

UNIVERSITÉ DU QUÉBEC

THÈSE

PRÉSENTÉE À

L'INSTITUT ARMAND-FRAPPIER

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN VIROLOGIE

PAR

OLIVIER NICOLAS

**CARACTÉRISATION MOLÉCULAIRE DU VIRUS
DE LA MOSAÏQUE DU NAVET**

MAI 1993

TABLE DES MATIÈRES

TABLE DES MATIÈRES	iii
LISTE DES ARTICLES	viii
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	x
LISTE DES ABRÉVIATIONS	xi
SOMMAIRE	xiv
INTRODUCTION	1
REVUE BIBLIOGRAPHIQUE.	3
1. BREF HISTORIQUE DE L'ÉTUDE DES VIRUS À ARN	4
2. LES VIRUS À ARN DES PLANTES	5
3. LE VIRUS DE LA MOSAÏQUE DU NAVET	7
3.1. GÉNÉRALITÉS	7
3.2. BIOLOGIE DU TUMV	8
3.2.1 Espèces infectées	8
3.2.2 Interactions TuMV-hôte.	11
3.2.3 Propagation	13
3.3 CONCLUSION	13
4. LES POTYVIRUS	14
4.1 GÉNÉRALITÉS	14
4.2 STRUCTURE DES POTYVIRUS	16
4.3 GÉNOME DES POTYVIRUS	17
4.4 STRATÉGIE D'EXPRESSION DES POTYVIRUS	20
4.5 CLIVAGE DE LA POLYPROTÉINE DES POTYVIRUS	21
4.5.1 Introduction	21
4.5.2 Protéase NIa	22
4.5.3 Protéase HC	25
4.5.4 Protéase P1	26
4.5.5 Régulation de la protéolyse	27
4.6 PROTÉINES DES POTYVIRUS	28

4.6.1	Introduction	28
4.6.2	Protéine de la capside	31
4.6.3	Protéine NIb	34
4.6.4	Protéines NIa-Pro et VPg	35
4.6.5	Protéines 6K1 et 6K2	37
4.6.6	Protéine CI	38
4.6.7	Protéine P3	40
4.6.8	Protéine HC	40
4.6.9	Protéine P1	42
4.7	RÉPLICATION DES POTYVIRUS	43
4.8	DÉPLACEMENT DES POTYVIRUS	44
4.8.1	Déplacement dans la plante	44
4.8.2	Transport par les vecteurs	45
4.8.3	Transport par les graines	45
4.9	ÉVOLUTION DES POTYVIRUS	46
5.	MODE D'ÉTUDE DES POTYVIRUS	49
5.1	CLONAGE ET SÉQUENÇAGE	49
5.2	TRANSCRITS INFECTIEUX	50
6.	CONCLUSION ET OBJECTIFS DE LA RECHERCHE	51
MATÉRIEL ET MÉTHODES		53
1	BACTÉRIES, VECTEURS, PLANTES ET VIRUS	54
1.1	SOUCHES D' <i>ESCHERICHIA COLI</i>	54
1.2	VECTEURS DE CLONAGE	54
1.3	PLANTES	55
1.4	VIRUS	55
2.	MILIEUX DE CULTURE, TAMPONS ET SOLUTIONS	55
2.1	MILIEUX DE CROISSANCE DES BACTÉRIES	55
2.2	ANTIBIOTIQUES	56
2.3	TAMPONS ET SOLUTIONS	57
3.	PRÉPARATION DES ADN	58
3.1	PLASMIDES	58

3.2	ADN SIMPLE BRIN	60
3.3	OLIGONUCLÉOTIDES	62
4.	MANIPULATION DES ADN	62
4.1	ÉLECTROPHORÈSE SUR GEL	62
4.1.1	Gel d'agarose	62
4.1.2	Gel alcalin	62
4.1.3	Gel de polyacrylamide	63
4.2	AMPLIFICATION PAR POLYMÉRISATION EN CHAÎNE	63
4.3	DIGESTION ENZYMATIQUE	65
4.4	PURIFICATION DES ADN APRÈS TRAITEMENT ENZYMATIQUE	65
4.5	CLONAGE DES ADN	66
4.5.1	Clonage par ligation	66
4.5.2	Clonage sans ligation (« clonage dC-dG »)	67
4.6	TRANSFORMATION D' <i>E. COLI</i>	68
4.7	SÉQUENÇAGE DES ADN	69
4.8	TRANSFERT, SONDÉS ET HYBRIDATIONS	69
4.8.1	Transfert sur nitrocellulose	69
4.8.2	Transfert d'ADN d'un gel sur membrane (« <i>Southern blot</i> »)	70
4.8.3	Synthèse des sondes	70
4.8.3.1	Sondes d'ADN double brins	70
4.8.3.2	Sondes d'ADN simple brin	70
4.8.4	Hybridation avec sonde ADN	70
5.	MANIPULATION DES ARN	71
5.1	EXTRACTION DE L'ARN VIRAL	71
5.2	ÉLECTROPHORÈSE DES ARN	72
5.3	SYNTHÈSE DES PREMIERS BRINS D'ADNc	72
5.4	TRANSCRIPTION <i>IN VITRO</i>	73
6.	ÉLECTROPHORÈSE DES PROTÉINES	74
7.	INOCULATION DES PLANTES	74

RÉSULTATS	75
1. PURIFICATION DU TUMV ET EXTRACTION DE L'ARN	76
2. CLONAGE ET SÉQUENÇAGE DU TUMV	82
2.1 INTRODUCTION AUX ARTICLES	82
2.2 ARTICLE 1: <i>Journal of General Virology</i> (1990), 71, 2769-2772. <i>Sequence of the 3'-terminal region of turnip mosaic virus RNA and the capsid protein gene</i>	84
2.2.1 Résumé	85
2.2.2 Contribution	85
2.2.3 Conclusion	86
2.3 ARTICLE 2: <i>Journal of Virological Methods</i> (1991), 32: 57-66. <i>The use of PCR for cloning of large cDNA fragments of turnip mosaic potyvirus</i>	88
2.3.1 Résumé	89
2.3.2 Contribution	89
2.3.3 Conclusion	89
2.4 ARTICLE 3: <i>Journal of General Virology</i> (1992), 73, 2785-2793. <i>The complete nucleotide sequence of turnip mosaic potyvirus RNA</i>	91
2.4.1 Résumé	92
2.4.2 Contribution	92
2.4.3 Conclusion	93
2.5 ARTICLE 4: <i>Virology</i> (1992), 190: 510-514. <i>Release of a 22-kDa protein derived from the amino-terminal domain of the 49-kDa NIa of turnip mosaic potyvirus in Escherichia coli</i>	94
2.5.1 Résumé	95
2.5.2 Contribution	95
2.5.3 Conclusion	96
3. OBTENTION DE TRANSCRITS DE TUMV	96
3.1 INTRODUCTION	96
3.2 ASSEMBLAGE D'UN CLONE COMPLET DE TUMV	98

3.2.1	Assemblage de la région 5' de TuMV	99
3.2.2	Assemblage de la région 3' de TuMV	102
3.2.3	Assemblage des régions 5' et 3'	106
3.3	TRANSCRIPTION <i>IN VITRO</i>	108
3.4	ÉVALUATION DES TRANSCRITS SYNTHÉTIQUES	111
DISCUSSION		112
1.	CARACTÉRISATION DU VIRUS DE LA MOSAÏQUE DU NAVET . . .	113
1.1	CLONAGE ET SÉQUENÇAGE	113
1.2	ANALYSE DE LA SÉQUENCE	118
2.	CONSTRUCTION D'UN CLONE COMPLET DE TUMV	130
CONCLUSION		137
REMERCIEMENTS		140
BIBLIOGRAPHIE		142
ANNEXES		157
	ANNEXE 1: AMORCES OLIGONUCLÉOTIDIQUES	158
	ANNEXE 2: AUTRE ARTICLE	160

LISTE DES ARTICLES

RÉSULTATS:

- ARTICLE 1: Sequence of the 3'-terminal region of turnip mosaic virus RNA and the capsid protein gene. *Journal of General Virology* (1990), 71: 2769-2772 84
- ARTICLE 2: The use of PCR for cloning of large cDNA fragments of turnip mosaic potyvirus. *Journal of Virological Methods* (1991), 32: 57-66 88
- ARTICLE 3: The complete nucleotide sequence of turnip mosaic potyvirus RNA. *Journal of General Virology* (1992), 73, 2785-2793 91
- ARTICLE 4: Release of a 22-kDa protein derived from the amino-terminal domain of the 49-kDa NIa of turnip mosaic potyvirus in *Escherichia coli*. *Virology* (1992), 190: 510-514 94

ANNEXES:

- The xylanase introns from *Cryptococcus albidus* are accurately spliced in transgenic tobacco plants. *Plant Molecular Biology* (1992), 18: 447-451 160

LISTE DES TABLEAUX

TABLEAU I :	Espèces d'importance économique sensibles au virus TuMV. . .	10
TABLEAU II :	Composition en bases des potyvirus	119
TABLEAU III:	Pourcentage d'homologie en nucléotides et en acides aminés au niveau de la région 3' non-codante et de la protéine de la capside de trois souches de TuMV	124

LISTE DES FIGURES

FIGURE 1	: Organisation du génome et de la polyprotéine des potyvirus	18
FIGURE 2	: Sites de clivages protéolytiques reconnus par les protéases NIa du PPV et du TEV	23
FIGURE 3	: Représentation des événements protéolytiques chez les potyvirus . .	29
FIGURE 4	: Carte génétique de trois potyvirus	30
FIGURE 5	: Comparaison génomique des picornavirus (poliovirus) et des virus du super-groupe « type-picornavirus » (potyvirus, népovirus et comovirus)	48
FIGURE 6	: Observation au microscope électronique en coloration négative d'une préparation de virus TuMV purifié	77
FIGURE 7	: Protéine de la capsidite du TuMV sur gel de polyacrylamide 12 % en présence de SDS (SDS-PAGE) et colorée au bleu de Coomassie. .	78
FIGURE 8	: Observation des corps d'inclusion cytoplasmiques induits par le virus TuMV	80
FIGURE 9	: Analyse de l'ARN purifié du TuMV sur gel dénaturant à 1 % d'agarose	81
FIGURE 10:	: Assemblage de la région 5' du TuMV conduisant au clone pKS4900	100
FIGURE 11	: Séquence du clone pKS2500 située vers l'amorce JF _{SalI7}	101
FIGURE 12	: Assemblage de la région 3' du TuMV conduisant au clone pTUM4100	103
FIGURE 13	: Assemblage de la région 3' du TuMV conduisant au clone pSK5000	104
FIGURE 14	: Assemblage final de la séquence ADN du TuMV conduisant au clone pTUM9900	107
FIGURE 15	: Représentation schématique de l'obtention des transcrits synthétiques du TuMV	109
FIGURE 16	: Analyse des transcrits synthétiques par électrophorèse sur gel dénaturant d'agarose	110
FIGURE 17	: Caractéristiques génomiques des potyvirus connus	121

LISTE DES ABRÉVIATIONS

3'-NC	: région 3' non-codante
5'-NC	: région 5' non-codante
6K1	: protéine 1 de 6 kDa
6K2	: protéine 2 de 6 kDa
A	: adénine
ADN	: acide désoxyribonucléique
ADNc	: ADN complémentaire
Ala	: alanine
Ampi	: ampicilline
Arg	: arginine
ARN	: acide ribonucléique
Asp	: acide aspartique
AT	: « <i>aphid transmissible</i> » (transmissible par les pucerons)
ATB	: antibiotiques
BaYMV	: « <i>barley yellow mosaic virus</i> »
Bet	: bromure d'éthidium
C	: cytosine
CaMV	: « <i>cauliflower mosaic virus</i> » (virus de la mosaïque du chou-fleur)
CI	: « <i>cylindrical inclusion</i> » (corps d'inclusions cylindriques)
Ci	: curie
CP	: « <i>capsid protein</i> » (protéine de la capside)
cv.	: cultivar
Cys	: cystéine
db	: double brins
DEPC	: diéthyl pyrocarbonate
DNase	: désoxyribonucléase
DTT	: dithiothreitol
EDTA	: éthylène diamine tétra-acétique
ELISA	: « <i>enzyme-linked-immunosorbent-assay</i> »
EMBL	: « <i>European Molecular Biology Library</i> »
G	: guanine
Gln	: glutamine
Glu	: acide glutamique
Gly	: glycine
GUS	: β -glucuronidase
HAV	: « <i>hypovirulence-associated virus</i> » (virus associé à l'hypovirulence du chancre du châtaignier)
HC	: « <i>helper component</i> » (protéine auxiliaire)
HC-Pro	: protéine HC à activité protéase
His	: histidine
ICTV	: « <i>International Committee for Taxonomy of Virus</i> » (Comité International de Taxonomie des Virus)
Ile	: isoleucine

IPTG	: isopropyl thio-galactoside
kDa	: kilo Dalton
KGB	: tampon potassium-glutamate
LB	: milieu de Luria-Bertani
Leu	: leucine
Lys	: lysine
MD	: marque déposée
Met	: méthionine
min	: minute
MMLV	: « <i>murine moloney leukemia virus</i> » (virus de la leucémie murine)
NAT	: « <i>non-aphid transmissible</i> » (non-transmissible par les pucerons)
NIa	: « <i>nuclear inclusion a</i> » (corps d'inclusion nucléaires a)
NIa-Pro	: protéine NIa à activité protéase
NIb	: « <i>nuclear inclusion b</i> » (corps d'inclusion nucléaires b)
NTBM	: « <i>nucleoside triphosphate binding motif</i> »
P1	: protéine P1 des potyvirus
P3	: protéine P2 des potyvirus
pb	: paire de bases
PCR	: « <i>polymerase chain reaction</i> » (réaction de polymérisation en chaîne)
PEG	: polyéthylène glycol
PeMoV	: « <i>pepper mottle virus</i> »
Phe	: phénylalanine
PM	: poids moléculaire
PNK	: polynucléotide kinase
Pol	: polymérase
PPV	: « <i>plum pox virus</i> »
Pro	: protéase
PSbMV	: « <i>pea seed-borne mosaic virus</i> »
PVY	: « <i>potato virus Y</i> » (virus Y de la pomme de terre)
RF	: « <i>replicative form</i> » (forme répliquative)
RNase	: ribonucléase
RTase	: « <i>reverse transcriptase</i> » (transcriptase inverse)
sb	: simple brin
SDS	: sodium-dodécyl-sulfate
SDS-PAGE	: « <i>SDS-polyacrylamide gel electrophoresis</i> » (électrophorèse en gel de polyacrylamide en présence de SDS)
Ser	: sérine
SMC	: site multiple de clonage
SMV	: « <i>soybean mosaic virus</i> »
SSC	: sodium sodium citrate
ssp.	: « <i>subspecies</i> » (sous-espèce)
T	: thymine
Taq	: « <i>Thermus aquaticus</i> »
TdT	: terminal-désoxy-transférase
TE	: tampon Tris-EDTA

Tet	: tétracycline
TEV	: « <i>tobacco etch virus</i> »
TMAC	: « <i>tetramethyl ammonium chloride</i> »
TuMV	: « <i>turnip mosaic virus</i> » (virus de la mosaïque du navet)
TMV	: « <i>tobacco mosaic virus</i> » (virus de la mosaïque du tabac)
TVMV	: « <i>tobacco vein mottling virus</i> »
Tyr	: tyrosine
U	: uracile
UV	: ultra-violet
V	: volume
Val	: valine
var.	: variété
VPg	: « <i>viral protein genome</i> »
X-gal	: 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

SOMMAIRE

Le virus de la mosaïque du navet (TuMV) est un virus phyto-pathogène extrêmement important. En effet, ce virus infecte un très grand nombre d'espèces appartenant principalement à la famille des crucifères, dont beaucoup sont d'importance économique; ce virus est également présent sur tous les continents. Au Québec et en Ontario, le TuMV est responsable d'une maladie grave affectant principalement le rutabaga (*Brassica napus* ssp. *napobrassica*).

Le TuMV appartient au groupe des potyvirus, qui est le plus important parmi ceux des virus de plantes. Il a un génome composé d'un unique ARN d'environ 10 000 bases, polyadénylé en 3', et qui possède une protéine (VPg) liée de façon covalente en 5'. Son virion a une forme de bâtonnet flexueux composé d'environ 2 000 copies de la protéine structurale principale qui est la protéine de la capsid (CP).

Dans l'objectif d'une caractérisation moléculaire du TuMV, et afin de pouvoir intégrer nos résultats à la compréhension des mécanismes moléculaires qui conduisent à la multiplication et la propagation des potyvirus, nous avons entrepris de cloner et de séquencer son génome. Nous avons utilisé une nouvelle stratégie pour amplifier et cloner l'ARN du TuMV. Celle-ci a consisté à l'amplification des premiers brins d'ADNc du TuMV par polymérisation en chaîne (PCR) avec l'ADN polymérase thermostable de *Thermus aquaticus* (Taq polymérase). Les produits d'amplification par PCR ont été clonés sans ligation dans le vecteur pUC9 dans une première étape, puis dans une deuxième étape, par ligation dans le vecteur pUC19, grâce à des sites de restriction introduits au cours de l'amplification. Certaines amorces spécifiques provenaient de la séquence réelle du TuMV alors que d'autres, dégénérées, provenaient d'homologies de séquence parmi quatre potyvirus séquencés. Le génome du TuMV a pu être ainsi cloné au complet grâce à sept amplifications successives par PCR. La séquence de l'extrémité 5' de l'ARN a été confirmée par une réaction d'extension d'amorce (« primer extension »).

Les séquences provenant des différents clones d'amplification ont été alignées et l'ARN du TuMV présente les caractéristiques suivantes: il possède un ARN de 9 163 bases suivies d'une queue polyadénylée d'au moins 55 résidus; cet ARN comporte une séquence 5' non-codante de 129 nucléotides (nt) suivie d'une séquence codant pour une polyprotéine de 3 163 acides aminés et d'une séquence 3'-NC de 212 nt. La polyprotéine a été analysée et neuf sites de clivage reconnus par les trois protéases virales ont été localisés. La protéine de la capsid a été séquencée à son extrémité aminée, ce qui a permis de localiser le premier site de clivage par la protéase NIa. Nous avons trouvé six autres sites potentiels de clivage par la NIa dont un, à l'intérieur même de la NIa, a été démontré chez *E. coli*. Ce site interne et celui situé en N-terminal de la NIa sont légèrement différents des autres du fait qu'en position -1 de la liaison peptidique hydrolysée, se trouve un acide glutamique plutôt qu'une glutamine. Les sites de clivage par les protéases P1 et HC ont également été localisés sur la polyprotéine. Les dix protéines potentielles pouvant donc être libérées après protéolyse par les trois protéases

virales ont été analysées et quelques hypothèses ont été élaborées quant à leur rôle probable sur la multiplication ou la propagation du virus.

Nous avons également construit un clone complet d'ADN du TuMV en aval d'un promoteur d'origine phagique. Ceci était une des étapes importantes conduisant à l'obtention d'un clone infectieux, c'est-à-dire d'un clone capable d'induire une infection du même type que celle obtenue avec l'ARN viral natif. La maîtrise des transcrits infectieux est un atout important pour l'étude moléculaire d'un virus tel que le TuMV car cela permet d'introduire des mutations sur l'ARN tout en conservant son caractère infectieux. Des transcrits synthétiques pratiquement identiques à l'ARN viral, mais comportant en 5' un analogue de coiffe méthylée, ont été synthétisés *in vitro*. Ces transcrits ont été testés par inoculation mécanique de deux espèces sensibles au TuMV.

INTRODUCTION

Le virus de la mosaïque du navet (TuMV) est un virus phyto-pathogène qui cause de sévères dégâts au Québec et en Ontario, principalement sur le rutabaga (*Brassica napus* ssp. *napobrassica*), qui est une espèce d'importance économique dans ces provinces. Le TuMV est classé parmi les potyvirus, groupe le plus important de virus pathogènes des plantes, de par ses caractéristiques morphologiques et génomiques, ainsi que de par les inclusions cytoplasmiques qu'il induit. Au début de ce travail, seuls quatre potyvirus avaient été entièrement clonés et séquencés parmi plus de 120 membres reconnus. Afin d'être capable d'analyser le rôle des différentes protéines codées par le TuMV, et à plus long terme, de mieux comprendre les mécanismes moléculaires intervenant dans la multiplication et la pathologie des potyvirus, deux objectifs ont été définis pour ce travail de doctorat:

- Cloner et séquencer complètement le génome du TuMV.
- Produire un ARN synthétique le plus identique possible à l'ARN viral.

L'intérêt de séquencer le TuMV était d'une part, de localiser les différentes protéines codées par son génome, et donc de commencer leur étude, après avoir produit des anticorps dirigés contre elles. Ceux-ci peuvent être ensuite utilisés *in vitro* ou *in vivo* pour analyser l'activité des différentes protéines virales. D'autre part, la production d'un ARN synthétique semblable au virus était la première étape conduisant à l'obtention d'un clone infectieux, c'est-à-dire d'un clone permettant de modifier de façon contrôlée le génome viral pour l'étude du rôle précis de ses régions codantes et non-codantes dans les mécanismes d'infection.

REVUE BIBLIOGRAPHIQUE

1. BREF HISTORIQUE DE L'ÉTUDE DES VIRUS À ARN.

L'année 1992 a permis de célébrer le centième anniversaire de la découverte du caractère « filtrant » du virus de la mosaïque du tabac (TMV) par Ivanowski (1892; références antérieures à 1900 citées dans Matthews, 1991); en effet, un broyat de plante infectée au TMV garde sa nature infectieuse après passage sur filtre de porcelaine. Depuis lors, le TMV est demeuré au coeur des grandes découvertes de la virologie.

Beijerinck (1898) montra ensuite que TMV diffuse à travers l'agar et qu'il peut survivre pendant deux ans dans des feuilles séchées de tabac. Ses conclusions furent que le TMV est une particule très petite, stable et soluble, capable de se multiplier dans les cellules vivantes. Stanley (1935) fut le premier à cristalliser un virus, le TMV, en utilisant les techniques de purification des protéines; le virus purifié et cristallisé garde son caractère infectieux. Bawden *et al.* (1936) montrèrent ensuite que le TMV est une nucléoprotéine dont la fraction nucléique contient du ribose. Enfin, Gierer et Schramm (1956) furent les premiers à dissocier les protéines de l'acide nucléique du TMV et à démontrer ainsi la nature infectieuse d'un ARN viral. Anderer *et al.* (1960) obtinrent la première séquence primaire d'une protéine virale qui était la protéine de la capsid (CP) du TMV. Dans les années 60 et 70, les techniques de microscopie électronique et de cristallographie et d'analyse aux rayons-X furent majoritairement utilisées pour l'étude structurale des virus de plantes, et il fallut attendre les années 80 pour que le génome des virus à ARN puisse être cloné et séquencé. Le premier virus ARN à avoir été séquencé au complet est le virus de la poliomyélite (poliovirus) (Racaniello et Baltimore, 1981a), suivi par la séquence du TMV, premier virus de plantes à ARN (Goellet *et al.*, 1982).

L'analyse de la séquence de ces virus a apporté beaucoup d'information sur leur stratégie d'expression ainsi que sur leurs protéines.

Une autre étape importante dans l'avancement des connaissances concernant les virus à ARN fut l'obtention de clones infectieux de ces virus, c'est-à-dire d'une copie ADN du génome qui conduit, elle-même ou après transcription *in vitro*, à une multiplication virale du même type que celle obtenue par l'ARN viral. Ainsi, Racaniello et Baltimore (1981*b*) obtinrent le premier clone infectieux du poliovirus, alors que Ahlquist *et al.* (1984) obtenaient celui du virus de la mosaïque du brome (BMV), un virus de plantes à génome segmenté. TMV fut quant à lui le deuxième virus à ARN de plantes pour lequel un clone infectieux a été obtenu (Dawson *et al.*, 1986). Les virus des plantes ont donc souvent été impliqués dans la découverte des principes fondamentaux concernant les virus à ARN, et il est évident que leur étude permettra encore de découvrir un certain nombre d'événements moléculaires propres à ces virus.

2. LES VIRUS À ARN DE PLANTES.

Les virus de plantes ont toujours suscité un grand intérêt parce qu'ils sont présents partout et qu'ils sont extrêmement bien adaptés à leur environnement, à leurs hôtes et aux vecteurs qui les véhiculent. L'étude de ces organismes a été, comme je l'ai montré dans la section précédente, d'un intérêt majeur dans l'avancement des connaissances en virologie.

En 1991, le Comité International de Taxonomie des Virus (ICTV) a classé 590 virus de plantes dans deux familles et 32 groupes (Matthews, 1991). Plus de 93 %

(550 membres) de tous ces virus ont un génome à ARN et les virus à ARN simple brin sont représentés à plus de 87 % (515 membres).

Il est reconnu que les virus à ARN se modifient génétiquement beaucoup plus rapidement que leur hôte du fait de leur taux élevé de mutations et des générations beaucoup plus courtes (Steinhauer et Holland, 1987). La polymérase du poliovirus par exemple, a une fréquence d'erreur estimée à $4.1 \times 10^{-3} \pm 0.6 \times 10^{-3}$, ce qui est environ 100 000 fois plus élevé que le taux d'erreur estimé des ADN polymérases (Ward et Flanagan, 1992). Les virus à ARN seraient donc capables de s'adapter très rapidement à tout changement du métabolisme de leur hôte ou aux conditions du milieu; ceci expliquerait pourquoi ces virus sont présents en aussi grand nombre, et comment ils infectent pratiquement tous les organismes vivants.

Les virus de plantes dont l'ARN est simple brin à polarité positive représentent 74 % du total des virus de plantes (437 membres). Ces virus ont développé cinq stratégies différentes, mais non-mutuellement exclusives, pour parvenir à la régulation de la synthèse de leurs protéines:

- Leur génome peut être segmenté et l'information génétique est portée par plusieurs brins d'ARN mono- ou multi-cistroniques.
- Il peut y avoir synthèse d'ARN sous-génomiques à partir de l'ARN génomique, et la régulation de l'expression se fait principalement au niveau de leur synthèse.
- L'ARN peut être traduit en une polyprotéine, qui est ensuite clivée à des sites spécifiques par certaines protéases pour donner les produits finaux.

- Il peut y avoir lecture au travers (« *read-through* ») du premier codon d'arrêt par un ARN de transfert suppresseur, et un ARN peut ainsi coder pour deux protéines en partant du même codon d'initiation.

- Il peut y avoir des cadres de lecture chevauchants et deux protéines peuvent être lues à partir du même ARN: les ribosomes exécutent un changement du cadre de lecture vers la fin de la traduction de la première protéine, ce qui conduit à la synthèse de la deuxième protéine.

Les virus à ARN se servent donc d'une ou de plusieurs de ces stratégies pour l'expression dans la cellule hôte des protéines nécessaires à leur réplication. Ce sont donc des organismes très complexes qui ont été capables de s'affranchir de la synthèse d'ADN.

3. LE VIRUS DE LA MOSAÏQUE DU NAVET.

3.1 GÉNÉRALITÉS.

Le virus de la mosaïque du navet (TuMV) a été décrit pour la première fois par Gardner et Kendrick (1921). Il a ensuite été classé dans le groupe du virus Y de la pomme de terre (PVY), c'est-à-dire parmi les potyvirus, du fait de ses caractéristiques morphologiques et sérologiques (Brandes et Berks, 1965). La particule virale du TuMV a la forme d'un bâtonnet flexible d'une longueur moyenne de 740 nm et d'un diamètre moyen de 14 nm (Edwardson et Christie, 1986). La protéine de la capsid (CP) du TuMV, sur gel dénaturant de polyacrylamide (SDS-PAGE), a un poids moléculaire (PM) d'environ 34K (Hiebert et McDonald, 1976). Les propriétés biochimiques de son acide nucléique

ont été déterminées pour la première fois par Hill et Shepherd (1972): le TuMV possède un unique ARN de 3.5×10^6 daltons, ce qui représente une taille d'environ 10 000 bases et correspond à 5 % de sa masse totale; de plus, sa teneur en adénosine est très élevée (34.5 ± 2.5 %). Lorsque cet ARN est traduit dans un lysat de réticulocyte de lapin, il apparaît un polypeptide précurseur d'au moins 200 kDa, qui est ensuite clivé en polypeptides de plus faible poids moléculaire si le temps d'incubation est prolongé (Shields et Wilson, 1987). Ce résultat, appuyé par le fait qu'aucun ARN sous-génomique n'est détecté, démontre donc que la stratégie du TuMV passe par la synthèse d'une unique polyprotéine qui subit une série de protéolyses.

3.2 BIOLOGIE DU TUMV.

3.2.1 Espèces infectées.

Le TuMV est un virus très répandu sur tous les continents puisqu'il a été détecté au Japon, en Chine, en Corée, au Royaume-Uni, en Europe centrale et de l'Ouest, au Canada, aux USA et en Afrique du Sud (Smith, 1972). Le TuMV infecte au moins 299 espèces végétales appartenant à 147 genres dans 39 familles, dont celle des crucifères est la plus représentée (43 % des espèces sensibles au TuMV sont des crucifères) (Edwardson et Christie, 1986). Son spectre d'hôte est extrêmement large pour un potyvirus, surtout dans la famille des crucifères, et ceci soulève naturellement la question de savoir ce qui définit la sensibilité d'une espèce à un potyvirus tel que TuMV. Le Tableau I indique quelques-unes des espèces d'importance économique sensibles au TuMV.

