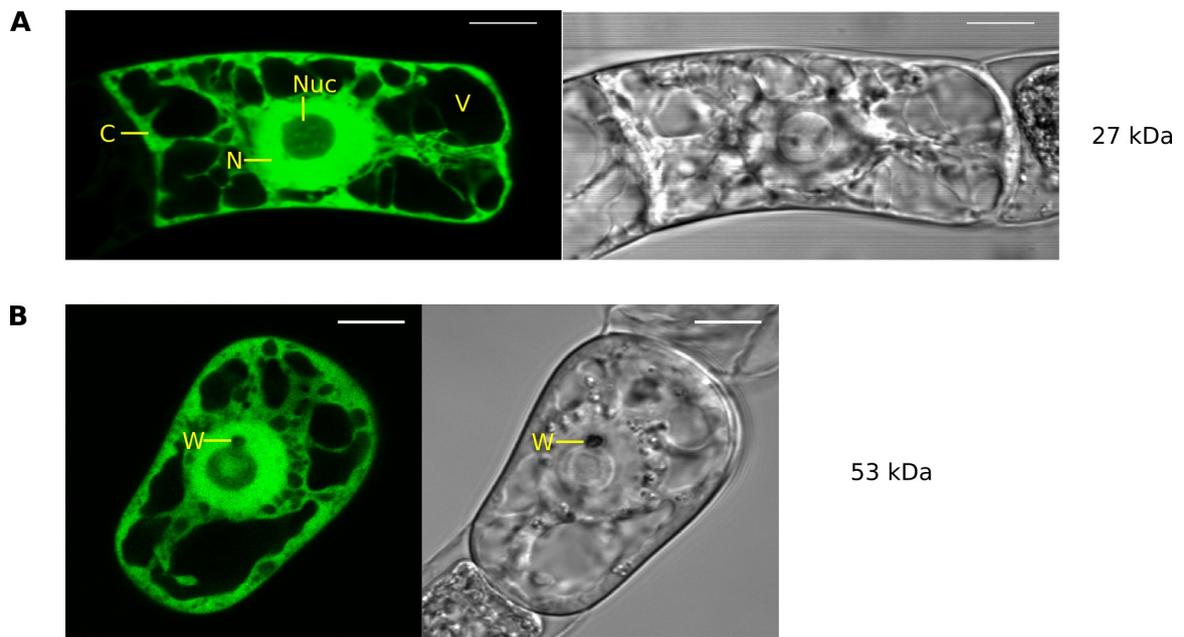


Figure 3.14 : Analyse par microscopie confocale de la localisation sub-cellulaire des protéines EGFP (A) et EGFP::P26 (B) exprimées transitoirement dans des cellules BY-2 de tabac. Les annotations jaunes permettent de localiser les compartiments sub-cellulaires : N : noyau ; C : cytoplasme ; V : vacuole ; Nuc : nucléole ; W : particule de tungstène. L'émission de fluorescence verte émise par la EGFP est présentée dans la partie gauche de la figure ; à droite, les mêmes images obtenues par contraste différentiel-interférentiel montrant les structures cellulaires. Les barres blanches correspondent à 10 μm . La masse moléculaire de la protéine fluorescente exprimée est précisée à droite des images.

35S. Ainsi, l'ensemble des séquences codant pour les protéines EGFP::P26 recombinantes a été placé, sous contrôle du promoteur 35S, dans le vecteur pCK entre la séquence activatrice de traduction Ω du TMV et le terminateur 35S. Un tel système permet la transcription *in vivo* des ARN messagers à partir de la matrice plasmidique, et ainsi de s'affranchir de la réplication virale, des erreurs de recopie et de l'apparition de formes délétées, amplifiées potentiellement lors des cycles multiples d'infection conduisant à la formation des lésions locales.

3.7.2 Expression transitoire en cellules BY-2

La lignée cellulaire de *Nicotiana tabacum* BY-2 est un outil de choix pour l'étude de la localisation sub-cellulaire de protéines *in vivo*. En effet, cette lignée cellulaire est constituée de grandes cellules facilement observables en microscopie optique et la technique de transfection est parfaitement reproductible. Cette technique de transfection, utilisée au laboratoire, m'a permis d'initier l'étude de la localisation sub-cellulaire de la protéine de fusion EGFP::P26 hors contexte viral. Chaque transfection a été répétée trois fois et une quinzaine de cellules fluorescentes ont été observées pour chaque construction.

Observations de la localisation sub-cellulaire de la protéine EGFP::P26

L'analyse de la localisation sub-cellulaire par microscopie confocale de la protéine EGFP::P26 (G26) est présentée sur la Figure 3.14 et comparée à celle de la protéine EGFP seule. Ces deux protéines sont exprimées à partir de plasmides pCK recombinants délivrés par biolistique dans les cellules BY-2. Dans cette expérience, la protéine EGFP, de 27 kDa, diffuse passivement dans le compartiment nucléaire, au centre de la cellule. Ainsi, le noyau fluoresce en vert, tout comme les travées cytoplasmiques où la protéine s'accumule. Par contre, les vacuoles et le nucléole, qui ne fluorescent pas, ne contiennent pas d'EGFP.

L'expression de la protéine G26 est également détectée et la protéine de fusion est localisée de manière identique à la protéine EGFP. Toutefois, comme nous l'avons décrit au paragraphe précédent, son poids moléculaire de 53 kDa n'est pas compatible avec une diffusion passive à travers les pores nucléaires. On peut noter également sur cette image un plus grand contraste du marquage nucléaire/cytoplasmique pour la protéine G26 que celui

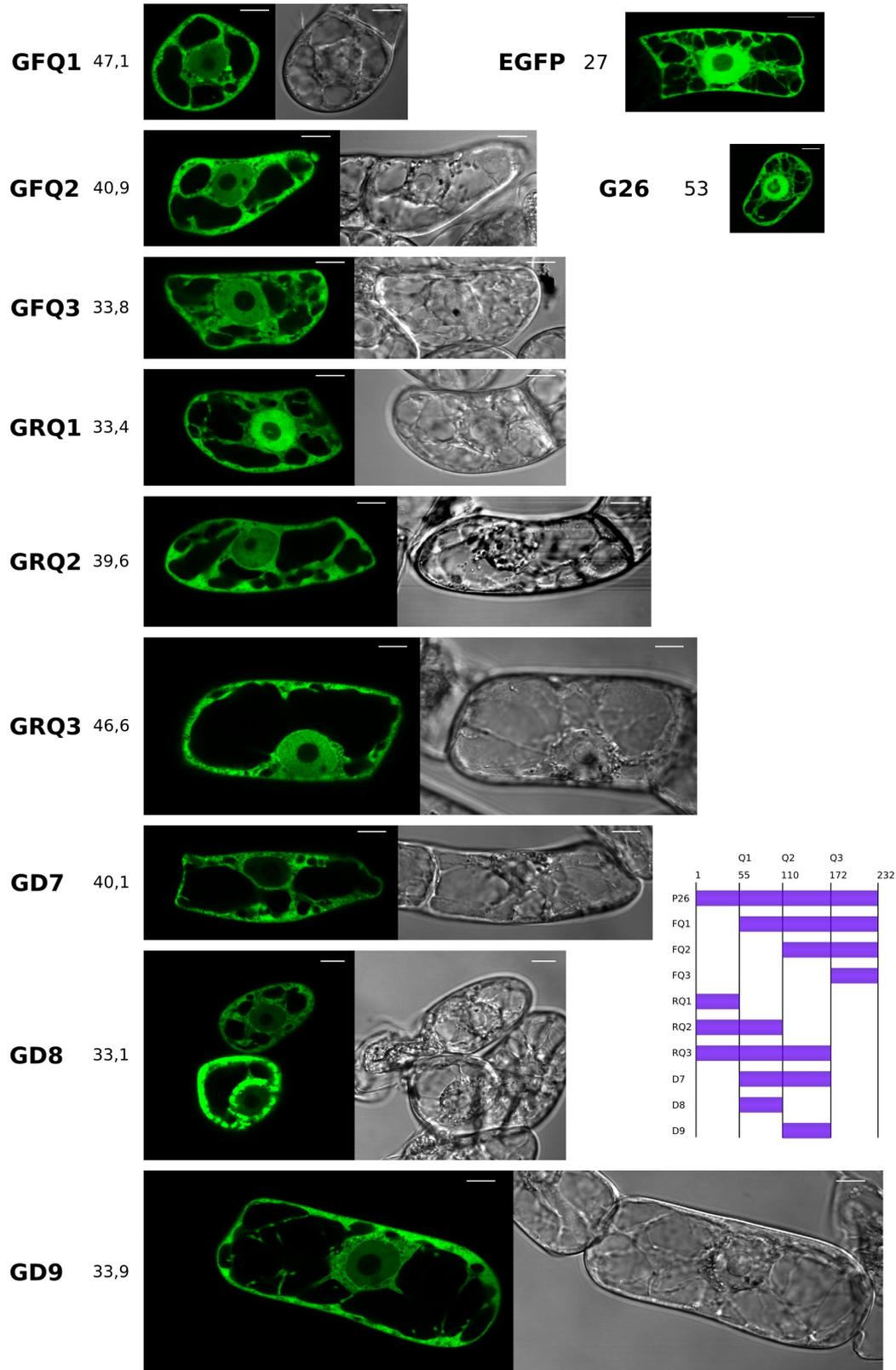


Figure 3.15 : Analyse par microscopie confocale de la localisation sub-cellulaire des protéines EGFP::P26 tronquées, exprimées transitoirement dans des cellules BY-2 de tabac. GFQ1) EGFP::FQ1 ; GFQ2) EGFP::FQ2 ; GFQ3) EGFP::FQ3 ; GRQ1) EGFP::RQ1 ; GRQ2) EGFP::RQ2 ; GRQ3) EGFP::RQ3 ; GD7) EGFP::D7 ; GD8) EGFP::D8 ; GD9) EGFP::D9 ; EGFP) EGFP ; G26) EGFP::P26. Les chiffres indiqués à droite des noms des constructions représentent la masse moléculaire des protéines de fusion. L'émission de fluorescence verte émise par la EGFP est présentée dans la partie gauche des images, à droite, le même champ obtenu par contraste différentiel-interférentiel montrant les structures cellulaires. Les barres blanches correspondent à 10µm.

observé dans le cas de la protéine EGFP seule. Cette observation conforte notre hypothèse de transport actif de la protéine P26 dans le noyau des cellules transfectées.

Analyse de la localisation sub-cellulaire de mutants de délétions de la protéine P26

L'ensemble des mutants de délétion de la protéine P26 fusionnés à la EGFP a été délivré par biolistique dans des cellules BY-2 et la localisation des protéines de fusion déterminée par microscopie confocale. Les résultats obtenus et caractéristiques de chaque protéine de fusion sont résumés sur la Figure 3.15. Grâce à cette approche, nous avons pu obtenir des résultats reproductibles qui nous ont permis de localiser le domaine d'adressage nucléaire dans la partie N-terminale de la protéine P26. En effet, la protéine G26, adressée dans le nucléoplasme, perd sa localisation nucléaire préférentielle lorsque le domaine N-terminal est délété (Figure 3.15, GFQ1, GFQ2 et GFQ3). De plus, le domaine N-terminal fusionné à la EGFP adresse la protéine de fusion préférentiellement dans le noyau des cellules transfectées (Figure 3.15, GRQ1). L'absence de localisation nucléaire des protéines GRQ2 et GRQ3, qui possèdent le domaine N-terminal, peut s'expliquer par la présence d'une séquence de rétention cytoplasmique ou d'export nucléaire exportant la protéine de fusion du noyau vers le cytoplasme, comme dans le cas de la protéine EGFP::P25 (Vetter *et al.*, 2004). En effet, la fusion du domaine interne de la protéine P26 (D7, D8 et D9) à la protéine EGFP ne localisent pas ou peu la fluorescence dans le compartiment nucléaire, particulièrement dans le cas des protéines GD8 et GD9 dont la masse moléculaire est inférieure à la limite de diffusion des NPC. L'effet de cette rétention cytoplasmique est également observable dans le cas de la protéine de fusion GFQ3 qui reste majoritairement cytoplasmique. Dans le cas de la protéine GRQ3, celle-ci s'accumule légèrement dans le noyau des cellules transfectées.

Au vu de ces localisations sub-cellulaires, la présence d'une séquence d'adressage nucléaire dans la partie N-terminale de la protéine P26 semble être confortée mais son activité pourrait être influencée par la présence d'éléments de rétention cytoplasmiques, d'un signal d'export nucléaire ou par une conformation tridimensionnelle de la protéine nécessitant l'intégralité de la séquence peptidique pour son repliement.

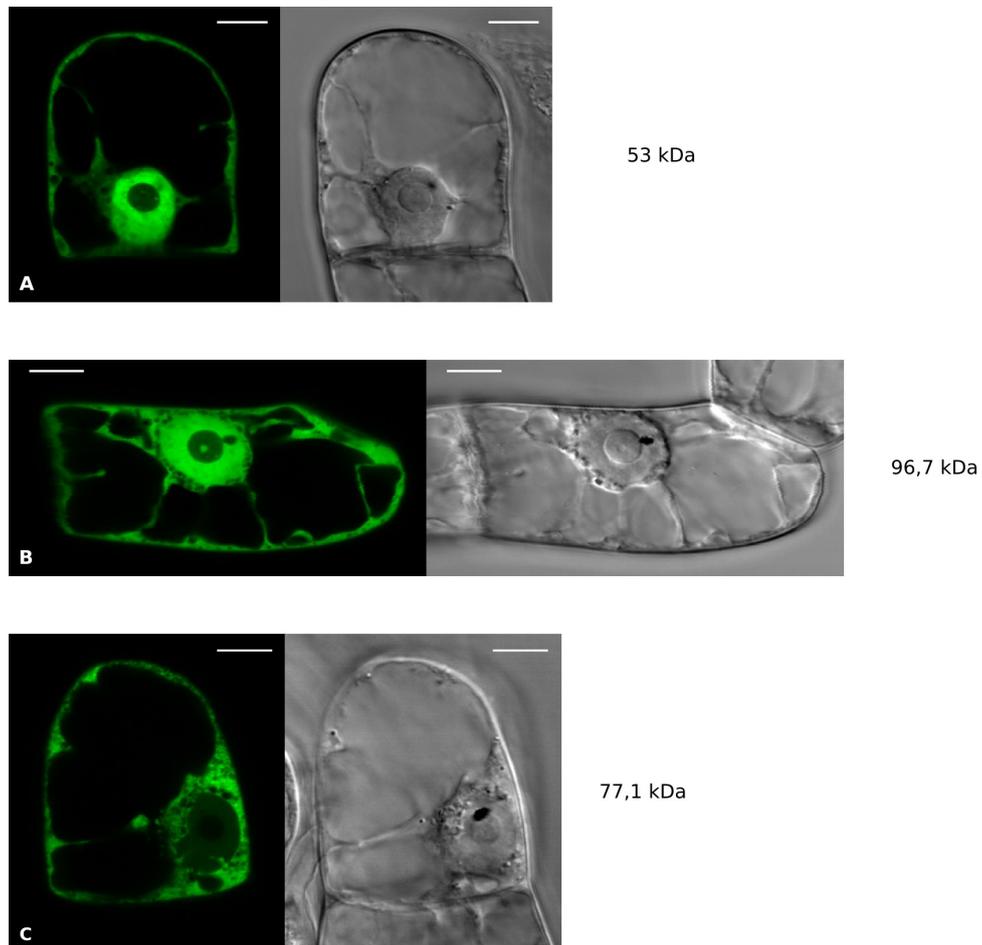


Figure 3.16 : Analyse par microscopie confocale de la localisation sub-cellulaire des protéines EGFP::P26 et EGFP::RQ1 fusionnées à la chalcone synthase (CHS), exprimées transitoirement dans des cellules BY-2 de tabac. Observations des constructions : A) EGFP::P26 ; B) EGFP::P26::CHS et C) EGFP::RQ1::CHS. L'émission de fluorescence verte émise par la EGFP est présentée dans la partie gauche des images, à droite, le même champ obtenu par contraste différentiel-interférentiel montrant les structures cellulaires. La barre blanche représente 10 μm. La masse moléculaire des protéines de fusion est précisée à droite de la figure.

Nous avons donc choisi d'augmenter la taille des protéines G26 et GRQ1 en fusionnant le gène de la chalcone synthase (CHS) à l'extrémité 3' de la séquence de la protéine P26 pour obtenir les constructions EGFP::P26::CHS et EGFP::RQ1::CHS.

Analyse de la localisation des protéines de fusion EGFP::P26::CHS et EGFP::RQ1::CHS

L'analyse de l'importation nucléaire de protéines de haut poids moléculaires permet de s'affranchir de la diffusion passive dans le noyau de protéines dont la masse moléculaire avoisine la limite de diffusion des NPC.

Comme précédemment, des cellules BY-2 ont été transfectées par les constructions recombinantes EGFP::P26::CHS et EGFP::RQ1::CHS et observées en microscopie confocale. Les résultats obtenus sont présentés sur la Figure 3.16. Comme auparavant lors de cette étude, nous avons observé la localisation nucléaire préférentielle de la protéine G26 (Figure 3.16 A). La fusion de la séquence de la chalcone synthase en C-terminal de la protéine P26 conduit à la synthèse d'une protéine de 96,7 kDa (Figure 3.16 B) qui reste capable de franchir la barrière nucléaire et de s'accumuler dans le nucléoplasme. La même fusion réalisée avec le domaine RQ1 induit la synthèse d'une protéine de 77,1 kDa qui n'est plus localisée dans le noyau (Figure 3.16 C). Ce résultat nous conforte donc sur la présence d'une séquence d'adressage nucléaire sur la séquence de la protéine P26 mais nous indique que la localisation nucléaire de la protéine GRQ1 est probablement due à une diffusion passive de cette protéine de 33 kDa dans le noyau.

Au vu de l'ensemble de ces données, la localisation nucléaire de la protéine P26 semble nécessiter un signal d'adressage qui pourrait être constitué de deux domaines distincts dont l'un serait présent à l'extrémité N-terminale. Une autre possibilité résiderait dans un adressage nucléaire de la protéine nécessitant l'intervention d'une protéine cellulaire qui reconnaîtrait la structure tridimensionnelle de la protéine P26, et servirait ainsi d'adaptateur pour son transfert vers le noyau.

Il nous faudra cependant déterminer si la protéine P26 possède une séquence d'export nucléaire, tout comme la protéine P25 du même virus (Vetter *et al.*, 2004). Pour cela, nous utiliserons la toxine Leptomycine B (LMB) qui inhibe l'export nucléaire en alkylant spéci-

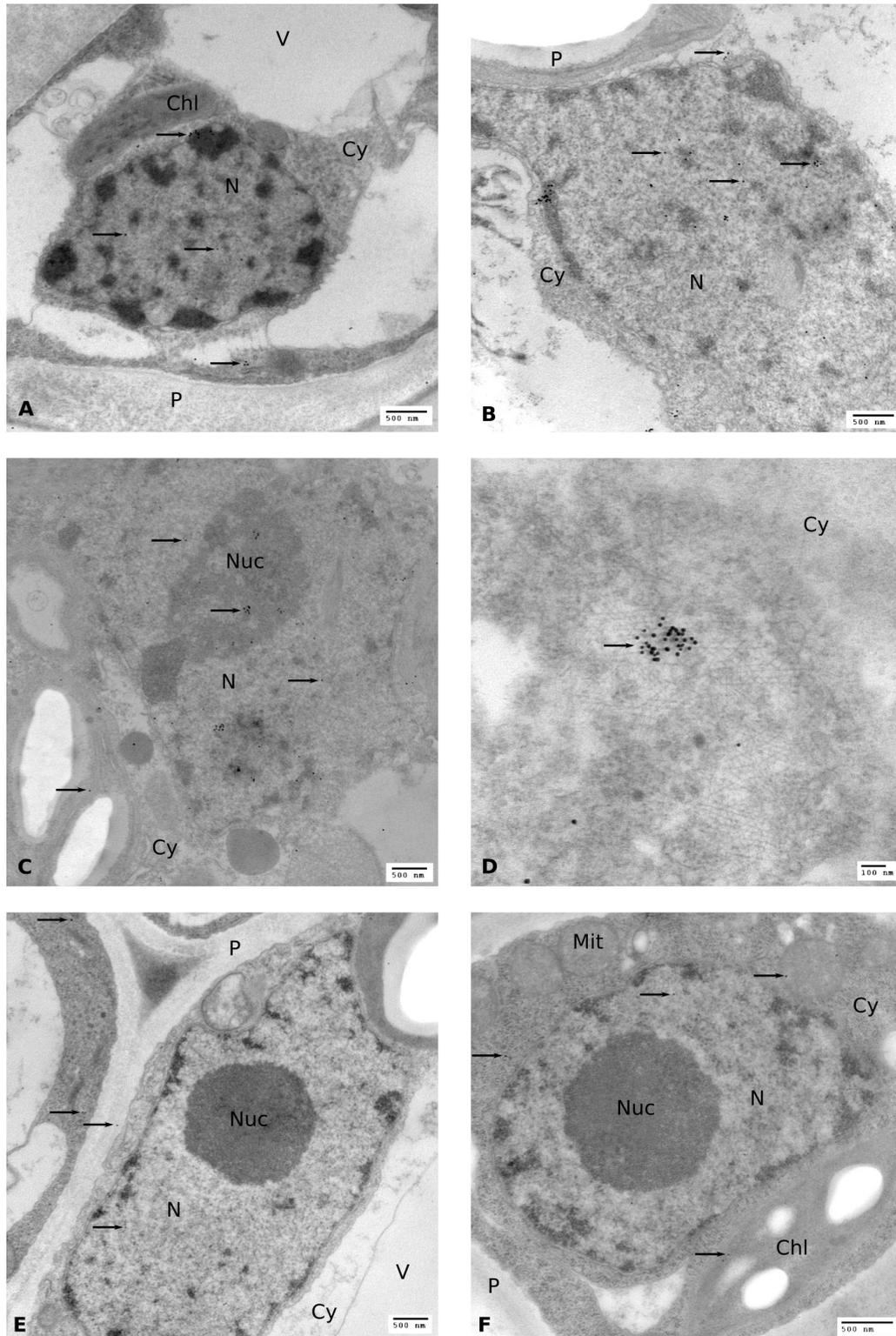


Figure 3.17 : Recherche par microscopie électronique de la localisation sub-cellulaire de la protéine P26HA en contexte viral dans des cellules de feuilles de *C. quinoa* infectées par la souche helper Stras12 supplémentée par le rep-P26HA (A à D) ou par le rep-P26 (E et F) qui sert de témoin. Des grains d'or colloïdal de 15 nm sont désignés par les flèches. A et B sont ont été photographiés à un grossissement de 20000 \times , C et E de 170000 \times , D de 60000 \times et F de 25000 \times . A) , B), C), E) et F) sont des photographies d'une cellule, centrées sur le noyau. D) est une photographie d'amas de virions présents dans le cytoplasme. Les échelles de longueur sont indiquées en bas à droite de chaque image. N : noyau; Nuc : nucléole; Mit : mitochondrie; Chl : chloroplaste; Cy : cytoplasme; P : paroi pecto-cellulosique; V : vacuole.

fièrement l'exportine CRM1 (Kudo *et al.*, 1998), et provoque l'accumulation dans le noyau des protéines nucléaires normalement exportées. L'effet de la LMB sur les cellules BY-2 transfectées par les protéines de fusion GD7, 8, 9, GRQ2 et GFQ1 permettra ainsi de préciser la localisation de ce domaine d'export, s'il existe, ou de démontrer l'existence d'une rétention cytoplasmique de la protéine P26. Cette dernière situation est démontrée dans le cas des protéines NF- κ B, famille de facteurs de transcription intervenant dans la réponse immunitaire des mammifères, qui sont séquestrées dans le cytoplasme par les protéines I κ B, les empêchant d'activer leurs gènes cibles dans le noyau (Li and Verma, 2002). L'élimination des domaines internes D7, D8 et D9 des protéines de fusion devra être également réalisée sur la construction EGFP::P26::CHS pour rechercher un éventuel effet synergique des séquences N et C-terminales.

3.7.3 Immunolocalisation de la protéine P26

Pour apporter une preuve supplémentaire de l'adressage nucléaire de la protéine P26, nous avons recherché cette protéine exprimée en contexte viral sous sa forme étiquetée HA (construction Rep-P26HA) sur des coupes ultrafines de lésions locales de feuilles de *C. quinoa*. Ces lésions ont été fixées et incluses en résine (voir Matériel et Méthodes) coupées et immunomarquées à l'aide d'un anticorps monoclonal de souris dirigé contre l'épitope HA. La révélation des complexes antigènes/anticorps a été réalisée à l'aide d'immunoglobulines complexées à des billes d'or colloïdal de 15 nm et dirigées contre les anticorps de souris.

Comme témoin, nous avons choisi d'utiliser des lésions locales dans lesquelles la protéine P26 est exprimée sous forme non étiquetée (Rep-P26). Un résultat représentatif des immunomarquages obtenus est présenté sur la Figure 3.17. L'analyse statistique effectuée sur 12 échantillons de chaque type d'infection nous a permis d'établir des proportions de localisation nucléaire et cytoplasmique de la protéine P26HA et du témoin. Si 26 % des grains d'or sont retrouvés dans le noyau des cellules témoin, nous observons 76 % de marquage nucléaire des cellules exprimant la protéines P26HA. Ce résultat confirme bien que la protéine P26 possède une localisation nucléaire. Dans de rares cas, nous avons pu observer la présence de grains d'or colloïdal sur des agrégats de virions dans le cytoplasme des cellules (Figure 3.17 D) pouvant suggérer une interaction directe ou indirecte de la protéine P26HA

avec les protéines majeure ou mineure de capsid. Nous avons donc recherché l'existence d'interactions directes en utilisant le système double hybride de levure.

3.8 Recherche de protéines interagissant avec la protéine P26 : utilisation du système double hybride de levure

La caractérisation des partenaires cellulaires et viraux qui interagissent avec la protéine P26 permettrait de préciser son rôle dans le cycle viral. Comme nous l'avons vu précédemment (paragraphe 2.3 page 35), nous avons choisi de rechercher les partenaires cellulaires et viraux de la protéine P26 par la technique du double-hybride.

3.8.1 Limites

La séquence de la protéine P26 a été fusionnée en phase en 3' de la séquence codant pour le domaine de fixation à l'ADN du facteur de transcription GAL4 (GAL4BD) et le vecteur recombinant obtenu (pGBT9-P26) introduit dans la souche de levure HF7C. Les clones obtenus, sélectionnés sur milieu SD dépourvu de tryptophane (SD-T) ont ensuite été repiqués sur milieu SD-T et milieu SD dépourvu de tryptophane et d'histidine (SD-TH) pour rechercher la capacité de la protéine P26 à autoactiver la transcription du gène rapporteur *His3*. Les résultats obtenus, présentés sur la Figure 3.18 A, indiquent que la protéine de fusion GAL4BD::P26 est capable d'activer la transcription du gène *His3* qui rend ainsi les levures prototrophes vis à vis de l'histidine. La souche HF7C transformée avec le vecteur exprimant le domaine GAL4BD seul n'est quant à elle pas capable de pousser sur ce milieu sélectif. De plus, le test d'activité β -galactosidase par "colony lift" nous a montré une coloration bleue aussi forte et apparaissant aussi rapidement que le contrôle positif utilisé au laboratoire. Ce contrôle est réalisé en exprimant la protéine p53 fusionnée à GAL4BD (pVA3) et TAg_{SV40} fusionné à GAL4AD (pTD1). Ces deux protéines interagissent *in vivo* (Li and Fields, 1993) et induisent la transcription du gène rapporteur *His3* ainsi que celle du gène de la β -galactosidase placé sous le contrôle d'un promoteur reconnu

par GAL4. L'activité β -galactosidase détectée dès 30 min dans la souche HF7C exprimant la protéine GAL4BD::P26 est identique à celle détectée pour les levures exprimant le facteur de transcription chimérique GAL4BD::p53/GAL4AD::TA_{gSV40}. Ainsi, l'utilisation de cette approche pour rechercher des facteurs cellulaires issus d'une banque d'ADNc (clonée dans le vecteur proie pGAD424) et interagissant avec la protéine P26 est vaine.

Cependant, un test identique effectué avec la protéine P26 clonée dans le vecteur pGAD424 en fusion avec le domaine d'activation de la transcription de GAL4 (GAL4AD::P26) ne permet pas la croissance des levures sur milieu sélectif (non montré). La recherche d'interactions entre des protéines virales, fusionnées au domaine GAL4BD, et la protéine P26, fusionnée au domaine GAL4AD, a été initiée par la technique de "mating" (voir Matériel et Méthodes, page 99).

3.8.2 Recherche d'interactions entre les protéines de capsid, la protéine P25 et la protéine P26

Les séquences codantes des protéines majeure et mineure de capsid et de la protéine P25 issue d'isolats A, B et A/P (voir chapitre 2) ont été clonées dans les vecteurs proie et appât (pGAD424 et pGBT9, voir matériel et méthodes) et introduits dans les souches de levure HF7C (vecteur pGBT9) et Y187 (vecteur pGAD424). Après sélection des souches transformées, "mating", l'ensemble des levures diploïdes obtenues a été repiqué à la fois sur milieu SD-TL pour visualiser leur croissance et sur milieu SD-TLH pour visualiser leur prototrophie vis à vis de l'histidine, traduisant l'activation des gènes rapporteurs par le facteur de transcription chimérique, créé par l'interaction entre protéines de fusion.

Les résultats obtenus sont résumés sur la Figure 3.18 B. L'ensemble des levures diploïdes contenant les deux vecteurs est capable de croître sur milieu SD-TL, à l'exception des diploïdes obtenus avec les souches dépourvues de l'un des deux vecteurs (lignes et colonnes "-"). L'analyse de la croissance de ces diploïdes sur milieu SD-TLH révèle à nouveau la forte activation de la transcription par la protéine P26 fusionnée au domaine GAL4BD qui est insensible à la présence de l'autre vecteur, contrairement à la situation décrite pour la protéine P25 (chapitre 2). Nous retrouvons également dans cette expérience les interactions



Figure 3.19 : Analyse en système dit simple hybride de l'activation de la transcription induite par les protéines de fusion GAL4BD::P26 sauvages et tronquées exprimées dans la souche de levure HF7C. Chaque photographie présente des colonies ayant poussé sur milieu gélosé sélectif. Le type de milieu gélosé utilisé et le plan de dépôt sont précisés. "0" correspond à des levures transformées par le vecteur pGBT9 vide.

entre les protéines P25 des isolats A et B décrites au chapitre 2. Deux autres interactions sont détectées, il s'agit de celles entre la protéine CP et elle-même et entre le domaine de translecture de la protéine mineure de capsid (RT) et la protéine CP. Il faut préciser que cette interaction RT/CP avait été prédite au laboratoire par Corine Schmitt (Tamada *et al.*, 1996b). Malheureusement, aucune interaction entre les protéines P25, CP ou RT et la protéine P26 n'est détectée. Ce résultat semble indiquer que la présence de la protéine P26HA sur les agrégats viraux (Figure 3.17 D) est artefactuelle ou nécessite un facteur cellulaire intermédiaire entre la capsid et la protéine P26. Une autre explication pourrait provenir de l'altération de la structure de la protéine CP en fusion avec GAL4BD. En effet, nous nous attendions à observer une forte interaction CP/CP, qui existe *in vivo*, et qui est faiblement observée ici. De plus, nous ne retrouvons pas l'interaction CP/RT lorsque la protéine CP est utilisée comme appât. Ces résultats sont donc à prendre avec précaution et des analyses de co-immunoprécipitation seront nécessaires pour approfondir nos observations en microscopie électronique.

3.8.3 Caractérisation du domaine de la protéine P26 responsable de l'activation de la transcription

Nous sommes conscients que l'activation de la transcription observée dans le système levure ne permet pas d'attribuer une telle fonction à la protéine P26 *in planta*. Toutefois, compte tenu de sa localisation nucléaire, siège des transcriptions cellulaires, et de son action spectaculaire sur le développement des symptômes, nous avons choisi de poursuivre son étude en système levure et ainsi caractériser en détail les acides aminés responsables du recrutement de l'ARN polymérase II et de ses cofacteurs. Pour cela, nous avons analysé la croissance de levures HF7C transformées par les vecteurs exprimant les protéines GAL4BD fusionnées aux domaines FQ1 à 3 et RQ1 à 3 de la protéine P26 (Figure 3.12 A). Comme précédemment, les plasmides pGBT9 recombinants ont été introduits dans la levure et la croissance des transformants analysée sur milieu minimum SD dépourvu de tryptophane et sur milieu dépourvu de tryptophane et d'histidine. Les résultats obtenus sont représentés sur la Figure 3.19. Si tous les transformants sont capables de croître sur milieu SD-T, seuls

A

	1	10	20	30	40	50	55	232
P26	MDIDHCMPVFD	MAYSDDNHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD	LNDNVE	
MQ1-A	<u>MAAA</u> HCM	PVFD	MAYSDDNHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD	LNDNVE
MQ1-B	MDID <u>AAA</u> P	VFD	MAYSDDNHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD	LNDNVE
MQ1-C	MDIDHC	<u>MAAA</u> D	MAYSDDNHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD	LNDNVE
MQ1-D	MDIDHC	MPV	<u>FAAA</u> S	DDNHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD
MQ1-E	MDIDHC	MPVFD	<u>MAAA</u> D	NHL	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD
MQ1-F	MDIDHC	MPVFD	MAYS	<u>DAAAL</u>	PYYIQRSTHN	VVRD	VDTG	FICYPLQVD
MQ1-G	MDIDHC	MPVFD	MAYS	DDN	<u>HAAAY</u>	IQRSTHN	VVRD	VDTG
MQ1-H	MDIDHC	MPVFD	MAYS	DDNHL	<u>PAAA</u> R	STHN	VVRD	VDTG
MQ1-I	MDIDHC	MPVFD	MAYS	DDNHL	PYYI	<u>QAAA</u> H	NVVRD	VDTG
MQ1-J	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	<u>STAA</u> V	RD	VDTG
MQ1-K	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	<u>VAAA</u>	VDTG
MQ1-L	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVR	<u>DAAAT</u>
MQ1-M	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	<u>VYAAA</u>
MQ1-N	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	VDTG
MQ1-O	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	VDTG
MQ1-P	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	VDTG
MQ1-Q	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	VDTG
MQ1-R	MDIDHC	MPVFD	MAYS	DDNHL	PYYIQR	STHN	VVRD	VDTG

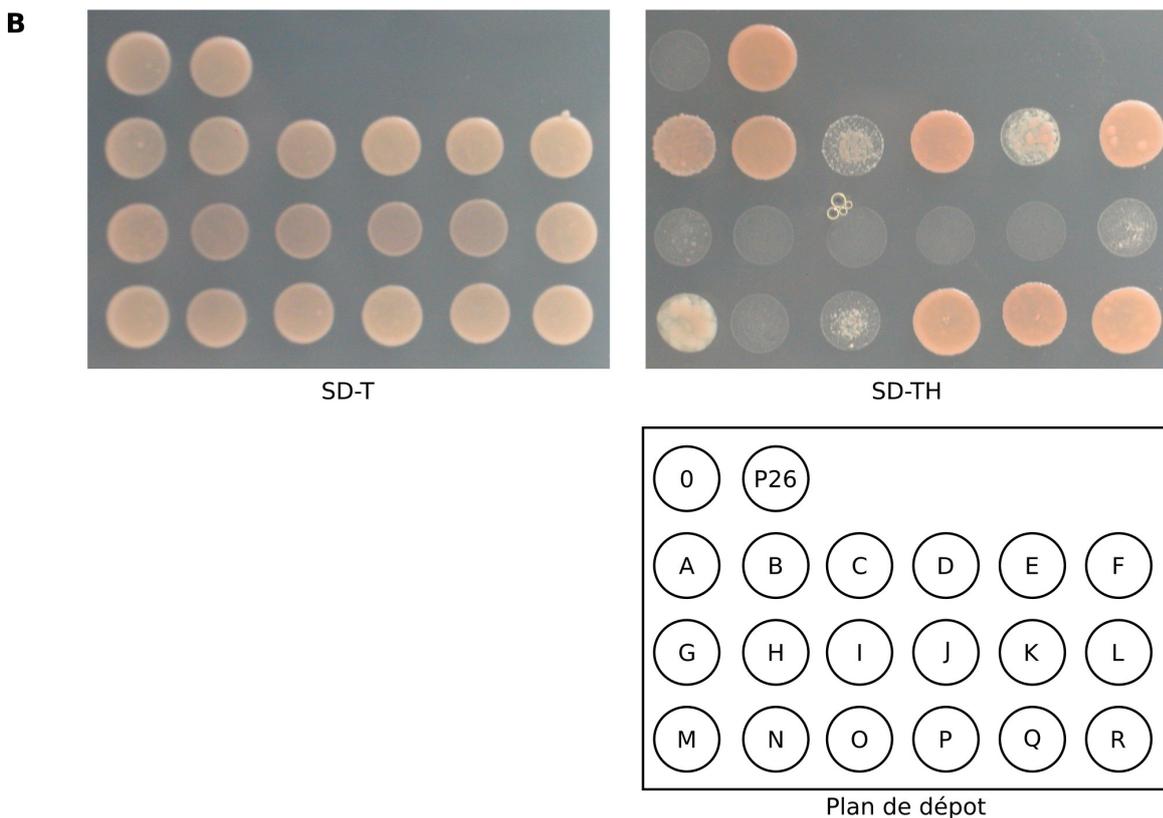


Figure 3.20 : Analyse en système dit simple hybride de l'activation de la transcription par les protéines de fusion GAL4BD::P26 et de ses formes mutantes exprimées dans la levure HF7C. A) Représentation de la séquence en acides aminés de l'extrémité N-terminale de la protéine P26 (domaine RQ1) et de l'ensemble des mutants obtenus après mutagenèse dirigée par alanine scanning. Les acides aminés modifiés sont soulignés en rouge. Le rectangle violet schématise la partie C-terminale de la protéine P26. B) Analyse de la croissance des levures transformées sur milieu SD-T et sur milieu SD-TH après 48h à 30°C. Le type de milieu gélosé ainsi que le plan de dépôt utilisé sont précisés. "0" correspond à des levures transformées par le vecteur pGBT9 vide. Chaque lettre représente le mutant MQ1-x correspondant.

ceux qui expriment les protéines de fusion GAL4BD::P26 et GAL4::RQ1 sont capables de pousser en absence d'histidine. Ceci démontre que les protéines P26 et surtout le domaine RQ1 peuvent recruter les facteurs de transcription à eux seuls. Les domaines RQ2 et RQ3 qui contiennent le domaine d'activation de la transcription RQ1 ne sont pas fonctionnels. La corrélation de cette perte de fonction à la perte d'adressage nucléaire des protéines de fusion GRQ2 et GRQ3 peut s'expliquer par la modification de la structure tridimensionnelle des protéines de fusion ou à leur instabilité *in vivo* qui n'a pas pu être testée pour les mêmes raisons que celles décrites dans le chapitre 2. J'ai donc choisi d'effectuer des expériences de mutagenèse dirigée par "overlap extension" sur la protéine P26 complète, au niveau des aminoacides présents dans le domaine RQ1.

L'ensemble des 54 acides aminés de ce domaine a été remplacé par des alanines, ceci par groupe de 3. Ainsi, dix huit mutants ont été obtenus (Figure 3.20 A) et fusionnés au domaine GAL4BD dans le vecteur pGBT9. Les levures HF7C, transformées par ces vecteurs recombinants, ont été sélectionnées et leur croissance testée à la fois sur milieu minimum dépourvu de tryptophane et sur milieu minimum dépourvu de tryptophane et d'histidine pour rechercher leur capacité à activer la transcription du gène rapporteur *His3*. Les résultats obtenus sont présentés sur la Figure 3.20 B. Nous avons ainsi pu observer que les levures exprimant les protéines mutées GAL4BD::MQ1-A, -B, -D, -F, -P, -Q et -R sont capables de croître sur milieu dépourvu d'histidine. Ces protéines mutées possèdent donc la même efficacité d'activation de la transcription du gène *His3* que la protéine de fusion sauvage. Les levures exprimant les protéines GAL4BD seule ou fusionnée aux protéines mutées MQ1-G, -H, -I, -J, -K et -N ne sont pas capables d'activer la transcription. Une situation intermédiaire est observée pour les fusions de GAL4BD avec les mutants MQ1-C, -E, -L, -M et -O qui permettent une faible activation de la transcription qui se traduit par une légère croissance des levures en absence d'histidine. Une fois encore, on ne peut cependant pas faire abstraction de la perte de stabilité de certaines des protéines de fusion mutées qui expliquerait l'absence de croissance observée.

Afin de rechercher l'incidence de l'activation de la transcription observée en système levure sur la fonction virale de la protéine P26, nous avons introduit 11 des 18 mutants MQ1 dans le vecteur d'expression viral Rep-3 en prenant soin de fusionner l'extrémité C-

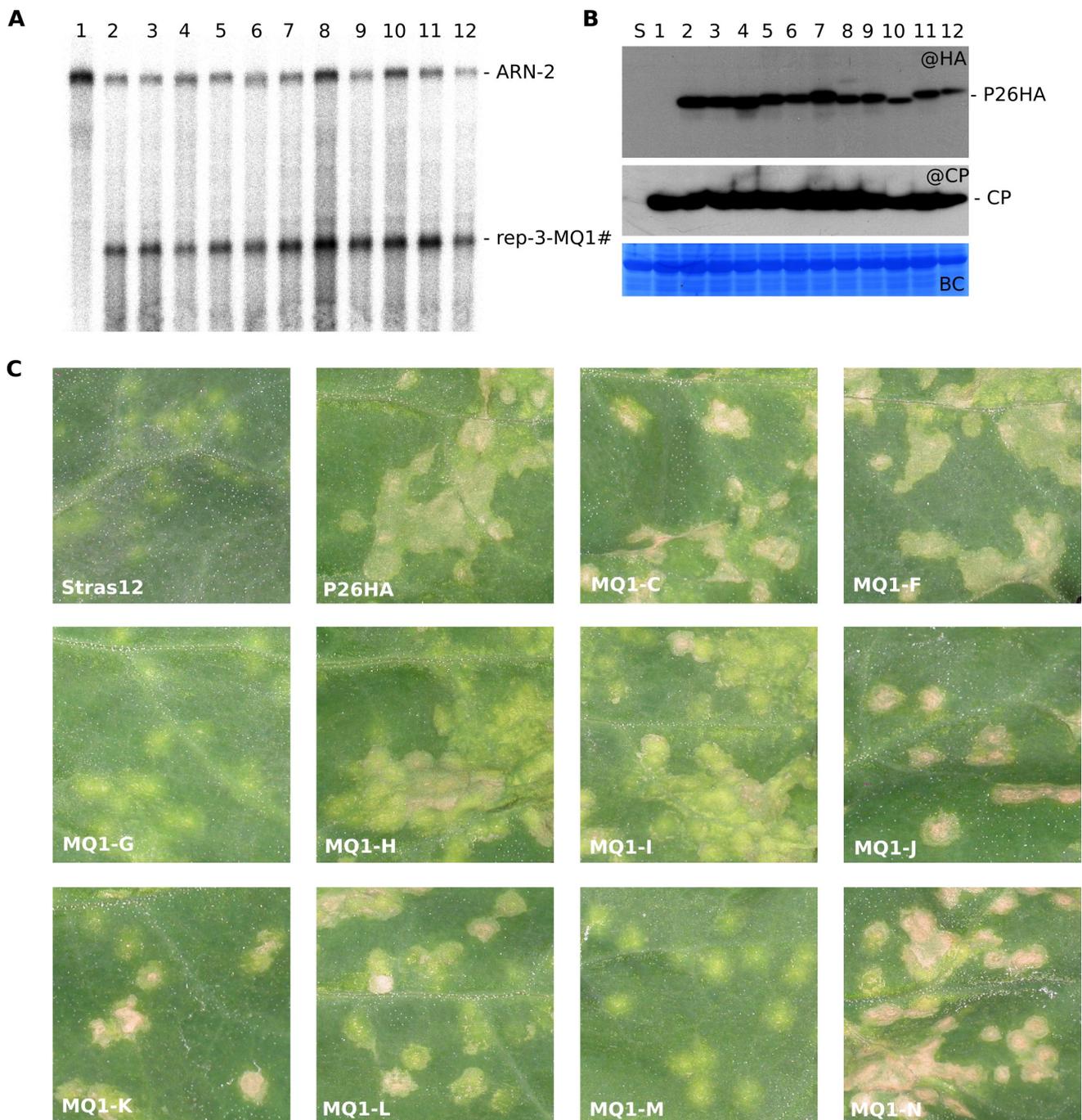


Figure 3.21 : Etude de l'influence sur la symptomatologie foliaire des constructions MQ1 et analyse de l'expression des protéines étiquetées et de la réplication des ARN viraux. A) Analyse par northern-blot de l'accumulation des ARN viraux de la souche helper Stras12 (1) supplémentée de rep-P26HA (2), rep-MQ1-C (3), rep-MQ1-F (4), rep-MQ1-G (5), rep-MQ1-H (6), rep-MQ1-I (7), rep-MQ1-J (8), rep-MQ1-K (9), rep-MQ1-L (10), rep-MQ1-M (11) et rep-MQ1-N (12). B) Détection par western-blot de l'accumulation des protéines P26 mutantes étiquetées HA. @HA : immunodétection de l'épitope HA ; @CP immunodétection de la protéine CP ; BC : coloration au Serva Blue. C) Symptômes observés 6 jours après inoculation mécanique sur feuilles de *C. quinoa*.

terminale des protéines à l'épitope HA, pour les rechercher par western-blot mais aussi pour caractériser leur localisation sub-cellulaire en microscopie électronique.

3.9 Caractérisation des propriétés biologiques de mutants de la protéine P26

Les réplicons exprimant les protéines étiquetées HA, P26, MQ1-C, -F, -G, -H, -I, -J, -K, -L, -M et -N ont été inoculés à des feuilles de *C. quinoa* en présence de la souche helper Stras12. Après 6 jours de virose, les lésions obtenues ont été caractérisées phénotypiquement, les ARN viraux extraits des lésions et recherchés par northern-blotting, et les protéines virales analysées par western-blotting (Figure 3.21). L'analyse par northern-blot de la présence des ARN viraux est présentée sur la Figure 3.21 A. Comme nous pouvons le constater, l'accumulation des différents réplicons exprimant les protéines P26 sauvage et mutées sont identiques. De même, la recherche des protéines P26HA par western-blot montre que ces protéines s'accumulent de manière quasi identique (Figure 3.21 B, panneau @HA), la charge en protéines totales (Figure 3.21 B, panneau BC) et en protéines de capsid (Figure 3.21 B, panneau @CP) étant comparable pour chaque échantillon testé. On peut également noter ici la modification de la mobilité électrophorétique de certaines protéines mutées qui pourrait être expliquée par des changements de charges et/ou de sites de modifications post-traductionnelles. L'observation des phénotypes des lésions obtenues permet de différencier deux catégories de mutants (Figure 3.21 C). D'une part les mutants qui restent capables d'induire des nécroses similaires à celles observées lorsque la protéine sauvage P26HA est exprimée ; c'est le cas des mutants MQ1-C, -F, -H, -I, -J, -K, -L et -N. D'autre part, les mutants qui n'induisent plus de nécrose et produisent un phénotype comparable à celui obtenu en présence de Stras12 seul ; c'est le cas des deux mutants MQ1-G et -M. Comme nous pouvons le voir, il n'existe pas de lien direct entre l'activation de la transcription et l'action nécrogène de la protéine P26. En effet, certains mutants toujours capables d'activer la transcription induisent des nécroses (MQ1-C, -F et -L) ou perdent cette action nécrogène (MQ1-M). A l'inverse, le mutant MQ1-G est quant à lui incapable

d'activer la transcription ni même d'induire des nécroses. Enfin, les mutants MQ1-H, -I, -J, -K et -N n'activent plus la transcription mais restent capables d'induire la mort cellulaire.

Afin de déterminer si l'induction de la réponse nécrotique de la plante fait intervenir la localisation de la protéine P26 dans le noyau ou dans le cytoplasme des cellules infectées, nous avons entrepris de caractériser la localisation sub-cellulaire des protéines P26 mutées par microscopie électronique. Les résultats ne sont pas encore disponibles au moment de la rédaction de ce manuscrit. L'analyse de la répartition du marquage entre noyau et cytoplasme nous permettra peut-être de corrélérer la localisation sub-cellulaire des mutants à la capacité ou non d'activer la transcription et celle d'induire ou pas des nécroses.

3.10 Conclusion

Cette étude de l'ARN-5 du BNYVV nous a permis de démontrer qu'il était effectivement impliqué dans l'augmentation du pouvoir pathogène d'un isolat viral ne le possédant pas initialement. Sa présence induit l'apparition de lésions nécrotiques sur feuilles de *C. quinoa* et *Tetragonia expansa* (non montré). Nous avons également prouvé que c'est bien l'expression du gène codant pour la protéine P26 qui est responsable de l'apparition du phénotype nécrotique, ceci grâce à l'expression du gène codant pour cette protéine P26 en contexte viral par l'intermédiaire d'un ARN dérivé de l'ARN-3. Cet ARN recombinant induit des symptômes nécrotiques identiques à ceux observés après inoculation de *C. quinoa* par un isolat A/P qui sont perdus si la protéine P26 n'est pas synthétisée (mutants P26HA-A, -D et -E) ou lorsqu'elle est tronquée (mutants P26FS, FQ1 à FQ3, RQ1 à RQ3, D7 à D9). Une autre approche a nécessité l'obtention d'un clone complet infectieux de l'ARN-5 qui permet d'obtenir des transcrits induisant des nécroses en présence de la souche helper Stras12 (isolat B). Cet ARN-5 est de plus encapsidé par la protéine CP de l'isolat B. Ce résultat démontre que les différents isolats du BNYVV sont encore assez proches pour être compatibles.

La séquence complète de l'ARN-5 nous a révélé l'existence d'une seule grande phase ouverte de lecture, codant pour la protéine P26. La présence de trois petites séquences codantes dans la partie 5' non codante de cet ARN pourrait indiquer la présence d'une régula-

tion fine du niveau de traduction de la séquence codante principale (Kozak, 2002). L'obtention d'un nouveau vecteur d'expression viral à partir de l'ARN-5, le Rep-5, conforte cette hypothèse. En effet, à réplication comparable, le taux d'expression du gène codant pour la protéine EGFP est 5 fois plus faible en réplicon-5 qu'en réplicon-3 (voir article en annexe au chapitre 3). Cette propriété pourrait être mise à profit pour exprimer des protéines qui sont cytotoxiques lorsqu'elles sont surexprimées. Ce réplicon-5 pourra être utilisé avantageusement en remplacement du réplicon-3 lorsqu'une instabilité est constatée, comme dans le cas des réplicons exprimant les protéines P26 fusionnées à l'EGFP. De plus, la possibilité d'exprimer n'importe quelle protéine à partir du Rep-5 en combinaison avec l'ARN-3 permet alors de produire massivement cette protéine dans les hôtes systémiques du BNYVV comme *B. macrocarpa*. Enfin, la co-expression de deux gènes codant pour deux protéines dans la même cellule infectée ouvre de nouvelles perspectives en recherche fondamentale, par des expériences de co-immunoprécipitation et l'analyse d'interactions protéine/protéine par exemple.

La protéine P26 ne présente aucune similarité de séquences avec des protéines connues, ni ne possède de motifs consensus particuliers au sein de sa séquence protéique. Après avoir démontré l'implication du premier codon d'initiation dans son expression en contexte viral, nous avons initié la caractérisation de sa localisation sub-cellulaire et la recherche de ses partenaires cellulaires pour tenter de lui attribuer une ou des fonctions.

Nous avons pu montrer que la protéine P26 était activement importée dans le nucléoplasme, grâce à sa fusion avec l'EGFP et la chalcone synthase, mais aussi par l'immuno-détection de la protéine étiquetée HA sur des coupes ultrafines de tissus infectés. L'étude de la localisation sub-cellulaire en cellules BY-2 des protéines P26 tronquées fusionnées à l'EGFP nous a permis de localiser le motif d'adressage nucléaire au voisinage du domaine RQ1, c'est à dire à l'extrémité N-terminale de la protéine. Toutefois, si le domaine RQ1 est nécessaire, il ne s'est pas montré suffisant pour localiser dans le noyau la protéine de fusion EGFP::RQ1::CHS. La protéine P26 pourrait donc contenir plusieurs motifs répartis sur sa séquence et jouant un rôle dans son import nucléaire. Ainsi, Burich and Lei (2003) ont montré que dans la protéine Mcm10, impliquée dans le complexe de pré-réplication du génome nucléaire et se trouvant dans le noyau des cellules eucaryotes tout au long du cycle

cellulaire, il existe deux NLS bipartites qui sont fonctionnels. De plus, Luo *et al.* (2004) ont montré que les trois NLS identifiés dans l'enzyme 5-lipoxygénase pourraient permettre une régulation fine du niveau d'import nucléaire de la protéine, fusionnée à la GFP. Toutefois, nos données, certes encore peu concluantes, peuvent s'expliquer par un bouleversement majeur de la structure tridimensionnelle des protéines de fusion qui expliquerait à la fois l'absence de localisation nucléaire des protéines GRQ2 et la localisation nucléaire partielle de la protéine GRQ3. L'absence d'activation de la transcription par les deux protéines GAL4BD::RQ2 et GAL4BD::RQ3 pourrait être expliquée de la même manière. Il nous reste à tester l'effet de la leptomycine B sur une éventuelle inhibition de l'export des protéines qui présentent une localisation cytoplasmique afin de déterminer si la protéine P26 possède ou pas la caractéristique d'effectuer une navette nucléo-cytoplasmique comme cela a été démontré pour la protéine P25 codée par l'ARN-3 du BNYVV (Vetter *et al.*, 2004). Ainsi, l'absence de séquence NLS conventionnelle nous permet de suspecter la présence d'un motif de localisation nucléaire, appelé NLM pour "nuclear localisation motif", au sein de la structure tridimensionnelle de la protéine P26 comme l'ont démontré Lombardo *et al.* (2000, 2002) pour les protéines VP1 et VP2 constitutives de la capsid du Parvovirus "Minute Virus of Mice" et qui sont impliquées dans l'import nucléaire du génome viral. Seule la résolution de la structure tridimensionnelle de la protéine P26 permettrait de trouver des motifs structuraux décrits dans les bases de données spécialisées (PDB, PROSITE, PRINTS et InterPro) et ainsi de cibler spécifiquement par mutagenèse dirigée les amino acides en position clé de la structure. De plus, on pourrait supposer l'implication d'un motif structural, reconnu par une protéine cellulaire différente des importines, qui interagirait avec la protéine P26 et la transporterait dans le noyau. Ces hypothèses ne peuvent être vérifiées qu'à l'aide de la recherche des partenaires cellulaires et/ou viraux de la protéine P26.