Le rutabaga et le chou sont les crucifères les plus cultivées au Québec et en

Ontario, mais ces deux espèces sont extrêmement sensibles à l'infection par le TuMV, ce qui cause des pertes de production allant jusqu'à 30 % (Shattuck et Stobbs, 1987). Sur 49 cultivars de rutabaga cultivés en Ontario, seuls quatre sont partiellement résistants à l'infection par le TuMV, c'est-à-dire que le virus se multiplie dans ces cultivars à un niveau très faible sans provoquer de symptômes (Shattuck et Stobbs, 1987). Des croisements génétiques ont montré que cette résistance partielle était due à un gène dominant unique, et il serait extrêmement intéressant d'isoler ce gène afin de comprendre le phénomène de résistance qu'il induit.

Les épidémies d'infection au TuMV ont sensiblement augmentées dans le sud de l'Ontario, sur les cultures de rutabaga ainsi que sur d'autres crucifères sensibles au TuMV (chou, brocoli, chou-chinois) et ce, depuis que le colza d'hiver (*Brassica napus* ssp. *oleifera*) a été introduit dans cette province. Cette observation suggère fortement que le colza est une espèce sensible au TuMV et que des mesures sanitaires devraient être prises pour éviter d'autres épidémies d'infection au TuMV (Stobbs et Shattuck, 1989).

TABLEAU I: Espèces d'importance économique sensibles au TuMV.

nom latin	nom commun
<i>Brassica campestris</i> ssp. <i>pekinensis</i>	chou chinois
<i>Brassica campestris</i> ssp. <i>rapifera</i>	navet
<i>Brassica napus</i> ssp. <i>oleifera</i>	colza (canola)
<i>Brassica napus</i> ssp. <i>napobrassica</i>	rutabaga
<i>Brassica oleracea</i> ssp. <i>botrytis</i>	chou-fleur
<i>Brassica oleracea</i> ssp. <i>italica</i>	broccoli
<i>Brassica rapa</i>	moutarde
<i>Cucurbita pepo</i>	courge
<i>Daucus carota</i>	carotte
<i>Lactuca sativa</i>	laitue

Le TuMV infecte de façon systémique *Arabidopsis thaliana* (communication personnelle du Dr. M. Fortin de l'Université McGill). Cette plante sauvage est d'un intérêt majeur en génétique: elle a en effet le plus petit génome connu parmi les plantes supérieures avec 145×10^6 paires de bases (pb), ce qui est 30 fois plus faible que le tabac (*Nicotiana tabacum* L.) (Arumuganathan et Earle, 1991); de plus, son cycle végétatif d'environ 15 jours et sa taille réduite permettent aux généticiens de tester en peu de temps un très grand nombre de plantes sur un espace de culture très réduit. *Arabidopsis thaliana* serait donc une espèce de choix pour l'étude des interactions « TuMV-hôte » ainsi que pour la recherche, par mutagenèse aléatoire de son génome, de gènes de résistance au TuMV. Ces gènes pourraient par la suite être isolés et transférés à d'autres plantes sensibles au TuMV mais d'intérêt économique. Le TuMV est donc un virus important de par sa répartition géographique et de par les espèces d'importance économique et biologique qu'il infecte.

3.2.2 Interactions TuMV-hôte.

Il y a différents types de réponses d'une plante à l'infection par un virus tel que le TuMV. Celle-ci peut être immunisée (« non-hôte »), dans quel cas il n'y a aucune multiplication du virus, ou bien la plante est un hôte pour le virus, auquel cas la multiplication du virus engendre différents types de réponse de la plante:

- L'extrême hypersensibilité: le virus ne se multiplie au point d'infection que dans la cellule initiale; celle-ci meurt, ce qui empêche la propagation du virus.

- La résistance (ou hypersensibilité): l'infection est limitée à une zone de cellules situées autour du point d'infection, ce qui est caractérisé par la formation de lésions locales nécrotiques et qui empêche la propagation du virus.
- La tolérance: il y a peu de symptômes et l'infection est latente, c'est-à-dire que le virus se multiplie et peut-être transmis à d'autres plantes.
- La sensibilité: la plante réagit avec des symptômes plus ou moins sévères qui peuvent affecter beaucoup sa croissance.

Les symptômes causés par le TuMV sont variés et, comme je l'ai mentionné pour le rutabaga (*Brassica napus* ssp. *napobrassica*), plusieurs cultivars d'une même espèce peuvent réagir différemment à l'infection virale. Voici quelques exemples des symptômes provoqués par le TuMV: lésions locales pour certains hôtes (*Nicotiana tabacum* L. var. SR1, *Chenopodium quinoa*, *Chenopodium amaranticolor*) et infection systémique avec mosaïques pour d'autres (*Brassica campestris* ssp. *rapifera* cv. « Just Right », *Brassica perviridis* var. *tendergreen*, *Lactuca sativa* L. var. *Calmar* et *Arabidopsis thaliana*). Les symptômes de mosaïques peuvent être suivis d'une déformation importante des feuilles et d'une baisse considérable de la croissance du plant infecté. De plus, des infections opportunistes peuvent conduire à la mort du plant virosé.

Les cellules infectées par le TuMV présentent au niveau du cytoplasme, par observation au microscope électronique, des inclusions cylindriques en forme enroulée (« *pinwheel* ») caractéristiques des potyvirus, et la morphologie exacte de ces inclusions dépend de la souche virale. La protéine formant ces inclusions cylindriques (protéine CI)

a un PM d'environ 70K (McDonald et Hiebert, 1975). Il y a au moins deux types de souches de TuMV qui se distinguent par la morphologie des inclusions cylindriques qu'elles engendrent ainsi que par leur spectre d'hôte (Stobbs et Van Schagen, 1987): le type I donne des lésions locales sur *Nicotiana glutinosa* et des symptômes de mosaïques sur plusieurs espèces du genre *Brassica* dont le rutabaga; le type II ne donne aucune infection sur ces espèces mais infecte par contre *Alliaria petiolata*, une autre crucifère.

3.2.3 Propagation.

A ce jour, aucune souche de TuMV transmissible par les graines n'a été rapportée, mais le virus peut survivre dans les parties souterraines de certaines espèces sauvages. C'est fort probablement de cette manière que le virus survit à la saison d'hiver dans les pays tempérés (Stobbs et Van Schagen, 1987). Le TuMV est transmis de façon non-persistante par 89 espèces de pucerons appartenant à 27 genres différents (Edwardson et Christie, 1986); ceci signifie que le virus reste viable dans l'appareil de succion de l'insecte pendant quelques minutes à quelques heures après la piqûre (Stobbs et Van Schagen, 1987). C'est au cours de cette courte période que le puceron transmet l'agent infectieux à des plantes saines, ce qui contribue à la propagation du virus. Aucune étude concernant les protéines du TuMV impliquées dans sa transmission par les pucerons n'a encore été faite.

3.3 CONCLUSION.

Le TuMV est donc un virus à ARN très important et très répandu. Sa biologie moléculaire est peu connue et, seules la caractérisation de son génome et la détermination

de ses protéines pouvaient permettre une meilleure compréhension des mécanismes qui contrôlent sa multiplication et sa transmission. Il faut donc bien analyser les données obtenues à partir de l'étude d'autres potyvirus avant de pouvoir intégrer nos résultats concernant le TuMV.

4. LES POTYVIRUS.

4.1 GÉNÉRALITÉS.

Le nom de ce groupe provient de celui de son membre type, le virus Y de la pomme de terre (PVY = *potato virus Y*) (Brandes et Wetter, 1959). Les caractéristiques de ce groupe étaient à l'origine basées sur la morphologie de la particule virale et sur les relations sérologiques avec le virus PVY. Les particularités majeures des potyvirus ont depuis, été définies plus précisément (Dougherty et Carrington, 1988; Riechmann *et al.*, 1992):

- La particule virale a la forme d'un bâtonnet flexible de 680 nm à 900 nm de long et de 12 nm à 15 nm de diamètre.
- Le génome se compose d'un unique ARN de polarité positive d'environ 10 000 nucléotides (nt), polyadénylé en 3', et qui possède à son extrémité 5' une protéine liée de façon covalente, la VPg (*viral protein genome*).
- L'infection par un potyvirus induit dans le cytoplasme la formation d'inclusions cylindriques de forme enroulée caractéristique.

Une étape importante a été franchie en 1986 pour l'étude des potyvirus lorsque deux de ses membres ont été séquencés au complet; il s'agit du *tobacco etch virus* (TEV) (Allison *et al.*, 1986) et du *tobacco vein mottling virus* (TVMV) (Domier *et al.*, 1986). La faculté de synthétiser des transcrits infectieux de potyvirus comme le TVMV (Domier *et al.*, 1989) a par la suite ouvert la possibilité d'étudier leur biologie au niveau moléculaire. La maîtrise du génie génétique permet maintenant l'utilisation de nouvelles approches expérimentales pour l'étude des différentes protéines virales *in vitro* ainsi que *in vivo*.

Le nombre de virus rattachés au groupe des potyvirus n'a cessé de croître depuis sa création; de 16 en 1959, il est passé à 120 en 1990, ce qui représente 27.5 % de tous les virus de plantes à ARN positif, et qui place les potyvirus au premier rang des virus pathogènes des plantes (Shukla et Ward, 1989*b*; Matthews, 1991).

Du fait de la grande diversité au sein du groupe des potyvirus, l'ICTV a proposé en 1989 la création de la famille des Potyviridae (Barnett, 1991) qui comprendrait quatre genres: le genre Potyvirus (virus type = PVY) regrouperait tous les virus de la famille dont la transmission se fait par les pucerons; le genre Baymovirus (virus type = *barley yellow mosaic virus*) regrouperait ceux transmis par les champignons; le genre Ryemovirus (virus type = *ryegrass mosaic virus*), ceux transmis par les acariens; et le genre Ipomovirus (virus type = *sweet potato mild mottle virus*) ceux transmis par la mouche blanche. Cette nouvelle classification doit être adoptée par le comité exécutif de l'ICTV pour être reconnue au niveau international. La taxonomie des potyvirus est complexe du fait du nombre croissant de ses membres et de leurs nombreuses souches. Ainsi, les

critères basés sur la taille du virion, la morphologie des corps d'inclusion, les réactions sérologiques, la symptomatologie et les vecteurs de transmission ne sont pas toujours suffisants pour distinguer deux potyvirus entre eux et il faut donc trouver des critères taxonomiques plus satisfaisants. Il est probable que dans cette optique, la séquence nucléotidique du génome des potyvirus soit de plus en plus utilisée, car elle seule permet de différencier deux virus très proches (Shukla et Ward, 1989a).

Ohki (1992) a émis l'hypothèse selon laquelle il ne faut pas considérer une infection par un potyvirus unique mais plutôt par une population de nombreux isolats dont un serait représenté de façon majoritaire; si les conditions du milieu changent, un des membres minoritaires de la population peut être sélectionné et conduire à l'apparition d'une nouvelle souche ou d'un nouveau virus, dépendant des nouvelles caractéristiques de celui-ci.

4.2 STRUCTURE DES POTYVIRUS.

La particule virale est composée de l'assemblage d'environ 2 000 unités protéiques de la capsid (CP), de l'ARN viral et de la protéine VPg. Le virion a la forme d'un bâtonnet flexible sans cavité centrale et ses dimensions, calculées à partir de microphotographies, varient de 680 nm à 900 nm pour la longueur et de 12 nm à 15 nm pour le diamètre, avec des variations pour un même virus dues sans doute aux conditions de préparation; le pas de l'hélice est d'environ 3.3 nm avec sept à huit protéines CP par tour d'hélice (Ward et Shukla, 1991). Des protéines CP produites chez *E. coli* ou *in vitro* sont capables de former sans la présence d'ARN viral, des « pseudo-particules » très semblables à celles natives, mais le mécanisme de l'encapsidation de l'ARN par les

protéines reste malgré tout encore inconnu (McDonald et Bancroft, 1977; Jagadish *et al.*, 1991). La flexibilité de la particule virale s'expliquerait par des interactions moléculaires de faible énergie entre les protéines de la capside des différents niveaux de l'hélice formant la particule virale, si on les compare à celles qui maintiennent la rigidité de la particule virale des tobamovirus (Shukla et Ward, 1989b).

4.3 GÉNOME DES POTYVIRUS.

Le génome des potyvirus, qui représente environ 5 % du poids moléculaire total du virion, se compose d'un ARN monocaténaire positif et aucun ARN sous-génomique n'a jamais été détecté (Figure 1). A ce jour, le plus long potyvirus séquencé est le *pea seed-borne mosaic virus* (PSbMV) avec 9 924 nt alors que le TVMV est le plus petit avec 9 471 nt (Domier *et al.*, 1986; Johansen *et al.*, 1991).

L'unique ARN des potyvirus est polyadénylé en 3' (Hari *et al.*, 1979) et la longueur de cette queue de résidus adénine (A) est très hétérogène; celle du *plum pox virus* (PPV) est constituée par exemple de 15 à 500 résidus A (Lafn *et al.*, 1988). La protéine VPg d'environ 23 kDa est liée de façon covalente à l'extrémité 5' de l'ARN viral (Hari, 1981; Siaw *et al.*, 1985). Pour le TVMV, de même que dans le cas du poliovirus (Rothberg *et al.*, 1978), une tyrosine de la VPg est liée par une liaison phosphodiester au premier nucléotide de l'ARN (Murphy *et al.*, 1991). Cette protéine n'est pas essentielle au caractère infectieux de l'ARN, mais son élimination par digestion aux protéases diminue beaucoup l'efficacité d'infection par rapport à l'ARN avec VPg (Hari, 1981; Riechmann *et al.*, 1989).

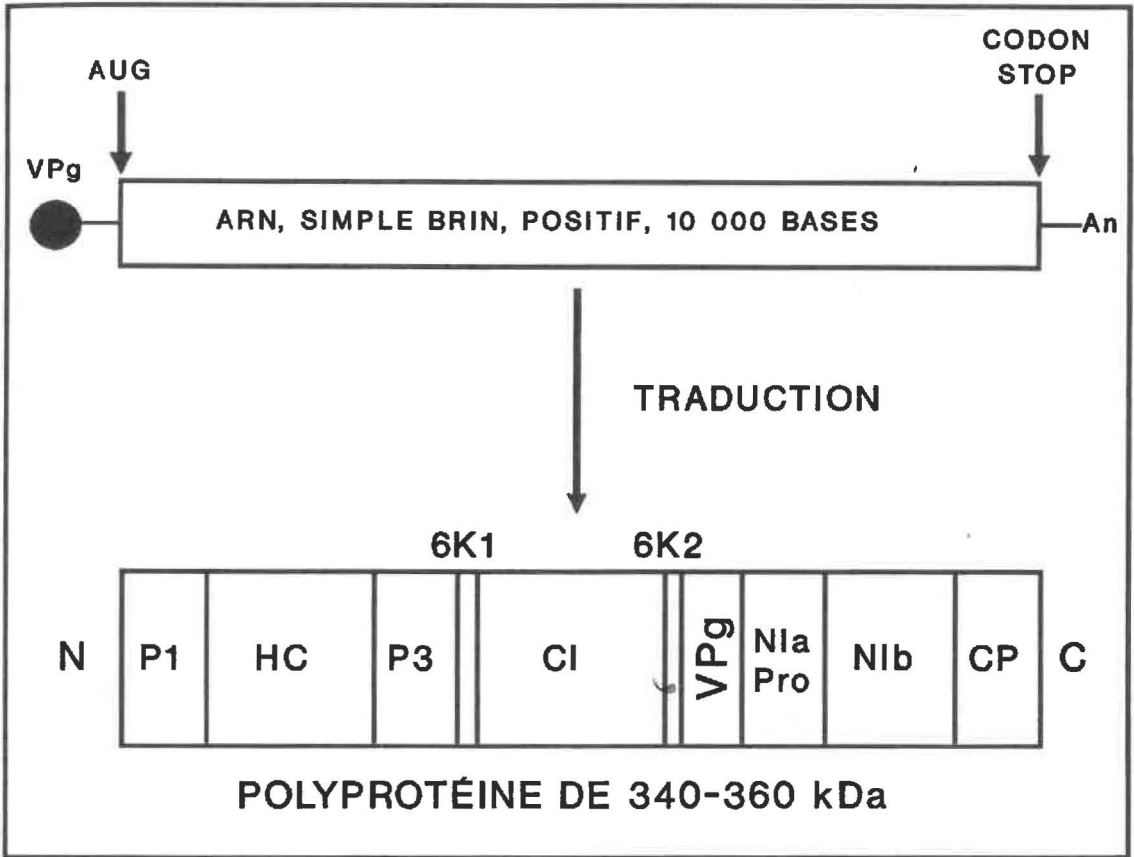


FIGURE 1: Organisation du génome et de la polyprotéine des potyvirus. La position de la VPg (cercle noir), du codon d'initiation (AUG) et de la queue polyadénylée (An) est indiquée sur l'ARN viral (rectangle supérieur). L'extrémité N-terminale (N) et C-terminale (C) de la polyprotéine (rectangle inférieur), ainsi que toutes les protéines connues des potyvirus, sont indiquées. Protéine P1 (P1); protéine auxiliaire (HC); protéine P3 (P3); protéines 1 et 2 de 6 kDa (6K1, 6K2); protéine formant les corps d'inclusions cylindriques (CI); protéines formant les corps d'inclusions nucléaires a et b (NIa, NIB); protéine de la capsid (CP).

On peut considérer trois régions sur l'ARN des potyvirus: une séquence non-codante en 5' (5'-NC), la séquence qui code pour la polyprotéine, et une séquence non-codante en 3' (3'-NC) suivie de la queue polyadénylée. La séquence 5'-NC des potyvirus connus a une longueur variant de 131 nt (virus de la mosaïque du soja, SMV) à 205 nt (TVMV); elle est composée à plus de 70 % de résidus A et uracile (U) (Jayaram *et al.*, 1992; Domier *et al.*, 1986). La région codante commence par un codon d'initiation AUG et a une longueur variant de 9 015 nt (TVMV) à 9 618 nt (PSbMV). Le premier codon d'arrêt rencontré sur le brin codant est suivi d'une séquence 3'-NC d'une longueur variant de 147 nt (*wheat streak mosaic virus*) à 475 nt (*johnson grass mosaic virus*) suivie de la queue polyadénylée (Ward et Shukla, 1991). Cette région 3'-NC qui est assez hétérogène entre différents virus, présente néanmoins quelques caractéristiques communes à tous les potyvirus: des segments riches en résidus A et U, quelques blocs de nucléotides conservés, et des structures secondaires identiques (Bryan *et al.*, 1992). Quand on compare deux souches d'un même potyvirus, la séquence 3'-NC est plus conservée que celle codant pour la capsid, ce qui est une indication de son importance probable dans les mécanismes de reconnaissance par la ou les enzymes impliquées dans la réplication (Wetzel *et al.*, 1991a). Il a été démontré que des mutations dans la région 3'-NC du TVMV ont un effet direct sur sa pathologie, sans que cela soit le résultat d'une baisse de production du virus; ceci serait attribuable aux structures secondaires de l'ARN dans cette région mais les mécanismes exacts conduisant à la formation des symptômes restent encore inconnus (Rodríguez-Cerezo *et al.*, 1991). Il a également été proposé que la séquence de la région 3'-NC soit utilisée comme marqueur dans la taxonomie des potyvirus (Ward et Shukla, 1991). Le génome des potyvirus ressemble donc beaucoup dans son organisation à celui

des picornavirus, virus qui infectent les animaux, mais aussi à deux autres groupes de virus de plantes à ARN bi-segmenté, les népovirus et les comovirus.

4.4 STRATÉGIE D'EXPRESSION DES POTYVIRUS.

La stratégie d'expression des potyvirus passe par la traduction de leur ARN en une polyprotéine de 340 à 360 kDa, qui est clivée par certaines protéases en polypeptides plus petits nécessaires à la multiplication et à la transmission du virus (Figure 1) (Riechmann *et al.*, 1992; Dougherty et Carrington, 1988). Cette stratégie d'expression est identique à celle des picornavirus, des comovirus et des népovirus. On retrouve chez ces trois groupes de virus ainsi que chez les potyvirus, une similitude dans l'organisation des protéines ayant une fonction équivalente sur leur polyprotéine respective. Il a ainsi été proposé que les trois groupes de virus de plantes (comovirus, népovirus, potyvirus) aient une origine commune avec les picornavirus, et ont de ce fait été classés dans le super-groupe des virus « type-picornavirus » (Goldbach, 1986).

L'ARN des potyvirus fonctionne comme un ARN messager eucaryote excepté qu'il est bloqué en 5' par la protéine VPg plutôt que d'être coiffé. L'initiation de la traduction chez les potyvirus, contrairement aux picornavirus, ne se ferait pas par entrée interne des ribosomes sur l'ARN mais plutôt par balayage (« *scanning* ») de l'ARN jusqu'au premier codon d'initiation placé dans le contexte le plus favorable, tel que défini par Lütcke *et al.* (1987) pour les ARN messagers des plantes (Riechmann *et al.*, 1991). Pourtant, la séquence 5'-NC du TEV fusionnée à un gène indicateur, stimule très fortement sa traduction *in vitro*, *in vivo* et indépendamment de la présence d'une coiffe (Carrington et

Freed, 1990). Ce résultat suggère que cette région peut être reconnue par des facteurs impliqués dans l'initiation de la traduction, ce qui pourrait ainsi stimuler la synthèse des protéines virales par rapport à celles de l'hôte. De plus, il est possible que la migration des ribosomes sur l'ARN viral jusqu'au codon d'initiation soit facilitée, du fait de la forte teneur en résidus A et U de la région 5'-NC qui peut déstabiliser les structures secondaires potentielles (Turpen, 1989). L'ARN des potyvirus est traduit en une polyprotéine et il n'y a pas d'ARN sous-génomiques; de ce fait, la régulation de l'expression des protéines ne peut se faire qu'au niveau des clivages de la polyprotéine, et il est ainsi extrêmement important de connaître les mécanismes de ces clivages protéolytiques pour comprendre la stratégie d'expression des potyvirus.

4.5 CLIVAGE DE LA POLYPROTÉINE DES POTYVIRUS.

4.5.1 Introduction.

Les premières évidences démontrant que l'ARN des potyvirus est traduit en une polyprotéine clivée par des protéases, ainsi que la localisation des protéines virales, proviennent des résultats de la traduction *in vitro* de l'ARN du *pepper mottle virus* (PepMoV) et du TVMV suivie d'une immunoprécipitation des produits de traduction (Dougherty et Hiebert, 1980*a,b,c* ; Hellmann *et al.*, 1980; Hellmann *et al.*, 1983). La Figure 1 indique une représentation schématique du génome et de la polyprotéine des potyvirus ainsi que la localisation et le nom des protéines connues à ce jour. Trois protéines virales ayant une activité protéolytique sur la polyprotéine ont été trouvées; il s'agit, par ordre chronologique de découverte, de la protéine NIa, de la protéine auxiliaire

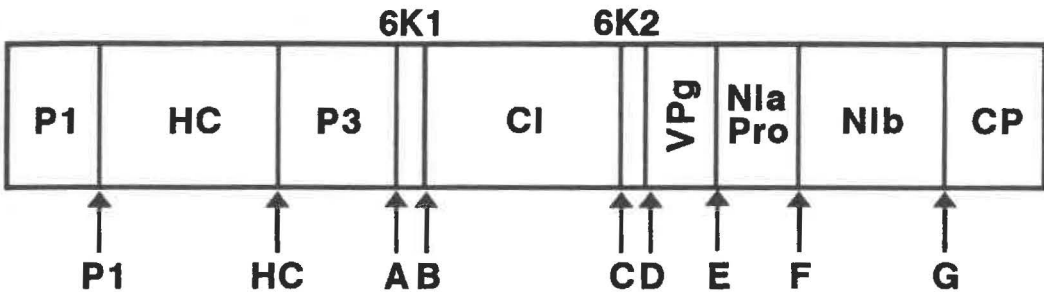
(HC) et de la protéine P1.

4.5.2 Protéase NIa.

Le virus TEV est un des rares potyvirus à induire la formation d'inclusions nucléaires « non-amorphes »; elles sont formées d'une quantité équimolaire de deux protéines, de 49 kDa et 54 kDa, qui ont été nommées respectivement NIa et NIb (Knuhtsen *et al.*, 1974; Hiebert *et al.*, 1984). La protéine NIa du TEV est une protéase responsable de plusieurs clivages sur la polyprotéine, et son activité protéolytique réside dans sa partie C-terminale; la partie N-terminale qui n'est pas indispensable à cette activité, constitue en fait la protéine VPg (Carrington et Dougherty, 1987a; Dougherty et Parks, 1991). Il a été démontré dans un système *in vitro* que la NIa du TEV reconnaît une séquence de sept acides aminés sur la polyprotéine, dont quatre sont absolument requis pour que se produise un clivage efficace; celui-ci se produit après le sixième acide aminé de l'heptapeptide qui est la glutamine (Gln) (Carrington et Dougherty, 1988; Dougherty *et al.*, 1988; Dougherty *et al.*, 1989a). Ce résultat fut confirmé pour la protéine NIa du PPV par expression chez *E. coli*, tout en montrant sa spécificité pour ses propres sites de clivage; en effet, la protéase NIa du PPV ne reconnaît pas un des sites de clivage du TEV (García *et al.*, 1989a,b).

La Figure 2 indique les différents sites de clivage reconnus par les protéases NIa respectives du TEV et du PPV.

FIGURE 2: Sites de clivages protéolytiques reconnus par les protéases NIa du *plum pox virus* (PPV) et du *tobacco etch virus* (TEV). La position des sites de clivages par les protéases P1 et HC sont indiqués respectivement par P1 et HC. Les sept sites de clivage reconnus par la NIa sont indiqués par une lettre (A à G). Les acides aminés (code de une lettre) constituant la séquence reconnue par la NIa sont numérotés par rapport au site de clivage (-6 à +1). Les acides aminés présents au moins quatre fois à chaque position sont encadrés. La liaison peptidique clivée par la NIa est indiquée par une flèche verticale.



Sites de clivage NIa

PPV

	-6	-5	-4	-3	-2	-1	+1
A	Q	V	V	V	H	Q	S
B	Q	A	V	Q	H	Q	S
C	E	C	V	H	H	Q	T
D	E	E	V	V	H	Q	G
E	E	E	V	D	H	E	S
F	E	F	V	Y	T	Q	S
G	N	V	V	V	H	Q	A

TEV

	-6	-5	-4	-3	-2	-1	+1
A	E	D	L	V	E	Q	A
B	E	I	I	Y	T	Q	S
C	E	T	I	Y	L	Q	S
D	E	P	V	Y	F	Q	G
E	E	D	L	T	F	E	G
F	E	L	V	Y	S	Q	G
G	E	N	L	Y	F	Q	S

Certains sites reconnus par la NIa sont clivés en *cis*, d'autres en *trans*, et l'efficacité de clivage varie d'un site à un autre (Carrington *et al.*, 1988; Dougherty et Parks, 1989; García *et al.*, 1989a,b,c). La protéase NIa du TEV est excisée de la polyprotéine par un mécanisme d'autoprotéolyse (protéolyse en *cis*) alors que la protéine CP l'est en *trans* (Carrington et Dougherty, 1987b; Dougherty *et al.*, 1989a). Le séquençage des extrémités N-terminales des protéines du PPV purifiées à partir de plantes infectées a permis de localiser six sites NIa sur sa polyprotéine (Figure 2, sites A, B, C, D, F et G) (García *et al.*, 1992; Martín *et al.*, 1990). Chez le TEV, un septième site NIa, situé à l'intérieur même de cette protéine a été démontré (Figure 2, site E) (Dougherty et Parks, 1991); ce clivage se produit cette fois-ci après l'acide glutamique (Glu), et il sépare la VPg de 21 kDa qui constitue la partie N-terminale de NIa de la partie C-terminale de 27 kDa (appelée NIa-Pro) qui garde la même activité protéolytique que la NIa complète.

La protéase NIa-Pro des potyvirus est homologue aux protéases à sérine de la famille de la trypsine et de la chymotrypsine, excepté que la sérine du site catalytique est remplacée par une cystéine (Bazan et Fletterick, 1988; Dougherty *et al.*, 1989a). Le site catalytique se compose en effet de trois acides aminés (histidine, acide aspartique (Asp) et cystéine) qui sont conservés dans toutes les protéines NIa des potyvirus séquencés (Dougherty *et al.*, 1989b). La protéase 3C des picornavirus est également une protéase du même type que la NIa des potyvirus, et les trois acides aminés du site actif sont identiques (Bazan et Fletterick, 1988; Palmenberg, 1990).