L'approche que nous avons utilisée pour rechercher les facteurs cellulaires et viraux interagissant avec la protéine p26 nous a permis d'observer, en système double hybride, que lorsque la protéine P26 est fusionnée au domaine de fixation à l'ADN du facteur de transcription Gal4, elle est capable de recruter l'ARN polymérase II et ainsi permettre l'expression du gène rapporteur *His3*, ce qui rend les levures prototrophes vis à vis de cet acide

aminé. Le domaine responsable de cette activation de la transcription a été localisé dans le domaine RQ1 et plus particulièrement entre les acides aminés 19 à 34 et 41 à 43. Nous disposons donc d'une collection de mutants de la protéine P26 qui n'activent plus la transcription et qui pourront nous servir à rechercher, en système double hybride, des partenaires cellulaires et viraux (P25, P26, ...) de la protéine P26 dans des banques d'ADNc, clonées dans le vecteur pGAD424.

L'activation de la transcription nécessite le domaine RQ1 isolé ou la protéine complète. Ces données apportent une preuve supplémentaire quant à l'importance de la structure tridimensionnelle de la protéine pour qu'elle soit active. Cette structure est notamment indispensable à l'induction des symptômes nécrotiques sur *C. quinoa*. Si les acides aminés 20 à 22 et 38 à 40, et dans une moindre mesure les acides aminés 26 à 31 où la nécrose est retardée, sont nécessaires, nous ne pouvons pas préciser s'ils ont un rôle de stabilisation de structure secondaire, s'ils sont directement reconnus par un facteur cellulaire déclenchant la mort cellulaire ou encore s'ils sont requis dans l'adressage nucléaire de la protéine P26 qui pourrait être impliquée dans l'induction de la mort cellulaire. Ainsi, la protéine AvrBs3 de *Xanthomonas campestris* (Szurek *et al.*, 2002) possède deux séquences NLS qui localisent la protéine dans le nucléoplasme. AvrBs3 interagit dans le noyau de cellules de poivron, avec la protéine Bs3 de l'hôte, ce qui induit la mort cellulaire (réponse HR). En absence de séquence NLS, aucune mort cellulaire n'est induite (Szurek *et al.*, 2002). Dans notre cas, l'absence de corrélation entre l'activation de la transcription et l'induction des nécroses nous laisse penser que les deux activités ne sont pas liées. Toutefois, il faut garder à l'esprit que la transcription est un phénomène nucléaire. En système double-hybride, les protéines de fusion sont adressées au nucléoplasme par l'intermédiaire des séquences NLS de GAL4. Ainsi, si certains de nos mutants peuvent activer la transcription dans la levure, il est possible que cette activation soit impossible en contexte viral et ceci en raison de l'absence d'adressage nucléaire des protéines mutées. Avant de conclure définitivement sur ces différentes hypothèses, il nous faut attendre les résultats de l'étude de l'immunolocalisation des mutants en contexte viral mais également étendre notre approche dans un premier temps aux mutants MQ1-A, B, D, E, O, P, Q, R et à d'autres variants ciblant le domaine D7 impliqué dans la rétention cytoplasmique ou l'export nucléaire de la protéine P26.

Chapitre 4

Conclusion générale et perspectives

Les résultats qui sont rapportés dans ce mémoire de thèse concernent l'analyse phylogénétique d'isolats de BNYVV et la caractérisation des propriétés moléculaires des protéines P25 exprimées par l'ARN-3 d'isolats de BNYVV mais aussi l'étude des propriétés biologiques de la protéine P26 codée par l'ARN-5.

L'analyse phylogénétique, effectuée sur un nombre important d'échantillons, nous a permis de préciser l'existence de nouveaux isolats viraux et a révélé qu'il existait une pathogénicité décroissante entre les isolats contenant un ARN-5, les isolats de type A et les isolats de type B. L'existence réelle de l'un de ces isolats a pu être démontrée grâce à l'obtention du clone ADNc complet pB55 permettant de synthétiser un ARN-5 infectieux. Cet ARN est amplifié et encapsidé en présence des ARN génomiques 1 et 2 d'un isolat B. La variabilité de la séquence de quatre aminoacides (tétrade) de la protéine P25, situés en aval de la séquence NLS, nous a permis de caractériser 11 variants majeurs. Cette variabilité a été reproduite par mutagenèse dirigée de l'ARN-3 d'un isolat B. Nous avons analysé l'effet de ces tétrades sur les propriétés biologiques de la protéine P25. Ainsi, ces amino-acides ne modifient pas les propriétés de navette nucléo-cytoplasmique de la protéine P25, n'affectent pas la réplication, ni la stabilité des ARN-3 *in planta*. Les symptômes obtenus sur plantes hôtes ont cependant révélé des différences phénotypiques, plus particulièrement sur *Tetragonia expansa*. L'expression des gènes des protéines P25 issues des variants de l'ARN-3 a révélé des différences de migration électrophorétique en conditions dénaturantes que nous avons pu attribuer, du moins en partie à des modifications de charge globale des protéines.

Toutefois, certains travaux de thèse effectués au laboratoire par Guillaume Vetter et qui ont consisté à réaliser une mutagenèse « Alanine et Aspartate scanning » des sites potentiels de phosphorylation avaient mis en lumière l'importance de la phosphorylation de la Thréonine 164 sur l'adressage nucléaire de la protéine P25. Un tel résultat démontre l'importance des modifications post-traductionnelles sur la fonction de la protéine P25 et suggère également qu'il pourrait exister des interactions entre le domaine C-terminal de la protéine et la séquence NLS N-terminale. Si de telles interactions existent, la tétrade se retrouverait donc au voisinage du domaine C-terminal contenant entre autres, le domaine AD1 responsable de l'activation de la transcription (Elodie Klein DEA 2003, Guillaume Vetter, Thèse 2003). D'une part, nous avons pu montrer que cette tétrade avait une influence sur l'activation de la transcription induite par les protéines GAL4BD::P25, observée en système « simple-hybride ». Cette fonction a pu être détectée uniquement dans le cas de protéines P25 provenant d'isolats ne contenant pas d'ARN-5. D'autre part, une fonction identique, plus forte, a été mise en évidence pour la protéine P26 codée par l'ARN-5. Cette protéine possède d'autres caractéristiques semblables à la protéine P25 : elle est adressée au nucléoplasme par une séquence NLS non classique (NLM) qui devra être caractérisée en détail et influence les symptômes foliaires.

Nous avons pu démontrer que le domaine N-terminal de la protéine P26 est impliqué dans l'activation de la transcription. Par contre, nous ne pouvons actuellement pas définir s'il existe un lien entre cette activation de la transcription, l'induction de nécroses, la structure tertiaire et la localisation nucléaire de la protéine P26. Les résultats de la localisation sub-cellulaire des mutants ponctuels du domaine RQ1 de la protéine P26 pourront en partie répondre à cette question.

Les similitudes de fonctions (activation de transcription, localisation nucléaire, rôle dans le développement des symptômes) qui existent entre les protéines P25 et P26 combinées au rôle accessoire de l'ARN-5 nous laisse penser que ces deux protéines pourraient adopter une structure et/ou une fonction proche qui serait redondante, conduisant à la perte de l'ARN-5 et qui pourrait ainsi valider notre hypothèse évolutive. Seule l'étude cristallographique de ces deux protéines pourra répondre à cette question. Dans cette optique, nous avons initié la purification de protéines P26 et P25 étiquetées HA à partir de plantes infectées (plantes

systemiques). L'obtention de ces protéines purifiées à l'homogénéité permettra d'une part de déterminer leur spectre de dichroïsme circulaire et ainsi de produire en systèmes hétérologues (bactéries, levures, insectes, plantes) des protéines recombinantes présentant les mêmes caractéristiques. Ces protéines seront cristallisées et leur structure déterminée par diffraction des rayons X. D'autre part, nous comptons analyser, par spectrométrie de masse, les protéines purifiées à partir de fractions cytoplasmiques ou nucléaires de cellules de plantes infectées afin de caractériser les modifications post-traductionnelles qui affectent ces deux protéines virales qui pourraient réguler leur localisation subcellulaire. Enfin, la mise au point des conditions de purification des protéines P25 et P26 étiquetées HA nous permettra de réaliser des expériences de co-immunoprécipitation de complexes protéiques. Nous pourrons ainsi identifier les partenaires cellulaires et viraux (nucléaires et cytoplasmiques) des deux protéines virale et les étudier. Cette dernière approche sera complétée par l'utilisation du système double hybride. Dans ce cas, nous utiliserons comme protéine de fusion appât, les protéines P25 provenant d'isolats A, B et P et un ou plusieurs mutants de la protéine P26. Ces protéines seront fusionnées au domaine de fixation des acides nucléiques de GAL4, dans le vecteur pGBKT7, qui permet la détection des protéines de fusion grâce à une étiquette myc.

Enfin, la construction de clones ADNc infectieux correspondant aux ARN viraux 1 à 4 de l'isolat «Pithiviers» a été initiée. Ces ARN infectieux permettront de réaliser des isolats chimériques pour tester leur comportement en conditions naturelles. L'isolat A/P pourra également servir à effectuer des expériences de co-infections avec l'isolat B. La descendance de ces infections mixtes sera analysée. Nous pourrons peut-être ainsi reproduire rapidement des phénomènes évolutifs observés lors de notre approche phylogénétique.

La caractérisation de la séquence d'échantillons mondiaux est poursuivie. Nous espérons ainsi découvrir de nouveaux variants du BNYVV, en particulier les isolats B/P et B/J, qui apporteraient une preuve supplémentaire à notre modèle d'évolution et pourquoi pas, isoler l'ancêtre commun à tous les isolats de BNYVV rencontrés, s'il existe.

Chapitre 5

Matériels et méthodes

5.1 ARN

5.1.1 Extraction à partir de tissus végétaux

L'extraction des ARN totaux à partir de tissus végétaux a été faite à l'aide de Trizol® (GibcoBRL™). L'extraction se déroule selon le protocole suivant. Les tissus de plantes (50 à 100 mg) sont broyés jusqu'à homogénéisation dans 0,5 ml de Trizol à l'aide d'un micro-broyeur en microtube de 1,5 ml. Une centrifugation optionnelle (utile lors du broyage de racines) est effectuée à 4°C pendant 10 min à 12000g pour éliminer la fraction non soluble. Une incubation de 5 min à température ambiante permet la dissolution des complexes nucléo-protéiques puis on ajoute 0,1 ml de chloroforme. Les tubes sont vigoureusement agités à la main pendant 15 sec et sont incubés à température ambiante pendant 2 à 3 min. La séparation de phase est effectuée par une centrifugation à 4°C pendant 15 min à 12000g maximum. Les ARN contenus dans la phase aqueuse sont précipités par addition de 0,25 ml d'isopropanol. Après 10 min d'incubation à température ambiante, les ARN sont récupérés par une centrifugation de 10 min à 4°C à 12000g maximum. Les ARN sont lavés par 1 ml d'éthanol à 70% et une centrifugation de 5 min à 4°C à 7500g maximum. Le culot est dissout dans 5 à 10 µl d'eau bidistillée stérile.

5.1.2 Analyse des ARN par la technique de northern-blot

Synthèse des sondes ribonucléotidiques

Les sondes ARN sont obtenues par transcription en présence d'UTP[α -³²P]. La transcription est réalisée à partir d'ADN plasmidique recombinant possédant les promoteurs des bactériophages T₃ ou T₇ de part et d'autre d'un fragment d'ADN d'intérêt. En utilisant l'ARN polymérase adéquate, il est possible d'obtenir l'ARN correspondant ou complémentaire à l'ADN inséré. Afin de limiter la transcription au seul insert, l'ADN plasmidique est au préalable linéarisé par une enzyme de restriction dont le site unique se trouve directement en aval du fragment d'ADN à transcrire.

Séparation des ARN sur gel d'agarose dénaturant

Les ARN sont séparés par électrophorèse en gel dénaturant (agarose 1 % ; HEPES-KOH 20 mM pH 7,8 ; EDTA 1 mM ; formaldéhyde 6 %). Les échantillons, en suspension dans le tampon de dénaturation (HEPES-KOH 20 mM pH 7,8 ; EDTA 1 mM ; formamide désionisée 50 % ; formaldéhyde 16 % ; glycérol 50 % ; bleu de bromophénol 0,01 %) sont dénaturés à 65°C pendant 5 min, puis déposés sur gel. L'électrophorèse est réalisée à 100 V dans le tampon de migration (HEPES-KOH 20 mM pH 7,8 ; EDTA 1 mM).

Détection des ARN par l'hybridation moléculaire

Cette méthode est basée sur la détection des ARN par hybridation moléculaire au moyen de sondes radioactives spécifiques.

Après séparation par électrophorèse en gel d'agarose dénaturant, les ARN sont transférés pendant une nuit par capillarité sur une membrane HYBOND® (Amersham™) dans du tampon SSC × 20 (NaCl 3 M ; citrate de sodium 300 mM). Après un bref lavage dans du SSC × 10, les ARN sont fixés sur la membrane par irradiation aux UV (254 nm, 120 mJoules) dans un Stratalinker® (Stratagene™). Les ARN cellulaires sont visualisés sous lumière UV grâce à l'addition de Bromure d'Ethidium (BET) dans le tampon de charge (0,005 µg / µl) et constituent le témoin de charge.

La membrane HYBOND est incubée pendant 2 h à 60°C dans 10 ml de tampon d'hy-

bridation ($\text{NaH}_2\text{PO}_4 / \text{Na}_2\text{HPO}_4$ 50mM pH6,5 ; formamide 50 % ; SSC $\times 5$; solution de Denhardt $\times 8$ [BSA 0,16% ; PVP 0,16% ; Ficoll 0,16%] ; SDS 0,1% ; ADN de sperme de saumon 250 $\mu\text{g} / \text{ml}$; ARN de levure 500 $\mu\text{g} / \text{ml}$). Le milieu d'hybridation est ensuite additionné de $1 \cdot 10^6$ à $5 \cdot 10^6$ cpm de chaque sonde radioactive. L'hybridation est effectuée à 55°C pendant 16h sous agitation douce.

Afin d'éliminer la sonde en excès et les hybridations aspécifiques, la membrane HYBOND est lavée deux fois à 65°C, pendant 20min dans une solution de SSC $\times 2$, SDS 0,1% et deux fois pendant 20min dans une solution SSC $\times 0,2$, SDS 0,1%, puis séchée et autoradiographiée. Le signal radioactif est quantifié à l'aide d'un phosphorimager FUJI BAS 1000.

5.1.3 Transcription inverse

La transcription inverse permet de produire un ADNc à partir d'ARN totaux ou polyadénylés. La réaction classique se fait dans un volume final de 20 μl comprenant 0,25 μl (5 U) de reverse transcriptase SuperScript II (Invitrogen), 1 μl d'ARN (1 $\mu\text{g} / \mu\text{l}$), 4 μl de tampon $\times 5$, 0,6 μl de DTT 100mM, 0,1 μl de RNaseout™ (inhibiteur de ribonucléases Invitrogen), 0,4 μl de dNTP 10mM, 4 μl d'un oligonucléotide complémentaire à la séquence de l'ARN ou un oligodT à 1 μM et 9,65 μl d'eau stérile. Le mélange réactionnel est incubé à 65°C pendant 5 min puis à 37°C pendant 1 h.

5.1.4 Obtention de clones ADNc complets d'ARN viraux par les techniques dites de 5' et 3' RACE

Après obtention d'un ADNc, à partir d'ARN totaux ou polyadénylés, par transcription inverse, et afin d'obtenir sa séquence complète, l'amplification rapide des extrémités d'ADNc (Rapid amplification of cDNA ends ou RACE) a été utilisée. C'est une technique basée sur la réaction de PCR (voir paragraphe 5.2.1, page 85) qui facilite le clonage de la séquence complète d'un ADNc.

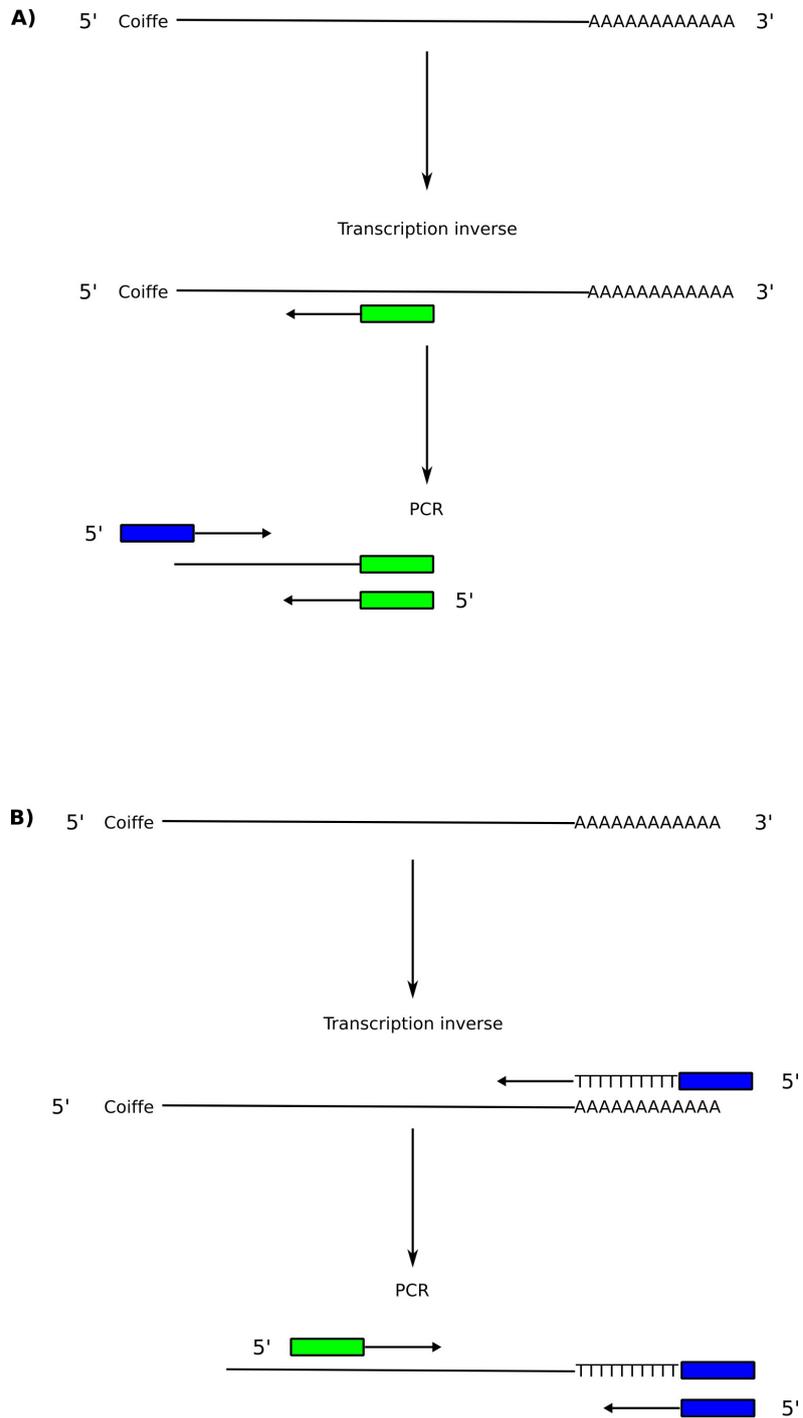


Figure 5.1 : Représentation schématique des techniques de 5' et 3' RACE. A) Schéma du protocole de 5' RACE. B) Schéma du protocole de 3' RACE. Les rectangles bleus représentent les séquences amorce adaptatrices, les rectangles verts représentent les séquences de l'ARN, ou leur complémentaires, auxquelles s'hybrident les amorces.

Obtention de la partie 5' de l'ARN

Le protocole de 5' RACE (Figure 5.1 A) est basé sur l'amplification par PCR de la partie 3' de l'ADNc mono-brin. Il est nécessaire de disposer d'une amorce spécifique d'une séquence interne de l'ARN afin d'obtenir la partie 5' complète de l'ARN par transcription inverse. L'ADNc est ensuite amplifié par PCR avec un couple d'amorce dont une est spécifique d'une séquence de l'ARN d'origine. L'autre contient la séquence du promoteur du bactériophage T₇ et s'hybride au début de l'ARN grâce aux séquences conservées en 5' des ARN viraux (voir table 2 dans l'introduction générale). Le produit de PCR est ensuite introduit dans un vecteur de clonage, comme décrit dans le paragraphe 5.2.8 (page 88).

Obtention de la partie 3' de l'ARN (kit Ambion : FirstChoice™ RML-RACE, référence 1700)

Le protocole de 3' RACE (Figure 5.1 B) est basé sur la production d'un ADNc à partir d'une amorce complémentaire à la séquence polyadénylée des ARN messagers complétée par une séquence adaptatrice. L'ADNc produit est ensuite amplifié par une réaction de PCR à l'aide d'une amorce spécifique de l'ARN et d'une amorce complémentaire à la séquence de l'adaptateur. La mise en oeuvre de ces étapes a été réalisée en utilisant les recommandations du fournisseur. Le fragment d'ADN obtenu est ensuite introduit dans un vecteur de clonage (voir le paragraphe 5.2.8, page 88).

5.1.5 Transcription *in vitro*

Le kit de transcription *in vitro* RiboMax de Promega™ a été utilisé en modifiant le protocole afin d'optimiser la production d'ARN coiffés en 5'. Pour une réaction de 20 µl final, le mélange réactionnel se compose de : 4 µl de tampon × 5 ; 1,5 µl de rATP 100 µM ; 1,5 µl de rCTP 100 µM ; 1,5 µl de rUTP 100 µM ; 0,06 µl de rGTP 100 µM ; 1 µl de DTT 100 µM ; 0,5 µl de RNaseOUT (inhibiteur de RNase, Invitrogen) ; 1 µl de T₇ ARN polymérase (New England Biolabs) ; 1 µl de mélange d'enzyme T₇ (Promega) ; 4 µl de CAP (m7Gppp de New England Biolabs) ; 2,5 µl d'eau stérile ; 1,5 µl de plasmide matrice linéarisé (1,5 µg). Ce mélange est incubé 30 minutes à 37° C puis 1,5 µl de rGTP 100 µM est ajouté. La réaction se

Table 5.1 : Composition du milieu réactionnel pour l'amplification de fragments d'ADN par PCR.

Composants	Volume
ADN (10 ng/μl)	2 μl
Oligonucléotide 1 (1 μM)	2 μl
Oligonucléotide 2 (1 μM)	2 μl
Tampon × 10	3 μl
MgCl ₂ (50 mM)	0,9 μl
dNTP (1 mM)	2 μl
Taq ADN polymérase (5 u/μl)	0,25 μl
H ₂ O qsp	30 μl

poursuit à 37°C pendant 3 à 4 heures. Les transcrits synthétisés sont analysés par dépôt de 1 µl de milieu réactionnel sur gel d'agarose 1% non dénaturant et visualisés par coloration du gel au BET.

5.1.6 Analyse de l'encapsidation des ARN viraux

Les ARN issus de matériel infecté sont extraits à l'aide du protocole Trizol (paragraphe 5.1.1, page 81) pour estimer la quantité totale d'ARN viraux nus et encapsidés. Une même quantité de matériel infecté est également soumise à une extraction dans un tampon Tris-MgCl₂ (10 mM-1 mM) et le broyat est incubé 30 min à 37°C pour favoriser l'action des ribonucléases endogènes. Les ARN viraux encapsidés, protégés de la dégradation, sont ensuite purifiés par extraction phénolique suivie d'une précipitation à l'éthanol (Jupin *et al.*, 1990).

5.2 ADN

5.2.1 Amplification d'un fragment d'ADN par PCR (Polymerase Chain Reaction)

Cette technique permet l'amplification *in vitro* d'un fragment d'ADN de manière exponentielle grâce à l'utilisation de l'ADN polymérase thermostable de *Thermus aquaticus* (appelée *Taq*). Elle nécessite l'utilisation de deux amorces oligodésoxyribonucléotidiques spécifiques de la séquence à amplifier. La réaction se déroule en 20 à 30 cycles dans un thermocycleur en programmant la répétition de 3 étapes : 30sec de dénaturation de l'ADN à 95°C, 30sec pour l'hybridation des amorces avec les brins matrice à une température compatible avec la température de fusion des amorces et de 30sec à 2min à 72°C pour l'élongation des amorces. La température de fusion des oligonucléotides a été déterminée en utilisant le logiciel Netprimer (voir paragraphe 5.4, page 93). Le temps d'élongation est déterminé sur la base de 30sec pour obtenir un fragment de 500pb, cette donnée pouvant varier en fonction de l'enzyme utilisée. Les conditions générales de l'amplification sont indiquées dans la table 5.1.

Table 5.2 : Composition du milieu réactionnel nécessaire à la digestion de 1 à 5 µg d'ADN par les enzymes de restriction.

Composants	Volume
ADN (1 µg / µl)	1 à 5 µl
BSA (2 mg / ml)	1 µl
Tampon × 10 spécifique de l'enzyme	2 µl
Enzyme (10U / µl)	0,2 µl
H ₂ O qsp	20 µl

Les amplicons obtenus sont analysés sur gel d'agarose et peuvent être introduits dans un vecteur (voir paragraphe 5.2.8, page 88).

5.2.2 Digestion par les enzymes de restriction

Ces endonucléases, qui coupent les molécules d'ADN bicaténares au niveau de séquences spécifiques, génèrent des fragments ayant soit des extrémités cohésives 5' ou 3' sortantes, soit des extrémités franches. Chaque enzyme de restriction présente une activité optimale dans un tampon spécifique et à une température donnée. En général, l'incubation se déroule pendant 1 h à 37°C.

Pour digérer 1 à 5 µg d'ADN, nous avons utilisé le milieu réactionnel de la table 5.2.

Les fragments de restriction sont ensuite analysés en gel d'agarose (voir paragraphe 5.2.4).

5.2.3 Déphosphorylation d'un vecteur

La déphosphorylation empêche un vecteur linéarisé de se refermer sur lui-même lors de l'étape de ligation ultérieure et favorise ainsi l'insertion du fragment d'ADN à cloner.

La déphosphorylation est réalisée par action de la phosphatase alcaline d'intestin de veau, qui élimine les groupements phosphates 5'-terminaux. Cette enzyme, thermostable, ne présente pas de spécificité étroite vis-à-vis d'un tampon d'incubation et peut donc être utilisée immédiatement après digestion par une enzyme de restriction. Classiquement, 1 µg d'ADN est incubé dans 50 µl de milieu réactionnel pendant 1 h à 37°C. L'enzyme est ensuite éliminée par extraction phénolique (voir paragraphe 5.2.5).

5.2.4 Analyse des fragments d'ADN par électrophorèse en gel d'agarose

Les fragments d'ADN obtenus par digestion enzymatique ou par amplification PCR sont analysés par électrophorèse en gel d'agarose de 0,8 à 2,5 % selon la taille des fragments à analyser. L'ADN, en solution dans le tampon de charge (Tris-HCl 2 mM pH 8, EDTA 10 mM, saccharose 5 %, bleu de bromo-phénol 0,01 %), est déposé sur le gel préparé dans

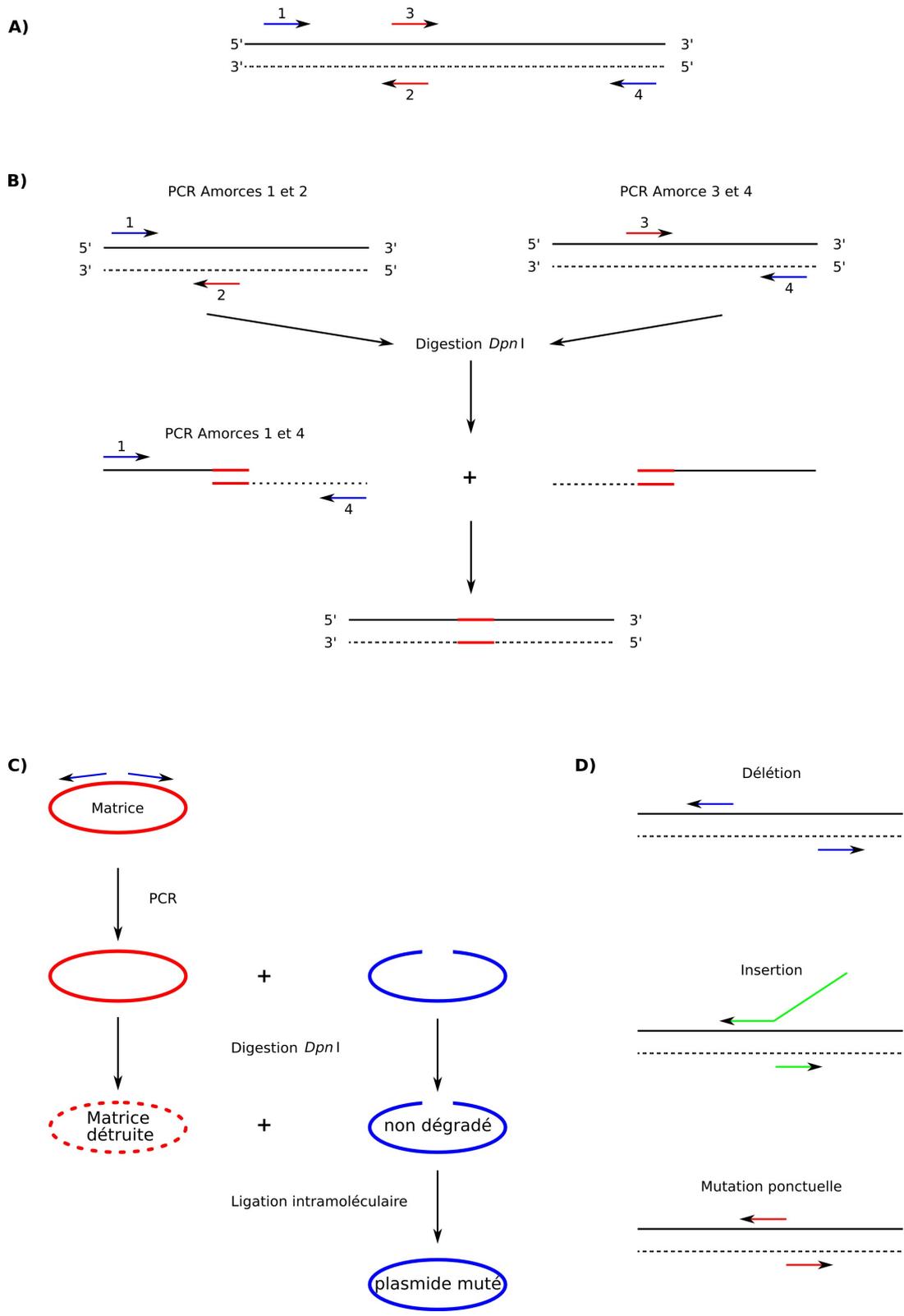


Figure 5.2 : Schémas des différents protocoles d'amplification et de mutagenèse par PCR. A) Position des quatres amorces nécessaires pour la mutagenèse par "overlap extension PCR". Les amorces rouges (2 et 3) sont mutagènes, contrairement aux amorces bleues (1 et 4). B) Schéma du protocole de mutagenèse par "overlap extension PCR". La séquence en rouge représente la mutation introduite. C) Schéma du principe de la mutagenèse par délétion. Le plasmide matrice est en rouge, le plasmide muté est en bleu. D) Présentation des différentes possibilités de mutagenèse réalisables en fonction des amorces utilisées, seule une partie de l'ADN circulaire est représenté.

du tampon TBE $\times 0,5$ (Tris-Borate 100mM pH 8,3 ; EDTA 2mM). Après électrophorèse à 100V dans le tampon TBE $\times 0,5$, le gel est coloré au BET (0,5 $\mu\text{g}/\text{ml}$) et les fragments d'ADN sont visualisés sous lumière UV.

5.2.5 Purification de fragments d'ADN.

Le milieu réactionnel à purifier est soumis à une extraction au phénol-Tris/chloroforme/alcool isoamylique (proportions respectives : 25/24/1) puis l'ADN est précipité par addition de 2 volumes d'éthanol absolu en présence de 200mM de NaCl et de 5 μg de glycogène. Après centrifugation à 4°C pendant 15 min à 15000rpm, le culot d'ADN est lavé à l'éthanol 70 %, séché et remis en suspension dans l'eau.

5.2.6 Mutagenèse dirigée par "overlap extension" PCR

Cette technique a été utilisée pour effectuer des expériences de mutagenèse dirigée. Elle consiste en l'extension de fragments chevauchant par amplification PCR ("overlap extension" PCR, Ho *et al.*, 1989; Horton *et al.*, 1989). Cette technique, que j'ai légèrement adaptée (Figure 5.2 B), nécessite quatre oligonucléotides amorces (de 1 à 4, Figure 5.2 A) dont deux sont mutagènes (2 et 3) et deux étapes successives d'amplification par PCR. Les amorces 2 et 3 possèdent une partie complémentaire chevauchante contenant la mutation que l'on veut introduire. Dans la première étape, deux amplifications par PCR sont réalisées avec les couples d'amorce 1+3 et 2+4 respectivement. Les deux fragments d'ADN obtenus, qui possèdent une région commune, sont soumis à une digestion par l'enzyme *DpnI*, qui clive l'ADN méthylé (voir paragraphe suivant), afin d'éliminer la matrice plasmidique. Après purification les produits sont utilisés comme matrice dans la seconde étape d'amplification par PCR en présence du couple d'amorce 1+4. Le produit final, généralement digéré par deux enzymes de restriction, peut être introduit dans un vecteur digéré par ces mêmes enzymes ou des enzymes produisant des extrémités compatibles.

Table 5.3 : Composition du milieu réactionnel nécessaire à la ligation de deux molécules d'ADN.

Composants	Volume
Vecteur	50 à 500 ng
Insert (rapport stoechiométrique de 1 pour 3)	50 à 500 ng
Tampon de ligation rapide × 4 (Fermentas)	5 µl
T4 ADN ligase (1 U/µl)	1 µl
H ₂ O qsp	20 µl

5.2.7 Mutagenèse dirigée par délétion

Cette technique qui est schématisée dans la Figure 5.2 C a été proposée initialement par Weiner *et al.* (1994). Les amorces utilisées peuvent permettre plusieurs types de mutagenèses comme l'insertion d'un site unique de restriction, des délétions de toutes tailles et des insertions de séquences pouvant aller jusqu'à 120 nucléotides (Figure 5.2 D). La méthode est basée sur une PCR longue, réalisant l'amplification du plasmide complet, ce qui la limite aux plasmides de petite ou moyenne taille, et nécessite une ADN polymérase avec une bonne activité correctrice comme l'enzyme *Pfu* commercialisée par Stratagen™. La PCR est effectuée dans un milieu réactionnel de 50 µl contenant une concentration 10 fois supérieure en amorces et en dNTP que dans le protocole décrit paragraphe 5.2.1 (page 85). 20 cycles d'amplification comportant une dénaturation de 30 sec à 94°C, l'hybridation des amorces de 30 sec à la température adaptée et l'élongation de 5 min à 72°C, sont effectués. L'ADN néo-synthétisé et les hybrides brin parents/nouveaux sont purifiés puis traités avec l'enzyme de restriction *DpnI* pour digérer les brins provenant de la matrice parentale méthylée. Après purification, l'ADN est phosphorylé, soumis à une ligation intra-moléculaire afin de transformer *E. coli*.

5.2.8 Ligation

La T₄ ADN ligase catalyse la formation d'une liaison phosphodiester entre les extrémités 3' hydroxyle et 5' phosphate des molécules d'ADN. Les fragments purifiés d'ADN peuvent être ligués dans les conditions indiquées dans la table 5.3.

La réaction est incubée 1 h à température ambiante ou une nuit à 4°C pour optimiser son rendement. Après extraction phénolique et précipitation à l'éthanol, les produits de ligation sont repris dans 5 à 10 µl d'eau.

5.2.9 Transformation d'organismes

Bactéries électrocompétentes

Préparation des bactéries électrocompétentes Les bactéries sont cultivées à 37°C dans un litre de milieu LB jusqu'à une DO_{600nm} de 0,5 à 0,8 à partir d'une colonie is-

sue d'un étalement sur milieu minimum. La culture est placée 30 min dans la glace puis, après des centrifugations de 15 min à 4000rpm, le culot bactérien est remis en suspension successivement dans 1 litre d'eau froide, puis dans ½ litre d'eau froide, ensuite dans 20 ml de glycérol 10% et finalement dans 2 à 3 ml de glycérol 10%. Les bactéries ainsi préparées peuvent être congelés dans la carboglace et stockées à -80°C pendant plusieurs mois.

Electrotransformation Pour transformer les bactéries électrocompétentes avec de l'ADN plasmidique, la technique d'électroporation a été utilisée. Cette méthode de transformation très efficace est réalisée à l'aide d'un électroporateur (BioRad) dans une cuve dont les électrodes sont distantes de 2 mm. Le contrôleur d'impulsions est réglé sur 25 µF de capacité et 2,5 kV / cm, la résistance étant ajustée à 200 Ω. 40 µl de bactéries compétentes sont décongelées dans la glace puis ajoutées à 1 µl de produits de ligation. Après passage du courant électrique, les bactéries sont resuspendues dans 0,7 ml de LB, incubées 30 min à 37° C puis étalées sur LB agar sélectif. Après une nuit de croissance à 37° C, les colonies sont analysées selon les protocoles décrits au paragraphe 5.2.10 (page 90).

Levures

La technique de transformation à l'acétate de lithium/polyéthylène-glycol et choc thermique a été utilisée. Chaque transformation nécessite 10 ml de culture en milieu liquide, resuspendue dans 50 µl de LiAc/TE (LiAc 0,1 M ; Tris 10 mM pH 7,5 ; EDTA 1 mM). Toutes les étapes sont effectuées à température ambiante.

Le protocole suivant est utilisé pour 5 transformations. Une préculture d'une nuit de 20 ml effectuée en milieu sélectif ou complet sert àensemencer 50 ml de milieu afin d'obtenir une absorbance à 600 nm de 0,15. La croissance des levures est réalisée sous agitation à 30°C jusqu'à atteindre une absorbance de 0,6. Les cellules sont centrifugées à 5000g pendant 3 min et lavées avec 2 ml d'eau stérile. Les cellules sont transférées en tubes de 2 ml et centrifugées 1 min à 4000g. Ce lavage à l'eau stérile est répété 2 fois puis est suivi de 2 lavages en tampon LiAc/TE. Les levures sont mises en suspension dans 200 µl d'AcLi/TE et réparties à raison de 50 µl dans des tubes de 1,5 ml contenant :

- 1 µl d'ADN plasmidique (0,1 mg / ml)

- 5 µl d'ADN de sperme de saumon (à 10 mg / ml, soniqué et bouilli)
- 350 µl de PEG 4000 40 % (dilué dans le LiAC/TE)

Après mélange par inversion, et les tubes sont incubés 30 min à 30°C. L'ADN plasmidique est introduit par un choc thermique de 20 min à 42°C. Les cellules sont diluées avec 700 µl d'eau et mélangées par inversion. Après centrifugation de 1 min à 4000 g, les cellules sont reprises dans 100 µl d'eau puis étalées sur une boîte de milieu sélectif et incubées 2-3 jours à 30°C.

5.2.10 Analyse des clones recombinants et purification de l'ADN plasmidique

Analyse par PCR

Les colonies recombinantes peuvent être analysées par PCR. Pour cela, les colonies ayant poussé sur milieu gélosé sont prélevées et transférées à la fois sur milieu sélectif solide ou liquide et dans un milieu réactionnel de PCR. La réaction de PCR est directement effectuée en utilisant l'ADN bactérien comme matrice. La culture sert à conserver les clones et amplifier le plasmide ultérieurement.

Préparation de petites quantités d'ADN plasmidique (minipréparation)

1,5 ml de milieu liquide sélectif estensemencé et incubé une nuit à 37°C sous agitation avec une colonie recombinante. L'ADN plasmidique est extrait selon la technique de la lyse alcaline décrite par Sambrook *et al.* (1989). Cette technique permet d'extraire un ADN plasmidique suffisamment pur pour être analysé par digestion enzymatique, par séquençage et servir de matrice pour la transcription *in vitro*.

Purification d'ADN plasmidique (midipréparation)

50 ml de culture sontensemencés par une colonie recombinante et incubés une nuit à 37°C sous agitation. L'ADN est extrait avec le kit JETSTAR (Genomed) selon une technique modifiée de lyse alcaline qui repose sur une purification sur colonne de silice afin d'obtenir un ADN de qualité en quantité suffisante.

Analyse de l'ADN plasmidique par séquençage automatique

L'analyse par séquençage d'ADN est basée sur la méthode de synthèse interrompue (Sanger *et al.*, 1977), méthode qui a été modifiée pour la Taq ADN polymérase (Murray, 1989) et adaptée à l'utilisation de didésoxyribonucléotides (ddNTP) couplés à des fluorochromes (Smith *et al.*, 1986). Les séquences sont obtenues par séparation électrophorétique capillaire sur un appareil ABI Prism® 3100 (Applied Biosystem).

5.3 Vecteurs de clonage et d'expression

5.3.1 Vecteurs de clonage

pBluescribe(+/-) Ces plasmides qui dérivent du plasmide pUC19 ont une taille de 3204 pb.

Ils portent le gène de résistance à l'ampicilline, l'origine de répllication du plasmide ColEI, la région intergénique du phage f1 (les deux plasmides diffèrent par l'orientation de cette région) et le gène *lacZ* codant pour le peptide α de la β -galactosidase ainsi qu'une cassette de clonage insérée au début de la séquence codante du gène *lacZ*. De part et d'autre de la cassette se trouvent les séquences promotrices des ARN polymérases des phages T₃ et T₇. Ces promoteurs permettent de synthétiser, en présence de l'ARN polymérase correspondante, les ARN complémentaires au fragment d'ADN inséré.

pGEM®-T Le vecteur pGEM®-T (Promega) a été développé pour cloner des produits de PCR. La cassette de clonage est coupée par *EcoRV* et une thymidine est ajoutée en 3' de chaque brin. Ces nucléotides supplémentaires au site d'insertion améliorent grandement l'efficacité de la ligation de produits de PCR dans le plasmide en évitant la recircularisation du vecteur. De plus, l'utilisation de polymérases thermostables sans activité exonucléasique 3' → 5' permet une compatibilité de séquence entre les thymines du vecteur et les adénines ajoutées naturellement aux extrémités 3' du produit de PCR. Ce vecteur a une taille de 3000 pb, porte le gène de résistance à l'ampicilline, la région intergénique du phage f1 et le gène *lacZ* codant pour le peptide α de la

β -galactosidase. On trouve de part et d'autre de la cassette de clonage les promoteurs des ARN polymérase T₇ et SP6.

5.3.2 Vecteurs d'expression *in vivo*

pCK-EGFP D'une taille de 3508 pb, ce vecteur dérivé du pUC9 est utilisé dans les expériences de bombardement de cellules BY-2. Il possède l'origine de répllication d'*E. coli* ainsi que le gène de résistance à l'ampicilline. En amont du gène codant pour l'EGFP se trouve le promoteur 35S suivi de la séquence Ω du TMV qui est un activateur de traduction (Sleat *et al.*, 1987), la séquence de l'EGFP est délimitée en aval par le terminateur 35S.

pGBT9 (Clontech®) Ce vecteur de 5,5 kb a été utilisé pour mettre en évidence l'activation de la transcription des protéines virales en système simple hybride dans la levure. Les domaines protéiques ont été clonés en fusion avec le domaine de fixation aux acides nucléiques (ou GAL4BD) du gène *gal4* présent dans ce plasmide. Les protéines de fusion sont exprimées par l'intermédiaire du promoteur constitutif de l'alcool déshydrogénase. La protéine de fusion est importée dans le noyau par l'intermédiaire d'un signal de localisation nucléaire intrinsèque au GAL4BD de la protéine GAL4. Ce plasmide est un vecteur navette qui se réplique de façon autonome aussi bien dans la levure (origine de répllication 2 μ) que dans *E. coli*. De ce fait, il porte un gène de résistance à l'ampicilline nécessaire à sa sélection dans *E. coli* et un gène de sélection *trp1* permettant aux souches de levures auxotrophe pour le tryptophane et possédant ce plasmide de croître sur milieu minimum dépourvu en cet acide aminé.

pGAD424 (Clontech®) Ce vecteur de 6,6 kb a été utilisé pour mettre en évidence les interactions protéine-protéine entre des protéines virales et cellulaires dans le système double-hybride de levure. Les domaines protéiques ont été clonés en fusion avec le domaine d'activation de la transcription (ou GAL4AD) du gène *gal4*. Les protéines de fusion sont exprimées par l'intermédiaire du promoteur constitutif de l'alcool déshydrogénase. La protéine de fusion est importée dans le noyau par l'intermédiaire d'un signal de localisation nucléaire intrinsèque au GAL4AD de la protéine GAL4. Ce plasmide est un vecteur navette qui se réplique de façon autonome aussi bien dans

la levure (origine de réplication 2μ) que dans *E. coli*. De ce fait, il porte un gène de résistance à l'ampicilline nécessaire à sa sélection dans *E. coli* et un gène de sélection *leu2* permettant aux souches de levures auxotrophe pour la leucine et possédant ce plasmide de croître sur milieu minimum dépourvu en cet acide aminé.

5.3.3 Vecteur de transcription/traduction *in vitro*

pIVEX® (In Vitro Expression, Roche) D'une taille de 3550 pb, ces vecteurs sont utilisés pour surexprimer des protéines dans un système procaryotique où la transcription et la traduction sont couplées. Ce système est commercialisé par la société Roche® sous le nom de RTS® (Rapid Translation System). Le vecteur pIVEX® 2.3 que nous utilisons pour surexprimer les protéines possède l'origine de réplication d'*E. coli* ainsi que le gène de résistance à l'ampicilline. En amont de la cassette de clonage se trouve un promoteur T₇ suivi d'une région RBS (Ribosome Binding Site) permettant la fixation des ribosomes. La cassette de clonage est délimitée en aval par le terminateur T₇.

5.4 Traitement des séquences

EMBOSS La suite logicielle EMBOSS (European Molecular Biology Open Software Suite) peut s'obtenir sur le site internet dédié et fournit les utilitaires de base pour l'exploitation des données de séquence.

Clustalw Le logiciel CLUSTALW (Thompson *et al.*, 1994) a été utilisé pour effectuer les alignements multiples des séquences nucléotidiques et protéiques.

Phylip La suite logicielle PHYLIP (Felsenstein, 1989) m'a permis de produire les données de reconstructions phylogénétiques.

SplitsTree La production d'arbres phylogénétiques a été effectuée à l'aide du logiciel SplitsTree4 (http://www-ab.informatik.uni-tuebingen.de/software/jsplits/welcome_en.html), développé par Daniel H. Huson et David Bryant (article en préparation).

Netprimer Logiciel disponible sur internet de Premier Biosoft (<http://www.premierbiosoft.com/netprimer/index.html>) permettant le calcul de toutes les caractéristiques des amorces utilisées entre autre pour la PCR.

Infobiogen Site internet du centre national de ressources informatiques appliquées à la génomique : <http://www.infobiogen.fr/>. Accès aux banques de données biologiques, à des ressources d'analyse et de traitement et à une documentation en ligne des outils de la bioinformatique.

5.5 Organismes

5.5.1 Virus

La souche helper (isolat B) contenant les ARN-1 et -2, appelée Stras12, provient d'ARN totaux issus de lésions locales de *C. quinoa* obtenues à partir d'inoculations mécaniques de transcrits des clones complets des ARN-1 et -2 (Quillet *et al.*, 1989). L'isolat d'origine ayant permis le clonage des ADNc complets, l'isolat français F2, provient de betteraves virosées de la région de Muret (Putz and Kuszula, 1978).

5.5.2 Bactéries

La souche MC1022 d'*E. coli* [*araD139* Δ (*ara leu*)7697 Δ (*lacZ*)M15 *galU galK rpsL*] (Casadaban and Cohen, 1980) est principalement utilisée pour les clonages et l'amplification de plasmide. Elle permet entre autre la sélection de type blanc/bleu des clones recombinants par la couleur en présence d'IPTG et de X-gal, un substrat chromogène de la β -galactosidase.

5.5.3 Plantes hôtes

Les plantes hôtes du BNYVV, utilisées au laboratoire, font partie principalement de la famille des *Chénopodiacées*.

Beta vulgaris Plante hôte du BNYVV répondant à l'infection virale par l'apparition de

Table 5.4 : Génotypes des souches de levures utilisées lors des tests de simple et double hybride.

Souche	Génotype	gène rapporteur	marqueur de transformation
HF7c	<i>MATα</i> ; <i>ura3-52</i> , <i>his3-200</i> , <i>ade2-101</i> , <i>lys2-801</i> , <i>trp1-901</i> , <i>leu2-3</i> , <i>112</i> , <i>gal4-542</i> , <i>gal80-538</i> , <i>LYS2::GAL1_{UAS}-GAL1_{TATA}-HIS3</i> , <i>URA3::GAL4_{17-mers(x3)}-CYC1_{TATA}-lacZ</i>	HIS3, lacZ	<i>trp1</i> , <i>leu2</i> , <i>cyh^r2</i>
Y187	<i>MATα</i> , <i>ura3-52</i> , <i>his3-200</i> , <i>ade2-101</i> , <i>trp1-901</i> , <i>leu2-3</i> , <i>112</i> , <i>gal4Δ</i> , <i>met⁻</i> , <i>gal80Δ</i> , <i>URA3::GAL1_{UAS}-GAL1_{TATA}-lacZ</i> , <i>MEL1</i>	lacZ, MEL1	<i>trp1</i> , <i>leu2</i>
CG-1945	<i>MATα</i> , <i>ura3-52</i> , <i>his3-200</i> , <i>ade2-101</i> , <i>lys2-801</i> , <i>trp1-901</i> , <i>leu2-3</i> , <i>112</i> , <i>gal4-542</i> , <i>gal80-538</i> , <i>cyh^r2</i> , <i>LYS2::GAL1_{UAS}-GAL1_{TATA}-HIS3</i> , <i>URA3::GAL4_{17-mers(x3)}-CYC1_{TATA}-lacZ</i>	HIS3, lacZ	<i>trp1</i> , <i>leu2</i> , <i>cyh^r2</i>
AH109	<i>MATα</i> , <i>trp1-901</i> , <i>leu2-3</i> , <i>112</i> , <i>ura3-52</i> , <i>his3-200</i> , <i>gal4Δ</i> , <i>gal80Δ</i> , <i>LYS2::GAL1_{UAS}-GAL1_{TATA}-HIS3</i> , <i>MEL1</i> , <i>GAL2_{UAS}-GAL2_{TATA}-ADE2</i> , <i>URA3::MEL1_{UAS}-MEL1_{TATA}-lacZ</i>	HIS3, ADE2, lacZ, MEL1	<i>trp1</i> , <i>leu2</i>

lésions locales chlorotiques vertes (ARN-1 et -2 seuls) à chlorotiques jaunes, voire nécrotiques selon l'isolat viral inoculé

Beta macrocarpa Plante hôte du BNYVV chez laquelle l'infection est systémique en présence d'ARN-3 combiné aux ARN-1 et -2. Les symptômes foliaires de mosaïques, gaufrage et jaunissement y sont caractéristiques.

Spinacea oleracea Plante hôte systémique du BNYVV.

Chenopodium quinoa Plante hôte du BNYVV répondant à l'infection virale par l'apparition de lésions locales chlorotiques vertes (ARN-1 et -2 seuls), jaunes (ARN-1 à -3 présents) ou nécrotiques (ARN-1 à -5 présents) sur les feuilles inoculées.

Tetragonia expansa Plante hôte du BNYVV, de la famille des *Aizoaceae*, répondant à l'infection virale par l'apparition de lésions locales visibles uniquement sous UV (ARN-1 et -2 seuls), jaunes (ARN-1 à -3 présents) ou nécrotiques (ARN-1 à -5 présents) sur les feuilles inoculées.

5.5.4 Levures (*Saccharomyces cerevisiae*)

Les différentes souches utilisées lors des tests de simple et double-hybride sont présentées dans la table 5.4 avec leur génotype. Toutes ces souches sont auxotrophes pour le tryptophane, la leucine et l'histidine.

HF7C Souche de levure utilisée pour les systèmes simple et double hybride, leur signe sexuel est *MATa* et leurs gènes rapporteurs sont *HIS3* et *lacZ*.

Y187 Souche utilisée pour le système double hybride, leur signe sexuel est *MAT α* et leurs gènes rapporteurs sont *lacZ* et *MEL1*.

CG1945 Souche utilisée pour le système double hybride, leur signe sexuel est *MATa* et leurs gènes rapporteurs sont *HIS3* et *lacZ*.