4.5.3 Protéase HC.

Il a été démontré dans un système *in vitro* et chez *E. coli* que la protéine HC de 56 kDa du TEV ne clive qu'à sa seule extrémité C-terminale, et que la réaction se fait uniquement en *cis*, c'est-à-dire de manière autocatalytique (Carrington *et al.*, 1989*a,b*). Le site de clivage reconnu par cette nouvelle protéase, appelée aussi HC-Pro, fut localisé entre deux résidus glycine (Gly) grâce au séquençage de l'extrémité N-terminale de la protéine située en C-terminal de HC. Le site actif de la protéase HC se situe en outre dans les 155 derniers acides aminés des 459 qui la composent, ce qui suggère que cette protéine a d'autres fonctions en N-terminal (Carrington *et al.*, 1989*a*). Par mutagénèse dirigée et traduction *in vitro*, les acides aminés essentiels à l'activité protéolytique de la HC du TEV furent démontrés comme étant la cystéine (position 345 de HC) et l'histidine (position 418) (Oh et Carrington, 1989). Ce résultat, appuyé du fait que ces deux résidus sont strictement conservés chez les autres potyvirus séquencés, suggère que HC est une protéase du type cystéine comme la papaine (Oh et Carrington, 1989). Les quatre acides aminés essentiels qui constituent le site de clivage autocatalytique de HC du TEV furent également déterminés, et tout changement de l'un d'eux affecte beaucoup l'efficacité de clivage. Il s'agit de la tyrosine en position -4, de la valine en position -2, de la glycine en position -1 et de la glycine en position +1 par rapport au site de clivage; ces résidus sont strictement conservés chez les autres potyvirus. (Carrington et Herndon, 1992). Un virus de champignon associé à l'hypovirulence du chancre du châtaignier (HAV), qui est donc très différent des potyvirus, puisqu'il a un génome à ARN double brin, code pourtant pour une protéase (p29) qui ressemble beaucoup à la protéase HC des potyvirus. En effet, la protéine p29 possède les mêmes acides aminés, cystéine et histidine, essentiels à son

activité autocatalytique, et le clivage se fait également entre deux résidus glycine (Choi *et al.*, 1991). Il ne faut donc pas exclure la possibilité que le HAV et les potyvirus aient une origine commune ou qu'il y ait eu recombinaison de leur génome à un moment donné (Koonin *et al.*, 1991). HC est donc une protéine qui a au moins deux fonctions: une activité autocatalytique qui permet sa libération rapide de la polyprotéine en cours d'élongation, et un rôle qui sera décrit plus loin, dans le transport du virus par les vecteurs.

4.5.4 Protéase P1.

Berger *et al.* (1989) montrèrent, par l'expression de la partie N-terminale de la polyprotéine du TVMV dans des plantes transgéniques suivies de la détection par « *Western blot* » de la protéine HC, que celle-ci est libérée de ses deux protéines adjacentes, supposant donc l'intervention d'au moins une protéase à chaque extrémité.

Carrington *et al.* (1990) démontrèrent, également par expression dans des plantes transgéniques de la partie N-terminale de la polyprotéine du TEV, qu'une troisième protéase, autre que les protéines NIa et HC, est responsable du clivage de la protéine de 35 kDa en N-terminal (protéine appelée NT par les auteurs, mais nommée également P1). La mutation de l'un des acides aminés essentiels de la protéine HC n'affecte pas le clivage à son extrémité N-terminale, alors qu'une délétion de 152 acides aminés dans la protéine P1 affecte beaucoup son efficacité. Verchot *et al.* (1991) ont par la suite démontré, par traduction *in vitro*, que la protéine P1 du TEV a une activité protéolytique sur la jonction P1-HC, c'est-à-dire à son extrémité C-terminale, et que la région responsable de cette activité est située dans sa partie C-terminale. Le séquençage de l'extrémité N-terminale

de la protéine HC du TVMV provenant du clivage d'une polyprotéine produite *in vitro*, a permis de déterminer le premier acide aminé de cette protéine, et donc, de localiser le site clivé par P1: entre la phénylalanine (position 256 de la polyprotéine) et la sérine (position 257) (Mavankal et Rhoads, 1991). Le changement de certains acides aminés de la P1 inhibant son activité, ont permis de déduire qu'il s'agissait d'une protéase du type sérine, qui a besoin pour son activité de facteurs non-viraux présents entre autre, dans le système de traduction *in vitro* de germe de blé, mais pas dans le lysat de réticulocyte de lapin (Verchot *et al.*, 1991). Les acides aminés qui constituent le site actif de la protéase P1, ainsi que ceux qui composent son site de clivage, n'ont pas encore été déterminés de façon méthodique.

4.5.5 Régulation de la protéolyse.

Si en 1992 la plupart des sites de clivages de la polyprotéine des potyvirus ont été démontrés *in vitro* ou *in vivo*, il y a encore peu d'information concernant la régulation de la protéolyse virale. Celle-ci est très importante puisqu'il faut, pour synthétiser une seule particule virale, une molécule d'ARN, une protéine VPg, et environ 2 000 copies de la protéine CP, ce qui signifie que l'ARN doit être traduit au moins 2 000 fois pour la synthèse d'une seule particule virale. On comprend donc pourquoi les potyvirus doivent avoir des mécanismes permettant, premièrement, une traduction efficace et complète de leur ARN, puisque la protéine CP se trouve à l'extrémité C-terminale de la polyprotéine, et deuxièmement, un système de régulation permettant une libération complète de la protéine CP de la polyprotéine, ainsi que le stockage ou la récupération des protéines non-structurales.

La Figure 3 indique les événements protéolytiques connus chez les potyvirus à partir de la polyprotéine, mais il est difficile de prédire exactement dans quel ordre se produisent ces événements, du fait que les précurseurs protéiques peuvent être difficilement isolés. On sait par exemple que le clivage autoprotéolytique par la protéase HC se fait dès le passage des ribosomes au site de clivage, alors que le clivage autoprotéolytique interne à la NIa n'est pas immédiat. En fait, plusieurs virologistes ont émis l'hypothèse selon laquelle, traduction, protéolyse, et réplication chez les virus à ARN, seraient des événements intimement reliés et concomitants (notes prises lors du « *Third Symposium on Plus-Strand RNA Viruses* », Clearwater, Floride, USA, 1992).

4.6 PROTÉINES DES POTYVIRUS.

4.6.1 Introduction.

Il a été montré dans la section précédente que la polyprotéine des potyvirus comportait au moins neuf sites de clivage protéolytiques reconnus par trois protéases virales, qui après clivage, libéraient dix protéines potentielles. Il faut également considérer que les précurseurs protéiques puissent jouer un rôle différent de celui des protéines qui en dérivent après protéolyse, ce qui complique l'étude du rôle spécifique des protéines potyvirales. Certaines protéines des potyvirus ont été détectées *in vivo*, d'autres *in vitro*, alors que certaines ne l'ont pas encore été. La Figure 4 montre la carte génétique des trois potyvirus qui ont été les mieux étudiés (PPV, TEV et TVMV), ainsi que la position des sites de clivage (démontrés ou potentiels) sur leur polyprotéine respective. Les protéines connues des potyvirus sont décrites ci-dessous.

FIGURE 3: Représentation des événements protéolytiques chez les potyvirus. La polyprotéine complète est représentée avec la position des différentes protéines. Les flèches indiquent les sites clivés par les protéases dont le nom est désigné dessous (NIa, HC et P1). Le type de clivage (en *cis* ou en *trans*) est également indiqué au dessus des flèches.

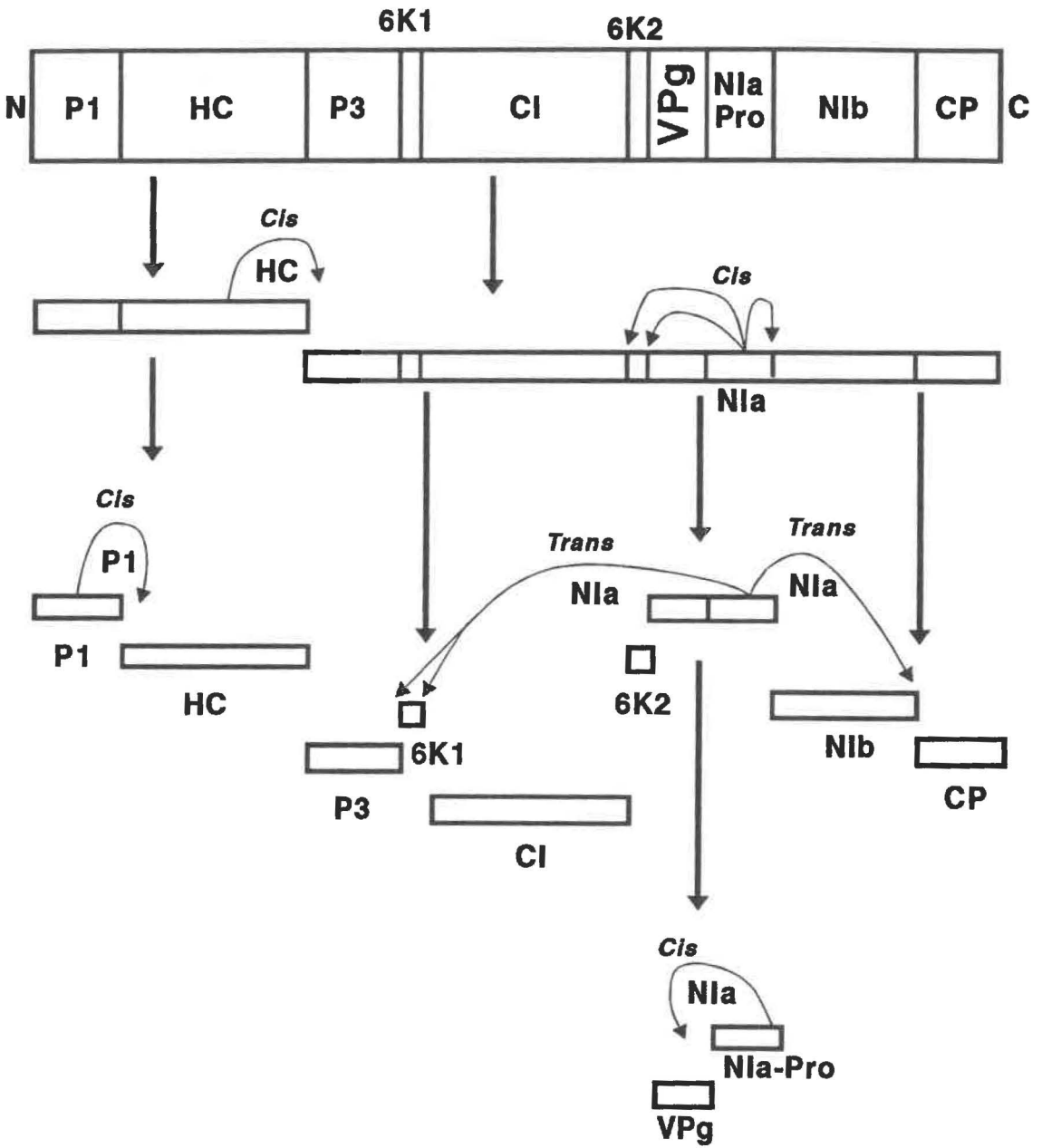
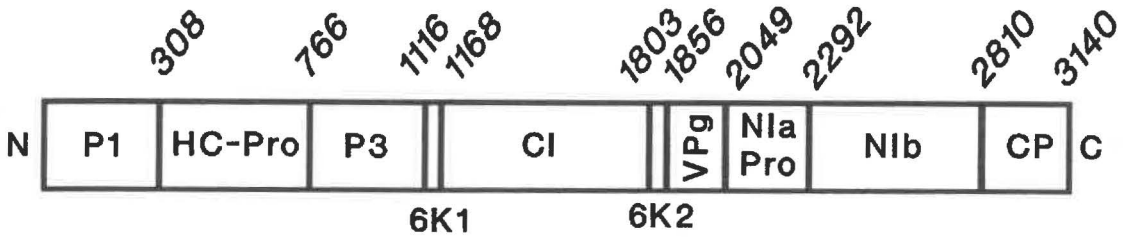
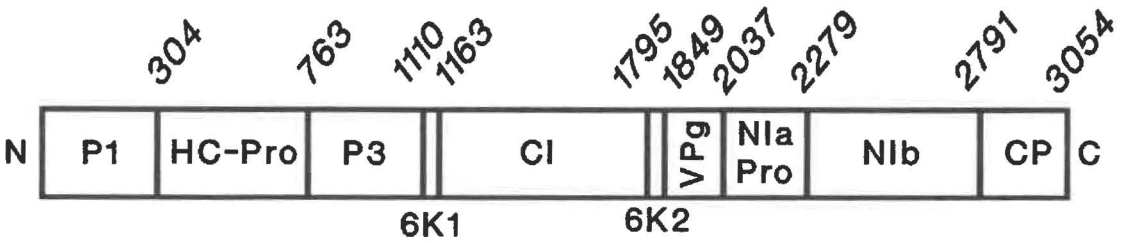


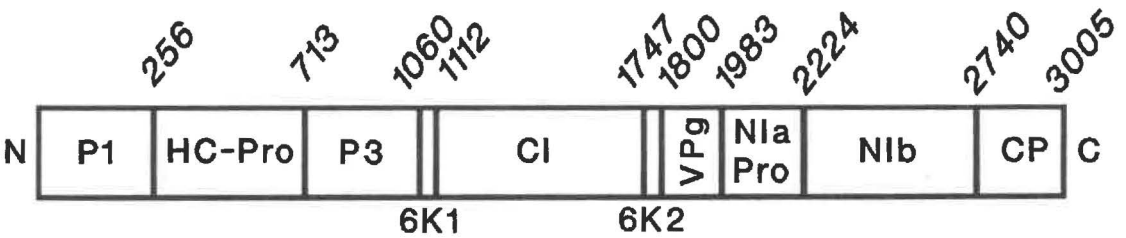
FIGURE 4: Carte génétique de trois potyvirus: le *plum pox virus* (PPV); le *tobacco etch virus* (TEV) et le *tobacco vein mottling virus* (TVMV). La polyprotéine de chaque virus est représentée avec les différentes protéines connues. Les nombres indiqués au dessus des sites de clivage représentent la position du dernier acide aminé composant la protéine dont le nom est indiqué vers la gauche de la barre verticale.



PLUM POX VIRUS (PPV)



TOBACCO ETCH VIRUS (TEV)



TOBACCO VEIN MOTTLING VIRUS (TVMV)

4.6.2 Protéine de la capside (CP).

Du fait que la protéine CP des potyvirus représente plus de 95 % du poids du virion et que sa région codante se situe en 3' de l'ARN viral, elle fut la première protéine des potyvirus à avoir été clonée et séquencée. En 1985, trois d'entre elles l'avaient été, et en 1992, ce nombre est passé à plus de 40, incluant les différentes souches de plusieurs virus (Bryan *et al.*, 1992; Ward et Shukla, 1991).

La protéine CP est assez hétérogène en taille puisqu'elle varie de 253 à 330 acides aminés, mais sa composition est suffisamment homogène pour servir de base au classement d'un virus parmi les potyvirus (Burger *et al.*, 1990; Fauquet *et al.*, 1986*a,b*; Shukla et Ward, 1989*a*). On trouve dans sa séquence de faibles homologies avec les protéines structurales d'autres groupes de virus de plantes à particule virale flexible (potexvirus, carlavirus et clostéovirus), ce qui suggère l'existence d'un ancêtre commun à ces différents groupes (Dinant *et al.*, 1991; Dolja *et al.*, 1991).

L'homologie en acides aminés au niveau de la protéine CP varie de 38 à 71 % entre deux potyvirus distincts, et de 90 à 99 % entre différentes souches du même virus (Shukla et Ward, 1988). Cette nette différence d'homologie entre virus distincts et souches virales favorise l'hypothèse d'une distribution bimodale des potyvirus (virus et souches) plutôt que celle d'un continuum qui regrouperait tous les potyvirus; cette valeur peut donc être utilisée pour comparer deux potyvirus afin de préciser s'il s'agit de deux virus distincts ou de deux souches d'un même virus, et servir ainsi à la taxonomie de ce groupe (Shukla et Ward, 1989*b*; Ward et Shukla, 1991).

Il y a des différences majeures de longueur et de séquence au niveau de la région N-terminale de CP, alors que les deux-tiers C-terminaux sont beaucoup plus conservés (Shukla et Ward, 1989b). En fait, la région N-terminale détient les épitopes immunodominants du virion puisque des anticorps dirigés spécifiquement contre elle peuvent être utilisés pour différencier deux virus très proches, alors que des anticorps dirigés contre la protéine CP complète vont reconnaître de nombreux potyvirus. Les techniques sérologiques basées sur l'utilisation d'anticorps anti-CP ont été jusqu'à présent les plus exploitées pour la détection et la différenciation des potyvirus, ce qui a conduit à quelques erreurs de classification. De ce fait, la séquence protéique de CP est maintenant souvent utilisée d'une part, pour différencier les potyvirus entre eux, et d'autre part, pour établir de façon précise la généalogie de ce groupe (Shukla et Ward, 1989a; Ward et Shukla, 1991).

Les régions C-, et principalement la N-terminale de la protéine CP sont fortement hydrophiles et exposées à la surface de la particule virale. Elle sont dégradées si le virus est mal conservé, ce qui explique parfois la présence de bandes plus petites que celle correspondant à la protéine CP intacte, sur gel dénaturant de protéines chargé par le virus purifié (Shukla et Ward, 1989b). Les extrémités terminales de CP peuvent être également éliminées par digestion partielle à la trypsine sans que cela ne modifie l'efficacité d'infection du virus traité ni son apparence en microscopie électronique; ceci signifie que la partie centrale de la protéine CP est probablement celle impliquée dans les interactions avec l'ARN, alors que les parties terminales situées vers l'extérieur de la particule virale, auraient d'autres rôles (Allison *et al.*, 1985; Shukla *et al.*, 1988). En effet, il a été montré que la protéine CP du TVMV a un rôle évident dans la transmission de ce virus par les

pucerons; une mutation à un seul acide aminé situé dans l'extrémité N-terminale de CP différencie la souche transmissible (AT) de celle non-transmissible (NAT) par ce vecteur (Atreya *et al.*, 1990). Ceci fut confirmé par la mutagenèse dirigée au niveau de l'acide aminé (une glycine) impliqué dans le phénotype AT. La comparaison de plusieurs souches AT et NAT de différents potyvirus montre également des différences de un ou deux acides aminés dans la même région N-terminale de CP (Atreya *et al.*, 1991).

L'hétéroencapsidation de deux potyvirus, un de phénotype AT et l'autre NAT, a confirmé de façon évidente que la protéine CP est impliquée dans la transmission par les pucerons (Bourdin et Lecoq, 1991). Il est probable, comme cela a été montré pour les tobamovirus, que la protéine CP des potyvirus soit également impliquée dans le déterminisme et la réponse de l'hôte, c'est-à-dire de la formation des symptômes (Saito *et al.*, 1987). Ceci pourrait expliquer par exemple, comment de petites différences de séquence au niveau de la capsid de deux souches d'un même virus peuvent affecter leur spectre d'hôte ainsi que les symptômes provoqués sur le même hôte.

L'infection d'une espèce végétale par une souche atténuée d'un potyvirus confère une protection vis-à-vis des souches virulentes de ce virus, phénomène qu'on appelle la prémunition ou la résistance induite, et qu'on retrouve pour d'autres virus de plantes. Les mécanismes de cette protection ne sont pas encore bien compris et sont probablement variables d'un groupe de virus à l'autre (Ponz et Bruening, 1986). Pourtant, il a été montré pour plusieurs groupes de virus dont les potyvirus, que des plantes transgéniques exprimant la protéine CP virale bénéficient d'une certaine résistance au virus codant pour cette protéine, et dans certains cas, à d'autres virus du même groupe (Beachy *et al.*, 1990;

Lawson *et al.*, 1990; Stark et Beachy, 1989). En fait, contrairement au mécanisme de résistance au TMV qui nécessite l'expression de la protéine CP du TMV (Powell *et al.*, 1990), la résistance aux potyvirus nécessite l'expression de l'ARN codant pour CP mais pas de la protéine elle-même; ceci suggère que chez les potyvirus, la résistance induite est le résultat d'interactions « ARN-ARN » qui interfèrent avec la réplication virale (Lindbo et Dougherty, 1992*a,b*; Van der vlugt *et al.*, 1992).

4.6.3 Protéine NIb.

La protéine NIb forme avec la NIa, des corps d'inclusion nucléaires stables de forme et de taille variables chez certains potyvirus comme le TEV (Hiebert *et al.*, 1984; Knuhtsen *et al.*, 1974). La NIb est la protéine la plus conservée en séquence et en longueur parmi les protéines des potyvirus connus avec environ 60 % d'homologie et une taille de 512 à 521 acides aminés (Figure 4). On retrouve dans la séquence de NIb quatre motifs trouvés chez les polymérases ARN-dépendante d'autres groupes de virus dont le motif (Gly-Asp-Asp) (Domier *et al.*, 1987; Poch *et al.*, 1989). Ces homologies ainsi que la localisation de la NIb sur la polyprotéine comparée à celle de la réplicase des picornavirus suggèrent fortement que NIb est la réplicase (ARN polymérase ARN-dépendante) des potyvirus (Habibi et Symons, 1989). Martín et García (1991) ont été les premiers à montrer une activité « ARN-polymérase » dans une plante infectée par un potyvirus; ils ont obtenu la synthèse d'ARN viral marqué par un isotope radioactif dans une fraction membranaire d'un extrait de plante infectée par le PPV. Cette fraction contenait la protéine NIb, la NIa, la CI et le précurseur (NIa-NIb) mais ils n'ont pas encore démontré l'activité enzymatique de la protéine purifiée NIb (communication

personnelle du Dr. M.T. Martín, Université Autonome de Madrid, Espagne). Il se peut que la réplication des potyvirus fasse intervenir aussi bien des protéines virales que celles de l'hôte, auquel cas la démonstration *in vitro* de l'activité polymérase de NIb ne pourra pas utiliser les mêmes procédures que pour la réplicase 3D du poliovirus, qui est une des ARN polymérases virales les mieux caractérisées (Dasgupta *et al.*, 1979). La fonction et le mécanisme du transport des protéines NIa et NIb vers le noyau sont encore inconnus, et le signal de transport de la NIb, s'il en est un, n'a pas encore été localisé (Restrepo *et al.*, 1990).

4.6.4 Protéines NIa-Pro et VPg.

La protéine NIa est constituée de deux domaines, la protéine VPg en N-Terminal et la NIa-Pro en C-terminal, qui peuvent être séparés par clivage autoprotéolytique comme cela a été décrit précédemment (chapitre 4.5.2) (Dougherty et Parks, 1991). La NIa-Pro possède de 241 à 246 acides aminés (Figure 4) et présente environ 50 % d'homologie parmi les virus connus. L'activité protéolytique de NIa-Pro a été démontrée pour plusieurs virus et elle est identique à celle de la protéine NIa non-clivée (Dougherty et Parks, 1991; García *et al.*, 1989a,b,c; Murphy *et al.*, 1990).

Une région située dans les 76 premiers acides aminés de la VPg du TEV contient un signal de transport de la NIa dans le noyau; ceci a été démontré par fusion de cette séquence à un gène marqueur, la β -glucuronidase (GUS), puis par visualisation histochimique de l'activité GUS après transfection de protoplastes ou transformation génétique du tabac par ces constructions (Carrington *et al.*, 1991; Restrepo *et al.*, 1990).

De même que pour la N1b, la fonction du transport de N1a vers le noyau reste inconnue, mais il est probable que de grandes quantités de ces deux protéines dans le cytoplasme puissent affecter la réplication virale, ce qui aurait conduit les potyvirus à élaborer un mécanisme efficace de transport nucléaire.

La protéine VPg des potyvirus possède de 183 à 194 acides aminés (Figure 4). Celle du TEV a un PM de 22K mais on retrouve des molécules d'ARN liées à la protéine N1a complète (Murphy *et al.*, 1990). Ceci suggère que le site de clivage interne de la N1a n'est pas optimal ni requis, puisque la protéine N1a de 49 kDa a la même activité que la N1a-Pro de 27 kDa; le site interne est probablement clivé vers la fin de l'infection.

Murphy *et al.* (1991) ont montré que la VPg du TVMV était liée par le résidu tyrosine (position 60 de la protéine) au premier nucléotide de l'ARN. Cet acide aminé est situé dans un motif très conservé chez tous les potyvirus séquencés, ce qui laisse présager que la VPg est liée à l'ARN par le même résidu pour tous les potyvirus. Il faut remarquer que la VPg du poliovirus qui a une taille beaucoup plus réduite (6 kDa), est également liée par un résidu tyrosine à l'ARN viral (Rothberg *et al.*, 1978).

La VPg n'est pas indispensable à l'infection puisque des transcrits synthétiques de certains potyvirus ainsi que le TEV traité par des protéases, sont capables par inoculation mécanique d'induire une infection virale du même type, mais de moindre efficacité, que celle obtenue par le virus complet (Hari, 1981; Maiss *et al.*, 1992). Par analogie avec la protéine VPg des picornavirus, pour laquelle il a été montré qu'elle sert d'amorce à la synthèse du brin négatif d'ARN (Tobin *et al.*, 1989), on peut émettre l'hypothèse selon laquelle la VPg des potyvirus serait impliquée dans la réplication virale. Il est possible que

la VPg des potyvirus, du fait de sa localisation à l'extrémité de l'ARN viral, soit impliquée dans plusieurs fonctions dont: la réplication de l'ARN, le transport par les vecteurs, la migration dans la plante et l'encapsidation de l'ARN.

4.6.5 Protéines 6K1 et 6K2.

Ces deux protéines d'environ 52 acides aminés chacune (Figure 4), et qui sont respectivement en position quatre et six sur la polyprotéine des potyvirus, n'ont jamais été détectées *in vitro*, ni *in vivo*. Pourtant certaines évidences suggèrent qu'elles sont libérées de la polyprotéine à un moment de l'infection; en effet, on trouve à chacune de leurs extrémités, des sites de clivage du type NIa qui ont tous été démontrés *in vitro* (Figure 2) (Carrington et Dougherty, 1987a,b; Carrington *et al.*, 1988; García *et al.*, 1992; Hellmann *et al.*, 1988; Rodríguez-Cerezo et Shaw, 1991). Bien que la 6K1 et la 6K2 aient la même taille, il n'y a pas d'homologie de séquence entre elles, mais on retrouve dans les deux cas une région centrale fortement hydrophobe. De plus, leur position respective à l'extrémité N-terminale de la protéine CI et de la VPg s'apparente à celle des protéines 2B et 3A des picornavirus qui précèdent respectivement la 2C (fonction hélicase) et la 3B (la VPg). Ces deux observations permettent d'envisager pour ces deux protéines de 6 kDa, un rôle probable dans la régulation de la réplication, par la fixation de leur région hydrophobe aux membranes cellulaires au niveau des complexes de réplication (Giachetti et Semler, 1991; Riechmann *et al.*, 1992).

Les sites CI-6K2 et 6K2-VPg sont clivés par la protéase NIa en *cis* alors que les sites P3-6K1 et 6K1-CI le sont en *trans* (Carrington et Dougherty, 1987b; Carrington *et al.*, 1988; García *et al.*, 1992). De plus, lorsque la 6K2 est présente sur le précurseur

peptidique de la NIa, le transport de celle-ci vers le noyau est inhibé (Restrepo-Hartwig et Carrington, 1992). On peut donc concevoir que l'une des fonctions de la 6K2 est de maintenir la NIa dans le cytoplasme à une concentration optimale qui peut être ajustée grâce au clivage protéolytique; dès que ces protéines sont séparées, la NIa (ou son précurseur NIa-NIb) est transportée vers le noyau grâce au signal de transport localisé dans sa région N-terminale (Restrepo *et al.*, 1990).

4.6.6 Protéine CI.

La formation de corps d'inclusion cytoplasmiques cylindriques, observables en microscopie électronique, est la propriété la plus caractéristique des potyvirus (Edwardson, 1992). Ces inclusions, contrairement à celles d'autres groupes de virus, ne sont pas formées de particules virales mais de l'agrégation d'une protéine unique (protéine CI) qui est la plus longue codée par les potyvirus avec 633 à 636 acides aminés (Figure 4) et un PM d'environ 70K. On peut classer les potyvirus en quatre types selon l'apparence des inclusions qu'ils induisent: type 1, rouleaux et « roues de chariot » ; type 2, « roues de chariot » et des lamelles allongées; type 3, « roues de chariot » avec rouleaux et lamelles allongées; type 4, rouleaux et lamelles courtes. Cette classification est assez limitée pour caractériser un virus puisque plusieurs souches d'un même virus peuvent appartenir à différents types de CI (Edwardson, 1992).

L'homologie de séquence de la CI des potyvirus connus est d'environ 57.5 % (Domier *et al.*, 1987) et on retrouve dans sa séquence des motifs du type « NTBM » (*nucleoside triphosphate-binding motif*) communs à de nombreuses protéines virales pour

lesquelles la propriété de fixer les nucléosides triphosphates a été démontrée (Gorbalenya *et al.*, 1988; Gorbalenya et Koonin, 1989; Habili et Symons, 1989). Les motifs « NTBM » de la CI sont également présents dans des protéines procaryotes et eucaryotes qui ont une activité hélicase, c'est-à-dire, responsable du déroulement et de la séparation des brins d'acides nucléiques bicaténaires (Gorbalenya *et al.*, 1988; Laín *et al.*, 1989b). La preuve de cette activité a été faite pour le PPV; il est ainsi le premier virus à ARN dont une des protéines a été démontrée comme ayant une activité hélicase. En effet, la protéine CI purifiée du PPV est capable *in vitro* de dérouler un ARN synthétique double brin, ce qui est la preuve de son activité hélicase (Laín *et al.*, 1990; Laín *et al.*, 1991). Cette activité enzymatique est dépendante de l'hydrolyse de l'ATP en ADP et se fait dans le sens 3' vers 5'.

D'après une étude d'immunolocalisation en microscopie électronique sur des plantes infectées au TEV, la CI serait associée aux membranes vers les plasmodesmes dans les premiers jours de l'infection puis sous forme d'inclusions cytoplasmiques plus tard au cours de l'infection (Baunoch *et al.*, 1991). Ce résultat suggère que la protéine CI est effectivement impliquée dans la réplication et qu'elle doit être associée à des complexes de réplication liés au système membranaire cellulaire, comme c'est le cas pour les picornavirus. Il se peut également que CI ait une fonction dans le transport du virus à travers la plante, hypothèse qui se base sur le fait qu'on la trouve localisée vers les plasmodesmes, perforations permettant la jonction entre les cellules (Baunoch *et al.*, 1991).

4.6.7 Protéine P3.

C'est une protéine de 347 à 358 acides aminés (Figure 4) d'un PM d'environ 38K et qui présente moins de 30 % d'homologie parmi les potyvirus connus (Domier *et al.*, 1987). C'est actuellement la protéine la moins connue des potyvirus, mais elle a pu être détectée dans des plantes infectées au TVMV grâce à des anticorps dirigés spécifiquement contre elle. Il a été suggéré que P3 était une protéine membranaire du fait qu'elle contient des motifs pouvant former des hélices transmembranaires; son rôle reste malgré tout totalement inconnu (Rodríguez-Cerezo et Shaw, 1991).

4.6.8 Protéine HC.

La taille de la protéine HC est de 457 à 459 acides aminés (Figure 4) et l'homologie parmi les potyvirus connus est d'environ 54 % (Domier *et al.*, 1987). La protéine HC catalyse de façon autocatalytique le clivage à son extrémité C-terminale comme cela a été décrit précédemment (chapitre 4.5.3); de ce fait elle est aussi appelée HC-Pro (Carrington *et al.*, 1989*a,b*). Le domaine responsable de l'activité protéolytique est localisé dans la deuxième moitié C-terminale de HC, et au moins une fonction supplémentaire est associée à sa région N-terminale. En effet, il fut démontré bien avant que la séquence des potyvirus soit connue, que la transmission des potyvirus par les pucerons dépend de la présence d'une protéine virale d'environ 51 à 58 kDa qui est la protéine HC, mais dont la forme active est un dimère de HC (Govier et Kassanis, 1974*a,b*; Govier *et al.*, 1977; Kassanis et Govier, 1971*a,b*; Thornbury *et al.*, 1985). Cette protéine est localisée dans le cytoplasme des cellules infectées, le plus souvent sous forme

d'inclusions amorphes (Baunoch *et al.*, 1990), et elle doit être présente en même temps que le virus pour que celui-ci soit transmis par les pucerons; le test qui permet de vérifier l'effet de HC sur la transmission consiste à nourrir des pucerons par une solution qui contient du saccharose, de la protéine HC purifiée et le virus (Berger et Pirone, 1986; Kassanis et Govier 1971*b*; Thornbury *et al.*, 1985).

Il existe des potyvirus dont certaines souches ne sont pas transmissibles par les pucerons (phénotype NAT) à moins que la protéine HC d'une souche transmissible (phénotype AT) ne soit ajoutée dans le test de transmission, ce qui démontre bien son rôle dans le transport par les vecteurs (Govier et Kassanis, 1974*b*; Thornbury *et al.*, 1985). Les comparaisons de séquences entre les souches AT et celles NAT d'un potyvirus ainsi qu'avec d'autres potyvirus de phénotype AT ont montré une différence de deux acides aminés dans la partie N-terminale de HC: une lysine (Lys) changée en acide glutamique (Glu) et une isoleucine (Ile) en valine (Val) (Thornbury *et al.*, 1990). La mutagenèse dirigée, combinée à l'utilisation de transcrits infectieux, ont permis de démontrer que le seul changement Lys en Glu dans la protéine HC du TVMV rend ce virus non-transmissible par les pucerons; ce changement devant probablement empêcher la formation du dimère actif de HC (Atreya *et al.*, 1992).

Il semble que la protéine HC permettrait au virus de se fixer au stylet du puceron vecteur, ce qui en faciliterait le transport (Berger et Pirone, 1986). Il faut noter, comme cela a été mentionné dans le chapitre concernant la protéine CP (chapitre 4.6.2), que celle-ci a également un rôle dans la transmission des potyvirus par les vecteurs; ainsi, certains potyvirus qui ont une protéine HC active dans le test de transmission, peuvent être

de phénotype NAT du fait du génotype de leur capsid (Atreya *et al.*, 1990).

Des mutations ont également été introduites dans une région de HC très riche en cystéines (Cys) qui constitue un motif de type « *zinc-finger* », motif présent dans les protéines impliquées dans la fixation des acides nucléiques (Robaglia *et al.*, 1989); ces mutations atténuent beaucoup la production d'ARN ainsi que les symptômes du virus recombinant (Atreya *et al.*, 1992). Ce résultat est en accord avec le fait que la protéine HC peut être purifiée par chromatographie sur colonne d'oligo (dT)-cellulose et qu'elle a de l'affinité pour certains polyribonucléotides; la présence d'un site type « *zinc-finger* » dans la séquence de HC suggère que cette protéine pourrait avoir également un rôle dans la structure de l'ARN viral (Thornbury *et al.*, 1985).

HC-Pro est donc une protéine multi-fonctionnelle: une activité protéase du type cystéine, un rôle déterminant dans la transmission du virus par les pucerons, et un rôle probable sur la conformation de l'ARN viral, et donc sur la stabilité du virus, la régulation de la réplication ou/et la formation des symptômes (Atreya *et al.*, 1992).

4.6.9 Protéine P1.

C'est la protéine des potyvirus la moins conservée en taille et en séquence avec 256 à 397 acides aminés (Figure 4) et moins de 15 % d'homologie sur l'ensemble de la protéine. On trouve par contre plus de 30 % d'homologie sur la moitié C-terminale, ce qui suggère une double fonction pour cette protéine, dont l'activité protéase qui a été décrite précédemment (chapitre 4.5.4) (Domier *et al.*, 1987; Laín *et al.*, 1989a).

La protéine P1 de 28 kDa du TVMV, mais pas celle de 34 kDa du TEV, présente quelques homologies avec la protéine de 30 kDa du TMV, qui est responsable du transport du virus de cellule à cellule (Deom *et al.*, 1987). Toutefois, il faut noter que les protéines impliquées dans le transport de cellule à cellule sont en général très différentes d'un groupe de virus à l'autre et même au sein d'un même groupe; cette observation rend compte de la spécificité des interactions d'un virus avec son hôte (Hull, 1989; Riechmann *et al.*, 1992).

La P1 du TVMV a été détectée dans une fraction de cellules enrichie en membranes du reticulum endoplasmique, mais sa localisation exacte, qui pourrait confirmer son rôle dans le transport de cellule à cellule, n'est pas encore connue (Rodríguez-Cerezo et Shaw, 1991). La protéine P1 a donc une fonction autoprotéolytique à son extrémité C-terminale, et a peut être d'autres fonctions non démontrées comme le transport de cellule à cellule.

4.7 RÉPLICATION DES POTYVIRUS.

Il n'y a pratiquement aucune information concernant la réplication des potyvirus et des hypothèses, basées sur les virus à ARN dont l'expression ressemble à celle des potyvirus, peuvent être faites. On peut ainsi considérer sur l'ARN des potyvirus, une région codant pour les facteurs impliqués dans la réplication, semblable à celle des picornavirus, qui comprend la région qui code pour la 6K1, la CI, la 6K2, la VPg, la NIa-Pro, la NIb et sans doute la P3 (Figure 4). Toutes ces protéines seraient impliquées dans la réplication mais la CI, la VPg et la NIb sont directement responsables de

l'initiation, de la polymérisation et de la séparation des brins d'ARN en cours d'élongation (Riechmann *et al.*, 1992). Il est possible que la réplication des potyvirus soit étroitement associée aux chloroplastes; en effet, de l'ARN viral négatif et double brin du TEV ont été détectés dans des fractions de tissus contenant les chloroplastes (Gadh et Hari, 1986). De plus, l'ARN du PVY ainsi que ses protéines HC et CP ont été détectés dans les chloroplastes d'une espèce infectée au PVY (Gunasinghe et Berger, 1991). On ne sait pas encore où se produit exactement la réplication des potyvirus, mais il est probable que les symptômes qu'ils provoquent sont le résultat de l'association de certains de leurs produits d'expression avec les chloroplastes.

4.8 DÉPLACEMENT DES POTYVIRUS.

4.8.1 Déplacement dans la plante.

Les virus des plantes pénètrent, dans la grande majorité des cas, par la perforation mécanique de la paroi cellulaire par des vecteurs (insecte, champignon, nématode) (Deom *et al.*, 1992). Une fois dans la cellule, si celle-ci n'est pas immunisée, le virus peut se multiplier et se propager dans la plante. Si le virus est incapable de se déplacer dans d'autres cellules, l'hôte est résistant au virus sinon, il y a infection plus ou moins importante de la plante. Le mouvement du virus à travers la plante représente donc un déterminant important de la virulence et de la pathologie.

Il y a deux phases dans le transport des virus: la première est le déplacement lent du virus de cellule à cellule par les plasmodesmes; la deuxième est le déplacement rapide sur de grandes distances à travers les tissus vasculaires de la plante.

Le diamètre normal des plasmodesmes est d'environ 3 nm, ce qui est plus étroit que le diamètre des particules virales filamenteuses tel que les potyvirus. Deux mécanismes ont été proposés pour expliquer le passage des virus à travers ces canaux: soit que le virus se déplace sous sa forme « ARN nu », soit sous sa forme encapsidée. Les protéines impliquées dans le mouvement de cellule à cellule pourraient intervenir au niveau des plasmodesmes afin d'augmenter leur diamètre (Deom *et al.*, 1992). On ne sait pour l'instant absolument rien du transport des potyvirus dans les plantes sauf que les protéines P1 et CI sont probablement impliquées dans ce mécanisme.

4.8.2 Transport par les vecteurs.

Les potyvirus sont transportés en grande majorité de façon non-persistante par les pucerons, c'est-à-dire que le virus reste infectieux pendant quelques minutes dans l'appareil de succion de l'insecte. C'est la combinaison d'au moins deux facteurs viraux, la protéine HC et la capsid, qui définit le caractère transmissible d'un potyvirus par un puceron. Il y a probablement d'autres facteurs puisque chaque potyvirus n'est transmissible que par une partie seulement des 3 700 espèces connues de pucerons (Matthews, 1991). On ne connaît pas encore les facteurs de transmission des quelques potyvirus qui ont d'autres vecteurs de transmission comme les champignons, les acariens et la mouche blanche.

4.8.3 Transport par les graines.

Moins de 15 % des potyvirus sont transmissibles par les graines de la plante hôte. Ce mode de transmission dépend beaucoup de l'espèce infectée et de son âge au moment

de l'infection. Les mécanismes par lesquels le virus infecte les graines sont encore peu connus; le virus doit entrer dans l'embryon et y demeurer stable pour être considéré comme transmissible par la graine (Matthews, 1991). Wang et Maule (1992) ont montré que le potyvirus PSbMV pénètre rapidement dans les embryons par les tissus adjacents, s'y multiplie et y persiste pendant la maturation de la graine. Le virus ainsi stabilisé dans les graines peut être transporté sur de grandes distances, contrairement à la transmission par les vecteurs qui est limitée dans l'espace.

4.9 ÉVOLUTION DES POTYVIRUS.

Les trois groupes de virus de plantes regroupés dans le super-groupe du « type-picornavirus » (chapitre 4.4) ainsi que les picornavirus, ont quatre caractéristiques communes: un ARN polyadénylé lié à une VPg; l'ARN traduit en une polyprotéine clivée par des protéases; des homologies significatives de séquence dans certaines protéines de la polyprotéine et, une similitude dans l'organisation de plusieurs protéines sur la polyprotéine. La Figure 5 montre l'organisation des protéines ainsi que les régions d'homologie sur les polyprotéines des potyvirus, des comovirus, des népovirus, et des picornavirus. L'hypothèse expliquant comment ces groupes de virus, qui infectent deux règnes eucaryotes différents (animal et végétal), peuvent avoir autant d'homologies dans la structure et l'organisation de leur génome, est la suivante: tous ces groupes auraient possiblement un ancêtre commun, et il y aurait eu à un moment donné des recombinaisons entre le génome de virus appartenant à différents groupes (Goldbach et Wellink, 1988; Lai, 1992). Si on compare ces trois groupes de virus de plantes (Figure 5), on remarque que les népovirus et les comovirus ont un génome bi-segmenté, mais que l'organisation

FIGURE 5: Comparaison génomique des picornavirus (poliovirus) et des virus de plantes du super-groupe « type-picornavirus » (potyvirus, népovirus et comovirus). La protéine VPg (cercle vide) et la queue polyadénylée (An) sont indiqués aux extrémités des ARN (rectangle). Les régions ayant la même fonction pour chacun des virus sont indiquées: la protéine ayant une fonction hélicase (triangle noir); la protéase du type cystéine (cercle noir); la polymérase (carré noir). Les régions ombragées représentent les domaines dans lesquels on retrouve une homologie significative de séquence ($> 20\%$). Figure adaptée suivant Golbach et Wellink (1988).

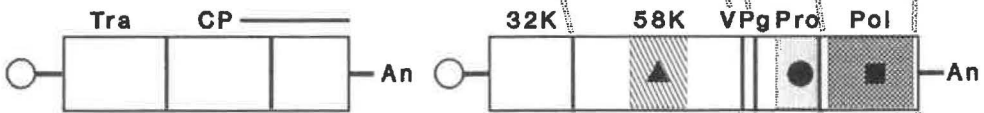
POTYVIRUS



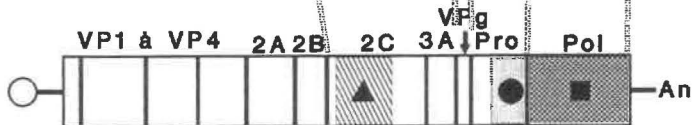
NÉPOVIRUS



COMOVIRUS



POLIOVIRUS



La taxonomie des potyvirus peut être abordée par calcul informatisé des niveaux d'homologie dans la séquence de la protéine CP (Shukla et Ward, 1989a; Ward et Shukla, 1991); des arbres généalogiques ont pu ainsi être élaborés, mais sans confirmation expérimentale. Il a également été démontré, par l'analyse informatique des séquences protéiques de la capsid, que tous les virus de plantes à ARN positif ayant une particule virale filamenteuse flexible (potyvirus, potexvirus, bymovirus, carlavirus et clostéovirus) forment une famille qui dérive probablement d'un ancêtre commun; tous les virus de plantes à ARN positif ayant une particule virale filamenteuse rigide (tobamovirus, tobravirus, furovirus et hordeivirus) forment quant à eux une autre lignée phylogénique distincte de la première et qui dérive d'un autre ancêtre commun (Dolja *et al.*, 1991). Ce résultat est donc en faveur de recombinaisons entre des virus appartenant à différents groupes. Cette hypothèse est bien représentée par les potyvirus, puisqu'ils ont une particule virale qui s'apparente à celle des virus filamenteux mais une stratégie d'expression et un génome qui s'apparentent à ceux des virus icosaédriques (népovirus et comovirus).

5. MODE D'ÉTUDE DES POTYVIRUS.

5.1 CLONAGE ET SÉQUENÇAGE.

En 1989, quatre potyvirus avaient été séquencés au complet; en 1992, ce nombre est passé à sept (sans compter les différentes souches du PPV, du PVY et du SMV). La méthode utilisée dans tous les cas pour le clonage, consistait en la construction d'une banque d'ADNc suivie du criblage d'un grand nombre de clones. Maiss *et al.* (1988) ont

par exemple, isolé plus de 70 clones de leur banque d'ADNc pour obtenir la séquence complète du génome du PPV, alors que Robaglia *et al.* (1989) en ont isolé 150 pour celle du PVY. Le problème était d'obtenir des clones représentant l'extrémité 5'-terminale de l'ARN viral, ce qui nécessitait le séquençage direct de l'ARN. De plus, lorsque Jayaram *et al.* (1992) ont cloné deux souches du virus SMV, ils ont dû construire une banque d'ADNc pour chacune des souches. La technique d'amplification par polymérisation en chaîne (PCR) n'avait pas encore été utilisée pour le clonage de potyvirus avant 1990.

Lorsque la séquence complète d'un potyvirus est connue, ses protéines sont localisées sur la polyprotéine et celles-ci peuvent être ensuite clonées, séparément ou au sein d'une polyprotéine, pour leur production en grande quantité dans un système d'expression (procaryote ou eucaryote) ou, pour l'étude de la protéolyse. Ces protéines et leurs précurseurs peuvent être également exprimés de façon stable par l'introduction par *Agrobacterium tumefaciens* dans le génome d'une plante, ou de façon transitoire, par transfection de protoplastes. Les produits viraux peuvent être également exprimés par la traduction *in vitro* de leur ARN ou de transcrits synthétiques.

5.2 TRANSCRITS INFECTIEUX.

Le principe des transcrits infectieux consiste à construire un clone représentant le génome complet d'un virus dans un vecteur plasmidique. Le plasmide est ensuite transcrit en ARN synthétique, et le caractère infectieux de cet ARN peut être vérifié, soit en l'inoculant de façon mécanique à des plantes hôtes, soit par transfection de protoplastes. Domier *et al.* (1989) ont obtenu le premier clone infectieux pour un potyvirus et la

difficulté principale résidait en la construction d'un clone complet du TVMV contenant un promoteur provenant d'un bactériophage (T3 ou T7) avec le minimum de séquence transcrite ne provenant pas du TVMV. L'efficacité d'infection des transcrits, qui doivent obligatoirement comporter un analogue de coiffe méthylée pour être infectieux, est extrêmement faible comparée à celle de l'ARN viral purifié (environ 100 à 500 fois moins efficace que l'ARN natif); ceci pourrait être attribué à la présence des séquences « non-virales », à des mutations, ou à l'absence de VPg (Domier *et al.*, 1989; Riechmann *et al.*, 1990). Cette efficacité d'infection par de l'ARN synthétique dépend de la qualité d'ARN, de l'hôte choisi pour l'inoculation, ainsi que de la technique d'inoculation utilisée (Riechmann *et al.*, 1990).

Une des étapes primordiales pour tester un clone potentiellement infectieux est la synthèse de grandes quantités d'ARN coiffé et polyadénylé. Maiss *et al.* (1992) ont pu éliminer cette étape pour le PPV en infectant directement des plantes sensibles par un clone complet du PPV placé sous le contrôle du promoteur 35S du virus de la mosaïque du chou-fleur (CaMV) et suivi du signal de polyadénylation de la nopaline synthase. Le plasmide, inoculé également de façon mécanique, est transcrit par l'hôte en ARN coiffé et polyadénylé, qui se comporte alors de la même manière que les transcrits synthétiques, en conduisant à la multiplication et à la production de virus.

6. CONCLUSION ET OBJECTIFS DE LA RECHERCHE.

Le groupe des potyvirus est donc le groupe le plus important des virus de plantes de par le nombre de membres qu'il comporte. C'est également le groupe de virus à ARN

positif qui code pour le plus grand nombre de protéines, dont sept ont au moins une fonction connue, et trois, encore inconnue.

En 1989, seuls quatre potyvirus avaient été clonés et séquencés au complet, et les recherches concernant trois de ces virus (le PPV, le TEV et le TVMV) ont procuré la majeure partie des informations concernant ce groupe viral.

Afin d'être capable de participer aux découvertes conduisant à la compréhension des mécanismes de la réplication et de la propagation des potyvirus, nous avons choisi de travailler sur le TuMV, un potyvirus d'intérêt agronomique, qui en 1989 n'était pas encore caractérisé. C'est dans la perspective de débiter la caractérisation moléculaire du virus TuMV que les deux objectifs énoncés ci-dessous ont été fixés pour ce travail de doctorat:

- Appliquer la technique d'amplification par PCR au clonage et au séquençage du génome de TuMV, ceci dans le but de localiser et de caractériser ses protéines.
- Construire un clone complet du TuMV, qui après transcription *in vitro*, puisse conduire à la synthèse d'un ARN le plus identique possible à l'ARN viral et qui si possible, puisse engendrer une infection du même type que celle obtenue par l'ARN purifié du TuMV.

MATÉRIEL ET MÉTHODES

1. BACTÉRIES, VECTEURS, PLANTES ET VIRUS.

1.1 SOUCHES D'*ESCHERICHIA COLI*.

Souches: **Génotype:**

JM 110 *rpsL (str^r), thr, leu, thi-1, lacY, galK, galT, ara, tonA, tsx, dam, dcm, supE44, Δ(lac-proAB), F'[traD36 proAB lacI^s lacZΔM15]*. Cette souche a permis de digérer certains plasmides par les enzymes de restriction sensibles aux méthylations.

XL1-Blue *recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac, F'[proAB⁺ lacI^s, lacZΔM15, Tn10(*tet^r*)]*. Cette souche a servi à la grande majorité des clonages ainsi qu'à la production de phage du type M13.

1.2 VECTEURS DE CLONAGE.

pUC9 (2680 pb): plasmide dérivant de pBR322. Il porte le gène de résistance à l'ampicilline (*Amp^r*). Le site multiple de clonage (SMC) se trouve dans le gène *lacZ*.

pUC18, pUC19 (2686 pb): plasmides dérivant de pUC9. *Amp^r*. pUC18 et pUC19 diffèrent par l'orientation du SMC dans le gène *lacZ*.

pBluescript^{MD} I KS (+), (pKS et pSK) (2961 pb) (Stratagene, La Jolla, CA): phagemides dérivant de pUC19. *Amp^r*. L'origine de répllication *f1 (+)* permet la synthèse de simple brin par coinfection à l'aide d'un phage auxiliaire. pKS et pSK diffèrent par l'orientation du SMC dans le gène *lacZ*. Il y a les promoteur de transcription des phages T3 et T7 de chaque coté du SMC.

M13mp18 RF et M13mp19 RF (7250 pb): forme répllicative (RF) de phage dérivant du M13. Même SMC que les plasmides pUC18 et pUC19 dans le gène *lacZ*.

1.3 PLANTES.

Plusieurs espèces végétales ont été utilisées pour la production du virus TuMV ainsi que pour les tests d'infection:

Brassica napus ssp. *napobrassica* cv. « Laurentien » (rutabaga)

Brassica campestris ssp. *rapifera* cv. « Just Right » (navet)

Brassica perviridis var. *tendergreen* (moutarde tendre)

1.4 VIRUS.

La souche de virus de la mosaïque du navet (TuMV) utilisée pour toutes les expériences incluses dans cette thèse provient d'un prélèvement dans la région de L'Assomption (Qué, Canada) de feuilles de rutabaga (*Brassica napus* ssp. *napobrassica* cv. « Laurentien ») ayant des symptômes très sévères de mosaïques caractéristiques d'une infection au TuMV. Le virus a été ensuite multiplié sur le rutabaga puis purifié dans le laboratoire du Dr. R. C. Sinha (Agriculture Canada, Ottawa, Ont, Canada) (Article 1).

2. MILIEUX DE CULTURE, TAMPONS ET SOLUTIONS.

2.1 MILIEUX DE CROISSANCE DES BACTÉRIES.

Milieu B (milieu de surface):

pH 7.5	8 g/l NaCl
6 g/l de bacto-agar	10 g/l Bacto-peptone

Milieu Luria-Bertani (LB):

pH 7.0	10 g/l NaCl
15 g/l de bacto-agar	5 g/l extrait de levure
pour les Pétris	10 g/l bacto-tryptone

Milieu SOB:

pH 7.0 0.5 g/l NaCl
 5 g/l extrait de levure
 20 g/l bacto-tryptone

Ajouter avant stérilisation 10 ml/l de KCl 250 mM) et après stérilisation 10 ml/l de solution filtrée de (1 M MgCl₂.6H₂O + 1 M MgSO₄.7H₂O).

Milieu SOC: milieu SOB avec 20 mM de glucose.

Milieu YT:

pH 7.5 5 g/l NaCl
 5 g/l extrait de levure
 8 g/l bacto-tryptone

Milieu de sélection « bleu/blanc » (X-gal/IPTG)

Milieu LB avec agar, autoclavé et encore liquide vers 55°C auquel on ajoute 40 mg/l de 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal, Boehringer Mannheim Canada Lté., Laval, Qué) à partir d'une solution concentrée à 20 mg/ml, et 0.5 mM d'isopropylthio-β-galactoside (IPTG, Boehringer Mannheim) à partir d'une solution concentrée à 100 mM.

2.2 ANTIBIOTIQUES.

Plusieurs antibiotiques (ATB) ont été utilisés pour la sélection des souches de *E. coli* ainsi que pour les souches transformées: l'ampiciline (Ampi) à 100 mg/l (Boehringer Mannheim) pour la sélection des bactéries transformées, et la tétracycline (Tet) à 15 mg/l (Sigma Chemical Company, St Louis, MO) pour la sélection de la souches XL1-Blue de *E. coli*. Ces ATB furent stérilisés sur filtre de 22 μm et ajoutés aux milieux liquides stériles ou aux milieux gélosés stériles mais encore liquides à une température de 55°C.

2.3 TAMPONS ET SOLUTIONS.

Acétate de potassium: (3 M K⁺, 5 M acétate)	Pour 100 ml	
	60 ml	d'acétate de potassium 5 M
	11.5 ml	d'acide acétique glacial
	28.5 ml	d'eau
SSC 20 X:	Pour un litre	
	175.3 g	NaCl
	88.2 g	citrate de sodium
	pH 7.0	(ajusté avec NaOH 10 N)
SSPE 20 X:	Pour un litre	
	175.34 g	NaCl
	27.6 g	NaH ₂ PO ₄ .H ₂ O
	7.4 g	EDTA
	pH 7.4	(ajusté avec NaOH 10N)
Tampon appariement 10 X: (méthode « dC-dG »)	100 mM	Tris. Cl pH 7.4
	1 M	NaCl
	2.5 mM	EDTA
Tampon KGB 10 X: (réactions enzymatiques)	1 M	glutamate de potassium
	250 mM	Tris-acétate pH 7.5
	100 mM	acétate de magnésium
	500 µg/ml	albumine sérique de boeuf (fraction V)
	5 mM	β-mercaptoéthanol
Tampon MOPS 10 X: (gels d'agarose)	0.2 M	MOPS [acide 3-(N-morpholino) propane sulfureux]
	50 mM	acétate de sodium
	10 mM	EDTA
	pH 7.0	(ajusté à pH 7.0 avec acide acétique)
Tampon TAE 50 X: (gels d'agarose)	Pour un litre	
	242 g	Tris base
	57.1 ml	acide acétique glacial
	100 ml	EDTA 0.5 M (pH 8.0)
Tampon TBE 5 X: (gels de polyacrylamide)	Pour un litre	
	54 g	Tris base
	27.5 g	acide borique
	20 ml	EDTA 0.5 M (pH 8.0)

Tampon TE pH 8.0:	10 mM 1 mM	Tris.Cl pH 8.0 EDTA pH 8.0
Tampon de charge 6 X pour ADN:	0.25 % 0.25 % 30 %	bleu de bromophéno xylène cyanol glycérol dans l'eau
Tampon de charge 2 X gel alcalin:	100 mM 20 % 0.02 % 2 mM	NaOH glycérol bleu de bromophéno EDTA
Tampon de charge 1.2 X pour ARN:	750 µl 150 µl 240 µl 100 µl 100 µl 80 µl	formamide désionisée tampon MOPS 10 X formaldéhyde à 37 % d'eau stérile glycérol 100 % bleu de bromophéno à 10 %
Tampon de charge 2 X pour protéines:	0.2 % 4 % 20 % 100 mM 200 mM	bleu de bromophéno sodium dodécyl sulfate (SDS) glycérol dans l'eau Tris.Cl pH 6.8 dithiothreitol (DTT)

3. PRÉPARATION DES ADN.

3.1 PLASMIDES.

La purification rapide de l'ADN plasmidique consiste en la lyse alcaline de 1.5 ml de culture bactérienne suivie d'une précipitation à l'éthanol, méthode décrite par Sambrook *et al.* (1989). Un kit de purification par colonne a également été utilisé en suivant le protocole qui l'accompagne (« Magic^{MD} Minipreps », Fisher Scientific Co. Lté, Montréal, Qué). La méthode de préparation des plasmides destinés à la séquence ou à

la transcription *in vitro* est basée sur la technique de purification au polyéthylène glycol (PEG) décrite par Sambrook *et al.* (1989). La technique modifiée utilisée est décrite ci-dessous.