AH109 Souche utilisée pour le système double hybride, leur signe sexuel est *MATa* et leurs gènes rapporteurs sont *HIS3*, *ADE2*, *lacZ* et *MEL1*.

5.6 Protéines

5.6.1 Analyse des protéines

Protéines végétales et de cellules BY-2

Pour réaliser l'analyse de protéines provenant de feuilles infectées, 3 disques de 6 mm de diamètre contenant des lésions locales sont broyées dans 100 µl de tampon de dissociation (Tris-HCl 125 mM pH 6,8 ; SDS 4 % ; β -mercaptoéthanol 10 % ; glycérol 20 % et Bleu de Bromo-Phénol 0,05 % ; selon Laemmli, 1970). Les cellules BY-2 de tabac transfectées sont sédimentées 2 min à 800 rpm et le culot est repris dans 50 à 100 µl de tampon de dissociation. Les protéines issues de plantes ou de cellules BY-2 de tabac sont dénaturées par chauffage à 95°C pendant 3 min et, après élimination des débris cellulaires par centrifugation, sont déposées sur un gel de polyacrylamide dénaturant et analysées par western-blot.

Protéines de levures

Pour chaque souche transformée nous avons utilisé la méthode d'extraction par Urée/SDS (Printen and Sprague, 1994). Une culture saturée de 5 ml dans un milieu sélectif approprié est utilisée pour ensemercer 50 ml de milieu complet. Pendant la croissance, des mesures d'absorbance à 600 nm sont effectuées jusqu'à atteindre 0,4-0,6. La culture est alors rapidement refroidie dans la glace et centrifugée 5 min à 1000 g à 4°C. Le culot est lavé deux fois par 50 ml d'eau froide puis congelé dans l'azote liquide et stocké à -80°C. Le tampon de dissociation stock (Urée 8 M ; SDS 5 % ; Tris-HCl 40 mM pH 6,8 ; EDTA 0,1 mM ; Bleu de bromophénol 0,4 mg / ml) est utilisé pour préparer extemporanément le tampon final en ajoutant 10 µl de β -mercaptoéthanol par ml de tampon stock ainsi qu'une aliquote d'un cocktail inhibiteur de protéase (Roche, "Protease inhibitor cocktail tablets"). Le tampon de dissociation est préchauffé à 60°C et sert à décongeler rapidement le culot (150 µl de tampon de dissociation par culot). Le mélange est transféré dans des micro-tubes de 1,5 ml contenant l'équivalent de 100 µl de billes de verre (425 – 600 µm, Sigma). Les échantillons sont chauffés 10 min à 70°C puis vortexés vigoureusement pendant 1 min. Les débris cellulaires sont sédimentés par une centrifugation de 10 min à 14000 rpm à 4°C. Le

Table 5.5 : Composition des gels de polyacrylamide pour l'analyse des protéines en conditions dénaturantes. APS : persulfate d'ammonium.

Composants	Gel de concentration	Gel de résolution				
		8 %	10 %	11 %	12,5 %	15 %
Acrylamide	4 %	8 %	10 %	11 %	12,5 %	15 %
bisacrylamide	0,1 %	0,20 %	0,25 %	0,27 %	0,31 %	0,37 %
Tris HCl pH 6,8	0,125 mM			-		
Tris HCl pH 8,8	-			0,375 mM		
SDS	0,1 %			0,1 %		
TEMED	0,2 %			0,1 %		
APS 25 %	0,2 %			0,1 %		

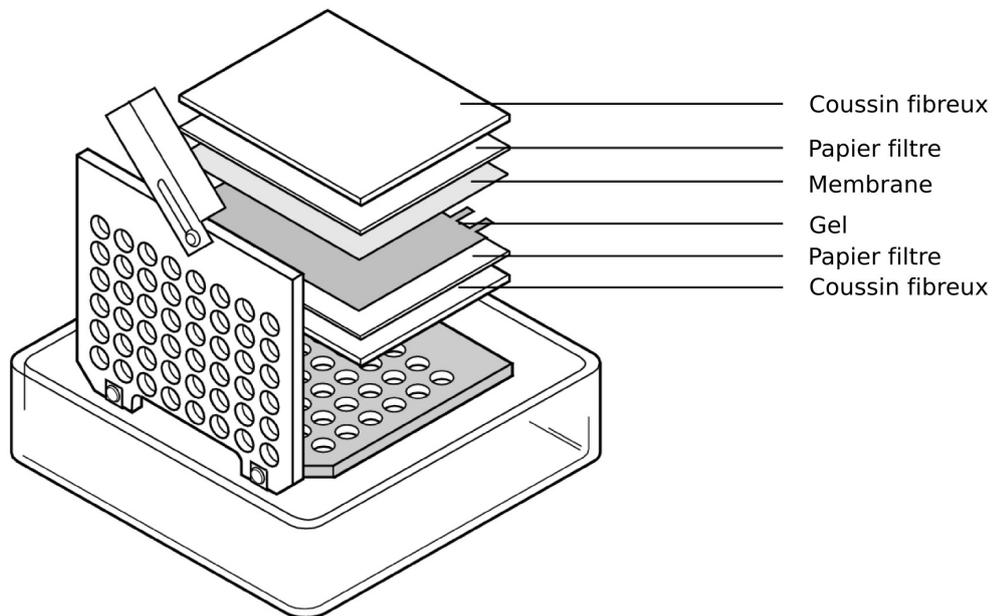


Figure 5.3 : Schéma du montage pour l'électrotransfert de protéines sur membrane après électrophorèse en gel de polyacrylamide. Extrait du manuel du "mini trans-blot cell" de Biorad utilisé au laboratoire.

surnageant récupéré est analysé sur gel de polyacrylamide en conditions dénaturantes (paragraphe 5.6.2) ou conservé à -80°C .

5.6.2 Immunodétection (Western-blot)

La technique consiste à séparer les protéines en gel de polyacrylamide dénaturant puis à les transférer sur membrane ImmobilonTM-PVDF (Millipore) pour la détection des protéines d'intérêt à l'aide d'anticorps spécifiques.

Electrophorèse en gel de polyacrylamide en conditions dénaturantes

Les protéines sont séparées par électrophorèse en gel de polyacrylamide en présence de SDS. On utilise un système de gels et de tampons discontinus, décrit par Laemmli (1970), qui permet de concentrer les échantillons en une bande très fine dans un "gel de concentration" puis de séparer les protéines selon leur masse moléculaire dans un "gel de séparation". La table 5.5 donne la composition des gels, dont le pourcentage varie en fonction de la taille des protéines à séparer.

Après une électrophorèse à 100V dans le tampon de migration (Tris-HCl 25 mM pH 8,3 ; glycine 192 mM ; SDS 0,1 %) le gel peut être coloré au bleu de Coomassie (Coomassie Brilliant Blue 0,25 % ; Méthanol 45 % ; acide acétique 10 % ; H₂O qsp) ou les protéines transférées sur membrane ImmobilonTM-PVDF.

Transfert sur ImmobilonTM-PVDF

Les protéines séparées sur gel dénaturant sont transférées sur membrane ImmobilonTM-PVDF par électrotransfert. Le gel et la membrane sont pris en "sandwich" entre deux papiers Whatman 3MM et deux coussins fibreux. Ce sandwich est maintenu par un système rigide en plexiglas (Figure 5.3). Les protéines sont transférées par l'application d'un courant continu de 500mA pendant 1h30min à 4°C dans du tampon de transfert (Tris-HCl 25mM pH 8,3 ; glycine 192mM).

Immunorévélation

Suite au transfert, la membrane est lavée en (PBS pH 7,4 + Tween20 1%) puis saturée en (PBS pH 7,4 + lait en poudre ou BSA 5% + Tween20 1%) pour occuper les sites de fixation libres de la membrane par des protéines de lait ou de la BSA. Ceci afin d'éviter, par la suite, la fixation aspécifique des immunoglobulines sur la membrane. Elle est ensuite incubée durant 12 h à 4°C en PBS/Tween/lait additionné d'anticorps dirigés contre la protéine recherchée. Après 3 lavages en PBS/Tween, on procède à une nouvelle incubation dans le tampon de saturation pendant 2 h à température ambiante en présence d'anticorps secondaire, dirigés contre les anticorps primaires, couplés par exemple à la peroxydase. L'excès d'anticorps secondaires et les interactions aspécifiques sont éliminés par 3 lavages en PBS/Tween, les complexes antigènes / anticorps sont révélés par chimiluminescence en utilisant le kit Lumi-Light^{PLUS} (Roche™). Pour cela la membrane est déposée sur 0,5 ml de substrat, et après 5 min de réaction à température ambiante, la membrane est placée sous autoradiographie pendant un temps pouvant varier de 5 sec à 15 min selon l'intensité du signal.

5.6.3 Interaction protéine-protéine dans le système double-hybride de levure

Pour rechercher l'existence d'interaction entre protéines, nous avons utilisé le système du "double hybride" qui est une méthode génétique pour tester l'interaction entre protéines dans la levure *Saccharomyces cerevisiae* (Fields and Song, 1989; Bartel *et al.*, 1993). Le principe de ce test repose sur la possibilité de scinder l'activité du facteur de transcription GAL4 en deux domaines : le domaine de fixation à l'ADN et le domaine d'activation de la transcription. Ces deux domaines gardent leurs propriétés et, lorsqu'ils sont rassemblés par une interaction, il y a reconstitution d'un facteur de transcription fonctionnel permettant l'expression d'un ou de plusieurs gènes rapporteurs. Le domaine de fixation à l'ADN de GAL4 est fusionné à une protéine X alors que le domaine d'activation de la transcription est fusionné à une protéine Y supposée interagir avec la protéine X. S'il existe une interaction entre les protéines X et Y, GAL4 sera reconstitué et l'on pourra observer l'expression du

ou des gènes rapporteurs. Dans certaines situations, la protéine d'intérêt X fusionnée au domaine de fixation à l'ADN de GAL4 peut recruter à elle seule les facteurs de transcription et induire la transcription du gène rapporteur sans nécessiter le domaine GAL4AD. Cette capacité à recruter les facteurs de transcription peut être due à une activité intrinsèque de la protéine (par exemple dans le cas d'activateurs de transcription, comme la protéine TIP d'*A. thaliana* selon Ren *et al.*, 2000). Une telle protéine ne permet donc pas de rechercher d'interactant par le système double hybride. Cependant, cette propriété peut être utilisée comme test "simple hybride" pour rechercher et caractériser les domaines fonctionnels des protéines activatrices de transcription.

La mise en pratique de ces tests est très simple. Pour un test d'interaction à petite ou moyenne échelle (de 2 à 15 partenaires) nous avons utilisé la méthode du "yeast mating". Pour le criblage d'une banque d'ADNc nous avons utilisé la méthode de co-transformation.

Méthode de test double-hybride par "yeast mating"

C'est une méthode utilisée pour introduire deux plasmides dans la même cellule hôte (Bendixen *et al.*, 1994). Le pré-requis pour ce protocole est de disposer de transformants de signe sexuel *MATa* possédant les plasmides dérivés de pGBT9 et de souches de signe sexuel *MAT α* contenant les plasmides dérivés de pGAD424. Une souche de signe *MATa* contenant un plasmide pGBT9 et une souche de signe *MAT α* contenant un plasmide pGAD424 sont cultivées ensemble et peuvent fusionner pour former une cellule diploïde contenant les deux plasmides et ainsi exprimer les protéines de fusion. Le test d'interaction double-hybride peut être ainsi mis en place par simple sélection des levures diploïdes.

La culture s'effectue dans 140 μ l de milieu YPD liquide, en plaques de 96 puits stériles. Pour chaque plasmide d'intérêt, quelques colonies transformées sont placées dans un micro-tube contenant 100 μ l d'eau stérile puis sont vortexé jusqu'à homogénéisation complète. Une fraction aliquote de 10 μ l de chaque souche est inoculée dans chaque puits. La plaque est recouverte d'un couvercle stérile et incubée dans un agitateur à 30°C et à 200 rpm pendant 6 à 18h. Les cellules diploïdes sont ensuite sélectionnées parallèlement sur milieu gélosé SD dépourvu de tryptophane et leucine (SD-TL) et SD dépourvu de tryptophane, leucine et histidine (SD-TLH). La croissance sur milieu SD-TL permet de visualiser la

présence de diploïdes et donc l'efficacité du mating, tandis que celle sur SD-TLH mesure l'activation du gène rapporteur *His3* et donc l'interaction entre les protéines d'intérêt.

Méthode de test double-hybride par co-transformation

Il est nécessaire d'obtenir au préalable une souche de levure AH109 contenant le plasmide pGBT9 exprimant la protéine d'intérêt (voir paragraphe 5.2.9, page 89). Cette souche est ensuite utilisée pour le criblage d'une banque d'ADNc d'*Arabidopsis thaliana* (Clontech®). Pour une bonne couverture de la banque, il est nécessaire de réaliser 10 transformations en utilisant le protocole décrit ci-dessous.

Une culture de 500ml de milieu complet (YPD) estensemencée pour que la mesure d'absorbance à 600nm atteigne 0,15 à 0,2. La culture est mise à pousser sous agitation à 30°C jusqu'à atteindre une mesure d'absorbance de 0,3 à 0,4. Les cellules sont alors sédimentées 10 min à 2000rpm et le culot repris dans 100ml de LiAc/TE (LiAc 0,1 M ; Tris 10mM pH 7,5 ; EDTA 1 mM). Après une incubation de 25 à 30 min à température ambiante, les cellules sont centrifugées pendant 10 min à 2000rpm, le culot est resuspendu dans 10ml de LiAc/TE.

Chaque transformation est réalisée en tube de 15 ml où sont mélangés :

- 1 ml de cellules
- 50µl de banque d'ADNc (1 µg / µl)
- 100µl d'ADN de sperme de saumon (à 10mg / ml, soniqué et bouilli)

Après 10 min d'incubation à température ambiante, 2,8ml de PEG 40% sont ajoutés et les cellules sont incubées 1 h à 30°C. 430µl de DMSO sont ajoutés avant d'effectuer un choc thermique de 20 min à 42°C. Les cellules sont refroidies 3 min dans la glace et centrifugées 7 min à 1500rpm. Les culots sont repris dans 30 ml de YPD et incubés 1 h à 30°C. Après centrifugation, les cellules sont lavées par 20 ml d'eau bidistillée stérile avant d'être mises en suspension dans 20 ml de milieu triplement sélectif SD dépourvu en histidine, tryptophane et leucine (SD-TLH). Les cellules sont laissées une nuit à 4°C. Après une centrifugation de 10 min à 2000rpm les cellules sont reprises dans 10 ml d'eau bidistillée stérile et sont étalées sur 25 boîtes de milieu gélosé SD-TLH. Ces étalements sont incubés pendant 3-4 jours à 30°C. Les colonies qui apparaissent sont repiquées, pour sauvegarde, sur

un milieu identique et sur un milieu gélosé SD dépourvu en histidine, tryptophane, leucine et adénine (SD-TLHA), plus sélectif. Ce dernier milieu permet d'éliminer les faux positifs et ainsi de sélectionner les levures exprimant des protéines recombinantes interagissant fortement.

5.6.4 Activation de la transcription dans le système simple-hybride de levure

Pour caractériser la capacité d'auto-activation de la transcription dans la levure des protéines ou de domaines de protéines fusionnés au GAL4BD, nous avons observé la croissance de levures HF7C transformées avec les plasmides pGBT9 recombinants. La croissance sur milieu gélosé SD-T permet de s'assurer de la présence du plasmide dans la souche de levure et une croissance sur milieu gélosé SD-TH permet de visualiser l'activation de la transcription du gène rapporteur *his3* par la protéine chimérique.

5.6.5 Activité ONPG (O-NitroPhényl- β -D-Galactopyranoside)

2 ml d'une culture de levures saturée effectuée en milieu SD sélectif sont utilisés pour ensemercer 8 ml de milieu complet YPD. Parallèlement, une solution d'ONPG à 4 mg/ml est préparée en tampon Z-buffer ($\text{Na}_2\text{HPO}_4 - 7\text{H}_2\text{O}$ 16,1 g/l; $\text{NaH}_2\text{PO}_4 - \text{H}_2\text{O}$ 5,50 g/l; KCl 0,75 g/l; $\text{MgSO}_4 - 7\text{H}_2\text{O}$ 0,246 g/l). Après 3 à 5 h de croissance à 28°C sous agitation (220 – 250 rpm) jusqu'à une mesure d'absorbance à 600 nm de 0,5 à 0,8, les cellules sont sédimentées par fraction aliquote de 1,5 ml puis lavées 2 fois par 1,5 ml de Z-buffer et finalement remis en suspension dans 300 μl de Z-buffer. Les cellules en suspension (100 μl) sont successivement congelées dans l'azote liquide et rapidement décongelées à 37°C. L'opération est répétée 3 fois pour maximiser la lyse des cellules. Le lysat est additionné de 700 μl de Z-buffer complété de β -mercaptoéthanol (0,27%). Au temps 0 de l'expérience, 160 μl d'ONPG (dissous extemporanément dans le Z-buffer) sont ajoutés au lysat et le mélange est incubé à 30°C jusqu'à apparition d'une couleur jaune. La réaction est stoppée par ajout de 0,4 ml de Na_2CO_3 1 M. Les débris cellulaires sont sédimentés 10 min à 14000 rpm et l'absorbance à 420 nm du surnageant est mesurée. On

calcule les unités enzymatiques de β -galactosidase (une unité est la quantité d'enzyme qui hydrolyse 1 μ mole d'ONPG en O-nitrophénol et D-galactose par minute et par cellule) par la formule : $1000 \cdot DO_{420} / (t \cdot V \cdot DO_{600})$ où

- t = temps écoulé durant l'incubation (en minutes)
- V = 0,1 ml · 5
- $DO_{600} = A_{600}$ pour 1 ml de culture

5.6.6 Activité β -galactosidase dans l'essai "colony lift"

Un filtre de nitrocellulose est déposé sur chaque boîte de Petri pour y transférer les colonies par capillarité. Après imbibition du filtre, celui-ci est congelé dans l'azote liquide. Chaque filtre de nitrocellulose est immédiatement placé sur un papier Whatmann imbibé de solution de Z-buffer/X-gal (β -mercaptoéthanol 0,27 % ; X-gal 33,4 mg/100 ml). Cette décongélation rapide perméabilise les levures et permet l'entrée du substrat. Les membranes sont incubées à 30°C jusqu'à apparition de la couleur bleue.

5.7 Transfection de cellules BY-2 de tabac par biolistique

5.7.1 Principe

Cette technique repose sur la capacité à introduire dans des cellules, par biolistique, de l'ADN ou de l'ARN adsorbé sur des particules de tungstène ou d'or. L'ADN utilisé doit comporter un promoteur et un terminateur de transcription de part et d'autre de la séquence à exprimer.

5.7.2 Cultures de cellules BY-2 de tabac

Les cellules de tabac (*Nicotiana tabacum* L., souche Bright Yellow 2 (BY-2), voir Nagata *et al.*, 1992) sont maintenues en culture liquide, à 27°C, sous agitation constante (130rpm) et en absence de lumière, dans du milieu Murashige et Skoog (Duchefa), com-

plémenté avec de la thiamine (1 mg/ml), du KH_2PO_4 (200 mg/ml), de l'acide 2,4-dichlorophénoxyacétique (2,4-D, à 0,2 mg/ml), du saccharose (3% poids/volume) et du myoinositol (100 mg/ml). Les cellules sont repiquées tous les sept jours par dilution de 1,5 ml de culture saturée dans 100 ml de milieu frais.

5.7.3 Préparation des cellules pour le bombardement

5 à 10 ml de cellules de tabac BY-2 d'une culture de 4 jours sont déposées sur un disque de papier filtre (diamètre de 3,5 cm) par légère aspiration. Le filtre est ensuite placé dans une boîte de Petri contenant du milieu de culture BY-2 additionné de 1,5% d'agar et 100 mM de mannitol et sorbitol. Les cellules sont plasmolysées par incubation de 2h à 24°C et à l'obscurité.

5.7.4 Précipitation de l'ADN plasmidique sur les particules de tungstène et bombardement

2 milligrammes de billes de tungstène sont mélangées à 1 ml d'éthanol absolu puis vortexées en continu pendant 3 min et laissées 30 min à température ambiante. Les billes sont alors séchées 5 min sous vide et resuspendues dans 33 μl de glycérol 50%, puis tout en vortexant, 10 μl d'ADN (1 $\mu\text{g}/\mu\text{l}$), 33 μl de CaCl_2 2 M et 15 μl de spermidine 0,5 M sont rajoutés. Le tout est incubé 30 min à température ambiante, centrifugé 1 min à 5000 g et lavé successivement par 150 μl d'éthanol 70% et 150 μl d'éthanol absolu, avant d'être finalement repris dans 30 μl d'éthanol absolu. Les particules sont conservées à 4°C et sont prêtes pour le bombardement. 6 μl de billes en suspension sont utilisées par tir. 4 bombardements répartis sur 2 boîtes de Pétri contenant les cellules BY-2 plasmolysées sont effectués pour chaque construction. Après bombardement, les filtres supportant les cellules bombardées sont transférés sur de nouvelles boîtes de Pétri avec un milieu de culture BY-2 et incubés de 6 à 24 h à 24°C et à l'obscurité.

5.7.5 Matériel

L'appareil utilisé, un canon à particules, est constitué d'une chambre de tir dans laquelle un vide partiel est créé (0,8 – 0,9 bar). Les billes de tungstène sont placées sur une grille métallique faisant partie de la buse située dans la chambre de tir. La buse est reliée à une bouteille d'hélium par l'intermédiaire d'une électro-vanne qui permet de libérer une impulsion de gaz de durée et de pression déterminées (25 ms, 7 bars respectivement). Le flux de gaz projette les particules de tungstène sur les cellules BY-2 placées à 11 cm de la buse sur un filtre de papier Whatman, déposé sur un milieu solide dans une boîte de Petri.

5.7.6 Application à l'analyse de protéines fusionnées à l'EGFP

Nous avons utilisé le vecteur pCK qui possède la promoteur 35S et le terminateur NOS entre lesquels a été placé le gène codant pour l'EGFP fusionné ou pas au gène codant pour la protéine d'intérêt. 6 à 24 h après bombardement, l'expression du gène codant pour la EGFP induit la fluorescence des cellules sous lumière UV (lampe à vapeur de mercure HBO103W, émission/excitation 488/505 – 545 nm) et permet leur prélèvement sous une loupe binoculaire. Ces cellules sont ensuite observées au microscope confocal.

5.8 Localisation *in situ* de protéines virales par microscopie

5.8.1 Préparation des échantillons et observation au microscope confocal

Préparation des échantillons

Echantillons de feuilles infectées Un morceau de feuille présentant des lésions locales est déposée entre lame et lamelle avec une goutte d'eau.

Echantillons de cellules BY-2 Les cellules BY-2 de tabac transformées par des constructions fusionnées à la EGFP sont sélectionnées à l'aide d'une loupe binoculaire couplée à une lampe à vapeur de mercure HBO103W (émission/excitation 488/505 –

Table 5.6 : Protocoles de fixation et d'inclusion d'échantillons pour l'observation d'immunomarquages par microscopie électronique.

Inclusion en résine EPON	Inclusion en résine LR-White	Durée	Température
Glutaraldéhyde 2 % ; Tampon phosphate pH 7,2		2 h dont infiltration sous vide de 15 min	4°C
Glutaraldéhyde 2 % ; acide picrique 10 %		16 h	4°C
OsO ₄ 0,2 % ; tampon phosphate pH 7,2		2 h	Température ambiante
Acétate d'uranyle 2 %		16 h	4°C
Lavages H ₂ O		4 de 10 minutes	Température ambiante
Déshydratations à l'éthanol : 30 % ; 50 % ; 70 %		10 min chacun	Température ambiante
Déshydratations à l'éthanol : 80 % ; 90 % ; 100 %		20 min chacun + 20 min dans l'éthanol 100 %	Température ambiante
Lavage à l'oxyde de propylène		20 min	Température ambiante
Lavage à l'oxyde de propylène/EPON (v/v)	Lavage à l'éthanol 100 %/LR-White (v/v)	1 h	Température ambiante
Lavage à l'EPON 100 %	Lavage au LR-White 100 %	1 h	Température ambiante
Pré-inclusion à l'EPON 100 %	Pré-inclusion au LR-White	16 h	20°C
Lavage à l'EPON	Lavage au LR-White	1 h	Température ambiante
Inclusion dans l'EPON	Inclusion dans le LR-White	72 h	60°C

545nm). Les cellules sélectionnées sont déposées entre lame et lamelle avec une goutte de milieu BY-2 liquide.

Observation au microscope confocal

Les observations en microscopie confocale ont été effectuées sur un microscope inversé (Axiovert) à balayage laser LSM 510 (Zeiss) équipé d'un laser à argon pouvant émettre à 488nm pour exciter la EGFP. La lumière émise par la EGFP est filtrée par un "band-pass" (BP) 505 – 530nm. L'acquisition des images est réalisée à l'aide du logiciel LSM 510 de Zeiss. L'autofluorescence des chloroplastes des feuilles observées est visualisée par l'observation des rayons réémis après excitation des pigments chlorophylliens par l'absorption d'une raie à 514nm émise par le laser à hélium, filtrés par un "long-pass" (LP) 650nm.

5.8.2 Préparation des échantillons et observations au microscope électronique

Fixation et inclusion des échantillons

La table 5.6 présente le mode opératoire utilisé pour la fixation de fragments de feuilles et leur inclusion en résine. Les lésions locales, découpées directement dans le glutaraldéhyde 2%, sont fixées successivement par le glutaraldéhyde mis en présence d'acide picrique, contrastées par l'osmium et l'acétate d'uranyle et progressivement déshydratées avant d'être incluses dans la résine.

Obtention de coupes ultrafines

Des coupes ultrafines d'une épaisseur de 90nm sont réalisées à l'aide d'un ultramicrotome Ultracut E Reichert-Jung, récoltées par flottaison sur de l'eau et placées sur des grilles de microscopie électronique recouvertes d'un film de Formvar (0,25% dans du 1,2-dichloroéthane). Les coupes sont ensuite séchées à l'air.

Immunomarquage à l'or colloïdal sur des coupes ultra-fines

La mise en évidence de la protéine est réalisée grâce à un sérum dirigé contre la protéine à détecter, puis avec un deuxième sérum anti-IgG couplé à des billes d'or colloïdal de 15 à 25 nm de diamètre. Seuls les antigènes présents à la surface des coupes sont accessibles aux anticorps. Les conditions d'incubation (agent bloquant, température d'incubation) sont choisies de manière à obtenir le meilleur rapport entre le marquage spécifique et le bruit de fond. Les traitements sont effectués par flottaison des grilles sur 20 µl de réactif. L'immunodétection est faite avec un anticorps primaire en tampon (PBS ; 1 % de BSA).

Observation des coupes

Les observations des échantillons sont effectuées sous une tension d'accélération de 75 kV, avec un microscope électronique à transmission Hitachi H600.

5.9 Milieux de culture

Ils se présentent sous forme liquide (voir compositions ci-dessous) ou solide. Dans ce dernier cas, 1,2 à 1,5 % d'agarose sont additionnés avant autoclavage. Ils sont ou non, selon les cas, supplémentés d'antibiotiques permettant la sélection des clones recombinants.

5.9.1 Milieu LB

Le milieu de culture le plus couramment utilisé est le milieu complet LB (Luria Bertani) : Bactotryptone 1 %, extrait de levure 0,5 %, NaCl 0,5 %, pH 7,4.

5.9.2 Milieu YPD

Le YPD est un milieu complet permettant la croissance non sélective des levures. Il est fourni sous forme de poudre par BD Biosciences Clontech (référence 630409 pour le milieu liquide, 630410 pour le milieu gélosé).

5.9.3 Milieu SD

Le SD est un milieu minimal permettant la croissance sélective des levures. Il est fourni sous forme de poudre par BD Biosciences Clontech (référence 630411 pour le milieu liquide, 630412 pour le milieu gélosé). Ce milieu est complété par des mélanges d'acides aminés complets sauf pour ceux nécessaires à la sélection et aux tests d'activation de la transcription :

- Dépourvu en tryptophane (référence 630413)
- Dépourvu en tryptophane et leucine (référence 630417)
- Dépourvu en tryptophane, leucine et histidine (référence 630419)
- Dépourvu en tryptophane, leucine, histidine et adénine (référence 630428)

Appendices du chapitre 1

Nom	Origine	Accession
SH1	Japon	AB018621
S44	Japon	AB018624
S42	Japon	AB018622
S12	Japon	AB018623
D104	Japon	AB018620
S	Japon	D84411
R83	Japon	-
T101	Japon	-
S113	Japon	-
O11	Japon	-
CY3	Chine	AB018625
CH2	Chine	AB018626
NM	Chine	S71490
Yao2	Chine	U25668
Kas3	Kazakhstan	AF197556
F75	France	AF197547
F13	France	X014197
FP	Pithiviers, France	AB018628
FC	Corbeille, France	AB018627
EP32A*	Lorcy (Ext. Pith.), France	AY696085
EP38*	Trinay (Ext. Pith.), France	AY696087
EP42*	Charmant (Ext. Pith.), France	AY696089
EP47A*	Rouvre (Ext. Pith.), France	AY696091
EP62*	Briand (Ext. Pith.), France	AY696092
EP2*	Neuville (Ext. Pith.), France	AY696082
EP24*	Chatillon (Ext. Pith.), France	AY696083
EP27*	Manhecourt (Ext. Pith.), France	AY696084
EP32B*	Andonville (Ext. Pith.), France	AY696086
EP37B*	Gironville (Ext. Pith.), France	AY734495
EP39B*	Yèvre la Ville (Ext. Pith.), France	AY696088
EP43B*	Préfontaines (Ext. Pith.), France	AY696090
EP47B*	Sceaux (Ext. Pith.), France	AY696092

Appendice 1.1 : Origine des séquences de l'ARN-2 utilisées dans la phylogénie. Les noms suivis d'une astérisque sont les séquences obtenues par l'INRA de Colmar.

Nom	Origine	Accession
Fb2*	Benfeld, France	AY696094
Fb7*	Benfeld, France	AY696095
H4*	Hanche, France	AY696097
RV1b*	Hanche, France	AY696107
RV2a*	Hanche, France	AY696108
RV2b*	Hanche, France	AY696109
G2*	Büttelborn (Hessen) - Allemagne	AY696096
A2*	Gänserndorf - Autriche	AY696072
A4*	Bruck/Leitha - Autriche	AY696075
B1*	Watervliet - Belgique	En cours
B2*	Mons - Haulchin - Belgique	AY696077
B3*	Momalle - Belgique	AY696079
NL3*	Oosterwaard - Pays-Bas	AY696105
S3*	Villalgorido del Jucar (Albacete) - Espagne	AY696110
S4*	Pozoamargo (Cuenca) - Espagne	AY696111
S5*	Torralba Cva. (Ciudad real) - Espagne	AY696112
S7*	Laguna de Negrillos (León) - Espagne	AY696113
S10*	Palazuelos (Burgos) - Espagne	AY696114
USA311*	Salinas, Imperial Valley, USA	AY696117
USA176*	Salinas, Imperial Valley, USA	AY696115
USA277*	Salinas, Imperial Valley, USA	AY696116
USA412*	Salinas, Imperial Valley, USA	AY696118
1EU*	Salinas, Imperial Valley, USA	AY696070
Ch*	Salinas, Imperial Valley, USA	AY696080
Rsf*	Salinas, Imperial Valley, USA	AY696106

Appendice 1.1 (suite) : Origine des séquences de l'ARN-2 utilisées dans la phylogénie. Les noms suivis d'une astérisque sont les séquences obtenues par l'INRA de Colmar.

Nom	Origine	Accession
F2	France	M36894
F72	Reigneville, France	AF197545
F76	France	AF197549
Hanches	Hanches, France	AY734504
RV1b*	Hanches, France	AY696166
RV2a*	Hanches, France	AY696167
RV2b*	Hanches, France	AY696168
H4*	Hanches, France	AY696156
Fb2*	Benfeld, France	AY696153
Fb7*	Benfeld, France	AY696154
F pith. 85	Pithiviers, France, 1985	AY734503
Pith.	Pithiviers, France	AY734505
NPi	France	-
EP64*	Tagot le Hangar (Ext. Pith.), France	AY696152
EP63*	Bergougnau (Ext. Pith.), France	AY696151
EP62*	Briand (Ext. Pith.), France	AY696150
EP61*	Mignères (Ext. Pith.), France	AY696149
EP47A*	Rouvre (Ext. Pith.), France	AY696147
EP44*	Pithiviers (Ext. Pith.), France	AY696146
EP43A*	Guigneville (Ext. Pith.), France	AY696144
EP42*	Charmant (Ext. Pith.), France	AY696143
EP39A*	Artenay (Ext. Pith.), France	AY696141
EP38*	Trinay (Ext. Pith.), France	AY696140
EP37A*	Villereaux (Ext. Pith.), France	AY696139
EP33*	Ormetrou (Ext. Pith.), France	AY696138
EP32A*	Lorcy (Ext. Pith.), France	AY696136
EP2*	Neuville (Ext. Pith.), France	AY696133
EP24*	Chatillon (Ext. Pith.), France	AY696134
EP27*	Manhecourt (Ext. Pith.), France	AY696135
EP32B*	Andonville (Ext. Pith.), France	AY696137
EP37B*	Gironville (Ext. Pith.), France	AY734501
EP39B*	Yèvre la Ville (Ext. Pith.), France	AY696142
EP43B*	Préfontaines (Ext. Pith.), France	AY696145
EP47B*	Sceaux (Ext. Pith.), France	AY696148
EP-R4	Yèvre la Ville (Ext. Pith.), France	En cours
EP-R31		En cours
EP-R36	Yèvre le Châtel (Ext. Pith.), France	En cours
EP-R13	Rougemenont (Ext. Pith.), France	En cours
EP-R14		En cours

Appendice 1.2 : Origine des séquences de l'ARN-3 utilisées dans la phylogénie. Les noms suivis d'une astérisque sont les séquences obtenues par l'INRA de Colmar.

Nom	Origine	Accession
B1*	Watervliet - Belgique	En cours
B2*	Mons - Haulchin - Belgique	AY696128
B3*	Momalle - Belgique	AY696130
S3*	Villalgorido del Jucar (Albacete) - Espagne	AY696169
S4*	Pozoamargo (Cuenca) - Espagne	AY696170
S5*	Torralba Cva. (Ciudad real) - Espagne	AY696171
S7*	Laguna de Negrillos (León) - Espagne	AY696172
S10*	Palazuelos (Burgos) - Espagne	AY696173
NL3*	Oosterwaard - Pays-Bas	AY696164
N7	Pays-Bas	AF197558
Angl.	Angleterre	-
G2*	Büttelborn - Allemagne (Hessen)	AY696155
A2*	Gänserndorf - Autriche	AY696123
A4*	Bruck/Leitha - Autriche	AY696126
I12	Italie	AF197549
Kas2	Kazakhstan	AF197553
S	Japon	D84412
R83	Japon	-
T101	Japon	-
S113	Japon	-
O11	Japon	-
NM	Chine	AJ239200
Rsf*	Salinas, Imperial Valley, USA	AY696165
Ch*	Salinas, Imperial Valley, USA	AY696132
1EU*	Imperial Valley, usa	AY696119
277USA*	Imperial Valley, usa	AY696120
412USA*	Imperial Valley, usa	AY696121

Appendice 1.2 (suite) : Origine des séquences de l'ARN-3 utilisées dans la phylogénie. Les noms suivis d'une astérisque sont les séquences obtenues par l'INRA de Colmar.

Nom	Origine	Accession
SH1	Japon	D63935
S43	Japon	D63936
R83	Japon	D63937
D5	Japon	D63759
S44	Japon	AB018612
S42	Japon	AB018610
S16	Japon	AB018608
S13	Japon	AB018607
S12	Japon	AB018606
S110	Japon	AB018600
S11	Japon	AB018605
T101	Japon	AB018604
U2	Japon	AB018602
D104	Japon	AB018599
CY3	Chine	AB018617
CY1	Chine	AB018616
CH2	Chine	AB018614
Kas2	Kazakhstan	AF197555
FC	Corbeille, France	AB018618
F72	Reigneville, France	U78293
F28	Moncharville, France	U78292
FP	Pithiviers, France	AB018619
EP-R3	Yèvre la Ville, France	En cours
EP-R4	Yèvre la Ville, France	En cours
EP-R13	Rougemont, France	En cours
EP-R14	Rougemont, France	En cours
EP-R34	Yèvre le Châtel	En cours
F pith.85	Yèvre la Ville, France 1985	AY734506
Hanches	Hanches, France	AY734507
RV1b*	Hanches, France	AY700058
RV2a*	Hanches, France	AY700059
H4*	Hanches, France	AY700057
EP42*	Pithiviers, France	AY700056
EP2*	Neuville, France	AY700052
EP24*	Chatillon, France	AY700053
EP37B*	Gironville, France	AY700054
EP39B*	Yevre la ville, France	AY700055
Angl.	Angleterre	En cours
U3	Japon	ABO18603
CH3	Chine	AB018615
H48	Japon	AB018613
Baotou	Chine	AJ236894

Appendice 1.3 : Origine des séquences de l'ARN-5 utilisées dans la phylogénie.

Appendice du chapitre 3

Use of a beet necrotic yellow vein virus RNA-5-derived replicon as a new tool for the expression of genes.

Laure Schmidlin, Didier Link, Jérôme Mutterer, Hubert Guilley and David Gilmer¹

Institut de Biologie Moléculaire des Plantes, 12 rue du Général Zimmer, 67084 Strasbourg cedex, France.

¹ Corresponding author: david.gilmer@ibmp-ulp.u-strasbg.fr

Running title: BNYVV RNA-5-derived replicon.

Word count:

Summary: 156

Main text: 1974

Number of Figures and Tables: 3

JGV n° 80720

Summary:

A new gene expression system based on beet necrotic yellow vein virus (BNYVV) RNA-5 has been constructed to allow the expression of recombinant proteins in virally infected cells. Replication and expression levels of the RNA-5-based replicon containing the green fluorescence protein (GFP) gene have been compared to those obtained with the well-characterized RNA-3-derived replicon (Rep-3). When RNA-3 and/or RNA-4 BNYVV RNAs were added to the inoculum, the expression levels of RNA-5-encoded GFP were considerably reduced. To a lesser extent, RNA-3-derived GFP expression was also affected by the presence of RNA-4 and -5. Both RNA-3 and RNA-5-derived molecules were able to express proteins within the same infected cells. Together with Rep-3, the RNA-5-derived replicon thus provides a new tool for the co-expression of different recombinant proteins. In *Beta macrocarpa* plant, Rep-5-GFP was able to move in systemic tissues in the presence of RNA-3 and provides a new expression system that is not restricted to the inoculated leaves.

Many phytoviruses have been used for the production of functional proteins in plants based on viral genome amplification within infected cells (for review see Pogue *et al.*, 1998; Gleba *et al.*, 2004; Pogue *et al.*, 2002; Awram *et al.*, 2002; Porta & Lomonossoff, 2002). Using a plant virus as a vector allows rapid expression of post-translationally modified functional proteins.

We have shown previously that *Beet necrotic yellow vein virus* (BNYVV) can be used as such a vector. BNYVV has five genomic RNAs but only RNAs 1 and 2 are required for virus propagation on leaves of *Chenopodium quinoa* (Quillet *et al.*, 1989). Thus, in principle, it should be possible to use one of the other three RNAs as a “replicon”, that is, a BNYVV replication-dependent vehicle for expression of a foreign gene inserted in place of the open reading frame (ORF) for the normally encoded protein. Indeed, BNYVV RNA-3 has already

been so used to express GFP (Erhardt *et al.*, 2000; Vetter *et al.*, 2004) or proteins encoded by BNYVV RNA-2 (Hehn *et al.*, 1995; Bleykasten-Grosshans *et al.*, 1997; Lauber *et al.*, 1998a). However, it has not proven possible to routinely express two foreign proteins from different RNA-3-based replicons; the replicons appear to compete with one another during replication so that one or the other is quickly eliminated from the infection site (Lauber *et al.*, 1999). This prompted us to examine the possibility of using one of the other small BNYVV RNAs as a replicon in the hope that, since RNA-3, -4 and -5 multiply together during normal viral replication, replicons derived from different molecules might co-exist as well.

Preliminary experiments with RNA-4 revealed that it has drawbacks for use as a replicon since it has *cis*-essential replication sequences extending more than 393 nucleotides from the 5'-terminus (Gilmer *et al.*, 1992; and additional observations). We therefore have concentrated on RNA-5, for which a full-length cDNA clone has recently become available (Link *et al.*, manuscript in preparation). In this study we have inserted the ORFs for several proteins such as GFP, the BNYVV RNA-2 proteins P13 and P15 (expressed from a dicistronic construct (Bleykasten-Grosshans *et al.*, 1997, Lauber *et al.*, 1998a) and the RNA-5 protein p26 modified by addition of a hemagglutinin tag, p26-HA several proteins, into an RNA-5 based replicon and compared their expression levels to those obtained with the well characterized RNA-3 replicon (Figure 1A). The effect on the behaviour of the replicons of the addition to the inoculum of various combinations of the other small BNYVV RNAs was also analysed.

Construction of the RNA-5 replicon was performed by replacing the p26-encoded sequence with a linker sequence (*Nco*I, *Eco*RV, *Bg*III and *Xba*I) while retaining the 5' proximal 438 nucleotides of RNA-5 and the 243 nucleotides preceding the RNA-5 polyA tail. The GFP sequence was inserted between the *Nco*I and *Bg*III restriction sites to produce Rep-5-GFP. Dicistronic Rep-5-P13-15 was obtained by inserting an *Nco*I-*Bg*III-digested PCR product,

obtained from the full-length RNA-2 cDNA clone, between the *NcoI* and *BglIII* restriction sites of Rep-5. Rep-3-GFP was from Erhardt *et al.*, 2000. GFP sequence was replaced by mRFP to produce Rep-3-RFP using *NcoI* and *BamHI* restriction sites. *C. quinoa* leaves were inoculated with Stras12 helper strain (i.e. RNA-1 and -2) together with different combination of the small BNYVV RNAs. Local lesions appeared 6 dpi and were analyzed for their phenotype, for GFP fluorescence (when appropriate), and for RNA and protein contents by northern and western blot.

Extended analysis of lesion phenotypes induced by different inocula was performed and is summarized in Table 1. When Stras12 was used alone or supplemented with Rep-3 and/or Rep-5 derived molecules, ~2 mm Stras12-typical green chlorotic local lesions were obtained. The presence of Rep-3-GFP induced the expression of GFP in infected tissues that produced a strong fluorescent ring under ultraviolet illumination. In the case of Rep-5-GFP, fluorescence within local lesions was high in the initially infected tissues (i.e. the centre of the lesion) and lowered in the periphery of the infected tissues.

The effects of the other small replicable RNAs were tested upon the symptom formation and GFP expression levels. When RNA-3 was added to the inocula, ~2 mm yellow chlorotic spots were observed (Table 1) that displayed a faint fluorescence when Rep-5-GFP was added to the inoculum. Lesion fluorescence was lowered when RNA-4 was also added to the inoculum. An identical approach was performed using full-length RNA-5HA. Expression of p26HA induced the necrosis of infected *C. quinoa* tissues and that result in the production of ~1 mm necrotic lesions. Such necrosis was drastically increased in the presence of RNA-4, leading to ~0,5 mm necrotic spots (Table 1). Rep-3-GFP expression in the presence of RNA-5HA was restricted to the periphery of the lesion and was not affected by the presence of RNA-4 (Table 1). Taken together, these findings suggest that, in *C. quinoa*, RNA-5-driven protein expression occurs later during the infection process and probably with lower efficiency than

that of RNA-3.

Combined tissues from three local lesions from the above inoculations were analyzed by Northern blot. After Trizol extraction, similar amounts of total RNAs (not shown) were separated on a denaturing gel, transferred onto nitrocellulose membrane and probed with riboprobes specific for each viral RNA (Lemaire *et al.*, 1988) or for GFP as described previously (Erhardt *et al.*, 2000; Lauber *et al.*, 1998b). Accumulation of the RNA was quantified with a Fuji MAS 1000 BioAnalyser. Results are summarized in Figure 1B. In each case, viral RNA-1 and -2 were detected in similar amounts (Figure 1B, a), as were also RNA-3 and Rep-3-GFP (Figure 1B-b, compare lanes 2 and 5). Rep-3-GFP accumulated similarly when inoculated alone or together with RNA-5HA or Rep-5-GFP (Figure 1B -a and -b, lanes 5, 7 and 11) but its accumulation was 6 times lower when RNA-4 was added to the inoculum (Figure 1B, a and b, lane 9). On the contrary, RNA-4 replication was not significantly affected by the presence of other small viral RNAs (Figure 1B-c, lanes 4, 9 and 10). The replication of RNA-5HA was affected when RNA-3 was added to the inoculum (Figure 1B-d, compare lanes 3 and 4) and it should be noticed that replacement of p26HA by the GFP did not modify the size of the RNA but indeed diminished accumulation of the RNA-5-derived viral vector (Figure 1B-d, compare lanes 3 and 6). Such a result may be due to a possible *cis* acting effect of the p26HA protein or the presence of replication regulatory elements within the p26 ORF as described previously for RNA-4 (Gilmer *et al.*, 1992). Interestingly, the addition of RNA-4 together with Rep-3-GFP did not alter the accumulation of RNA-5HA (Figure 1B-d, compare lanes 4 and 9). Quantification of the detected signals revealed that Rep-5-GFP accumulated 4 to 6 fold less than Rep-3-GFP. Moreover, Rep-5-GFP replication was drastically affected by the presence of other small BNYVV RNAs (Figure 1B-a, compare lanes 5 and 6, lanes 7 and 8, lanes 9 and 10 and Figure 1B-d, lanes 6, 8 and 10). Such lower replication levels explain the low accumulation of GFP fluorescence observed within infected tissues (Table 1).

To confirm that the GFP expression patterns were correlated with the observed RNA replication levels, immunodetection of virally expressed proteins were performed on single local lesions. The lesions were ground in 50 μ l of Laemmli buffer, centrifuged, heat denatured and equivalent amounts of protein were separated by SDS-PAGE before transfer onto immobilon membrane (Niesbach-Klosgen *et al.*, 1990; Lauber *et al.*, 1998b; Erhardt *et al.*, 2000). Total protein content was estimated by staining the membrane with Serva Blue G and appeared to be equal in each sample (Figure 1C, panel MS). After each immunodetection, the immobilon membrane was striped by a 5 min incubation in 0.2 N NaOH and extensive washes in water. Signals were acquired using a Chemidoc analyzer and quantified with Quantity One software (Bio-Rad). In each case, similar amounts of BNYVV p21 coat protein were detected (Figure 1C, panel CP) permitting comparison of the accumulation level of other virally expressed proteins. Such comparisons revealed that the relative expression levels of RNA-3-encoded p25 (Figure 1C, panel p25, lanes 2, 4, 8 and 10) paralleled the accumulation levels of full-length RNA-3 described above. On the other hand, the accumulation of RNA-5-encoded p26HA did not exactly correlate with the lowered replication of RNA-5HA observed in the presence of RNA-3 and -4 (compare Figure 1C, panel p26HA, lanes 3 and 4). This result suggests that the p26HA protein possesses a high stability *in vivo*.

The spatial and temporal patterns of GFP expression from Rep-3 and Rep-5 were distinct as different accumulation patterns of fluorescence were observed within infected tissues (Table 1). This observation was confirmed by the immunodetection of GFP within local lesions. Thus, inoculation with Rep-3-GFP produced GFP fluorescence in infected tissues (Table 1) and strong signals were observed on western blot using GFP-specific antibody (Figure 1C, panel GFP, lanes 5, 7, 9 and 11). However, only low amounts of GFP protein were immunodetected when Rep-5-GFP was inoculated (Figure 1C, panel GFP, lanes 6, 8 and 10) reflecting the low fluorescence observed within local lesions obtained in these conditions

(Table 1). For similar RNA accumulation levels (Figure 1B, panel a, lanes 6 and 9), quantitative comparison of GFP protein levels (Figure 1C, panel GFP, lanes 6 and 9) indicated that protein expression from Rep-5 was 4 to 5 times lower than from Rep-3. Importantly, when both replicons were co-inoculated together with Stras12, both were able to replicate (Figure 1C, panels b and d, lane 11) and to express GFP protein (Table 1 and Figure 1C, panel GFP, lane 11; and see also Figure 2E).

The lower accumulation level of Rep-5 was further confirmed by the use of replicons (namely Rep-3-P13-15 and Rep-5-P13-15) expressing the dicistronic P13-15 sequence from RNA-2. When inoculated with Stras12, both replicons were detected but Rep-5-P13-15 accumulated about 6 times less than Rep-3-P13-15 (Figure 1D, compare lanes 13 and 14).

We reported earlier that RNA-3 is required for long-distance movement of the virus in systemic hosts and Rep-3 derivatives are not able to systemically invade the plant, restricting their utilization to the inoculated tissues (Lauber *et al.*, 1998b; unpublished observations). *Beta macrocarpa* mechanically inoculated with Stras12 supplemented with RNA-3 and Rep-5-GFP permitted the detection of fluorescence in infected leaves (Figure 2A) but also in upper leaves (Figure 2B and C) and in rootlets (Figure 2D), indicating that Rep-5 can produce recombinant protein in systemically infected tissues.

To further demonstrate that both replicons were able to replicate and express two proteins in the same infected cell, we used *C. quinoa* leaves inoculated with Stras12 supplemented with Rep-3-RFP and Rep-5-GFP. Fluorescence monitoring using specific band pass filter permitted the detection of both GFP (Figure 2E, lane a) and RFP (Figure 2E, lane b) in the same infected cells (Figure 2E, lane c). Such result allows us to conclude that Rep-3 and Rep-5 RNAs replicate together within infected cells and thus allow co-expression of recombinant proteins for interaction studies *in planta* (e.g. FRET).

In this paper, we have shown that Rep-5 represents a useful new tool for the expression of

recombinant proteins, in particular together with RNA-3 when using BNYVV systemic host. On local host crop, both replicons allow the co-expression of proteins. While Rep-3 derivatives permit high levels of protein synthesis in infected tissues, use of Rep-5 may be preferable in situations where over-expression leads to undesired side effects in infected cells.

Acknowledgments:

We thank Malek Alioua for sequence analysis; Daniele Scheidecker and Audrey Schirmer for technical support; Pierre Pfeiffer for providing mRFP-containing vector; Ken Richards for critical reading of this manuscript. L. S. was supported by SES-Advanta under the CIFRE programme. D. L. was supported by ITB and Region Alsace under a Bourse Régionale convention. The Inter-Institut confocal microscopy plate-form was co-financed by : CNRS, Université Louis Pasteur, Region Alsace and ARC.

References

- Awram, P., Gardner, R. C., Forster, R. L. & Bellamy, A. R.** (2002). The potential of plant viral vectors and transgenic plants for subunit vaccine production. *Adv Virus Res* **58**, 81-124.
- Bleykasten-Grosshans, C., Guilley, H., Bouzoubaa, S., Richards, K. E. & Jonard, G.** (1997). Independent Expression of the First Two Triple Gene Block Proteins of Beet Necrotic Yellow Vein Virus Complements Virus Defective in the Corresponding Gene but Expression of the Third Protein Inhibits Viral Cell-to-Cell Movement. *Mol Plant Microbe Interact* **10**, 240-246.
- Erhardt, M., Morant, M., Ritzenthaler, C., Stussi-Graud, C., Guilley, H., Richards, K., Jonard, G., Bouzoubaa, S. & Gilmer, D.** (2000). P42 movement protein of Beet necrotic yellow vein virus is targeted by the movement proteins P13 and P15 to punctate bodies associated with plasmodesmata. *Mol Plant Microbe Interact* **13**, 520-8.

- Gilmer, D., Richards, K., Jonard, G. & Guilley, H.** (1992). cis-active sequences near the 5'-termini of beet necrotic yellow vein virus RNAs 3 and 4. *Virology* **190**, 55-67.
- Gleba, Y., Marillonnet, S. & Klimyuk, V.** (2004). Engineering viral expression vectors for plants: the 'full virus' and the 'deconstructed virus' strategies. *Curr Opin Plant Biol* **7**, 182-8.
- Hehn, A., Bouzoubaa, S., Bate, N., Twell, D., Marbach, J., Richards, K., Guilley, H. & Jonard, G.** (1995). The small cysteine-rich protein P14 of beet necrotic yellow vein virus regulates accumulation of RNA 2 in cis and coat protein in trans. *Virology* **210**, 73-81.
- Lauber, E., Bleykasten-Grosshans, C., Erhardt, M., Bouzoubaa, S., Jonard, G., Richards, K. E. & Guilley, H.** (1998a). Cell-to-cell movement of beet necrotic yellow vein virus: I. Heterologous complementation experiments provide evidence for specific interactions among the triple gene block proteins. *Mol Plant Microbe Interact* **11**, 618-25.
- Lauber, E., Guilley, H., Tamada, T., Richards, K. E. & Jonard, G.** (1998b). Vascular movement of beet necrotic yellow vein virus in Beta macrocarpa is probably dependent on an RNA 3 sequence domain rather than a gene product. *J Gen Virol* **79**, 385-93.
- Lauber, E., Jonard, G., Guilley, H. & Gilmer, D.** (1999). Effects of structural modifications upon the accumulation in planta of replicons derived from beet necrotic yellow vein virus RNA 3. *Arch Virol* **144**, 1201-8.
- Lemaire, O., Merdinoglu, D., Valentin, P., Putz, C., Ziegler-Graff, V., Guilley, H., Jonard, G. & Richards, K.** (1988). Effect of beet necrotic yellow vein virus RNA composition on transmission by *Polymyxa betae*. *Virology* **162**, 232-5.
- Niesbach-Klosgen, U., Guilley, H., Jonard, G. & Richards, K.** (1990). Immunodetection in vivo of beet necrotic yellow vein virus-encoded proteins. *Virology* **178**, 52-61.