Une culture bactérienne de 50 ml (Milieu LB + antibiotiques) à 37°C est centrifugée 15 min à 4 000 rpm à 4°C dans un tube FALCON 50 ml. Le culot est lavé avec 10 ml de solution STE (0.1 M NaCl, 10 mM Tris.Cl pH 8, 1 mM EDTA pH 8) puis resuspendu vigoureusement dans 2 ml de solution TEG froide (25 mM Tris pH 8, 10 mM EDTA, 50 mM glucose) et laissé sur glace pour 5 min. On ajoute 4 ml d'une solution fraîche de (0.2 M NaOH, 1 % SDS) et le tube est retourné plusieurs fois puis laissé 10 min à température ambiante. On ajoute 2 ml d'acétate de potassium (3 M K⁺, 5 M acétate), le tube est retourné plusieurs fois puis laissé sur glace pour 10 min. Le contenu du tube est transféré dans un tube COREX et soumis à une centrifugation de 15 min à 4 000 rpm. Le surnageant est filtré à travers un filtre de nylon très fin (66 µm de porosité) et versé dans un tube COREX. On ajoute 3.9 ml d'isopropanol à 100 % et le mélange est fortement agité au Vortex, laissé 10 min sur glace puis soumis à une centrifugation de 15 min à 8 000 rpm. Le surnageant est jeté et le culot d'acides nucléiques est rincé avec 2 ml d'éthanol à 70 % froid puis séché à l'air. Le culot est repris dans 300 µl de tampon TE et versé dans un tube Eppendorf de 1.5 ml. On ajoute 300 µl de LiCl 5 M et le tube est agité au Vortex pour quelques secondes puis soumis à une centrifugation à 12 000 g pour 10 min. Le surnageant est transféré dans un autre tube à centrifugation et 300 µl d'isopropanol froid sont ajoutés et le tube agité. On centrifuge 10 min à 12 000 g et le surnageant est enlevé. Le culot est rincé avec 100 µl d'éthanol à 70 % et séché à l'air. Il est ensuite dissout dans 500 µl de tampon TE contenant

20 $\mu\text{g/ml}$ de RNase A et laissé 1/2 heure à température ambiante avant d'y ajouter 500 μl de NaCl 1.6 M dans 13 % de PEG. Le tube est agité fortement, laissé quelques minutes sur glace et centrifugé à 12 000g pour 15 min. Le surnageant est enlevé délicatement et les traces de PEG sont éliminées par une deuxième centrifugation de 1 min. Le culot est repris vigoureusement dans 200 μl de tampon TE. On procède à une extraction par un volume de phénol, une extraction par un volume d'un mélange phénol-chloroforme 1:1 (V/V) et une dernière par un volume de chloroforme. On ajoute un dixième de volume d'acétate de sodium 3M et deux volumes d'éthanol 95 % froid. Après agitation vigoureuse, le tube est laissé 10 min sur glace puis centrifugé 10 min à 12 000 g. Le surnageant est jeté et le culot est rincé avec 100 μl d'éthanol à 70 % puis séché à l'air. L'ADN plasmidique est resuspendu dans 100 μl de tampon TE et la concentration en ADN est mesurée par spectrophotométrie à 260 nm (Sambrook *et al.*, 1989).

3.2 ADN SIMPLE BRIN.

La purification d'ADN simple brin (sb) destiné à la séquence et préparé à partir de 1.5 ml de culture bactérienne infectée par un phage du type M13 est dérivée de celle décrite par Sambrook *et al.* (1989); le protocole utilisé est décrit ci-dessous.

On inocule 2 ml de milieu YT avec la bactérie hôte (XL1-Blue) à partir d'une colonie isolée et la culture est agitée à 37°C pour une nuit. Cette culture est diluée cinquante fois avec du milieu YT et on verse 5 ml de la dilution dans un Erlenmeyer stérile de 100 ml. Chaque flacon est inoculé avec une plage de phage bien isolée à l'aide d'une pipette Pasteur. Les cultures sont agitées fortement à 37°C pour six à huit heures.

Un volume de 1.5 ml de culture est transféré en tube Eppendorf et centrifugé 5 min à 12 000 g. Le surnageant contient du virus pur qui constitue un stock de phage gardé à 4°C alors que le culot contient la bactérie hôte avec la forme répliquative (RF) du virus qui peut être purifiée et analysée comme un plasmide. Le surnageant est récupéré et centrifugé une deuxième fois à 12 000 g pour 5 min. On transfère 1.2 ml de surnageant dans un tube de 1.5 ml puis 300 μ l d'une solution de 20 % PEG et 2.5 M NaCl sont ajoutés; le tube est agité au Vortex et laissé 15 min à température ambiante pour la précipitation du phage puis centrifugé 10 min à 12 000 g. Le surnageant est prélevé et les traces de PEG sont éliminées. Le culot est resuspendu dans 100 μ l de tampon TES (20 mM Tris.Cl pH 7.5, 10 mM NaCl, 0.1 mM EDTA) et 50 μ l de phénol saturé au tampon TE pH 8.0 sont ajoutés. Le mélange est agité au Vortex pour 40 sec puis centrifugé 3 min à 12 000 g. La phase aqueuse est récupérée et versée dans un autre tube de 1.5 ml. Un volume de 50 μ l de chloroforme-isoamyl alcool 24:1 (V/V) est ajouté et le tube est agité fortement au Vortex pour 20 sec puis centrifugé 3 min à 12 000 g. Le surnageant (90 μ l) est versé dans un tube de 1.5 ml auquel sont ajoutés 9 μ l d'acétate de sodium 3 M (pH 5.2) et 200 μ l d'éthanol à 95 % froid. Le tube est agité fortement pour quelques secondes au Vortex et l'ADN est précipité à -70°C pour 20 min puis récupéré par une centrifugation de 30 min à 12 000 g. Le culot d'ADN viral est lavé avec 100 μ l d'éthanol à 70 % froid suivi d'une centrifugation de 10 min à 12 000 g. Il est ensuite séché à l'air pour quelques minutes puis resuspendu dans 20 μ l de tampon TE. L'ADN sb peut être visualisé en chargeant 1 à 3 μ l de la solution précédente sur un gel de 1 % d'agarose et être utilisé directement pour l'appariement d'une amorce en vue du séquençage enzymatique.

3.3 OLIGONUCLÉOTIDES.

Tous les oligonucléotides (Annexe 2) ont été synthétisés à l'aide d'un synthétiseur automatique d'oligonucléotides (Gene Assembler, Pharmacia Canada Inc., Baie d'Urfée, Qué) au laboratoire du Dr. F. Shareck (Centre de recherche en microbiologie appliquée, Institut Armand-Frappier).

4. MANIPULATION DES ADN.

4.1 ELECTROPHORÈSE SUR GEL.

4.1.1 Gel d'agarose.

Ce type de gel (0.4 à 2 % d'agarose de grade électrophorèse, BIO-RAD Laboratories, Mississauga, Ont) a permis de visualiser avec une bonne résolution tous les ADN double brins (db) dont la taille était supérieure à 300 pb ainsi que les ADN simple brin (sb) provenant de la préparation de phages dérivés du M13. Le mode de préparation décrit par Sambrook *et al.* (1989) a été utilisé (tampon TAE 1 X) à l'exception du fait que le bromure d'éthidium (Bet) était directement ajouté à 0.5 $\mu\text{g}/\text{ml}$ à l'agarose liquide avant le montage. Les ADN pouvaient être directement observés et photographiés à l'aide d'une lampe à ultra-violet (UV). Le marqueur de 1 kpb (Gibco Canada, Burlington, Ont) ainsi que celui de 100 pb (Pharmacia) ont été utilisés pour estimer la taille des fragments d'ADN.

4.1.2 Gel alcalin.

Ce type de gel a permis d'analyser les produits de synthèse d'ADNc sb marqués

au [^{32}P]. Le principe de ce gel (Sambrook *et al.*, 1989) consiste à préparer un gel de 1 % d'agarose dans (50 mM NaCl, 1 mM EDTA pH 8.0) puis à équilibrer ce gel pour 45 min dans le tampon alcalin d'électrophorèse (30 mM NaOH, 1 mM EDTA). Un volume de tampon de charge 2 X pour gel alcalin est ajouté aux ADN sb puis le mélange est chargé sur le gel, qui est soumis à une électrophorèse dans le tampon alcalin pour 5 à 6 heures à 40 V et à 4°C. Le gel est ensuite neutralisé 45 min dans de l'acide trichloroacétique à 7 % puis rincé à l'eau distillée. Il est ensuite séché entre deux couches de papier absorbant, enveloppé dans du « Saran Wrap^{MD} » puis exposé dans une cassette à un film photographique « X-OMAT » (Kodak) à -70°C.

4.1.3 Gel de polyacrylamide.

Les électrophorèses sur gel de polyacrylamide (PAGE) (acrylamide de grade électrophorèse, BIO-RAD) ont permis de visualiser les ADN db dont la taille était inférieure à 1000 pb. Le mode de préparation utilisé est celui décrit par Sambrook *et al.* (1989) qui consiste au montage d'un gel de polyacrylamide tamponné par le TBE, au chargement du gel par le mélange (ADN-tampon de charge 6 X), et à l'électrophorèse dans le TBE 1 X. Le gel est ensuite coloré au Bet pour 20 min puis analysé et photographié aux UV.

4.2 AMPLIFICATION PAR POLYMÉRISATION EN CHAÎNE (PCR).

L'enzyme thermorésistante utilisée était l'ADN polymérase de *Thermus aquaticus* (Taq polymérase, BIO/CAN Scientific Inc., Mississauga, Ont). Les réactions de 100 μl

étaient incubées en tube Eppendorf de 500 μl et recouvertes par 50 μl d'huile de paraffine.

Chaque mélange d'amplification comportait les composants suivants:

10 μl	de tampon 10 X (0.5 M KCl, 0.1 M Tris pH 9.0, 1 % Triton X-100)
1 μl	de mélange des 4 dNTP à 20 mM
1 μl	de TMAC $5 \cdot 10^{-3}$ M (chlorure-ammonium-tétraméthyle)
10 pmole	de chaque amorce oligonucléotidique
1 à 4 mM	de MgCl_2 à partir d'une solution à 25 mM
1 μl	ADN simple brin dilué ou 1 ng d'ADN db
2.5 unités	de <i>Taq</i> polymérase
eau	quantité suffisante pour 100 μl

Les tubes étaient ensuite maintenus dans un bloc chauffant contrôlé par ordinateur (DNA Thermal Cycler, Perkin Elmer Cetus Corp.) et les conditions d'amplification programmées pour chaque nouvelle paire d'amorces. Les conditions standards utilisées étaient les suivantes (X représente un temps de 1 min/ kpb à amplifier; Y représente une température d'appariement de 45°C à 60°C, dépendant des amorces utilisées):

- Cycles 1-2 : dénaturation de 30 sec à 94°C, appariement de 30 sec à 37°C, polymérisation de X min à 72°C.
- Cycles 3-34 : 30 sec à 94°C, 30 sec à la température Y, X min à 72°C.
- Cycle 35 : 30 sec à 94°C, 30 sec à la température Y, 6 min à 72°C.

A la fin des cycles d'amplification, l'huile de paraffine était extraite par 100 μl de chloroforme et la phase aqueuse récupérée après 2 min de centrifugation à 12 000 g et conservée à 4°C jusqu'à son utilisation. Une partie des réactions d'amplification pouvait être analysée directement par électrophorèse sur gel d'agarose. Les ADN provenant des amplifications ont été traités comme la plupart des ADN db.

4.3 DIGESTION ENZYMATIQUE.

Les ADN db ont été digérés par une ou plusieurs enzymes de restriction (Pharmacia) selon la même procédure; l'ADN est dilué dans un volume de 10 à 20 μ l de tampon KGB à la concentration optimale (0.5 X à 2 X) définie pour chaque enzyme (Sambrook *et al.*, 1989) et à la température optimale de l'enzyme. Le temps d'incubation dépendait de la quantité d'ADN à digérer mais n'excédait pas quatre heures.

4.4 PURIFICATION DES ADN APRÈS TRAITEMENT ENZYMATIQUE.

Le traitement des ADN après une réaction enzymatique dépendait des applications pour lesquelles ils étaient destinés:

- Pour être analysé par électrophorèse sur gel, le tampon de charge 6 X était directement ajouté à la réaction enzymatique et le mélange chargé sur gel.
- Pour subir une autre réaction enzymatique dont la température ou la concentration optimale en tampon KGB était différente de la première, le tampon KGB était ajusté à la concentration optimale de la deuxième enzyme, celle-ci était ajoutée et le mélange réactionnel incubé à sa température optimale d'incubation.
- Pour être cloné, l'ADN était purifié directement à partir de la réaction enzymatique ou après électrophorèse sur gel d'agarose et prélèvement de la bande d'ADN à l'aide du kit de purification d'ADN « GeneClean^{MD} » (BIO/CAN) pour les fragments supérieurs à 300 pb, et du kit de purification d'ADN « MerMaid^{MD} » (BIO/CAN) pour ceux inférieurs à 300 pb, en suivant le protocole recommandé pour chacun d'eux.

- Pour être transcrit en ARN, les plasmides linéarisés étaient extraits trois fois au phénol-chloroforme, extraits une fois au chloroforme, précipités à l'éthanol puis dilués à une concentration de $1 \mu\text{g}/\mu\text{l}$ avec du tampon TE stérile.

4.5 CLONAGE DES ADN.

4.5.1 Clonage par ligation.

Chaque réaction de ligation consistait en $10 \mu\text{l}$ de tampon 1 X « One-Phor-All Buffer Plus » (Pharmacia) avec 0.5 mM d'ATP et l'ADN ligase du phage T4 (Pharmacia).

Deux types de ligations ont été utilisées:

- Ligation d'extrémités franches: le vecteur et l'ADN à cloner sont ajoutés dans une proportion de un pour un dans la réaction finale en présence de deux unités de ligase. La réaction est incubée 16 heures à 16°C .
- Ligations d'extrémités cohésives: le vecteur et l'ADN à cloner sont ajoutés dans un rapport molaire (insert: vecteur) supérieur à trois en présence de 0.5 unités de ligase. La réaction est incubée de une à quatre heures à 20°C .

Tous les ADN devant être clonés et qui comportaient un ou deux sites distincts de restriction ont été digérés, purifiés et ligués tel que décrit précédemment. Les bactéries étaient directement transformées par le produit des ligations puis étalées sur milieu de sélection. La sélection « bleu-blanc » des bactéries poussant sur X-gal/IPTG ainsi que l'analyse aléatoire d'une dizaine de colonies blanches ont été utilisées lorsque le clonage se faisait dans un vecteur non recombinant. Quand le vecteur de clonage avait déjà une

insertion dans son SMC, les bactéries issues de la transformation étaient analysées par hybridation *in situ*. Les plasmides provenant de celles donnant un signal positif après hybridation, étaient analysés par cartographie de restriction.

4.5.2 Clonage sans ligation (« clonage dC-dG »).

Certaines amorces utilisées pour l'amplification par PCR n'avaient pas de sites de restriction, ce qui a impliqué l'utilisation de la méthode décrite ci-dessous pour leur clonage dans le vecteur pUC9; une queue de résidus cytosine (C) est ajoutée aux extrémités 3' des produits d'amplification grâce à la terminale-désoxy-transférase (TdT) (Gibco), puis ces ADN sont clonés par appariement au vecteur pUC9 possédant une queue de résidus guanosine (G) dans son site de restriction *Pst*I (« 3'-oligo(dG)-tailed pUC9 », Pharmacia). La méthode utilisée pour l'adjonction d'une queue de C est celle-ci: les produits d'amplification sont analysés et purifiés après électrophorèse sur gel d'agarose puis dilués dans un petit volume de tampon TE avant d'être ajoutés au mélange réactionnel suivant:

10 μ l	de tampon 5 X (0.5 M cacodylate de potassium pH 5.2, 10 mM CoCl ₂ , 1 mM DTT)
1 μ l	de dCTP 1 mM
14 unités	de TdT
volume d'eau	pour compléter à 50 μ l

La réaction est incubée 5 min à 37°C puis stoppée en ajoutant 2 μ l d'EDTA 500 mM. L'ADN est ensuite purifié tel que décrit précédemment. La réaction d'appariement se fait avec un rapport molaire vecteur-insertion environ égal à cinq

dans 100 μ l de tampon d'appariement 1 X. Le mélange est incubé 5 min à 67°C puis deux heures à 57°C. Des bactéries compétentes sont alors directement transformées par un volume de la solution d'appariement correspondant à 5-20 ng de vecteur, puis étalées sur boîtes de Pétri (milieu LB avec Amp^r) en présence de X-gal et d'IPTG. Les plasmides isolés à partir des colonies blanches sont analysés par cartographie de restriction à l'aide de l'enzyme *Pst*I.

4.6 TRANSFORMATION D'*E. COLI*.

Les bactéries compétentes ont été préparées selon la méthode d'Hanahan (1983) et le protocole de congélation et de transformation de ces bactéries est décrit par Sambrook *et al.* (1989). Les bactéries compétentes étaient conservées par volume de 200 μ l en tube à congélation de 1.5 ml à -70°C et deux protocoles distincts ont été utilisés après leur transformation suivant la nature du vecteur:

- Pour les bactéries transformées par un plasmide, 800 μ l de milieu SOC sont ajoutés après le choc thermique, et le mélange est agité à 37°C pour une heure. Les bactéries sont alors centrifugées 30 sec à 12 000 g et resuspendues dans 200 μ l de milieu SOB. On étale 20 μ l ou 180 μ l des bactéries sur milieu LB avec les antibiotiques appropriés et on incube une nuit à 37°C.
- Pour étaler les bactéries transformées par la forme répliquative (RF) d'un dérivé du phage M13, on ajoute après le choc thermique, 3 ml de milieu B de surface gardé liquide à 45°C, et 200 μ l d'une culture en phase exponentielle de bactéries XL1-Blue (milieu YT). On agite brièvement au vortex et on coule le mélange à la surface de boîtes de Pétri

(milieu LB) préchauffées à 37°C et on incube une nuit à 37°C.

4.7 SÉQUENÇAGE DES ADN.

Tous les plasmides et les ADN sb ont été séquencés par la méthode enzymatique (Sanger *et al.*, 1977) en utilisant l'ADN polymérase du phage T7 (avec les kits de séquence « T7 Sequencing^{MD} Kit » et « Auto Read^{MD} sequencing kit», Pharmacia). Toutes les séquences ont été réalisées avec la collaboration du service de séquençage de l'Institut Armand-Frappier, sous la direction du Dr. P. Tijssen (Centre de recherche en virologie). La grande majorité des réactions de séquences étaient marquées au [³⁵S] et la lecture des nucléotides se faisait à partir d'autoradiographies. Quelques-unes des réactions de séquence ont utilisé des amorces couplées à des fluorochromes et ont été chargées sur un séquenceur automatique (« Automated Laser Fluorescent ALF DNA sequencer », Pharmacia); la lecture de ces séquences s'est alors faite à partir de graphiques traités et dessinés par ordinateur.

4.8 TRANSFERT, SONDÉS ET HYBRIDATIONS.

4.8.1 Transfert sur nitrocellulose.

La recherche de clones recombinants a été faite dans certains cas par le transfert des colonies bactériennes sur membrane de nitrocellulose (BA85, Schleicher & Schuell, Keene, NH), suivi de la dénaturation et de la fixation de l'ADN tel que décrit par Sambrook *et al.* (1989). Les filtres ont ensuite été hybridés par une sonde marquée au [³²P] (sonde double brins ou sonde d'ADNc).

4.8.2 Transfert d'ADN d'un gel sur membrane (« *Southern blot* »).

Le transfert d'ADN sur une membrane de nylon (« Zeta-Probe^{MD} », BIO-RAD) à partir d'un gel non-dénaturant d'agarose a été fait en milieu alcalin: les ADN sont directement transférés par capillarité (Southern, 1975) par du NaOH 0.4 M pendant une nuit sur la membrane imbibée d'eau distillée; après transfert, celle-ci est rincée dans du SSC 2 X puis séchée deux heures à 80°C avant d'être hybridée.

4.8.3 Synthèse des sondes.

4.8.3.1 Sondes ADN double brins.

Les ADN employés comme sonde ont été purifiés à partir de gels d'agarose puis marqués au [³²P] par la procédure d'« *oligolabelling* » (« Oligolabelling^{MD} kit », Pharmacia) en suivant le protocole recommandé. Ces sondes étaient dénaturées 4 min à 100°C puis placées sur glace, avant d'être ajoutées au tampon d'hybridation.

4.8.3.2 Sondes ADN simple brin.

La synthèse des sondes d'ADNc à partir d'ARN viral est décrite plus loin (chapitre 5.3). Celles-ci étaient ajoutées sans dénaturation au tampon d'hybridation.

4.8.4 Hybridation avec sonde ADN.

Les membranes de nitrocelluloses (hybridation *in situ*) et de nylon (« *Southern blots* ») ont été hybridées et lavées suivant les protocoles décrits par Sambrook *et al.*

(1989). La pré-hybridation et l'hybridation se faisaient à 42°C dans le tampon d'hybridation contenant: 50 % de formamide désionisée, du SSPE 6 X, et des agents bloquants tels que la solution Blotto 0.05 X ou la solution de Denhardt 5 X. Les lavages des membranes se faisaient dans des conditions de plus en plus stringentes, en commençant par des lavages à température ambiante par (SSC 2 X, SDS 0.1 %), et en finissant par des lavages à 68°C dans (SSC 0.1 X, SDS 0.1 %).

5. MANIPULATION DES ARN.

5.1 EXTRACTION DE L'ARN VIRAL.

La méthode utilisée est basée sur celle décrite par Maiss *et al.* (1988): le virus purifié (0.5 mg à 1 mg) est prélevé à partir d'une solution stock conservée à -70°C, puis est digéré pendant 1/2 heure à 37°C par 200 µg/ml de protéinase K (Sigma) dans le tampon réactionnel suivant (10 mM Tris-HCl pH 7.6, 100 mM NaCl, 0.5 % SDS). On procède ensuite à deux extractions par un volume de phénol, et une extraction par un volume de phénol-chloroforme 1:1 (V/V). L'ARN est ensuite précipité une nuit à -20°C en ajoutant un dixième de volume d'acétate de sodium 3 M (pH 5.2) et deux volumes d'éthanol à 95 %. Le précipité est récupéré par une centrifugation de 1/2 heure à 12 000 g. Le culot d'ARN est ensuite rincé avec 100 µl d'éthanol à 70 % puis resuspendu dans 25 à 50 µl d'eau stérile. Une fraction de cette solution est analysée par électrophorèse sur gel dénaturant d'ARN afin d'estimer la quantité et l'intégrité de l'ARN viral.

5.2 ÉLECTROPHORÈSE DES ARN.

Le protocole utilisé est celui décrit par Fourney *et al.* (1988) qui consiste à préparer un gel d'agarose à 1 % dans le tampon (MOPS 1 X, 0.66 M de formaldéhyde). Les ARN à analyser sont dilués dans 5 μ l d'eau auxquels on ajoute 25 μ l de tampon de charge 1.2 X à ARN et 1 μ l de Bet à 1 mg/ml. Les ARN sont ensuite dénaturés 15 min à 65°C puis chargés directement sur gel. L'électrophorèse se déroule dans le tampon MOPS 1 X, à voltage constant (50 V à 100 V) et pour trois à quatre heures. Les ARN peuvent être directement visualisés et photographiés sur une lampe à UV. Les ARN de 0.24-9.5 kb (Gibco) ont servi de marqueurs de poids moléculaire.

5.3 SYNTHÈSE DES PREMIERS BRINS D'ADNc.

Les premiers brins d'ADNc ont été synthétisés par la transcriptase-inverse murine (M-MLV RTase, « SuperScript^{MD} » et « SuperScript^{MD} RNase H⁻ », Gibco) en suivant le protocole vendu avec l'enzyme. Moins de 1 μ g d'ARN étaient utilisés pour chaque réaction, avec comme amorce, 500 ng d'oligo (dT)₁₂₋₁₈ (Pharmacia) ou 40 ng d'amorces aléatoires (dN₆) (Pharmacia). Après une heure d'incubation à 37°C, 5 ng/ μ l de RNase A (Sigma) et 0.1 unité/ μ l de RNase H (Pharmacia) étaient ajoutés, avant de poursuivre l'incubation pour 30 min. Les premiers brins d'ADNc étaient ensuite purifiés par le kit de purification « GeneClean^{MD} » (BIO/CAN) et été conservés à -70°C dans 100 μ l de tampon TE pH 8.0.

Pour synthétiser une sonde radioactive simple brin à partir d'ARN viral, on ajoutait 10 μ Ci de dCTP [α -³²P] (ICN Biomedical Canada Lté, St Laurent, Qué) en abaissant

à 0.1 mM la concentration en dCTP dans le mélange réactionnel de synthèse d'ADNc amorcé à l'oligo (dT)₁₂₋₁₈. Les nucléotides non-incorporés étaient éliminés par la procédure de purification « GeneClean^{MD} » (BIO/CAN).

Pour ajouter une queue polyadénylée à l'extrémité 3' des premiers brins d'ADNc, la TdT (Gibco) a été utilisée comme suit: les ADNc, synthétisés à partir de 1 µg d'ARN, sont purifiés et ajoutés à la réaction d'élongation qui contient 100 µl de tampon cacodylate 1 X (100 mM cacodylate de potassium pH 5.2, 2 mM CoCl₂, 0.2 mM DTT), 0.01 mM de dATP et 14 unités de TdT. La réaction est incubée 5 min à 37°C et stoppée en ajoutant 2 µl d'EDTA 0.5 M. Les ADNc allongés en A sont purifiés par la procédure « GeneClean^{MD} » et conservés à -70°C avant leur utilisation pour l'amplification de l'extrémité 5' des ARN.

5.4 TRANSCRIPTION *IN VITRO*.

Nous avons utilisé pour chaque réaction de synthèse, l'ARN polymérase du phage T7 (« MEGAscript^{MD} T7 », Ambion, Austin, TX) pour transcrire environ 10 µg d'ADN linéarisé, en présence de l'analogue de coiffe (« m⁷G(5')ppp(5')G », Ambion), tel que décrit dans le protocole vendu avec l'enzyme de T7. Après quatre heures d'incubation à 37°C, la réaction était stoppée en ajoutant 1 unité/µl de DNase I suivi de 1/2 heure d'incubation à 37°C. Les ARN synthétiques étaient extraits trois fois au phénol-chloroforme 1:1 (V/V), une fois au chloroforme, puis précipités par un volume d'acétate de potassium 3M et deux volumes d'éthanol à 95 %. Après deux heures à -70°C et 30 min de centrifugation à 12 000 g, le culot d'acides nucléiques était lavé avec 100 µl d'éthanol à 70 % et repris dans 20 µl d'eau stérile. La concentration des ARN était

calculée par spectrophotométrie à 260 nm (Sambrook *et al.*, 1989). L'intégrité des transcrits était analysée par électrophorèse sur gel dénaturant d'ARN.

6. ÉLECTROPHORÈSE DES PROTÉINES.

Les protéines ont été analysées par électrophorèse sur gel de polyacrylamide en présence de SDS (SDS-PAGE) tel que décrit par Sambrook *et al.* (1989). Les protéines étaient diluées dans le tampon de charge 2 X pour protéines, dénaturées 4 min à 95°C puis chargées sur le gel SDS-PAGE.

7. INOCULATION DES PLANTES.

Des plants de navet (*Brassica campestris* ssp. *rapifera* cv. « Just Right ») ainsi que de moutarde tendre (*Brassica perviridis* var. *tendergreen*) ont été inoculés 15 jours post-germination, avec 2 µg de transcrits synthétiques coiffés (50 plants par espèce), avec 500 ng d'ARN de TuMV purifié (20 plants par espèce), ou avec 100 ng de virus purifié (10 plants par espèce). Le protocole d'inoculation était le suivant: une demi-feuille par plante est frottée à l'aide de la poudre de « Carborundum »; les ARN ou le virus sont dilués dans 50 µl de tampon TE pH 8.0 et étalés directement à l'aide d'un « Coton-tige^{MD} » sur la demi-feuille blessée; celle-ci est ensuite vaporisée avec de l'eau. L'apparition des symptômes fut suivie jusqu'à un mois après l'infection.

RÉSULTATS

1. PURIFICATION DU TuMV ET EXTRACTION DE L'ARN.

L'origine de la souche virale de TuMV est décrite dans le chapitre « matériel et méthodes ». Le virus a été, aussi bien dans le cas des préparations purifiées que des feuilles infectées, conservé à -70°C jusqu'à son utilisation ultérieure, car il est extrêmement instable à température ambiante.

Après purification, le TuMV a été observé directement en microscopie électronique par coloration négative (Figure 6). Il se présente sous forme de bâtonnets flexibles dont la longueur est d'environ 700 nm à 800 nm, ce qui est caractéristique des potyvirus. On remarque la présence de nombreuses particules virales plus courtes que la moyenne ainsi que de certaines plus longues; les bâtonnets plus courts seraient dus au fractionnement du virus lors de sa préparation plutôt qu'à l'encapsidation d'ARN incomplets; ceux d'environ deux fois la taille moyenne du virus, correspondraient à l'agrégation bout à bout de plusieurs particules virales, phénomène rapporté par Hill et Shepherd (1972).

Le virus purifié a également été chargé sur gel dénaturant de protéine (SDS-PAGE) afin de vérifier le poids moléculaire (PM) de la protéine structurale principale qui est la protéine de la capsid (CP) (Figure 7). Celle-ci a un PM d'environ 34K, ce qui correspond à la taille attendue pour une protéine CP de potyvirus. On observe également la présence d'une bande de plus faible PM qui est probablement un produit de dégradation, comme cela se produit pour la plupart des protéines CP des potyvirus (Shukla et Ward, 1989b).

FIGURE 6: Observation au microscope électronique en coloration négative d'une préparation de virus TuMV purifié (photo prise par le Dr. R. C. Sinha, Agriculture Canada, Ottawa, Ont, Canada). La barre représente une longueur d'environ 200 nm.