Pogue, G. P., Lindbo, J. A., Dawson, W. O. & Turpen, T. H. (1998). Tobamovirus transient expression vectors: tools for plant biology and high-level expression of foreign protein in plants. In *Plant Molecular Biology Manual*, pp. 1. Edited by S. B. Gelvin & R. A. Schilperoort. Dordrecht: Kluwer Academic Publishers.

Pogue, G. P., Lindbo, J. A., Garger, S. J. & Fitzmaurice, W. P. (2002). Making an ally from an enemy: plant virology and the new agriculture. *Annu Rev Phytopathol* **40**, 45-74.

Porta, C. & Lomonossoff, G. P. (2002). Viruses as vectors for the expression of foreign sequences in plants. *Biotechnol Genet Eng Rev* **19**, 245-91.

Quillet, L., Guilley, H., Jonard, G. & Richards, K. (1989). In vitro synthesis of biologically active beet necrotic yellow vein virus RNA. *Virology* **172**, 293-301.

Vetter, G., Hily, J. M., Klein, E., Schmidlin, L., Haas, M., Merkle, T. & Gilmer, D. (2004). Nucleo-cytoplasmic shuttling of the beet necrotic yellow vein virus RNA-3-encoded p25 protein. *J Gen Virol* **85**, 2459-69.

Legend to figures:

Figure 1: Analysis of RNA (B and D) and protein (C) expression capacities of BNYVV replicons in different inocula combination. (A) Diagram of coding capacities of RNA-3 and RNA-5HA molecules and their derivative replicons. Open reading frames are indicated by labelled rectangles and non coding regions by lines. Black boxes refer respectively to p6,8 and p4,8 genes. Drawings are not to scale. (B) Accumulation levels of BNYVV RNAs within local lesions of *C. quinoa* leaves 6 dpi with Stras12 alone (1) or supplemented with RNA-3 (2); RNA-5HA (3); RNA-3 and RNA-4 plus RNA-5HA (4); Rep-3-GFP (5); Rep-5-GFP (6); Rep-3-GFP plus RNA-5HA (7); Rep-5-GFP plus RNA-3 (8); Rep-3-GFP plus RNA-4, -5HA (9); Rep-5-GFP plus RNA-3, -4 (10); Rep-3-GFP plus Rep-5-GFP (11). Membranes were probed with specific riboprobes complementary to sequences of RNA-1, -2 and GFP (a),

RNA-3 (b), RNA-4 (c) or RNA-5HA (d). Note that Rep-3-GFP, Rep-5-GFP and RNA-5HA co-migrate. (C) Immuno-detection of BNYVV expressed proteins within local lesions of *C. quinoa* leaves 6 dpi with the above constructs. MS is an image of the immobilon membrane stained with Serva blue G. (D) Analysis of the accumulation levels of RNA-3 and -5-derived replicons expressing the dicistronic P13-15 sequence. Accumulation level of BNYVV RNAs within local lesions of *C. quinoa* leaves 6 dpi with Stras12 alone (12), supplemented with Rep-3-P13-15 (13) or with Rep-5-P13-15 (14). Membranes were probed with riboprobes complementary to RNA-1 and -2. The RNA-2-specific riboprobe was chosen so as to detect the P13-15 sequence expressed in both replicons. Similar amounts of total RNA were loaded. The identities of RNAs and specifically immuno-detected proteins are to the right.

Figure 2: Rep-5-GFP moves systemically within *Beta macrocarpa* in the presence of Stras12 supplemented with RNA-3 (A to D) and is co-expressed with mRFP in the same *C. quinoa* infected cells (E). (A) Leaf of *B. macrocarpa* inoculated with Stras12 supplemented with RNA-3 and Rep-5-GFP. Detection of RNA-5-driven-GFP expression within systemic infected leaves (B and C) and in rootlets (D). Photographs were taken under UV illumination with a x4 magnification binocular equipped with a Nikon Coolpix 5000 digital camera. (A to D), red fluorescence is due to chlorophyll. (E) Co-expression of mRFP and GFP. Visualization of the fluorescence pattern within (1) *C. quinoa* local lesions produced by Stras12 supplemented with Rep-3-RFP and Rep-5-GFP. (2) CLSM observations of cellular fluorescence pattern within the aforesaid lesions. (a) Expression pattern of GFP, (b) expression pattern of mRFP, (c) merged images. Bars correspond to 0.5 mm for the first column and 5 μ m for the CLSM images. Pictures of GFP and mRFP expression patterns were acquired with a Nikon E800 microscope equipped with a Nikon DXM1200 camera and processed with Nikon ACT1 software, CLSM images with a LSM510 Meta Zeiss confocal microscope using a 63X, 1.2 water immersion objective.

Table 1: Summary of *C. quinoa* local lesions sizes, phenotypes and their GFP expression patterns (when applicable), 6 dpi with Stras12 (a) supplemented with different inoculums composition carrying GFP sequence on RNA-3- and/or RNA-5-derived replicons (Rep-). CS, chlorotic spots; YS, yellow spots; Nec, necrotic spots. NA, not applicable

Inoculums: Stras12 supplemented with:	Lesion diameter	Symptoms	GFP expression	
			Intensity	Pattern
-	~2 mm	CS	NA	NA
RNA-3	~2 mm	YS	NA	NA
RNA-5HA	~1 mm	Nec	NA	NA
Rep-3-GFP	~2 mm	CS	Strong	Fluorescent rings
Rep-5-GFP	~2 mm	CS	Faint	Fluorescent rings
Rep-3-GFP + RNA-5HA	~1 mm	Nec	Strong	Front of infection
RNA-3 + Rep-5-GFP	~2 mm	YS	Faint	Lesion center
RNA-3 + RNA-4 + RNA-5HA	~0.5 mm	Nec	NA	NA
Rep-3-GFP + RNA-4 + RNA-5HA	~1 mm	Nec	Strong	Front of infection
RNA-3 + RNA-4 + Rep-5-GFP	~2 mm	YS	Faint	Lesion center
Rep-3-GFP + Rep-5-GFP	~2 mm	CS	Strong	Entire lesion

Figure 1

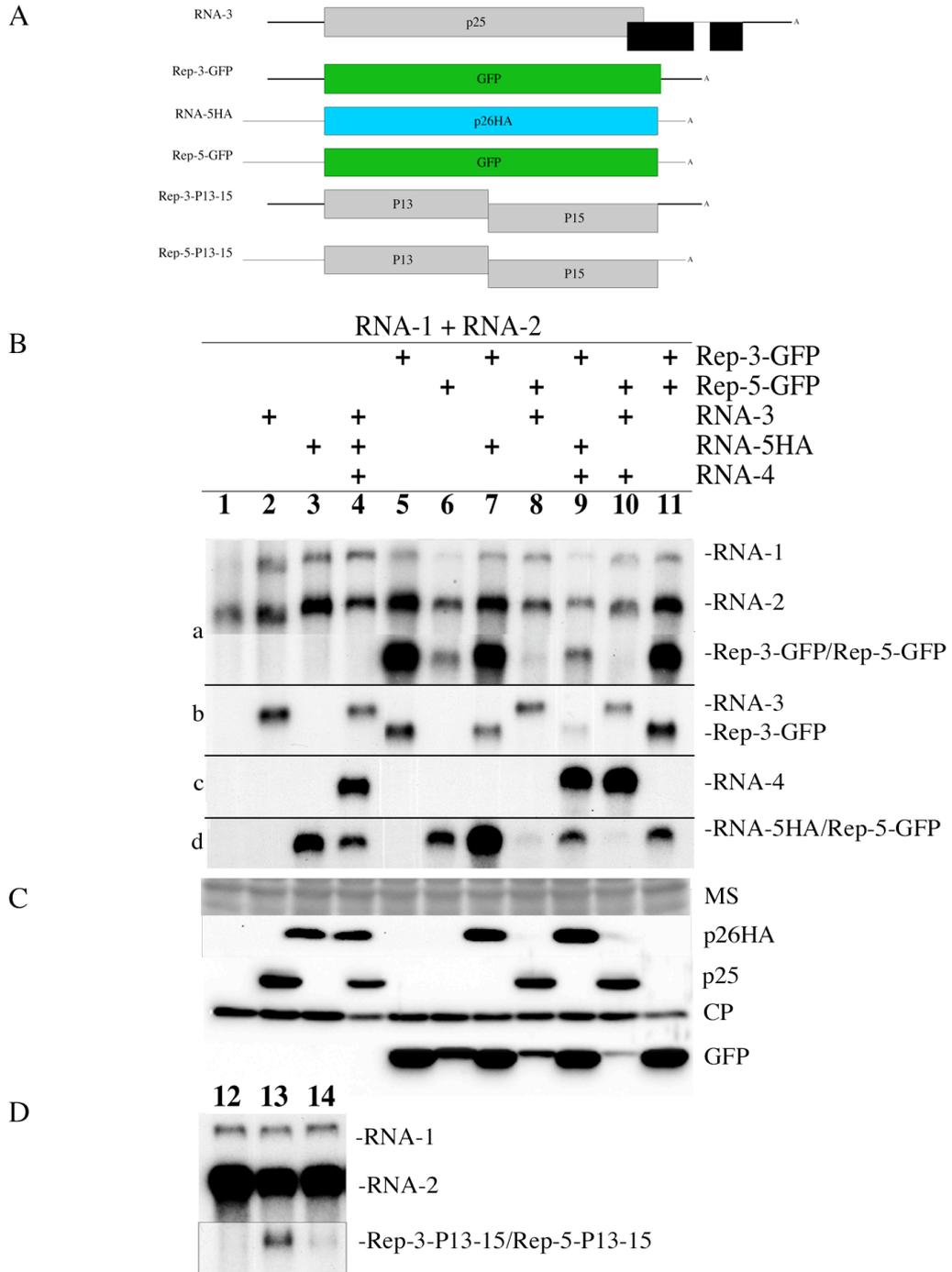
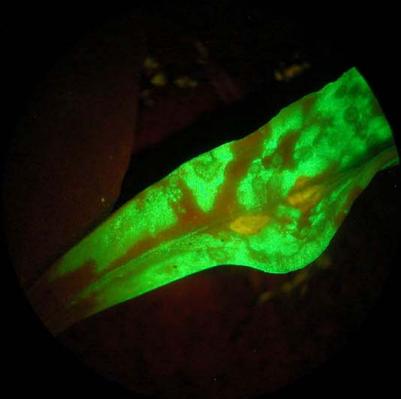
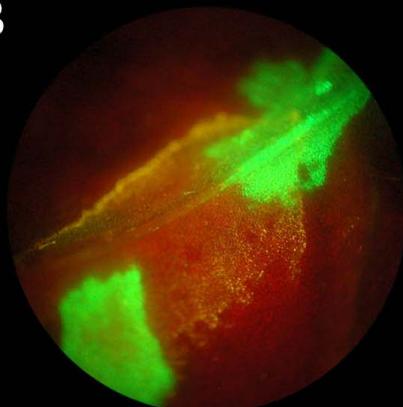


Figure 2

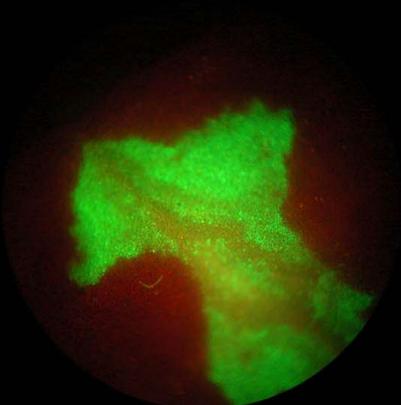
A



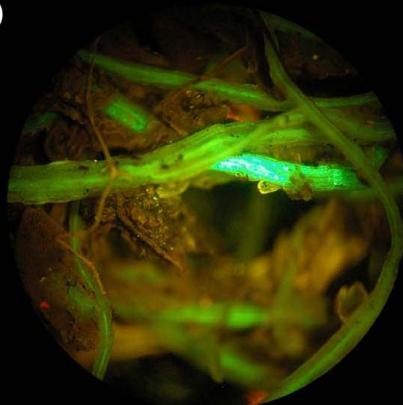
B



C



D

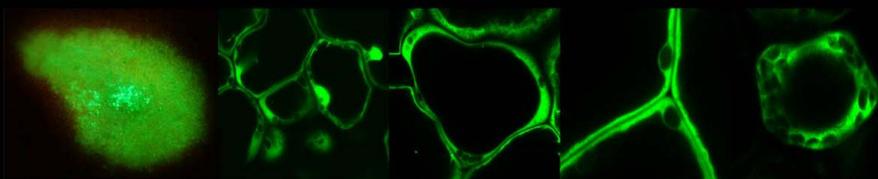


E

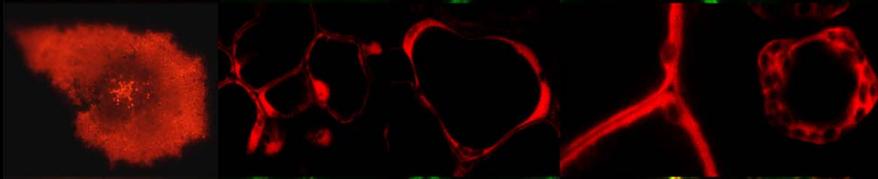
1

2

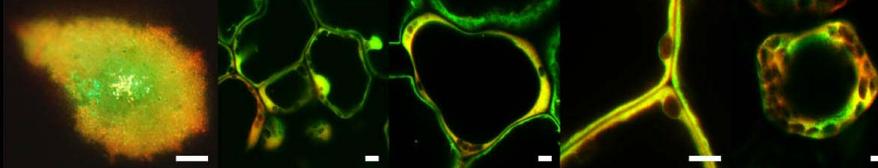
a



b



c



Bibliographie

- Abe H. and Tamada T.** (1986) Association of beet necrotic yellow vein virus with isolates of polomyxa betae keskin *Ann. Phytopath. Soc. Japan* **52**, 235–247. 6
- Adams M.** (2002) *Advances in Botanical Research* vol. 36 chap. Fungi San Diego, California, USA : Academic Press 47–64 advances in Botanical Research incorporating Advances in Plant Pathology ; Plant virus vector interactions. 8
- Balmori E., Gilmer D., Richards K., Guilley H. and Jonard G.** (1993) Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. *Biochimie* **75**, 517–521. 5
- Bartel P., Chien C., Sternglanz R. and Fields S.** (1993) Elimination of false positives that arise in using the two-hybrid system. *Biotechniques* **14**, 920–924. 98
- Baulcombe D., Chapman S. and Cruz S.S.** (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053. 55, 63
- Beals C., Sheridan C., Turck C., Gardner P. and Crabtree G.** (1997) Nuclear export of NF-ATc enhanced by glycogen synthase kinase-3. *Science* **275**, 1930–1934. 49
- Bednenko J., Cingolani G. and Gerace L.** (2003) Nucleocytoplasmic transport : navigating the channel. *Traffic* **4**, 127–35. 44, 47
- Belov G., Evstafieva A., Rubtsov Y., Mikitas O., Vartapetian A. and Agol V.** (2000) Early alteration of nucleocytoplasmic traffic induced by some RNA viruses. *Virology* **275**, 244–248. 52
- Ben-Efraim I. and Gerace L.** (2001) Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. *J. Cell Biol.* **152**, 411–417. 49
- Bendixen C., Gangloff S. and Rothstein R.** (1994) A yeast mating-selection scheme for detection of protein-protein interactions. *Nucleic Acids Res* **22**, 1778–1779. 99
- Bischoff F. and Ponstingl H.** (1991) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* **354**, 80–82. 49
- Bischoff F., Klebe C., Kretschmer J., Wittinghofer A. and Ponstingl H.** (1994) Ran-GAP1 induces GTPase activity of nuclear Ras-related Ran. *Proc. Natl. Acad. Sci. USA* **91**, 2587–2591. 49

-
- Bischoff F., Krebber H., Smirnova E., Dong W. and Ponstingl H.** (1995) Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J.* **14**, 705–715. 49
- Blanc S., Froissart R., Gratraud P., Uzest M. and Michalakis Y.** (2003) Caractérisation génétique et biologique d'une quasi-espèce virale (CaMV) in *IXèmes rencontres de virologie végétale* 64. 27
- Bogerd H., Fridell R., Madore S. and Cullen B.** (1995) Identification of a novel cellular cofactor for the Rev/Rex class of retroviral regulatory proteins. *Cell* **82**, 485–494. 46
- Boulikas T.** (1993) Nuclear localization signals (NLS). *Crit. Rev. Eukaryot. Gene Expr.* **3**, 193–227.
- Bouzoubaa S., Guilley H., Jonard G., Richards K. and Putz C.** (1985) Nucleotide sequence analysis of rna-3 and rna-4 of beet necrotic yellow vein virus, isolates f2 and g1 *J. Gen. Virol.* **66**, 1553–1564. 5
- Bouzoubaa S., Ziegler V., Beck D., Guilley H., Richards K. and Jonard G.** (1986) Nucleotide sequence of beet necrotic yellow vein virus rna-2 *J. Gen. Virol.* **67**, 1689–1700. 4
- Bouzoubaa S., Quillet L., Guilley H., Jonard G. and Richards K.** (1987) Nucleotide sequence of beet necrotic yellow vein virus rna-1 *J. Gen. Virol.* **68**, 615–626. 4
- Bouzoubaa S., Niesbach-Klösigen U., Jupin I., Guilley H., Richards K. and Jonard G.** (1991) Shortened forms of beet necrotic yellow vein virus RNA-3 and -4 : internal deletions and a subgenomic RNA. *J. Gen. Virol.* **72**, 259–266. 5, 64
- Burich R. and Lei M.** (2003) Two bipartite NLSs mediate constitutive nuclear localization of Mcm10. *Curr. Genet.* **44**, 195–201. 75
- Campbell E. and Hope T.** (2003) Role of the cytoskeleton in nuclear import. *Adv. Drug Deliv. Rev.* **55**, 761–771. 50
- Campbell R., Tour O., Palmer A., Steinbach P., Baird G., Zacharias D. and Tsien R.** (2002) A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* **99**, 7877–7882. 54
- Canova A.** (1959) Appunti di patologia della barbatietola *Informatore Fitopatologico* **9**, 390–396. 1
- Casadaban M. and Cohen S.** (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* **138**, 179–207. 94
- Chalfie M., Tu Y., Euskirchen G., Ward W. and Prasher D.** (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805. 53, 54
- Chare E. and Holmes E.** (2004) Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission *J. Gen. Virol.* **85**, 3149–3157. 25

- Chow C., Rincón M., Cavanagh J., Dickens M. and Davis R.** (1997) Nuclear accumulation of NFAT4 opposed by the JNK signal transduction pathway. *Science* **278**, 1638–1641. 49
- Cody C., Prasher D., Westler W., Prendergast F. and Ward W.** (1993) Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* **32**, 1212–1218. 53
- Cronshaw J., Krutchinsky A., Zhang W., Chait B. and Matunis M.** (2002) Proteomic analysis of the mammalian nuclear pore complex. *J. Cell Biol.* **158**, 915–927. 44
- Cruz S., Chapman S., Roberts A., Roberts I., Prior D. and Oparka K.** (1996) Assembly and movement of a plant virus carrying a green fluorescent protein overcoat. *Proc. Natl. Acad. Sci. USA* **93**, 6286–6290. 55, 63
- Cubitt A., Heim R., Adams S., Boyd A., Gross L. and Tsien R.** (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem. Sci.* **20**, 448–455. 53
- Dang C. and Lee W.** (1989) Nuclear and nucleolar targeting sequences of c-erb-A, c-myc, N-myc, p53, HSP70, and HIV tat proteins. *J. Biol. Chem.* **264**, 18019–18023.
- Davis L.** (1995) The nuclear pore complex *Annu. Rev. Biochem.* **64**, 865–896. 44
- Doye V. and Hurt E.** (1997) From nucleoporins to nuclear pore complexes. *Curr. Opin. Cell Biol.* **9**, 401–411. 44
- Driskel B., Doss P., Littlefield L., Walker N. and Verchot-Lubicz J.** (2004) Soilborne wheat mosaic virus movement protein and RNA and Wheat spindle streak mosaic virus coat protein accumulate inside resting spores of their vector, *Polymyxa graminis*. *Mol. Plant Microbe Interact.* **17**, 739–48. 6
- Dunoyer P., Pfeffer S., Fritsch C., Hemmer O., Voinnet O. and Richards K.** (2002) Identification, subcellular localization and some properties of a cysteine-rich suppressor of gene silencing encoded by peanut clump virus. *Plant J.* **29**, 555–67. 5
- Eck R. and Dayhoff M.** (1966) *Atlas of Protein Sequence and Structure 1966* Silver Spring, Maryland : National Biomedical Research Foundation. 15
- Enninga J., Levy D., Blobel G. and Fontoura B.** (2002) Role of nucleoporin induction in releasing an mRNA nuclear export block. *Science* **295**, 1523–1525. 51
- Epel B., Padgett H., Heinlein M. and Beachy R.** (1996) Plant virus movement protein dynamics probed with a GFP-protein fusion. *Gene* **173**, 75–79. 55
- Erhardt M., Dunoyer P., Guilley H., Richards K., Jonard G. and Bouzoubaa S.** (2001) Beet necrotic yellow vein virus particle localize to mitochondria during infection *Virology* **286**, 256–262. 55
- Fahrenkrog B. and Aebi U.** (2003) The nuclear pore complex : nucleocytoplasmic transport and beyond. *Nat. Rev. Mol. Cell Biol.* **4**, 757–766. 44

-
- Favre D., Studer E. and Michel M.** (1994) Two nucleolar targeting signals present in the N-terminal part of Semliki Forest virus capsid protein. *Arch. Virol.* **137**, 149–155.
- Felsenstein J.** (1981) Evolutionary trees from DNA sequences : a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376. 16
- Felsenstein J.** (1985) Confidence limits on phylogenies : an approach using the bootstrap *Evolution* . 16
- Felsenstein J.** (1989) Phylip - phylogeny inference package (version 3.2) *Cladistics* **5**, 164–166. 93
- Felsenstein J.** (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.* **266**, 418–427. 16
- Fields S. and Song O.** (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246. 98
- Fischer U., Huber J., Boelens W., Mattaj I. and Lührmann R.** (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475–483. 46
- Fornerod M., Ohno M., Yoshida M. and Mattaj I.** (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051–1060. 48
- Fraile A., Escriu F., Aranda M., Malpica J., Gibbs A. and García-Arenal F.** (1997) A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J. Virol.* **71**, 8316–8320. 25, 27
- Fridell R., Partin K., Carpenter S. and Cullen B.** (1993) Identification of the activation domain of equine infectious anemia virus rev. *J. Virol.* **67**, 7317–7323.
- Fritz C. and Green M.** (1996) HIV Rev uses a conserved cellular protein export pathway for the nucleocytoplasmic transport of viral RNAs. *Curr. Biol.* **6**, 848–854. 46
- Gilmer D., Bouzoubaa S., Hehn A., Guilley H., Richards K. and Jonard G.** (1992a) Efficient cell-to-cell movement of beet necrotic yellow vein virus requires 3' proximal genes located on RNA 2. *Virology* **189**, 40–47. 5
- Gilmer D., Richards K., Jonard G. and Guilley H.** (1992b) cis-active sequences near the 5'-termini of beet necrotic yellow vein virus RNAs 3 and 4. *Virology* **190**, 55–67. 61
- Gustin K.** (2003) Inhibition of nucleo-cytoplasmic trafficking by RNA viruses : targeting the nuclear pore complex. *Virus Res.* **95**, 35–44. 51, 52
- Gustin K. and Sarnow P.** (2001) Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. *EMBO J.* **20**, 240–249. 51, 52
- Görlich D. and Kutay U.** (1999) Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–60. 47, 48

- Görner W., Durchschlag E., Martinez-Pastor M., Estruch F., Ammerer G., Hamilton B., Ruis H. and Iler C.** (1998) Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes Dev.* **12**, 586–597. 49
- Haeberlé A., Stussi-Garaud C., Schmitt C., Garaud J., Richards K., Guilley H. and Jonard G.** (1994) Detection by immunogold labelling of P75 readthrough protein near an extremity of beet necrotic yellow vein virus particles. *Arch. Virol.* **134**, 195–203. 4, 7, 52
- Harveson R., Rush C. and Wheeler T.** (1996) The spread of beet necrotic yellow vein virus from point source inoculations as influenced by irrigation and tillage *Phytopathology* **86**, 1242–1247. 8
- Heese-Peck A. and Raikhel N.** (1998) A glycoprotein modified with terminal N-acetylglucosamine and localized at the nuclear rim shows sequence similarity to aldose-1-epimerases. *Plant Cell* **10**, 599–612. 44
- Hehn A., Bouzoubaa S., Jonard G., Guilley H. and Richards K.** (1994) Artificial defective interfering RNAs derived from RNA 2 of beet necrotic yellow vein virus *Arch. Virol.* **135**, 143–151. 57
- Hehn A., Fritsch C., Richards K., Guilley H. and Jonard G.** (1997) Evidence for in vitro and in vivo autocatalytic processing of the primary translation product of beet necrotic yellow vein virus RNA 1 by a papain-like proteinase. *Arch. Virol.* **142**, 1051–1058. 4
- Heijbroek W., Musters P. and Schoone A.** (1999) Variation in pathogenicity and multiplication of beet necrotic yellow vein virus (BNYVV) in relation to the resistance of sugar-beet cultivars *Eur. J. Plant Pathol.* **105**, 397–405. 11
- Heim R. and Tsien R.** (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178–182. 54
- Heim R., Prasher D. and Tsien R.** (1994) Wavelength mutations and posttranslational autooxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* **91**, 12501–12504. 53
- Heinlein M. and Epel B.** (2004) Macromolecular transport and signaling through plasmodesmata. *Int. Rev. Cytol.* **235**, 93–164. 55
- Heinlein M., Epel B., Padgett H. and Beachy R.** (1995) Interaction of tobamovirus movement proteins with the plant cytoskeleton. *Science* **270**, 1983–1985. 55
- Henry M. and Silver P.** (1996) A novel methyltransferase (Hmt1p) modifies poly(A)⁺-RNA-binding proteins. *Mol. Cell. Biol.* **16**, 3668–3678. 50
- Hinshaw J., Carragher B. and Milligan R.** (1992) Architecture and design of the nuclear pore complex. *Cell* **69**, 1133–1141. 44