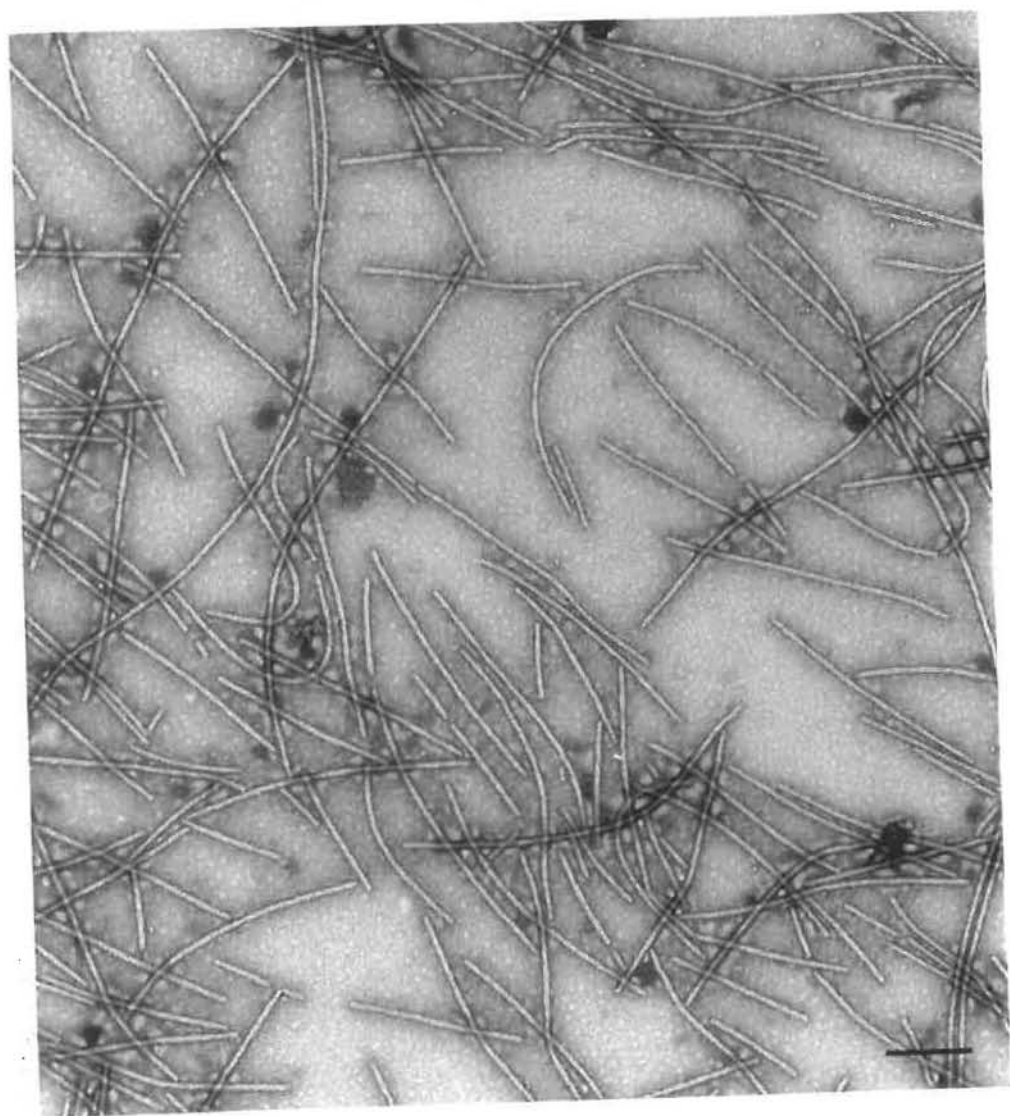
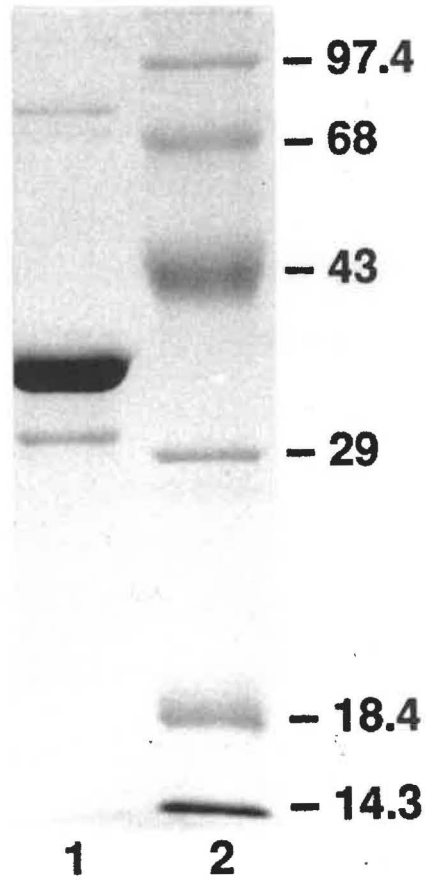


FIGURE 7: Protéine de la capsidite du TuMV sur gel de polyacrylamide 12 % en présence de SDS (SDS-PAGE) et colorée au bleu de Coomassie. Puits 1, environ 4 μ g de préparation virale de TuMV. Puits 2, marqueurs protéiques de poids moléculaire (BIO-RAD).



Nous avons procédé à l'infection mécanique de plusieurs espèces dont le navet (*Brassica campestris* ssp. *rapifera* cv. « Just Right ») et la moutarde tendre (*Brassica perviridis* var. *tendergreen*) par un extrait de feuilles de rutabaga infectées au TuMV. Les symptômes apparaissaient après trois semaines, aussi bien sur les feuilles infectées que sur les nouvelles feuilles, ce qui démontre bien le caractère systémique de l'infection par le TuMV dans ces deux espèces. Les feuilles de moutarde tendre, qui présentaient les symptômes de mosaïques, ont subi des coupes ultra-fines et été visualisées au microscope électronique par le service de microscopie électronique de l'Institut Armand-Frappier. On observe la présence de corps d'inclusion cytoplasmiques, absents des cellules de plantes non-infectées, de forme caractéristique enroulée et de lamelles allongées (Figure 8). D'après ces observations, la souche de TuMV que nous avons isolée serait du type 2 ou du type 3 en ce qui concerne les corps d'inclusion cytoplasmiques, comme cela a été défini au chapitre 4.6.6 de la « revue bibliographique ».

L'ARN du TuMV a été extrait à partir des préparations purifiées de virus et a été conservé à -70°C pour éviter sa dégradation. Sa taille et son intégrité ont été vérifiées par électrophorèse sur gel dénaturant d'agarose après chaque nouvelle extraction (Figure 9). On observe une bande peu dégradée d'environ 9.5 à 10 kb correspondant bien à la taille attendue pour l'ARN d'un potyvirus. Cet ARN a servi pour toutes les manipulations conduisant à son clonage et son séquençage.



kb

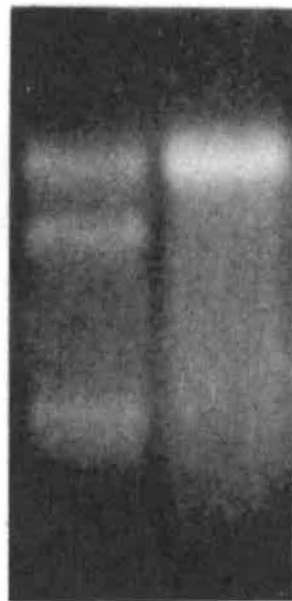
M

1

9.5-

7.5-

4.4-



2. CLONAGE ET SÉQUENÇAGE DU TuMV.

2.1 INTRODUCTION AUX ARTICLES.

Le premier objectif de ce travail de doctorat était d'obtenir la séquence complète de l'ARN du TuMV par la méthode la plus efficace; ceci en vue d'étudier la biologie moléculaire de ce virus de plante, entre autre, par le clonage de ses différentes protéines et la production d'anticorps dirigés contre elles. Jusqu'en 1989, la seule méthode utilisée pour cloner un virus à ARN consistait à construire une banque d'ADN complémentaire (ADNc) double brin (db) dans un phage lambda ou dans un plasmide, puis à procéder au « criblage » de cette banque jusqu'à ce que le génome au complet ait été cloné et séquencé. Ainsi, les quatre potyvirus (PPV, PVY, TEV et TVMV) ont été clonés par cette méthode (Allison *et al.*, 1986; Domier *et al.*, 1986; Laín *et al.*, 1989a; Maiss *et al.*, 1989; Robaglia *et al.*, 1989). Nous avons également envisagé de procéder à la synthèse d'une banque d'ADNc db de TuMV, et le Dr. M.F. Tremblay a tenté de construire celle-ci dans le phage « LambdaZap^{MD} » (Stratagene) au laboratoire du Dr. J.-F. Laliberté. Après l'analyse de plusieurs milliers de plages de lyse par hybridation *in situ*, seuls deux clones positifs ont été trouvés et séquencés: un de 500 pb (pSK500), l'autre de 1 200 pb (pSK1200); ce résultat peu satisfaisant nous a conduit à considérer une autre stratégie pour le clonage du TuMV. En effet, la découverte de la technique d'amplification par polymérisation en chaîne (PCR) (Saiki *et al.*, 1985), combinée à l'utilisation d'une ADN polymérase thermorésistante provenant de *Thermus aquaticus* (Taq polymérase) (Saiki *et al.*, 1988), nous ont permis d'envisager une méthode alternative et la mise au point de l'amplification de l'ARN de TuMV par PCR.

Les résultats du clonage par PCR de TuMV, ainsi que la description des clones provenant de la banque d'ADNc initiale, sont rapportés dans trois articles (Articles 1, 2 et 3) qui sont inclus dans cette thèse, car ils contiennent la majeure partie des résultats obtenus au cours de ce travail de doctorat. L'article 4 résulte de l'analyse de la séquence protéique du TuMV (Article 3) et démontre la première activité biologique d'une des protéines clonée du TuMV; en effet, nous avons montré chez *E. coli* que la protéine NIa du TuMV avait une activité protéolytique sur un site de clivage interne.

2.2 ARTICLE 1:

**TREMBLAY, M.-F., O. NICOLAS, R.C. SINHA,
C. LAZURE et J.-F. LALIBERTÉ**

**Sequence of the 3'-terminal region of turnip mosaic
virus RNA and the capsid protein gene.**

Journal of General Virology (1990), 71: 2769-2772.

2.2.1 Résumé.

Une séquence de 1 801 nucléotides provenant de la région 3' de l'ARN du TuMV a été clonée en partie par PCR; elle possédait un long et unique cadre de lecture sans codon d'initiation. Les séquences en acides aminés de trois différentes régions de la protéine CP purifiée de TuMV, incluant l'extrémité N-terminale, ont été obtenues par son séquençage direct; ces séquences partielles correspondaient bien à celle prédite à partir de la séquence codante. Ce résultat suggérait que la protéine CP de TuMV était un produit issu de la maturation d'une polyprotéine plus grande, stratégie d'expression commune des potyvirus. Le site potentiel de clivage de CP avec la protéine située à son extrémité N-terminale correspondait au dipeptide glutamine-alanine (Gln-Ala), un site fréquemment utilisé dans la protéolyse des potyvirus. La séquence codant pour la protéine CP se composait de 864 bases, ce qui correspondait à une protéine de 288 aminoacides et un PM calculé de 33.2K. La région 3' non-codante (3'-NC) adjacente comptait 667 nucléotides (nt) suivis de la queue polyadénylée. La séquence protéique de CP ressemblait beaucoup à celle d'autres potyvirus, la plupart des différences se trouvant localisées dans sa région N-terminale.

2.2.2 Contribution.

Ma contribution à cet article a consisté à la mise au point de l'amplification par PCR d'une région située en 3' de l'ARN du TuMV, puis au clonage des produits de PCR dans le vecteur pUC9. J'ai procédé à deux amplifications à partir du premier brin d'ADNc de l'ARN du TuMV amorcé par l'oligo(dT)₁₂₋₁₈; ces amplifications ont été réalisées grâce à la même amorce spécifique FT₂ (Annexe 1), complémentaire à l'ARN

viral, et à deux amorces dégénérées, homologues à la séquence virale. FT₂ provenait du séquençage d'un clone de 1 200 pb (pSK1200) issu de la banque d'ADNc db faite par le Dr. M.-F. Tremblay. L'amorce dégénérée JF₁ (Annexe 1) était basée sur la séquence protéique de l'extrémité N-terminale de la protéine CP du TuMV, alors que l'amorce dégénérée JF₃ (Annexe 1), était basée sur une homologie de séquence dans la région C-terminale de la protéine NIb de quatre potyvirus (PPV, PVY, TEV et TVMV). Les produits PCR de 400 pb et 700 pb, issus respectivement des amplifications par les paires d'amorces (FT₂-JF₁) et (FT₂-JF₃), ont été clonés dans le vecteur pUC9 puis été sous-clonés dans les vecteurs M13mp18 et M13mp19 en utilisant les sites de restriction *Pst*I conservés du vecteur pUC9. Les différents clones ont ensuite été séquencés avec la collaboration du service de séquençage de l'Institut Armand-Frappier.

2.2.3 Conclusion.

Nous avons donc été capables de cloner et de séquencer la région 3' de l'ARN du TuMV qui comprend la séquence complète codant pour sa capsid, ainsi que la région 3'-NC. Ce résultat a été obtenu en grande partie grâce à la technique d'amplification par PCR, car les deux clones provenant de la banque initiale d'ADNc de TuMV ne recouvraient pas la totalité de la séquence de la protéine CP. Les deux clones PCR ont permis d'apporter l'information manquante. La séquence 3'-NC, obtenue par le séquençage des deux clones de la banque d'ADNc, paraissait anormalement longue (667 nt) comparée à celle des autres potyvirus, ce qui nous a conduit par la suite à amplifier et à cloner la région 3' du TuMV pour vérification (Article 3).

L'objectif suivant était de cloner le génome de TuMV au complet, en ayant toujours recours au PCR, mais en amplifiant cette fois-ci des fragments d'ARN beaucoup plus longs que ceux obtenus et décrits dans l'Article 1. Ceci pour une raison importante: nous étions limités par le nombre d'amorces dégénérées pouvant être synthétisées sur la base des homologies provenant de la séquence complète des quatre potyvirus disponibles en 1990; certaines de ces régions d'homologie étaient séparées par plus de 1 200 bases, ce qui impliquait donc la nécessité d'amplifier des fragments d'ADN d'au moins 1 200 pb.

Nous avons ainsi travaillé à la mise au point d'une méthode d'amplification par PCR, puis de clonage de grands fragments d'ARN du TuMV. Le résultat de ce travail a été publié (Article 2).

2.3 ARTICLE 2:

NICOLAS, O. et J.-F. LALIBERTÉ

**The use of PCR for cloning of large cDNA
fragments of turnip mosaic potyvirus.**

Journal of Virological Methods (1991), 32: 57-66.

2.3.1 Résumé.

Cet article décrit une méthode dans laquelle le premier brin d'ADNc de l'ARN du virus TuMV est synthétisé, puis amplifié de façon enzymatique par l'ADN polymérase de *Thermus aquaticus*. Les amorces dégénérées qui ont servi aux amplifications, provenaient de régions d'homologie dans les séquences des protéines CI et NIB de quatre potyvirus. Des réactions d'amplification utilisant ces amorces, en combinaison avec des amorces spécifiques, permirent d'amplifier des fragments de 1.2 kpb et de 3.3 kpb. Ces fragments furent clonés dans le vecteur pUC9 par adjonction enzymatique d'une queue de résidus cytosines (méthode d'appariement « dC-dG »). La séquence de ces clones démontra qu'ils provenaient bien de l'ARN de TuMV, et environ 4.4 kb de génome viral fut ainsi cloné et séquencé.

2.3.2 Contribution.

J'ai réalisé toutes les expériences décrites dans cet article, à l'exception du séquençage enzymatique de l'ADN, qui a été fait en collaboration avec le service de séquence de l'Institut Armand-Frappier.

2.3.3 Conclusion.

Nous avons donc amplifié et cloné deux régions très longues de l'ARN du TuMV qui se chevauchaient sur environ 100 bases, ce qui était une étape nécessaire à la réalisation de l'objectif principal de ce travail de doctorat. Plus de 4 400 nt du génome de TuMV ont ainsi été obtenus par le séquençage ultérieur de ces deux régions. Ceci nous

a permis d'obtenir, avec celle publiée auparavant (Article 1), la séquence de plus de 5 000 bases du TuMV, ce qui représentait environ la moitié 3'-terminale de son génome.

Afin d'atteindre l'objectif principal de ce travail de doctorat, qui était la séquence complète du TuMV, il restait donc à amplifier et à cloner la première moitié de son ARN, et à amplifier également sa région 3' pour en vérifier la séquence, puisque celle provenant de la banque initiale nous paraissait anormalement longue. Ces travaux ont été réalisés et ont été publiés (Article 3).

2.4 ARTICLE 3:

NICOLAS, O. et J.-F. LALIBERTÉ

**The complete nucleotide sequence of
turnip mosaic potyvirus RNA.**

Journal of General Virology (1992), 73: 2785-2793.

2.4.1 Résumé.

Le génome complet à ARN du virus TuMV a été amplifié par sept amplifications PCR successives à partir de l'ADNc sb. Les produits de PCR ont été clonés dans les vecteurs de la série pUC9. L'assemblage des différentes séquences nous a permis d'obtenir les données suivantes: l'ARN du TuMV a une longueur de 9 830 nt, sans compter la queue polyadénylée; il contient une séquence de 9 489 bases qui code pour une longue polyprotéine de 3 163 acides aminés d'un PM calculé de 358K. La région 5'-NC, qui précède le premier codon d'initiation, comporte 129 nt et a une forte teneur (70 %) en adénine et uracile. Sa structure secondaire théorique est caractérisée par une boucle en épingle à cheveux, qui a une variation d'énergie libre estimée de -69.9 kJ/mol. Le premier codon de terminaison est suivi d'une région 3'-NC de 209 bases riche en A et U, et d'une queue polyadénylée. Sept sites potentiels de clivage par la protéase NIa ont également été localisés sur la polyprotéine du TuMV. Leurs séquences de sept acides aminés sont en accord avec celles des autres potyvirus, excepté pour le clivage 6K2-NIa, qui est caractérisé par la présence de l'acide glutamique (Glu) au lieu de la glutamine (Gln) en position -1 du site peptidique hydrolysé. L'analyse des différentes protéines potentielles du TuMV a montré qu'elles étaient très similaires à celles correspondantes des autres potyvirus connus.

2.4.2 Contribution.

J'ai réalisé toutes les expériences incluses dans cet article, à l'exception du séquençage enzymatique des clones obtenus, qui a été fait en collaboration avec le service de séquence de l'Institut Armand-Frappier.

2.4.4 Conclusion.

Le premier objectif prévu au début de ce travail de doctorat était le clonage et le séquençage complet de l'ARN d'une souche canadienne de TuMV. Cet objectif a été atteint assez rapidement grâce à l'utilisation de l'amplification par PCR, qui était en 1989 une nouvelle technique fort prometteuse. Il y a peu de différence au niveau de la séquence protéique de TuMV comparée à celle des autres potyvirus connus, mais cela nous a permis d'amorcer l'étude biologique de ses protéines. TuMV a la deuxième plus longue séquence publiée de potyvirus et, nous avons trouvé sur sa polyprotéine de 3 163 acides aminés, deux sites de clivage NIa légèrement, mais significativement, différents de ceux qui étaient connus. En effet, un premier site de clivage potentiel par la protéase NIa a été localisé, par analogie aux autres potyvirus, à sa propre extrémité N-terminale; ce site était assez similaire aux autres sites potentiels NIa du TuMV, excepté qu'en position -1 du site de clivage se trouvait un résidu Glu plutôt que Gln. Ce résultat était observé pour la première fois chez les potyvirus, mais se retrouva plus tard chez le PSbMV (Johansen *et al.*, 1991). Ceci nous a permis de trouver un site potentiel de clivage du même type, qui n'avait pas encore été mentionné dans la littérature, à l'intérieur même de la protéine NIa. Nous avons donc émis l'hypothèse selon laquelle ce site interne de clivage à la NIa pouvait être effectivement clivé par la NIa. La démonstration de cette hypothèse a fait l'objet d'une étude dont les résultats ont été publiés (Article 4).

2.5 ARTICLE 4:

**LALIBERTÉ, J.-F., O. NICOLAS, H. CHATEL,
C. LAZURE et R. MOROSOLI**

Release of a 22-kDa protein derived from the amino-terminal domain of the 49-kDa NIa of turnip mosaic potyvirus in Escherichia coli.

Virology (1992), 190: 510-514.

2.5.1 Résumé.

Les régions du génome de TuMV qui codent pour la protéine NIa ainsi que pour son précurseur 6K2-NIa ont été introduites dans le vecteur pET-11d afin d'obtenir une expression élevée chez *E. coli*. L'analyse des protéines produites chez *E. coli*, sur gel dénaturant de protéines (SDS-PAGE) suivi de leur immunodétection après transfert sur membrane de nitrocellulose, a montré que la protéine NIa était soumise à un clivage interne qui libérait un polypeptide de 22 kDa. Le séquençage de l'extrémité N-terminale du peptide de 22 kDa a été obtenue, et il correspondait à l'extrémité aminée de la protéine NIa. La mutagenèse dirigée a permis de confirmer que l'hydrolyse était bien due à l'activité protéolytique de la NIa, et que le clivage interne se faisait après un acide glutamique. Cet acide aminé est présent au sein d'un heptapeptide de NIa qui correspond bien à la séquence consensus reconnue par la NIa chez le TuMV. La fusion de la protéine 6K2 avec la NIa inhibait partiellement le clivage protéolytique interne à NIa. Ces résultats nous ont suggéré que la protéine de 22 kDa était fort probablement la protéine VPg du TuMV.

2.5.2 Contribution.

J'ai contribué principalement à la formulation des hypothèses qui ont conduit à l'élaboration de cet article plutôt qu'à la partie technique, et ce, à partir de l'analyse de la séquence protéique publiée du TuMV (Articles 3). En effet, le site de clivage 6K2-NIa trouvé sur la polyprotéine du TuMV, par analogie aux autres potyvirus connus, était différent du fait qu'en position -1 de la liaison peptidique hydrolysée se trouvait un acide glutamique plutôt qu'une glutamine; c'était la première fois qu'un tel site était trouvé à

cette position. En cherchant d'autres sites du même type, nous avons localisé un site potentiel de clivage par la NIa à l'intérieur même de cette protéase. Ce résultat était assez surprenant car il n'avait été mentionné pour aucun autre potyvirus, même s'il était déjà connu que la NIa comportait deux domaines d'activité (Murphy *et al.*, 1990). Nous devions donc démontrer que ce site de clivage interne à la NIa pouvait effectivement être clivé, et par la même occasion, montrer pour la première fois l'activité biologique d'une protéine clonée du TuMV.

2.5.3 Conclusion.

Cet article a donc permis de montrer, par expression de précurseurs protéiques chez *E. coli*, que notre hypothèse de la présence d'un site de clivage interne à la NIa était exacte. Nous avons également montré que le site de clivage 6K2-NIa était clivé par la NIa chez *E. coli*. Cet article est la première démonstration de l'activité biologique d'une des protéines clonées du TuMV (protéase NIa). Ces résultats constituent une des étapes conduisant à la compréhension des mécanismes de multiplication de ce virus; ils indiquent également que les études concernant le TuMV peuvent apporter de nouvelles informations sur la biologie moléculaire des potyvirus.

3. OBTENTION DE TRANSCRITS DE TUMV.

3.1 INTRODUCTION.

La deuxième partie de ce travail de doctorat ne pouvait être entreprise qu'à la condition que le génome du TuMV ait été séquencé au complet, ce qui a été réalisé et

été décrit ci-dessus (Article 3). L'objectif consistait à construire un clone d'ADN représentant la séquence complète du génome de TuMV, puis à transcrire ce clone en ARN; il fallait également tester le pouvoir infectieux des transcrits, c'est-à-dire examiner leur capacité à induire une infection du même type que celle obtenue par l'ARN viral natif. Les étapes permettant la réalisation d'un tel objectif étaient les suivantes:

- Assembler différents clones d'ADN en conservant sur l'un des deux brins la séquence complète du TuMV, suivie d'une queue polyadénylée, ceci avec le moins de séquence « non-virale » possible.
- Transcrire le clone complet en ARN, en ajoutant un analogue de coiffe méthylée, puis vérifier la nature des transcrits.
- Tester les transcrits synthétiques sur une espèce sensible au TuMV, suivi de l'observation de l'apparition des symptômes.

Dans le but de transcrire l'ADN en ARN, nous avons introduit un promoteur d'origine phagique en 5' de la séquence virale, ainsi qu'un site de restriction unique en 3'. Les transcrits synthétiques devaient également comporter en 5', un analogue de coiffe méthylée, ceci afin d'augmenter les chances de succès de la démonstration de leur caractère infectieux, comme cela a été indiqué pour d'autres potyvirus (Domier *et al.*, 1989; Gal-On *et al.*, 1991; Riechmann *et al.*, 1990).

Les clones obtenus pour la construction d'un clone complet de TuMV sont nommés par trois lettres suivies d'un nombre: les trois lettres représentent le vecteur de clonage (pTUM pour pUC9, pUC18 ou pUC19; pKS pour pBluescript^{MD} I SK; pSK pour

pBluescript^{MD} I SK), alors que le nombre représente la taille approximative du fragment d'ADN porté par le vecteur recombinant. Toutes les amorces (Annexe 1) dont les deux premières lettres commencent par JF ou PR ont une séquence homologue à l'ARN viral; les amorces dont le nom commence par FT ou ON (sauf FT_{oxk} et FT_o) sont complémentaires à la séquence virale. La position des sites de restriction et des amorces est basée sur la numérotation de la séquence du TuMV (Article 3).

3.2 ASSEMBLAGE D'UN CLONE COMPLET DE TUMV.

Les sites de restriction (*Bam*HI, *Kpn*I, *Sal*I et *Sma*I) sont absents de la séquence ADN du TuMV, mais ils font partie des sites multiples de clonages (SMC) des vecteurs de clonage les plus utilisés; ils ont de ce fait servi à l'assemblage des différents clones, par leur introduction dans les amorces servant à l'amplification de l'ARN viral. La construction du génome complet du TuMV s'est faite en trois étapes:

- Assemblage des 4 900 premiers nucléotides de la région 5' du génome de TuMV, avec l'introduction d'un promoteur du phage T7, menant au clone pKS4900.
- Assemblage des 5 000 derniers nucléotides du TuMV avec une queue polyadénylée suivie de deux sites de restriction uniques, menant au clone pTUM5000.
- Assemblage complet de l'ADNc de TuMV conduisant au clone pTUM9900.

Toutes les étapes de clonage ont été vérifiées par l'élaboration des cartes de restriction appropriées, ainsi que par le séquençage des régions de jonctions et des amorces. Toutes les réactions d'amplification par PCR ont été réalisées à partir

d'ADNc sb de TuMV, amorcé avec les amorces aléatoires d(N)₆ ou l'oligo (dT)₁₂₋₁₈, et purifié tel que décrit dans le chapitre « matériel et méthodes ». La synthèse des premiers brins d'ADNc a été vérifiée dans certains cas, après marquage radioactif, par électrophorèse sur gel alcalin d'agarose. La recherche des clones recombinants a été faite par hybridation *in situ*, soit à l'aide d'une sonde préparée à partir du fragment à cloner, soit à l'aide d'une sonde simple brin préparée à partir d'ARN viral.

3.2.1 Assemblage de la région 5' de TuMV.

Les différentes étapes du clonage de la région 5' de TuMV sont représentées sur la Figure 10. L'amorce JF_{SalI7} possède la séquence du site *SalI*, celle du promoteur T7 et des 21 premières bases du TuMV (Annexe 1). L'amorce FT_{S2} est complémentaire à l'ARN du TuMV au niveau d'un site *EcoRV* (position 2 500). L'amplification par PCR de l'ADNc du TuMV, avec les amorces JF_{SalI7} et FT_{S2}, a permis d'obtenir un fragment d'environ 2 500 pb, qui digéré par *SalI* et *EcoRV* puis cloné par ligation dans les deux sites identiques du vecteur pKS, a conduit au clone pKS2500 (Figure 10). La séquence du promoteur T7 ainsi que des premières bases de TuMV ont été confirmées par le séquençage des extrémités du clone pKS2500 (Figure 11).

Un site *BamHI* artificiel a été introduit dans la séquence en ADN de TuMV, en modifiant un seul nucléotide en position 4 863 (un T modifié en C) sans que cela ne modifie l'acide aminé codé par le codon impliqué; ce changement a été obtenu grâce à une mutation dans les amorces JF_{S3} et FT_{S0} (Annexe 1). L'introduction d'un tel site unique de restriction *BamHI* vers le milieu du génome de TuMV (position 4 859), a permis l'assemblage final de deux clones d'environ 5 000 pb (pKS4900 et pTUM5000).

FIGURE 10: Assemblage de la région 5' du TuMV conduisant au clone pKS4900. Les principaux sites de restriction ainsi que le nom des vecteurs sont indiqués. (pKS = pBluescript^{MD} I KS). Les rectangles hachurés représentent la séquence ADN du TuMV, alors que les traits pleins représentent les vecteurs. Les amorces servant aux amplifications sont marquées entre parenthèses après « PCR ». La taille des fragments d'ADN amplifiés par PCR sont indiquées en paires de bases (pb). Les flèches représentent les étapes de ligation.