-
- Ho S., Hunt H., Horton R., Pullen J. and Pease L.** (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59. 87
- Horton R., Hunt H., Ho S., Pullen J. and Pease L.** (1989) Engineering hybrid genes without the use of restriction enzymes : gene splicing by overlap extension. *Gene* **77**, 61–68. 87
- Hugo S., Henry C. and Harhu V.** (1996) The role of alternative hosts of polomyxa betae in transmission of beet necrotic yellow vein virus (bnyvv) in england *Plant Pathology* **45**, 662–666. 7
- Hübner S., Smith H., Hu W., Chan C., Rihs H., Paschal B., Raikhel N. and Jans D.** (1999) Plant importin alpha binds nuclear localization sequences with high affinity and can mediate nuclear import independent of importin beta. *J. Biol. Chem.* **274**, 22610–22617. 48
- Izaurrealde E., Kutay U., von Kobbe C., Mattaj I. and Görlich D.** (1997) The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J.* **16**, 6535–6547. 49
- Jans D. and Hübner S.** (1996) Regulation of protein transport to the nucleus : central role of phosphorylation. *Physiol. Rev.* **76**, 651–685. 49
- Jin L. and Nei M.** (1990) Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.* **7**, 82–102. 15
- Jukes T. and Cantor C.** (1969) *Mammalian Protein Metabolism* chap. Evolution of protein molecules New York : Academic Press 21–132. 15
- Jupin I., Quillet L., Ziegler-Graff V., Guilley H., Richards K. and Jonard G.** (1988) In vitro translation of natural and synthetic beet necrotic yellow vein virus rna-1 *J. Gen. Virol.* **69**, 2359–2367. 4
- Jupin I., Richards K., Jonard G., Guilley H. and Pleij C.** (1990) Mapping sequences required for productive replication of beet necrotic yellow vein virus RNA 3. *Virology* **178**, 273–280. 61, 85
- Jupin I., Guilley H., Richards K. and Jonard G.** (1992) Two proteins encoded by beet necrotic yellow vein virus RNA 3 influence symptom phenotype on leaves. *EMBO J.* **11**, 479–488. 5, 19, 29, 30, 32
- Kalderon D., Richardson W., Markham A. and Smith A.** (1984a) Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* **311**, 33–38. 45
- Kalderon D., Roberts B., Richardson W. and Smith A.** (1984b) A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499–509. 45
- Kanyuka K., Ward E. and Adams M.** (2003) Polomyxa graminis and the cereal viruses it transmits : a research challenge *Mol. Plant Pathology* **4**, 393–406. 6

- Keskin B.** (1964) [polymyxa betae n.sp., a parasite in the roots of beta vulgaris tournefort, particularly during the early growth of the sugar beet] *Arch. Mikrobiol.* **49**, 348–74. 6
- Kiguchi T., Saito M. and Tamada T.** (1996) Nucleotide sequence analysis of rna-5 of five isolates of beet necrotic yellow vein virus and the identity of a deletion mutant *J. Gen. Virol.* **77**, 575–580. 6
- Kimura M.** (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120. 15
- Kluge A. and Farris J.** (1969) Quantitative phyletics and the evolution of anurans *Systematic Zoology* **18**, 1–32. 15
- Kobayashi T., Kamitani W., Zhang G., Watanabe M., Tomonaga K. and Ikuta K.** (2001) Borna disease virus nucleoprotein requires both nuclear localization and export activities for viral nucleocytoplasmic shuttling. *J. Virol.* **75**, 3404–12.
- Koenig R. and Büttner G.** (2004) Strategies for the detection of potential beet necrotic yellow vein virus genome recombinations which might arise as a result of growing A type coat protein gene-expressing sugarbeets in soil containing B type virus. *Transgenic Res.* **13**, 21–8. 24
- Koenig R. and Lennefors B.L.** (2000) Molecular analyses of european a, b and p type sources of beet necrotic yellow vein virus and detection of the rare p type in kazakhstan *Arch. Virol.* **145**, 1561–1570. 10, 11
- Koenig R., Jarausch W., Li Y., Commandeur U., Burgermeister W., Gehrke M. and Lüddecke P.** (1991) Effect of recombinant beet necrotic yellow vein virus with different RNA compositions on mechanically inoculated sugarbeets. *J. Gen. Virol.* **72**, 2243–6. 5, 29
- Koenig R., Lüddecke P. and Haeblerlé A.** (1995) Detection of beet necrotic yellow vein virus strains, variants and mixed infections by examining single-strand conformation polymorphisms of immunocapture rt-pcr products *J. Gen. Virol.* **76**, 2051–2055. 10
- Koenig R., Haeblerlé A.M. and Commandeur U.** (1997) Detection and characterization of a distinct type of beet necrotic yellow vein virus rna 5 in a sugarbeet growing area in europe *Arch. Virol.* **142**, 1499–1504. 6
- Koonin E. and Dolja V.** (1993) Evolution and taxonomy of positive-strand RNA viruses : implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28**, 375–430. 4
- Kozak M.** (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**, 1–34. 58, 61, 75
- Kruse M., Koenig R., Hoffmann A., Kaufmann A., Commandeur U., Solovyev A., Savenkov I. and Burgmeister W.** (1994) Restriction fragment length polymorphism analysis of reverse transcription-pcr products reveals the existence of two major strain groups of beet necrotic yellow vein virus *J. Gen. Virol.* **75**, 1835–1842. 10

- Kubota S., Siomi H., Satoh T., Endo S., Maki M. and Hatanaka M.** (1989) Functional similarity of HIV-I rev and HTLV-I rex proteins : identification of a new nucleolar-targeting signal in rev protein. *Biochem. Biophys. Res. Commun.* **162**, 963–970.
- Kudo N., Wolff B., Sekimoto T., Schreiner E., Yoneda Y., Yanagida M., Horinouchi S. and Yoshida M.** (1998) Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Exp. Cell. Res.* **242**, 540–547. 68
- Laemmli U.** (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage t4 *Nature* **227**, 680–685. 96, 97
- Lauber E., Guilley H., Richards K., Jonard G. and Gilmer D.** (1997) Conformation of the 3'-end of beet necrotic yellow vein benevirus RNA 3 analysed by chemical and enzymatic probing and mutagenesis. *Nucleic Acids Research* **25**, 4723–4729.
- Lauber E., Guilley H., Tamada T., Richards K. and Jonard G.** (1998) Vascular movement of beet necrotic yellow vein virus in beta macrocarpa is probably dependent on an rna 3 sequence domain rather than a gene product *J. Gen. Virol.* , 385–393. 34
- Lemaire O., Merdinoglu D., Valentin P., Putz C., Ziegler-Graff V., Guilley H., Jonard G. and Richards K.** (1988) Effect of beet necrotic yellow vein virus RNA composition on transmission by *Polymyxa betae*. *Virology* **162**, 232–235. 6, 7
- Li B. and Fields S.** (1993) Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. *FASEB J.* **7**, 957–963. 69
- Li Q. and Verma I.** (2002) NF-kappaB regulation in the immune system. *Nat. Rev. Immunol.* **2**, 725–734. 68
- Lin X., Sun B., Liang M., Liang Y.Y., Gast A., Hildebrand J., Brunnicardi F., Melchior F. and Feng X.H.** (2003) Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding. *Mol. Cell* **11**, 1389–1396. 50
- Liu J., Lee L., Ye Y., Qian Z. and Kung H.** (1997) Nucleolar and nuclear localization properties of a herpesvirus bZIP oncoprotein, MEQ. *J. Virol.* **71**, 3188–3196.
- Liu Q. and Dreyfuss G.** (1995) In vivo and in vitro arginine methylation of RNA-binding proteins. *Mol. Cell. Biol.* **15**, 2800–2808. 50
- Lombardo E., Ramírez J., Agbandje-McKenna M. and Almendral J.** (2000) A beta-stranded motif drives capsid protein oligomers of the parvovirus minute virus of mice into the nucleus for viral assembly. *J. Virol.* **74**, 3804–3814. 43, 76
- Lombardo E., Ramírez J., Garcia J. and Almendral J.** (2002) Complementary roles of multiple nuclear targeting signals in the capsid proteins of the parvovirus minute virus of mice during assembly and onset of infection. *J. Virol.* **76**, 7049–7059. 76
- Lucy A., Guo H., Li W. and Ding S.** (2000) Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus. *EMBO J.* **19**, 1672–1680. 43

- Luo M., Pang C., Gerken A. and Brock T.** (2004) Multiple nuclear localization sequences allow modulation of 5-lipoxygenase nuclear import. *Traffic* **5**, 847–54. 76
- Mancuso V., Hope T., Zhu L., Derse D., Phillips T. and Parslow T.** (1994) Posttranscriptional effector domains in the Rev proteins of feline immunodeficiency virus and equine infectious anemia virus. *J. Virol.* **68**, 1998–2001.
- Matz M., Fradkov A., Labas Y., Savitsky A., Zaraisky A., Markelov M. and Lukyanov S.** (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **17**, 969–973. 54
- Michael W., Choi M. and Dreyfuss G.** (1995) A nuclear export signal in hnRNP A1 : a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* **83**, 415–22. 46
- Michael W., Eder P. and Dreyfuss G.** (1997) The K nuclear shuttling domain : a novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J.* **16**, 3587–3598. 45, 46
- Michel Y., Borman A., Paulous S. and Kean K.** (2001) Eukaryotic initiation factor 4G-poly(A) binding protein interaction is required for poly(A) tail-mediated stimulation of picornavirus internal ribosome entry segment-driven translation but not for X-mediated stimulation of hepatitis C virus translation. *Mol. Cell. Biol.* **21**, 4097–109. 58
- Miyamoto Y., Hieda M., Harreman M., Fukumoto M., Saiwaki T., Hodel A., Corbett A. and Yoneda Y.** (2002) Importin alpha can migrate into the nucleus in an importin beta- and Ran-independent manner. *EMBO J.* **21**, 5833–5842. 48
- Miyanishi M., Kusume T., Saito M. and Tamada T.** (1999) Evidence for three groups of sequence variants of beet necrotic yellow vein virus RNA 5. *Arch. Virol.* **144**, 879–892. 19, 21
- Moore M. and Blobel G.** (1993) The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature* **365**, 661–3. 47
- Morales F., Ward E., Castaño M., Arroyave J., Lozano I. and Adams M.** (1999) Emergence and partial characterization of rice stripe necrosis virus and its fungus vector in south america *Eur. J. Plant Pathol.* **105**, 643–650. 3
- Moroianu J., Blobel G. and Radu A.** (1997) RanGTP-mediated nuclear export of karyopherin alpha involves its interaction with the nucleoporin Nup153. *Proc. Natl. Acad. Sci. USA* **94**, 9699–9704. 52
- Morozov S., Dolja V. and Atabekov J.** (1989) Probable reassortment of genomic elements among elongated RNA-containing plant viruses. *J. Mol. Evol.* **29**, 52–62. 5
- Murray V.** (1989) Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nucleic Acids Res* **17**, 8889. 91

-
- Möller A. and Schmitz M.L.** (2003) Viruses as hijackers of PML nuclear bodies. *Arch. Immunol. Ther. Exp. (Warsz)* **51**, 295–300. 51
- Nachury M. and Weis K.** (1999) The direction of transport through the nuclear pore can be inverted. *Proc. Natl. Acad. Sci. USA* **96**, 9622–9627. 49
- Nagata T., Nemoto Y. and Hasezawa S.** (1992) Tobacco by-2 cell line as the "hela" cell in the cell biology of higher plants *Int. Rev. Cyto.* **132**, 1–30. 102
- Nakielny S. and Dreyfuss G.** (1999) Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**, 677–690. 46, 47, 48, 52
- Niesbach-Klößen U., Guillely H., Jonard G. and Richards K.** (1990) Immunodetection in vivo of beet necrotic yellow vein virus-encoded proteins. *Virology* **178**, 52–61. 4
- Oparka K., Boevink P. and Cruz S.S.** (1996) Studying the movement of plant viruses using green fluorescent protein *Trends in plant science* **1**, 412–418. 55, 63
- Ormö M., Cubitt A., Kallio K., Gross L., Tsien R. and Remington S.** (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395. 53
- Ossareh-Nazari B., Bachelier F. and Dargemont C.** (1997) Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* **278**, 141–144. 48
- Pang K., Lee E. and Knecht D.** (1998) Use of a fusion protein between GFP and an actin-binding domain to visualize transient filamentous-actin structures. *Curr. Biol.* **8**, 405–408. 54
- Paragas J., Talon J., O'Neill R., Anderson D., García-Sastre A. and Palese P.** (2001) Influenza B and C virus NEP (NS2) proteins possess nuclear export activities. *J. Virol.* **75**, 7375–7383.
- Paraskeva E., Izaurralde E., Bischoff F., Huber J., Kutay U., Hartmann E., Lührmann R. and Görlich D.** (1999) CRM1-mediated recycling of snurportin 1 to the cytoplasm. *J. Cell Biol.* **145**, 255–264. 48
- Percipalle P., Clarkson W., Kent H., Rhodes D. and Stewart M.** (1997) Molecular interactions between the importin alpha/beta heterodimer and proteins involved in vertebrate nuclear protein import. *J. Mol. Biol.* **266**, 722–732. 52
- Prasher D., Eckenrode V., Ward W., Prendergast F. and Cormier M.** (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229–233. 53
- Printen J. and Sprague G.** (1994) Protein-protein interactions in the yeast pheromone response pathway : Ste5p interacts with all members of the MAP kinase cascade. *Genetics* **138**, 609–619. 96

-
- Putz C. and Kuszula M.** (1978) La rhizomanie de la betterave sucrière en alsace. recherche d'une nouvelle méthode de purification du beet necrotic yellow vein virus *Annales de phytopathologie* **10**, 247–262. 94
- Quillet L., Guilley H., Jonard G. and Richards K.** (1989) In vitro synthesis of biologically active beet necrotic yellow vein virus RNA. *Virology* **172**, 293–301. 31, 94
- Rao A., Luo C. and Hogan P.** (1997) Transcription factors of the NFAT family : regulation and function. *Annu. Rev. Immunol.* **15**, 707–747. 49
- Ren T., Qu F. and Morris T.** (2000) HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. *Plant Cell* **12**, 1917–1926. 99
- Richards K. and Tamada T.** (1992) Mapping functions on the multipartite genome of beet necrotic yellow vein virus *Annu. Rev. Phytopathol.* **30**, 291–313. 3, 29
- Robbins J., Dilworth S., Laskey R. and Dingwall C.** (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence : identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615–623. 45
- Rodriguez M., Dargemont C. and Stutz F.** (2004) Nuclear export of RNA. *Biol. Cell* **96**, 639–655. 45
- Rose A. and Meier I.** (2001) A domain unique to plant RanGAP is responsible for its targeting to the plant nuclear rim. *Proc. Natl. Acad. Sci. USA* **98**, 15377–15382. 44
- Rowland R. and Yoo D.** (2003) Nucleolar-cytoplasmic shuttling of PRRSV nucleocapsid protein : a simple case of molecular mimicry or the complex regulation by nuclear import, nucleolar localization and nuclear export signal sequences. *Virus Res.* **95**, 23–33.
- Rush C.** (2003) Ecology and epidemiology of benyviruses and plasmodiophorid vectors *Annu. Rev. Phytopathol.* **41**, 567–592. 7, 8, 9
- Ryabova L., Park H.S. and Hohn T.** (2004) Control of translation reinitiation on the cauliflower mosaic virus (CaMV) polycistronic RNA. *Biochem. Soc. Trans.* **32**, 592–6. 58
- Saitou N. and Nei M.** (1987) The neighbor-joining method : a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–25. 15
- Sambrook J., Fritsch E. and Maniatis T.** (1989) *Molecular cloning : a laboratory manual* Cold Spring Harbor Laboratory, New-York USA : Cold Spring Harbor Laboratory Press 2nd edn. 90
- Sandri-Goldin R.** (1998) ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding viral intronless RNAs through an RGG motif. *Genes Dev.* **12**, 868–879.

- Sanger F., Nicklen S. and Coulson A.** (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467. 91
- Scholten O. and Lange W.** (2000) Breeding for resistance to rhizomania in sugar beet : A review *Euphytica* **112**, 219–231. 6, 9
- Schröder H., Rottmann M., Wenger R., Bachmann M., Dorn A. and Muller W.** (1988) Studies on protein kinases involved in regulation of nucleocytoplasmic mRNA transport. *Biochem. J.* **252**, 777–790. 50
- Shah S., Tugendreich S. and Forbes D.** (1998) Major binding sites for the nuclear import receptor are the internal nucleoporin Nup153 and the adjacent nuclear filament protein Tpr. *J. Cell Biol.* **141**, 31–49. 52
- Shirako Y., Suzuki N. and French R.** (2000) Similarity and divergence among viruses in the genus Furovirus. *Virology* **270**, 201–207. 2
- Shoya Y., Kobayashi T., Koda T., Ikuta K., Kakinuma M. and Kishi M.** (1998) Two proline-rich nuclear localization signals in the amino- and carboxyl-terminal regions of the Borna disease virus phosphoprotein. *J. Virol.* **72**, 9755–9762.
- Siomi H. and Dreyfuss G.** (1995) A nuclear localization domain in the hnRNP A1 protein. *J. Cell Biol.* **129**, 551–60. 45
- Siomi H., Shida H., Nam S., Nosaka T., Maki M. and Hatanaka M.** (1988) Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. *Cell* **55**, 197–209.
- Siomi M., Eder P., Kataoka N., Wan L., Liu Q. and Dreyfuss G.** (1997) Transportin-mediated nuclear import of heterogeneous nuclear RNP proteins. *J. Cell Biol.* **138**, 1181–1192. 45
- Sleat D., Gallie D., Jefferson R., Bevan M., Turner P. and Wilson T.** (1987) Characterisation of the 5'-leader sequence of tobacco mosaic virus RNA as a general enhancer of translation in vitro. *Gene* **60**, 217–225. 92
- Smith L., Sanders J., Kaiser R., Hughes P., Dodd C., Connell C., Heiner C., Kent S. and Hood L.** (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674–679. 91
- Sodeik B.** (2000) Mechanisms of viral transport in the cytoplasm. *Trends Microbiol.* **8**, 465–472. 50
- Stark G., Kerr I., Williams B., Silverman R. and Schreiber R.** (1998) How cells respond to interferons. *Annu. Rev. Biochem.* **67**, 227–264. 51
- Steven A., Trus B., Putz C. and Wurtz M.** (1981) The molecular organization of beet necrotic yellow vein virus *Virology* **113**, 428–438. 3

- Stoffler D., Fahrenkrog B. and Aebi U.** (1999) The nuclear pore complex : from molecular architecture to functional dynamics. *Curr. Opin. Cell Biol.* **11**, 391–401. 44
- Ström A. and Weis K.** (2001) Importin-beta-like nuclear transport receptors. *Genome Biol.* **2**, 3008.1–3008.9. 45, 49
- Szurek B., Rossier O., Hause G. and Bonas U.** (2002) Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol. Microbiol.* **46**, 13–23. 77
- Takeshita M., Shigemune N., Kikuhara K., Furuya N. and Takanami Y.** (2004) Spatial analysis for exclusive interactions between subgroups i and ii of cucumber mosaic virus in cowpea *Virology* **328**, 45–51. 28
- Tamada T.** (1975) Beet necrotic yellow vein virus *Commonwealth Mycological Institute / Association of Applied Biologists Descriptions of Plant Viruses* . 2
- Tamada T. and Abe H.** (1989) Evidence that beet necrotic yellow vein virus rna-4 is essential for efficient transmission by the fungus *polymyxa betae* *J. Gen. Virol.* **70**, 3391–3398. 6, 7
- Tamada T. and Kusume T.** (1991) Evidence that the 75k readthrough protein of beet necrotic yellow vein virus rna-2 is essential for transmission by the fungus *polymyxa betae* *J. Gen. Virol.* **72**, 1497–1504. 7
- Tamada T., Shirako Y., Abe H., Saito M., Kiguchi T. and Harada T.** (1989) Production and pathogenicity of isolates of beet necrotic yellow vein virus with different numbers of rna components *J. Gen. Virol.* **70**, 3399–3409. 3, 6
- Tamada T., Kusume T., Uchino H., Kiguchi T. and Saito M.** (1996a) Evidence that beet necrotic yellow vein virus rna-5 is involved in symptom development of sugarbeet roots in *Proceedings, 3rd Symposium of the International Working Group on Plant Viruses with Fungal Vectors*, edited by **A.S. of Sugar Beet Technologists** Denver 49. 6
- Tamada T., Schmitt C., Saito M., Guilley H., Richards K. and Jonard G.** (1996b) High resolution analysis of the readthrough domain of beet necrotic yellow vein virus readthrough protein : a KTER motif is important for efficient transmission of the virus by *Polymyxa betae*. *J. Gen. Virol.* **77**, 1359–1367. 4, 7, 71
- Tamada T., Uchino H., Kusume T. and Saito M.** (1999) Rna 3 deletion mutants of beet necrotic yellow vein virus do not cause rhizomania disease in sugar beets *Phytopathology* **89**, 1000–1006. 5, 29, 30, 34, 52
- Tateno Y., Takezaki N. and Nei M.** (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. *Mol. Biol. Evol.* **11**, 261–277. 17
- Thompson J., Higgins D. and Gibson T.** (1994) Clustal w : improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice *Nucleic Acids Research* **22**, 4673–4680. 21, 93

-
- Torrance L. and Mayo M.** (1997) Proposed re-classification of furoviruses. *Arch. Virol.* **142**, 435–439. 2
- Vetter G., Hily J., Klein E., Schmidlin L., Haas M., Merkle T. and Gilmer D.** (2004) Nucleo-cytoplasmic shuttling of the beet necrotic yellow vein virus RNA-3-encoded p25 protein. *J. Gen. Virol.* **85**, 2459–2469. 19, 30, 43, 52, 55, 66, 67, 76
- Wang P., Palese P. and O’Neill R.** (1997) The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J. Virol.* **71**, 1850–1856.
- Ward B. and Lazarowitz S.** (1999) Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay. *Plant Cell* **11**, 1267–1276.
- Weidman M., Sharma R., Raychaudhuri S., Kundu P., Tsai W. and Dasgupta A.** (2003) The interaction of cytoplasmic RNA viruses with the nucleus. *Virus Res.* **95**, 75–85. 43, 51
- Weiner M., Costa G., Schoettlin W., Cline J., Mathur E. and Bauer J.** (1994) Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene* **151**, 119–123. 88
- Weis K.** (1998) Importins and exportins : how to get in and out of the nucleus. *Trends Biochem. Sci.* **23**, 185–189. 47, 48
- Wen W., Meinkoth J., Tsien R. and Taylor S.** (1995) Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463–473. 46
- Whittaker G. and Helenius A.** (1998) Nuclear import and export of viruses and virus genomes. *Virology* **246**, 1–23. 50
- Wolff T., Unterstab G., Heins G., Richt J. and Kann M.** (2002) Characterization of an unusual importin alpha binding motif in the borna disease virus p10 protein that directs nuclear import. *J. Biol. Chem.* **277**, 12151–12157.
- Wood L., Irvin B., Nucifora G., Luce K. and Hiebert S.** (2003) Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. *Proc. Natl. Acad. Sci. USA* **100**, 3257–3262. 50
- Wu J.** (1986) Jackknife, bootstrap and other resampling plans in regression analysis *Annals of Statistics* **14**, 1261–1295. 17
- Wu J., Matunis M., Kraemer D., Blobel G. and Coutavas E.** (1995) Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. *J. Biol. Chem.* **270**, 14209–14213. 49
- Yang J., Liu J. and DeFranco D.** (1997) Subnuclear trafficking of glucocorticoid receptors in vitro : chromatin recycling and nuclear export. *J. Cell Biol.* **137**, 523–538. 49

-
- Yokoyama N., Hayashi N., Seki T., Panté N., Ohba T., Nishii K., Kuma K., Hayashida T., Miyata T. and Aebi U.** (1995) A giant nucleopore protein that binds Ran/TC4. *Nature* **376**, 184–8. 49
- Youvan D. and Michel-Beyerle M.** (1996) Structure and fluorescence mechanism of GFP. *Nat. Biotechnol.* **14**, 1219–1220. 53
- Ziegler V., Richards K., Guilley H., Jonard G. and Putz C.** (1985) Cell-free translation of beet necrotic yellow vein virus : readthrough of the coat protein cistron *J. Gen. Virol.* **66**, 2079–2087. 4
- Ziegler-Graff V., Bouzoubaa S., Jupin I., Guilley H., Jonard G. and Richards K.** (1988) Biologically active transcripts of beet necrotic yellow vein virus RNA-3 and RNA-4 *J. Gen. Virol.* **69**, 2347–2357. 29, 31

Etude de la variabilité moléculaire et du pouvoir pathogène d'isolats naturels du virus des nervures jaunes et nécrotiques de la betterave : caractérisation des protéines codées par les ARN-3 et -5 de différents isolats.

Une étude phylogénétique réalisée sur des isolats viraux mondiaux de BNYVV a permis d'affiner la classification de ce Benyvirus en 5 groupes distincts. Cette nouvelle classification repose sur le type de protéine de capsid, la nature des acides aminés hypervariables en position 67 à 70 de la protéine P25 et enfin sur la présence d'un ARN-5 génomique surnuméraire de type Européen ou Asiatique. L'incidence des onze séquences hypervariables a été testée après mutagenèse dirigée de la protéine P25 provenant du clone infectieux disponible au laboratoire. Ainsi, nous avons démontré que ces quatre acides aminés influencent les capacités de la protéine P25 à dimériser et modulent ses fonctions biologiques sur plante hôte hypersensible.

Quant à l'ARN-5 de type Européen, un clone ADNc infectieux a pu être produit et son influence testée sur le pouvoir pathogène du virus. L'étude de la protéine P26 codée par cet ARN-5 génomique a été initiée. Son action nécrogène sur plante hôte a été démontré. La localisation nucléaire de la protéine P26 a pu être observée à la fois en contexte viral (immuno-marquages et fusion GFP) et hors contexte viral (fusion GFP). La recherche de partenaires viraux et cellulaire en système double-hybride a révélé la forte capacité intrinsèque de la protéine à activer la transcription dans la levure. Le domaine responsable de cette activation de la transcription a pu être localisé par mutagenèse dirigée dans la partie N-terminale de la protéine, domaine également impliqué dans l'adressage nucléaire de la protéine.

L'ensemble des résultats nous permet de proposer un nouveau modèle évolutif en se basant sur le postulat d'une pathogénicité décroissante des isolats de BNYVV, qu'ils aient 4 ou 5 ARN génomiques de polarité positive.

Molecular epidemiology and pathogenicity studies of *Beet necrotic yellow vein virus* isolates: characterization of RNA-3 and -5-encoded proteins.

Based on the CP sequence, P25 variability at aa positions 67-70, and the presence of a fifth RNA, a new BNYVV classification into 5 groups has been proposed. The effect of the hypervariable sequence of P25 has been reproduced by mutagenesis and tested upon the biological properties of the infectious clone. We have demonstrated an important role of the hypervariable sequence upon dimerization of the P25 protein and the induced symptoms on plant leaves.

A full-length cDNA clone of RNA-5 has been produced and the encoded P26 protein studied. P26 was located in the nuclear compartment of infected and transfected cells, by transmission electron microscopy and GFP fusion protein, respectively. A two-hybrid screen for viral and cellular partners revealed strong transcription activation activity located within the N-terminal part of the P26 protein. Such a domain was involved in the nuclear targeting of P26 protein. On the basis of our results and observations, we proposed a new evolutionary model for this multi-component positive-stranded RNA virus.