PCR (FT₅₂ - JF_{SaIT7})
=2500 pb

PCR (FT₅₀ - JF₆₁)
=2400 pb

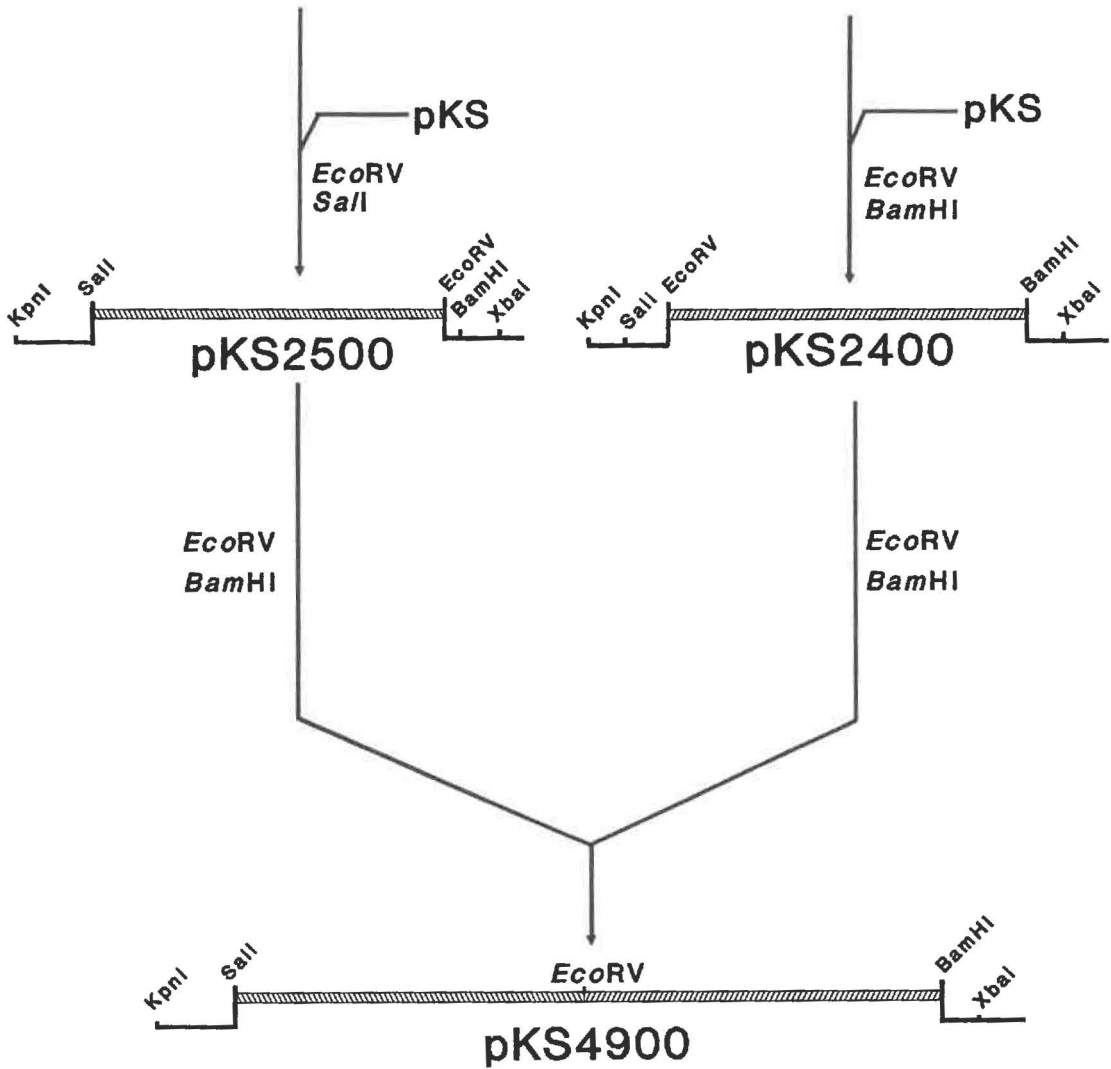


FIGURE 11: Séquence du clone pKS2500 située vers l'amorce JF_{SalI}T7. L'oligonucléotide FT₃₀ (Annexe 1) a servi d'amorce pour le séquençage enzymatique du clone pKS2500. L'autoradiographie du gel de séquence a été tournée de 180 degrés afin de lire la séquence réelle du TuMV. La position du site de restriction *SalI*, du promoteur de transcription T7 ainsi que des premiers nucléotides du TuMV est indiquée par des flèches.

G A T C



Séquence de TuMV

Premier nucléotide transcrit

Premier nucléotide de TuMV

Promoteur de la T7 ARN polymérase

Site SalI

Les amorces FT₃₀ et JF₃₁ ont permis d'amplifier un fragment d'environ 2 400 pb, qui après digestion par *EcoRV* et *BamHI* puis ligation au vecteur pKS digéré par ces enzymes, a conduit à l'obtention du clone pKS2400 (Figure 10). Ce clone a été séquencé à ses extrémités pour confirmer qu'il n'y avait pas eu d'erreur introduite par la *Taq* polymérase ni dans la synthèse des amorces. La dernière étape de l'assemblage de la première moitié du génome de TuMV a consisté à purifier le fragment de 2 400 pb, provenant de la digestion du clone pKS2400 par *EcoRV* et *BamHI*, puis à cloner cet ADN dans le plasmide pKS2500 digéré par les mêmes enzymes. Ceci a conduit, après ligation et transformation, au clone pKS4900 (Figure 10). La carte de restriction de pKS4900 a permis de confirmer que tous les sites de restriction utilisés étaient conservés.

3.2.2 Assemblage de la région 3' de TuMV.

Les différentes étapes du clonage de la région 3' de TuMV sont représentées sur les Figures 12 et 13. L'amorce FT_{0xx} (Annexe 1), qui comportait de 5' vers 3', un site *KpnI* et un site *SmaI* suivi de 17 résidus dT, a été synthétisée. L'amplification de l'ADNc du TuMV (amorcé à l'oligo (dT)₁₂₋₁₈) par les amorces FT_{0xx} et JF₃, a conduit à l'obtention d'un fragment d'environ 1 900 pb; sa digestion par les enzymes *XbaI* (position 8 457) et *KpnI* (de l'amorce FT_{0xx}), a libéré un fragment de 1 400 pb, qui purifié et ligué au vecteur pUC19 (*XbaI* / *KpnI*), a conduit au clone pTUM1400 (Figure 12). Plusieurs clones pTUM1400 ont été séquencés, et l'un d'eux, qui comportait environ 55 résidus dA, a servi pour les autres étapes d'assemblage.

FIGURE 12: Assemblage de la région 3' du TuMV conduisant au clone pTUM4100. Les principaux sites de restriction ainsi que le vecteur pUC19 sont indiqués. Les rectangles hachurés représentent la séquence ADN du TuMV alors que les traits pleins représentent les vecteurs. Les amorces utilisées pour les amplifications sont marquées entre parenthèses après « PCR ». La taille des fragments d'ADN amplifiés sont indiquées en paires de bases (pb). Les flèches représentent les étapes de ligation.

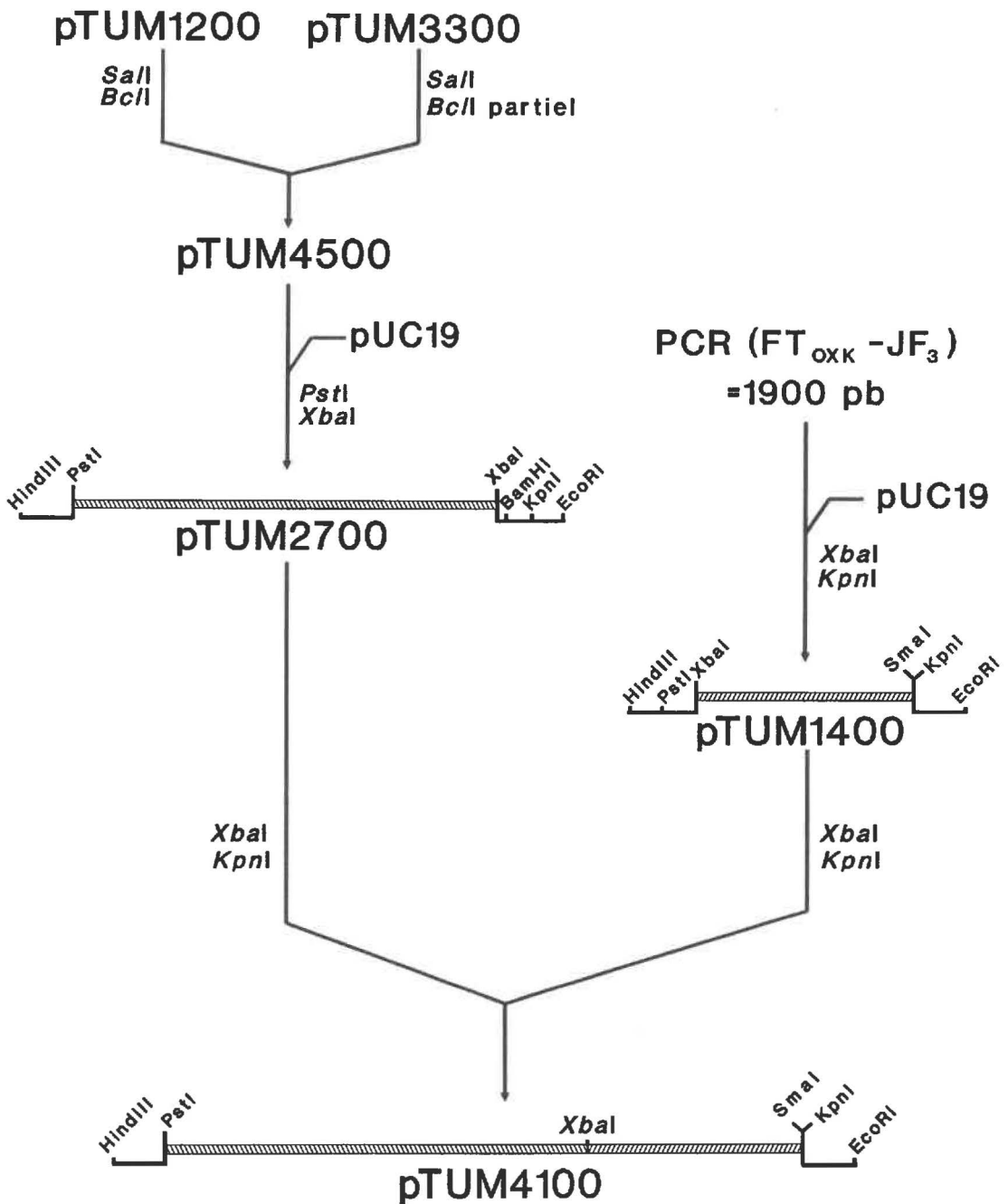
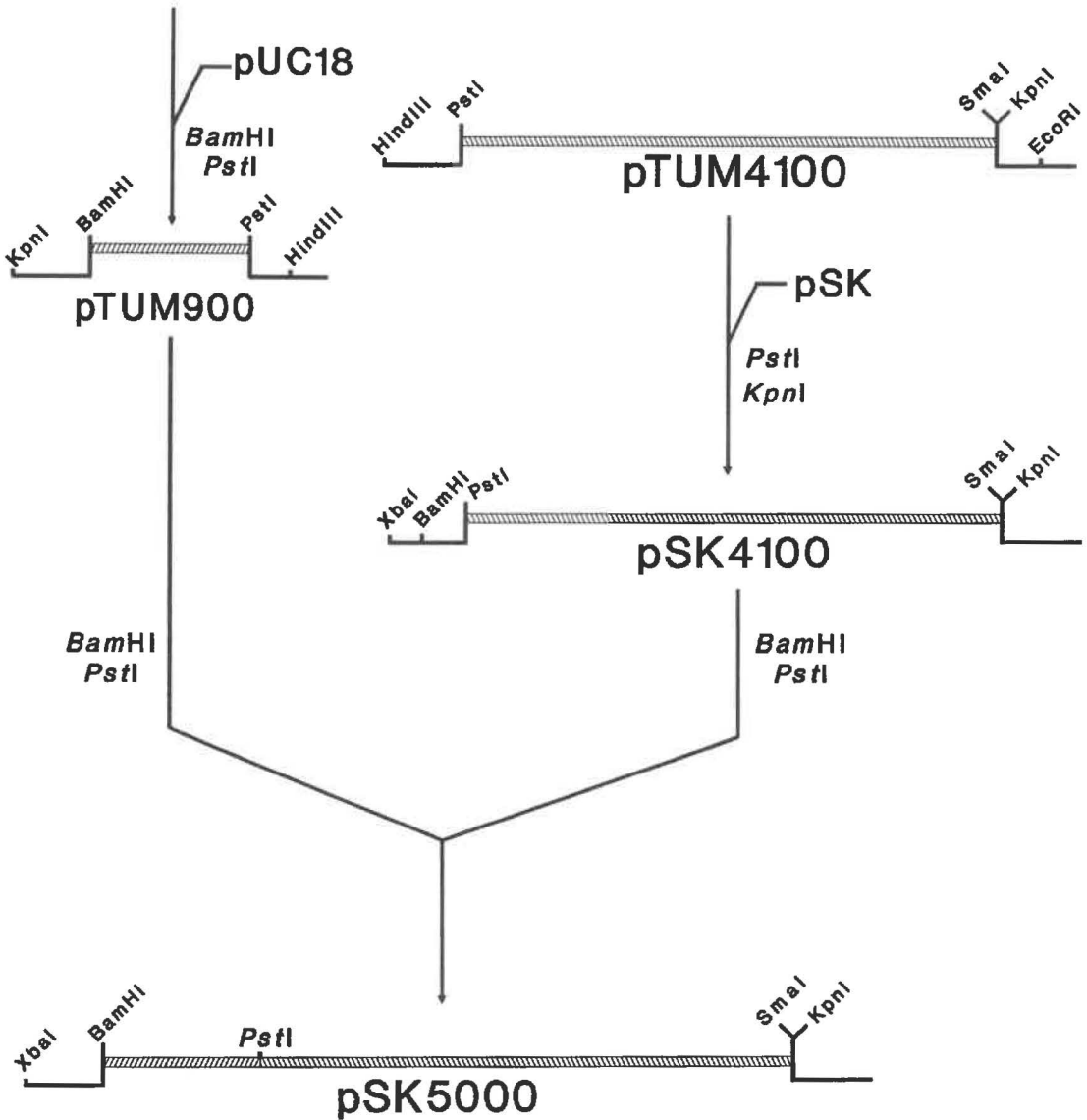


FIGURE 13: Assemblage de la région 3' du TuMV conduisant au clone pSK5000. Les principaux sites de restriction ainsi que le nom des vecteurs sont indiqués (pSK = pBluescript^{MD} I SK). Les rectangles hachurés représentent la séquence ADN du TuMV, alors que les traits pleins représentent les vecteurs. Les amorces utilisées pour les amplifications sont marquées entre parenthèses après « PCR ». La taille des fragments d'ADN amplifiés sont indiquées en paires de bases (pb). Les flèches représentent les étapes de ligation.

PCR (ON₉-JF₅₃)
=1150 pb



Les deux clones pTUM1200 et pTUM3300, dont l'obtention a été décrite précédemment (Article 2), ont été assemblés comme suit: le clone pTUM1200 a été digéré par *SalI* (site du vecteur pUC9) et par *BclI* (position 7 982 du TuMV), puis le vecteur a été purifié sur gel d'agarose; le clone pTUM3300 a été linéarisé par *SalI* (site du vecteur pUC9), puis partiellement digéré par *BclI*, du fait de la présence d'un deuxième site *BclI* en position 7 839 de TuMV. L'ADN d'environ 3 300 pb provenant de cette double digestion a été purifié sur gel d'agarose et cloné dans le vecteur pTUM1200 (*SalI* / *BclI*). Le clone issu de cette ligation a été appelé pTUM4500 (Figure 12), et la présence des deux sites de restriction *BclI* a été confirmée par l'analyse de restriction du plasmide pTUM4500 sur gel de polyacrylamide.

Le clone pTUM4500 a été digéré par *PstI* (position 5 768 du TuMV) et *XbaI* (position 8 457). Le fragment d'environ 2 700 pb issu de cette double digestion a été purifié et cloné dans le vecteur pUC19 (*PstI* / *XbaI*) pour conduire au clone pTUM2700 (Figure 12). Celui-ci a ensuite été digéré par *XbaI* (site de TuMV) et par *KpnI* (site de pUC19); la bande d'environ 5 400 pb, issue de cette double digestion, a été purifiée sur gel d'agarose et a servi de vecteur dans une réaction de ligation avec la bande de 1 400 pb (*XbaI* / *KpnI*) provenant du clone pTUM1400. Le clone résultant de cette ligation a été appelé pTUM4100 (Figure 12). La carte de restriction de ce clone a été effectuée afin de vérifier que tous les sites connus étaient bien conservés. Dans le but de récupérer un site de clonage *BamHI*, absent du clone pTUM4100, mais nécessaire à l'étape suivante, la bande de 4 100 pb (*PstI* / *KpnI*) de pTUM4100 a été liguée au vecteur pBluescript^{MD} I SK pour obtenir le clone pSK4100 (Figure 13).

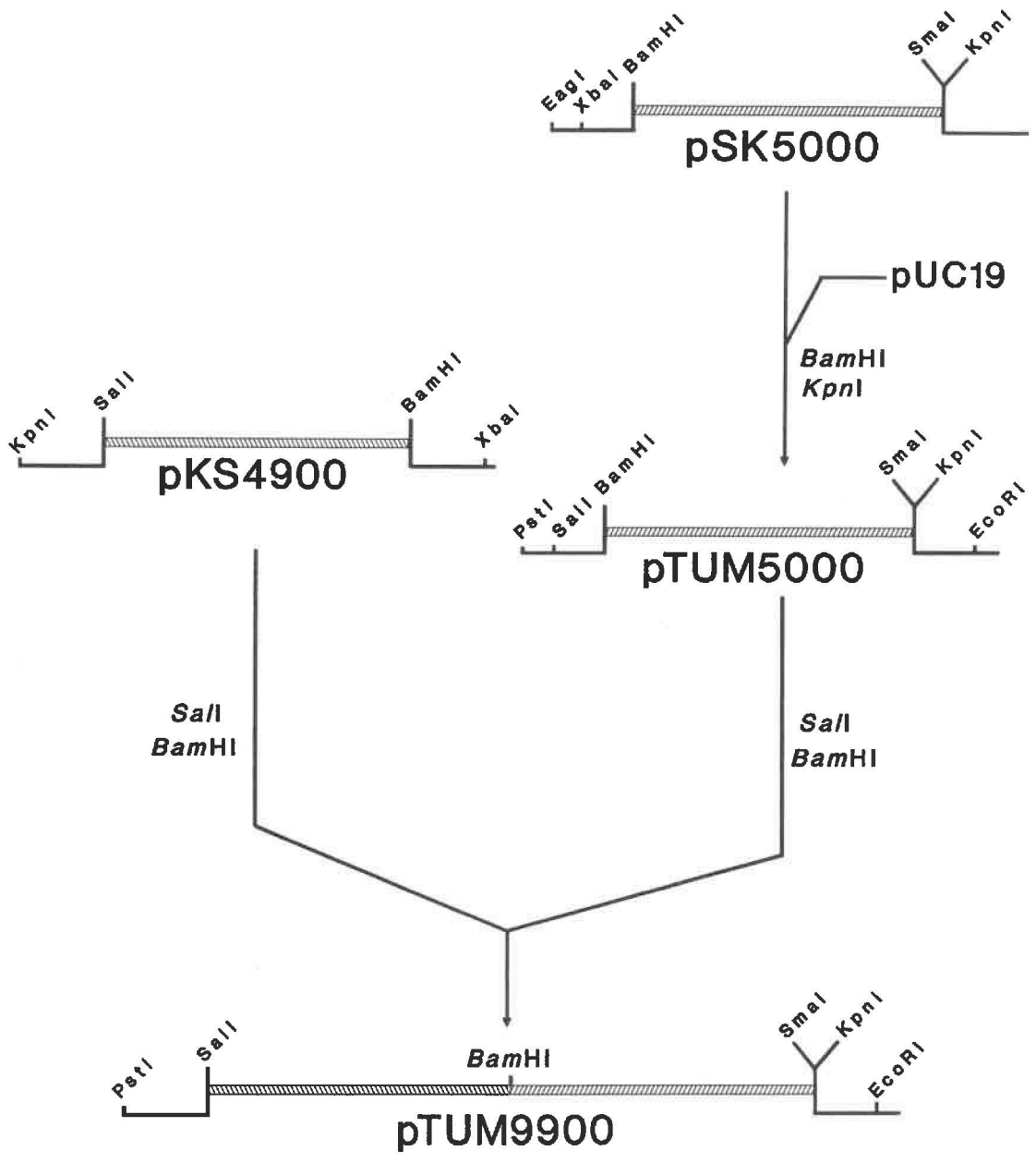
Les amorces ON₉, complémentaire à l'ARN viral, et JF₃₃, qui possède un site artificiel *Bam*HI, ont été utilisées pour amplifier un ADN d'environ 1 180 pb, qui digéré par *Pst*I et *Bam*HI a libéré un fragment d'environ 900 pb. Celui-ci a été cloné dans le vecteur pUC19 digéré par les mêmes enzymes de restriction, pour conduire à l'obtention du clone pTUM900 (Figure 13); celui-ci a été séquencé aux deux extrémités pour vérification.

Le vecteur pSK4100 a été digéré par *Bam*HI (site du vecteur) et par *Pst*I (site de TuMV), et l'ADN d'environ 7 100 pb, issu de cette double digestion, a été purifié sur gel d'agarose. Il a ensuite servi de vecteur pour la ligation de la bande de 900 pb (*Bam*HI / *Pst*I) du clone pTUM900 (Figure 13). Le clone pSK5000 résultant de cette ligation incluait donc toute la séquence du TuMV à partir du site artificiel *Bam*HI jusqu'à la queue polyadénylée suivie des sites *Sma*I et *Kpn*I.

3.2.3 Assemblage des régions 5' et 3'.

Les étapes de cet assemblage sont schématisées sur la Figure 14. Afin de récupérer un site de restriction *Sal*I nécessaire à l'assemblage final, l'ADN de 5 000 pb (*Bam*HI / *Kpn*I) du clone pSK5000 a été ligué au vecteur pUC19, résultant en un clone appelé pTUM5000 (Figure 14). La carte de restriction de pTUM5000 a permis de vérifier qu'aucune erreur de clonage n'avait été faite. L'ADN de 4 900 pb, provenant de la digestion de pKS4900 par *Sal*I et *Bam*HI, a été cloné dans le vecteur pTUM5000 digéré par les mêmes enzymes, pour conduire au clone complet pTUM9900 (Figure 14). La carte de restriction de pTUM9900 a confirmé que tous les sites de restriction étaient conservés.

FIGURE 14: Assemblage final de la séquence ADN du TuMV conduisant au clone pTUM9900. Les sites de restriction principaux ainsi que le nom des vecteurs sont indiqués. Les rectangles hachurés représentent la séquence ADN du TuMV alors que les traits pleins représentent les vecteurs. Les flèches représentent les étapes de ligation.



3.3 TRANSCRIPTION *IN VITRO*.

Le plasmide pTUM9900 a été digéré par *SmaI*, ce qui laissait donc sur le brin codant, une queue polyadénylée suivie de trois résidus cytosine (Figure 15). L'ADN linéarisé a ensuite été purifié et dilué à une concentration de $1 \mu\text{g}/\mu\text{l}$, puis été vérifié par électrophorèse sur gel d'agarose. Cet ADN a été transcrit *in vitro*, en présence de l'analogue de coiffe méthylée « $\text{m}^7\text{G}(5')\text{ppp}(5')\text{G}$ », par l'ARN polymérase du phage T7 (Figure 15). Dans le meilleur des cas, le rendement de la synthèse d'ARN coiffé de 9 900 bases fut de $20 \mu\text{g}$ pour $10 \mu\text{g}$ de plasmide linéarisé. Il a donc fallu procéder à plusieurs réactions de synthèse pour pouvoir effectuer un essai d'inoculation mécanique avec les transcrits. Afin de démontrer que les ARN synthétiques provenaient bien de la transcription du clone complet pTUM9900, celui-ci a été linéarisé par trois enzymes (*SmaI*, *BamHI* et *EcoRV*), puis les produits de la transcription de ces plasmides ont été analysés par électrophorèse sur gel dénaturant d'agarose (Figure 16). On peut estimer la taille des trois produits de transcription à respectivement, 9 900 bases (puits 2), 4 900 bases (puits 3) et 2 500 bases (puits 4), ce qui correspond aux résultats attendus. La transcription par l'ARN polymérase T7 commence toujours par une guanine, résidu qui a été introduit dans l'amorce JF_{saIT7} entre le promoteur de transcription et la séquence du TuMV (Annexe 1). Afin de confirmer que les transcrits synthétiques commençaient bien par la séquence attendue, l'extrémité 5' des transcrits « non-coiffés » a été clonée tel que décrit précédemment (Article 3). Le séquençage des clones obtenus a effectivement montré que les transcrits commençaient par un résidu G, immédiatement suivi de la séquence du TuMV.

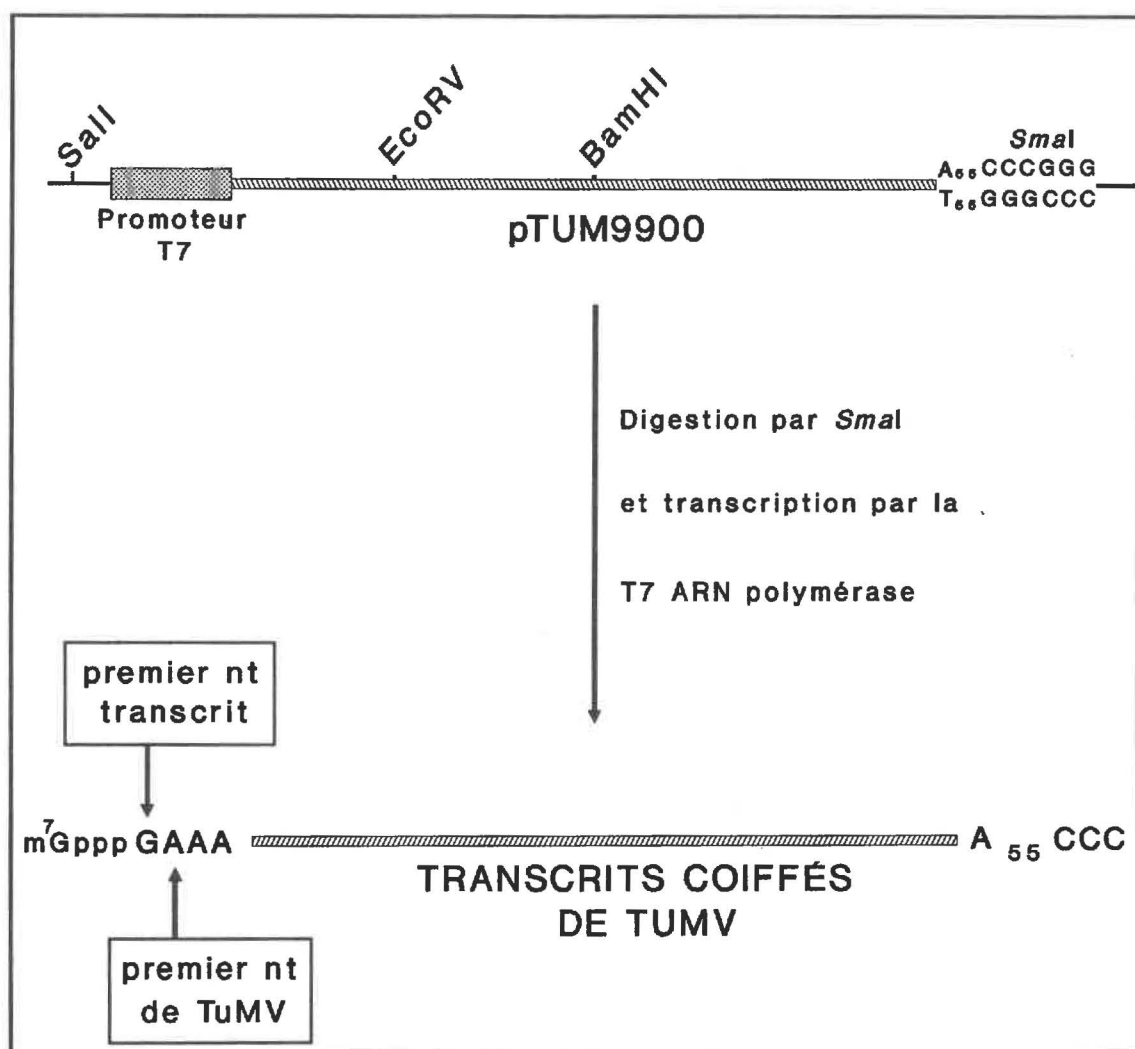


FIGURE 15: Représentation schématique de l'obtention des transcrits synthétiques du TuMV. Les principaux sites de restriction sont indiqués. La séquence théorique aux extrémités des ARN synthétiques est indiquée.

FIGURE 16: Analyse des transcrits synthétiques par électrophorèse sur gel dénaturant d'agarose. La taille des ARN du marqueur de poids moléculaire (Puits M) est indiquée sur la gauche. ARN purifié (1 μg) du TuMV (Puits 1). ARN coiffé (1 μg) issu de la transcription *in vitro* du clone pTUM9900 linéarisé par *Sma*I (Puits 2), par *Bam*HI (Puits 3) et par *Eco*RV (Puits 4).

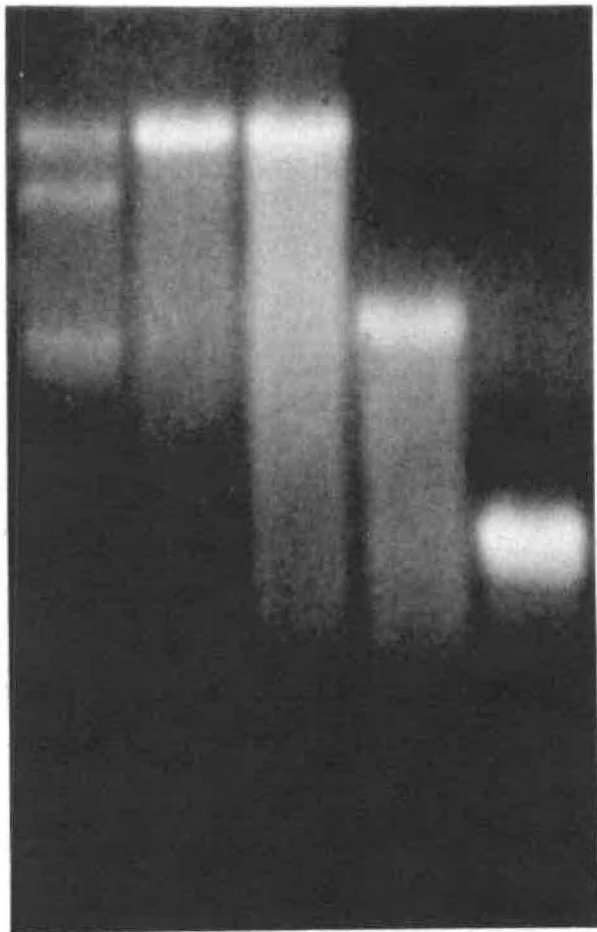
kb M 1 2 3 4

9.5-

7.5-

4.4-

2.4-



3.4 ÉVALUATION DES TRANSCRITS SYNTHÉTIQUES.

Les transcrits ont été testés par inoculation mécanique de deux espèces sensibles au TuMV qui donnent des symptômes de mosaïques très marqués; il s'agit du navet (*Brassica campestris* ssp. *rapifera* cv. « Just Right ») et de la moutarde tendre (*Brassica perviridis* var. *tendergreen*). Le protocole utilisé pour l'infection est décrit dans le chapitre « matériel et méthodes ».

Un mois après l'inoculation, aucune des 50 plantes des deux espèces infectées avec environ 2 μ g d'ARN synthétique coiffé n'a développé de symptômes, alors que les 10 plantes infectées avec 100 ng de virus purifié avaient des symptômes de mosaïques après seulement dix jours. Il faut cependant noter qu'aucune des 20 plantes infectées avec 500 ng d'ARN purifié de TuMV n'a développé de symptômes caractéristiques d'une infection au TuMV, même un mois post-infection.

DISCUSSION

1. CARACTÉRISATION DU VIRUS DE LA MOSAÏQUE DU NAVET.

1.1 CLONAGE ET SÉQUENÇAGE.

Une souche du virus de la mosaïque du navet (TuMV) a été purifiée et son ARN a été isolé. Sa longueur ainsi que son intégrité ont été vérifiées par électrophorèse sur gel d'agarose.

Notre objectif initial, pour débiter la caractérisation moléculaire du TuMV, était de cloner et de séquencer la protéine de sa capsid (CP) (Article 1). La maîtrise de l'amplification par polymérisation en chaîne (PCR) a été primordiale pour atteindre cet objectif, car les clones issus de la banque initiale d'ADN complémentaire (ADNc) ne recouvraient pas complètement la séquence codante de CP. De plus, c'est la première fois que le PCR servait à amplifier le génome d'un potyvirus, méthode qui est maintenant fréquemment utilisée dans la plupart des laboratoires concernés par ces virus.

Après avoir amplifié par PCR deux régions de 400 pb et de 700 pb de l'ARN du TuMV (Article 1), nous avons mis au point une méthode efficace pour amplifier et cloner de grandes régions de ce virus, ceci dans le but de le cloner au complet (Article 2). Les premières amorces utilisées pour les amplifications ne possédaient pas de site de restriction, ce qui nécessitait de cloner les fragments d'amplification dans le vecteur pUC9 par la méthode d'appariement « dC-dG ». L'avantage de cette méthode est que les fragments d'ADN pouvaient être clonés au complet, sans avoir besoin de traiter les sites de restriction internes par des méthylases. Par contre, son désavantage résidait dans le fait qu'il était pratiquement impossible de séquencer les insertions en utilisant les amorces

universelles du vecteur pUC9; ce phénomène est probablement dû à la forte stabilité des fragments riches en « G-C » présents entre l'insertion et le vecteur, ce qui peut empêcher le passage des polymérases utilisées pour le séquençage. Dans le but d'éliminer ces séries de « G-C », il fallait donc sous-cloner les insertions en utilisant leurs sites de restriction internes. Les amorces FT₈ et JF₉ (Annexe 1) ont été les premières comportant la séquence d'un site de restriction, ce qui a permis le clonage du fragment de 3 300 pb, amplifié grâce aux amorces FT₈ et JF₉, par ligation dans les vecteurs M13mp18 et M13mp19, et son séquençage par les amorces des vecteurs. Par la suite, toutes les amorces dégénérées et spécifiques comportaient la séquence d'au moins un site de restriction vers leur extrémité 5'; les produits d'amplification étaient clonés directement par ligation après purification sur gel d'agarose.

Il est probable que les amorces dégénérées dessinées pour l'amplification du TuMV puissent être utilisées pour le clonage d'autres potyvirus, car elles correspondent à des régions très conservées sur la polyprotéine. De plus, les conditions d'amplification que nous avons utilisées autorisent un appariement imparfait des amorces dégénérées; en effet, l'amorce JF₁₃ (Annexe 1) a été utilisée avec succès pour l'amplification d'un fragment de 1 100 pb, alors qu'elle présente une homologie imparfaite avec la séquence réelle du TuMV; elle présente dans sa région 3' quatre nucléotides en ligne [G-T-(AT)-C] qui ne sont pas homologues à la séquence réelle du TuMV qui est [C-A-C-A]. Le fait de procéder à deux cycles d'amplification avec une température d'appariement de 37°C donne certainement la possibilité aux amorces dégénérées de se fixer sur la matrice ADN, même si certains nucléotides ne sont pas complémentaires.

L'introduction de sites de restriction dans les amorces permet également de cloner directement les produits d'amplification dans certains vecteurs d'expression, tels que ceux de la série pET utilisés pour l'expression d'une protéine chez *E. coli* (Article 4).

L'intérêt de travailler avec du virus purifié, comme nous avons procédé pour le TuMV, est que l'on diminue considérablement la possibilité d'amplifier de façon non-spécifique du matériel « non-viral ». Au cours d'une expérience préliminaire, une région de 700 pb du TuMV a été amplifiée par des amorces spécifiques à partir d'ARN total provenant de feuilles d'une plante infectée au TuMV; le produit d'amplification de 700 pb a été obtenu avec quelques bandes non-spécifiques, ce qui indique qu'il est également possible d'amplifier un potyvirus à partir d'ARN total d'une plante infectée. D'ailleurs, après publication de notre travail, certaines équipes ont mis au point des tests très sensibles de détection par PCR de potyvirus, avec des amorces dégénérées et/ou spécifiques (Kohnen *et al.*, 1992; Langeveld *et al.*, 1991; Wetzel *et al.*, 1991b).

Sept clones de 130 pb à 3 300 pb, issus de réactions distinctes d'amplification, ont permis de recouvrir et de séquencer la totalité du génome du TuMV (Article 3) qui est à ma connaissance, le premier virus de plante dont la séquence complète a été obtenue grâce au PCR. C'est donc une méthode alternative très avantageuse par rapport à la construction classique d'une banque d'ADNc, et ceci pour les raisons suivantes:

- La quantité d'ARN viral nécessaire à l'amplification est très faible, puisque environ 1 μ g d'ARN purifié du TuMV a été suffisant pour procéder au clonage complet de son génome.

- Les produits d'amplification peuvent être visualisés par électrophorèse sur gel d'agarose et manipulés facilement; ainsi, nous avons obtenu de 100 ng à 1 000 ng d'ADN pour chacune des réactions d'amplification.
- Le nombre de clones nécessaires pour recouvrir le génome viral au complet est limité, et sept clones PCR furent suffisants pour cloner entièrement celui du TuMV.
- La taille approximative des fragments à amplifier, ainsi que leur localisation sur le génome, est connue, ce qui permet un criblage facile des clones recombinants par cartographie de restriction.

Nous avons également utilisé le PCR pour amplifier la région 5' de l'ARN viral; ceci fut avantageux puisque les autres équipes ont dû séquencer directement l'ARN, n'ayant pas de clone représentant l'extrémité 5'-terminale. Comme une protéine VPg est liée de façon covalente à l'ARN viral et que cette protéine est digérée à la protéinase K, il est difficile de préciser la nature exacte du premier nucléotide des potyvirus (Allison *et al.*, 1986; Domier *et al.*, 1986; Laín *et al.*, 1989a; Maiss *et al.*, 1989; Robaglia *et al.*, 1989). Cependant, Turpen (1989) a montré, par le séquençage direct de l'extrémité 5' de l'ARN du PVY, que sa séquence était identique, avec ou sans la VPg, ce qui indique que la séquence du TuMV que nous avons obtenue est probablement complète (Article 3) puisque la digestion et l'extraction de la VPg n'élimine pas de nucléotides en 5'. Néanmoins, il semble y avoir une population minoritaire de virus dont l'ARN commence par six adénines au lieu de cinq (Article 3, Fig. 3), ce qui pourrait être la conséquence du mode de réplication impliquant la VPg comme amorce à la synthèse d'ARN, modèle proposé par Tobin *et al.* (1989).

Le problème majeur de l'amplification par PCR est le taux élevé d'erreur de l'ADN polymérase de *Thermus aquaticus* (*Taq* polymérase), enzyme qui n'a pas d'activité exonucléase 3' vers 5' lui permettant de corriger les erreurs en cours de polymérisation (Tindall et Kunkel, 1988). Plusieurs études ont montré que le taux d'erreur de la *Taq* polymérase est d'environ 2.0×10^{-4} par base et par cycle de réplication, et que la grande majorité des mutations qu'elle introduit intervient au niveau des résidus A et T, changés respectivement en G et C; de plus, la fidélité de l'enzyme dépend aussi des conditions d'amplification ainsi que de la matrice ADN (Keohavong et Thilly, 1989; Tindall et Kunkel, 1988).

Afin d'éliminer les risques d'erreurs de séquence inhérents à l'utilisation de la *Taq* polymérase, nous avons procédé au séquençage d'au moins trois clones indépendants provenant de chaque amplification. Cette stratégie nous a permis de lever certaines ambiguïtés obtenues par la séquence de seulement deux clones, et d'obtenir également la séquence « majoritaire » du TuMV (Article 3), c'est-à-dire la séquence la plus représentée dans la population d'ARN de départ. Nous avons noté une différence moyenne de trois nucléotides pour 1 000 bases entre plusieurs clones provenant de la même amplification. Par contre, la plupart des mutations survenaient sur le troisième nucléotide d'un codon, si bien que la séquence en acides aminés restait le plus souvent inchangée d'un clone à l'autre. Ainsi, la majorité des mutations observées entre plusieurs clones indépendants ne seraient pas dues à la *Taq* polymérase, mais plutôt à l'hétérogénéité de la population d'ARN dans les tissus infectés, conséquence de la grande fréquence d'erreur de l'ARN polymérase virale. Une alternative pour éviter d'avoir à séquencer plusieurs clones provenant de la même réaction PCR, serait de séquencer

directement les produits d'amplification, ce qui aurait pour effet de masquer l'hétérogénéité de la population d'ARN ainsi que les erreurs d'incorporation dues à la *Taq* polymérase (Bloch, 1991).

1.2 ANALYSE DE LA SÉQUENCE.

Après le séquençage de tous les clones provenant des amplifications par PCR, nous avons pu aligner une séquence totale de 9 830 nucléotides pour l'ARN du TuMV. Le tableau II indique sa composition en bases comparée à celle de la plupart des potyvirus dont la séquence est connue. La composition de l'ARN du TuMV est assez similaire à celle des autres potyvirus et il a une forte teneur en adénine, ce qui confirme les résultats expérimentaux de Hill et Shepherd (1972). Sa teneur en (A + U) (54.8 %) est légèrement inférieure à celle des autres potyvirus (56.6 à 58.8 %), et on ne pouvait pas exclure la possibilité que cette faible teneur soit une conséquence du clonage par PCR, ce qui confirmerait que la *Taq* polymérase modifie le plus souvent les résidus A et T en G et C (Keohavong et Thilly, 1989). Or, comme il est peu probable que la majorité des changements introduits par la *Taq* polymérase soient intervenus de façon aléatoire sur la troisième base des codons, la composition en bases du TuMV est sans doute naturellement moins riches en (A et U) que les autres potyvirus. La teneur en résidus (A + U) de l'ARN des potyvirus est légèrement supérieure à 50 % (Tableau II), et il est possible que ce soit une caractéristique importante permettant sa répllication ou son encapsidation. Cette teneur pourrait par exemple faciliter la « fonte » des structures secondaires de l'ARN au cours de la polymérisation par l'ARN polymérase virale, et aider à la séparation des formes répllicatives par les protéines ayant une activité hélicase.

TABLEAU II: Composition (en pourcentage) en bases des potyvirus. *turnip mosaic virus* (TuMV); *pea seed-borne mosaic virus* (PSbMV); *plum pox virus* (PPV); *potato Y virus* (PVY); *pepper mottle virus* (PeMoV); *soybean mosaic virus* (SMV); *tobacco vein mottling virus* (TVMV).

Virus\bases	adénine	guanine	uracile	cytosine
TuMV	32.0	24.3	22.8	20.9
PSbMV	32.0	23.5	26.0	18.5
PPV	31.3	23.0	25.3	20.4
PVY	31.1	23.4	26.8	18.7
PeMoV	31.7	23.1	27.1	18.1
SMV	32.0	24.0	26.0	18.0
TVMV	32.0	22.9	26.1	19.0

La Figure 17 résume les caractéristiques génomiques du TuMV et de tous les potyvirus séquencés à ce jour. Le TuMV a le deuxième plus long génome (9 830 bases) mais également la deuxième plus longue polyprotéine (3 163 acides aminés). En fait, la principale différence de longueur parmi les potyvirus se trouve localisée au niveau de la protéine P1; il y a par exemple, une différence de 107 acides aminés entre celle du TuMV et celle du TVMV, ce qui représente 90 % de la différence de longueur totale entre ces deux virus. Si on considère que la P1 est effectivement la protéine impliquée dans le transport du virus de cellule à cellule (« revue bibliographique », chapitre 4.6.9), on peut imaginer que le caractère évolutif d'un potyvirus soit en rapport avec la propriété de sa protéine P1 à être adaptée à de nombreux hôtes, ce qui pourrait se traduire par la grande variation de taille observée pour cette protéine.

TuMV a la plus courte séquence 5' non-codante (5'-NC) avec 129 bases (Figure 17). Cette région, ou son ARN complémentaire, comporte possiblement les signaux permettant aux enzymes impliquées dans la réplication de reconnaître spécifiquement l'ARN viral. De plus, il a été montré (communication de P. Dallaire, étudiant au laboratoire du Dr. J.-F. Laliberté) que cette séquence potentialise nettement la traduction *in vitro* et *in vivo* d'un gène indicateur (β -glucuronidase) qui y est fusionné, résultat qui confirme celui obtenu par Carrington et Freed (1990) pour le TEV. La séquence 5'-NC du TuMV est capable de former une boucle potentielle en épingle à cheveux qui a une variation d'énergie libre estimée de -69,9 kJoule/mole (Article 3, Fig. 5). Cette structure a probablement un rôle dans la réplication et/ou la traduction de l'ARN. Des expériences de délétions et de mutagénèse dirigée dans la région 5'-NC permettraient de confirmer son rôle exact.

Virus	Long. tot.	5'-NC	Région codante/ A.A.	3'-NC	PM (K)
PSbMV	9924	143	9618/ 3206	163	364
TuMV	9830	129	9489/ 3163	212	358
PPV	9786	146	9420/ 3140	220	355
PVY	9704	184	9189/ 3063	331	348
PeMoV	9640	167	9204/ 3068	269	348
SMV	9588	131	9198/ 3066	259	349
TEV	9496	144	9162/ 3054	190	346
TVMV	9471	205	9015/ 3005	251	340

FIGURE 17: Caractéristiques génomiques des potyvirus connus. De la gauche vers la droite sont indiqués: le nom du virus, la taille totale du génome en nucléotides (Long. tot.), la taille de la région 5' non-codante (5'-NC), la taille de la région codante avec le nombre d'acides aminés (A.A.) de la polyprotéine, la taille de la région 3' non-codante (3'-NC), et le poids moléculaire (PM) de la polyprotéine (K).

La séquence 3'-non-codante (3'-NC) de 667 nt qui provenait exclusivement des deux clones de la banque initiale d'ADNc (Article 1) était un résultat assez surprenant; en effet, la longueur moyenne pour la région 3'-NC des potyvirus est d'environ 240 nt avec un maximum de 475 nt (*Johnson grass mosaic virus*) et un minimum de 189 nt (TEV) (Ward et Shukla, 1991). De plus, la séquence de la région 3' de deux autres souches du TuMV provenant de Chine et du Japon, qui ont été publiées postérieurement à celle présentée dans cette thèse (Article 3), comportait une région 3'-NC de 212 nt, ce qui était en contradiction avec notre premier résultat (Ling-Jie *et al.*, 1990; Sano *et al.*, 1992). Afin de vérifier que cette séquence ne provenait pas d'un artefact de clonage de la banque d'ADNc, l'amplification de l'extrémité 3' du génome du TuMV a été réalisée suivant la procédure de Frohman *et al.* (1988), ce qui a conduit effectivement à corriger la séquence du TuMV (Article 3). Afin de comprendre d'où provenait la longue séquence initiale de 667 nt, des recherches d'homologies avec la banque de données « EMBL » (*European Molecular Biology Library*) ont été réalisées (Sano *et al.*, 1992); une homologie de 89.8 % a été trouvée entre les 467 dernières bases de cette séquence et le gène codant pour la protéine ribosomale S12 du chloroplaste du soja, alors que les 203 premières bases présentaient plus de 92 % d'homologie avec celles des souches asiatiques du TuMV (Lin-Jie *et al.*, 1990; Sano *et al.*, 1992). Il est intéressant de constater que la séquence que nous avons obtenue soit celle d'un gène chloroplastique et elle pourrait ne pas être la conséquence d'un artefact de clonage, mais plutôt le résultat d'une recombinaison entre le virus et l'ARN chloroplastique de l'espèce infectée; ce phénomène a déjà été rapporté une fois pour un lutéovirus (Mayo et Jolly, 1991). Si notre premier résultat avait été corroboré par d'autres clones du même type, nous aurions ainsi

confirmé que certaines étapes de la réplication des potyvirus se produisent dans les chloroplastes, hypothèse basée sur le fait que des protéines et ARN de potyvirus ont déjà été détectés dans ces organites (Gadh et Hari, 1986; Gunasinghe et Berger, 1991). Toutefois, l'ARN recombiné que nous avons séquencé doit être très faiblement représenté, car nous avons amplifié plusieurs fois l'extrémité 3' de l'ARN du TuMV par PCR sans jamais noter de différence significative de taille ni de séquence entre les produits d'amplification. La souche que nous avons séquencée a donc une région 3'-NC de 212 nt suivie d'une queue polyadénylée (Article 3).

Les trois souches de TuMV dont la séquence en 3' est connue (souche canadienne et deux souches asiatiques), ont une capsid de 288 acides aminés et une séquence 3'-NC de 212 nucléotides (Article 3, Ling-Jie *et al.*, 1990; Sano *et al.*, 1992). Le tableau III indique les pourcentages d'homologie entre ces différentes souches, en nucléotides et en acides aminés. On remarque que la séquence nucléotidique 3'-NC (Tableau IIIa) est nettement plus conservée que celle qui code pour la capsid (Tableau IIIb); ceci est une indication qu'il doit y avoir une pression de sélection très forte sur cette région 3'-NC et qu'elle doit avoir un rôle extrêmement important dans la multiplication virale. En effet, comme cette séquence correspond à une des extrémités de l'ARN et que la réplication commence théoriquement par celle-ci, il est probable qu'elle comporte des signaux (séquences et/ou structures secondaires) permettant l'initiation de la réplication. Bryan *et al.* (1992) ont ainsi trouvé dans la séquence 3'-NC de 44 potyvirus différents, une séquence conservée du type [5'-G(U/C)GAGGN-XXX-N'CCUC-3'] (X est une des quatre bases, N et N' sont des bases complémentaires) capable de former une boucle potentielle en épingle à cheveux; ce motif est retrouvé en position 9 783-9 797 de l'ARN du TuMV.

TABLEAU III: Pourcentage d'homologie en nucléotides et en acides aminés (entre parenthèse) au niveau de la région 3' non-codante (TABLEAU IIIa) et de la protéine de la capsid (TABLEAU IIIb) de trois souches de TuMV: Can. = souche du Canada; Jap. = souche du Japon; Chi. = souche de Chine.

Souches	TuMV-Can.	TuMV-Jap.
TuMV-Jap.	95.3	-
TuMV-Chi.	94.8	99.5

TABLEAU IIIa

Souches	TuMV-Can.	TuMV-Jap.
TuMV-Jap.	89.6 (95.1)	-
TuMV-Chi.	88.7 (94.5)	98.7 (98.6)

TABLEAU IIIb

Afin de démontrer le rôle des régions terminales des potyvirus, il faudrait analyser l'efficacité de réplication d'ARN recombinants contenant différentes mutations dans ces régions, après transfection de protoplastes.

D'après les travaux de Shukla et Ward (1988; 1989a), il apparaît que deux potyvirus distincts ont une homologie de 38 à 71 % au niveau de la capsidie alors que plusieurs souches du même virus ont une homologie supérieure à 90 %. Ce résultat est effectivement confirmé pour le TuMV, puisque la capsidie de la souche canadienne que nous avons séquencée présente plus de 94 % d'homologie avec celle de deux souches d'Extrême-Orient (Tableau IIIb) et beaucoup moins avec celle d'autres potyvirus (56.2 % avec le TVMV; 66.1 % avec le PVY; 70.4 % avec le PPV). Plus de 90 % des différences entre les trois souches de TuMV sont localisées dans la région N-terminale de la protéine, ce qui confirme que cette région est exposée à la surface de la particule virale et qu'elle est soumise à plus de variabilité. Ceci serait dû au fait qu'elle n'est pas indispensable à l'encapsidation et qu'elle a un rôle dans le transport par les pucerons (Allison *et al.*, 1985; Shukla *et al.*, 1988). A cet égard, comme la présence du tripeptide (Asp-Ala-Gly) à l'extrémité N-terminale de la protéine CP des potyvirus est critique pour la transmission par les pucerons (phénotype AT, *aphid transmissible*)(Atreya *et al.*, 1990; Atreya *et al.*, 1991), la souche de TuMV que nous avons séquencée ne semble pas avoir ce phénotype. En effet, le génotype de cette souche (Asp-Ala-Asp) est identique à celui d'un autre potyvirus de phénotype NAT (*non-aphid transmissible*), la souche N du *soybean mosaic virus* (SMV) (Eggenberger *et al.*, 1989). Les deux souches asiatiques de TuMV sont également différentes entre elles; celle de Chine (Asp-Ala-Gly) serait de phénotype AT, alors que celle du Japon (Asp-Thr-Gly) serait de phénotype NAT (Ling-Jie *et al.*, 1990;

Sano *et al.*, 1992). Il faudrait pouvoir procéder à un test d'infectivité de notre souche de TuMV par les pucerons pour confirmer notre résultat, mais il a déjà été rapporté un cas d'un isolat de TuMV (isolat 31) qui avait perdu son caractère transmissible par certains pucerons (Edwardson et Christie, 1986). Il est possible que le virus ait perdu son caractère transmissible par les pucerons du fait qu'il a été multiplié de nombreuses fois de façon artificielle en laboratoire, ce qui a donc éliminé la sélection positive pour le phénotype AT. Une seule mutation sur la deuxième base du codon de la glycine aurait suffi à changer cet acide aminé (phénotype AT) en acide aspartique (phénotype NAT).

Il paraît peu probable que les trois souches de TuMV soient apparues de façon indépendantes en des régions si éloignées, ni qu'elles se soient distinguées avant la dérive des continents. Une des explications de la propagation du TuMV dans différentes régions pourrait être qu'il a été transporté, volontairement ou involontairement, par des plantes infectées ou sous une forme purifiée, puisqu'il est, ni transmissible par les graines, ni stable dans les vecteurs qui le véhiculent.

La protéine de la capsid a servi jusqu'à présent à comparer les différents potyvirus afin d'établir la distance génétique entre eux. Il serait peut-être avantageux, lorsqu'elles seront disponibles, de comparer la séquence de la protéine P1 d'un grand nombre de potyvirus, car c'est la protéine la plus variable en longueur et en séquence. Ceci permettrait de comprendre les origines de cette protéine dont le rôle est vraisemblablement de véhiculer le virus de cellule à cellule.

La séquence de la polyprotéine du TuMV a été analysée, et tous les sites potentiels de clivage par les trois protéases virales ont été localisés (Article 3). Ceux des

protéases P1 et HC ont été trouvés d'après les travaux de Carrington et Herndon (1992) et de Mavankal et Rhoads (1991). Le site reconnu par la P1 se situerait après la Phe³⁶² et celui reconnu par HC, après la Gly⁸²⁰. Alors que la protéine HC du TuMV est constituée de 458 acides aminés, ce qui est très semblable aux autres potyvirus (« revue bibliographique », Figure 4), la protéine P1 comporte 363 acides aminés, ce qui est significativement plus élevé que les potyvirus connus (Figure 4). La signification d'une telle différence au niveau de la protéine P1 a été abordée au début de ce chapitre.

La protéine HC du TuMV a déjà été testée pour la transmission par les pucerons (Sako et Ogata, 1981); elle permet son transport, mais pas celui d'autres potyvirus, ce qui suppose qu'il y a des interactions spécifiques entre la particule virale et la protéine HC. On retrouve dans sa région N-terminale, un acide aminé très important (Lys⁴¹⁴) conservé pour tous les potyvirus dont la protéine HC est active pour le transport par les pucerons (Atreya *et al.*, 1992; Thornbury *et al.*, 1990). De plus, cet acide aminé est présent au sein même d'un motif de type « *zinc finger* » très riche en résidus cystéine (Cys³⁸⁹-X₈-Cys-X₇-Cys-X₅-Cys-X₄-Cys-X₂-Cys⁴²⁰) tel que proposé par Robaglia *et al.* (1989). Cette région serait impliquée dans le maintien de la structure de l'ARN, ce qui aurait une conséquence sur la régulation de la réplication. On retrouve dans la partie C-terminale de la protéine HC du TuMV, les acides aminés essentiels (Cys⁷⁰⁶ et His⁷⁷⁹) à son activité autoprotéolytique (Oh et Carrington, 1989) ainsi que ceux qui composent son site de clivage (Tyr⁸¹⁷-Arg-Val-Gly-Gly⁸²¹) (Carrington et Herndon, 1992).

Nous avons localisé sur la polyprotéine du TuMV, sept sites potentiels de clivage par la protéase NIa (Article 3, Fig. 6), dont le site NIb-CP a été le premier démontré

grâce au séquençage de l'extrémité aminée de la capside (Article 1). Deux des sites (jonctions 6K2-VPg et VPg-NIa) sont particuliers puisqu'ils présentent un résidu Glu en position -1 au lieu d'une Gln pour les cinq autres sites. Le TuMV et le PSbMV sont les seuls potyvirus à avoir un tel site en position 6K2-VPg, et malgré qu'ils soient les deux plus longs potyvirus séquencés, ils n'ont que 58 % d'homologie au niveau de la capside, et sont donc probablement peu apparentés.

Nous avons démontré pour la première fois chez *E. coli*, que le site interne à la NIa était reconnu et clivé de façon protéolytique en une protéine de 22 kDa, qui représente le domaine N-terminal de NIa (la VPg), et une autre protéine non-détectée de 27 kDa (la NIa-Pro), qui représente son domaine C-terminal; celle-ci est probablement dégradée par des enzymes bactériennes (Article 4). On retrouve dans la VPg du TuMV un motif très conservé, dont la Tyr¹⁹⁶ qui serait l'acide aminé lié de façon covalente à l'ARN viral (Murphy *et al.*, 1991).

Ce travail a donc confirmé de façon évidente pour le TuMV et chez *E. coli*, ce que Dougherty et Parks (1991) avaient obtenu *in vitro* pour le TEV, à savoir qu'il existe un site interne de clivage dans la NIa. Par contre, nous avons apporté une nouvelle information en montrant que l'inhibition du clivage à la jonction 6K2-NIa (Article 4, Fig. 3), réduisait considérablement le clivage interne à la NIa. Or il a été montré que la 6K2 du TEV empêchait la NIa de migrer vers le noyau (Restrepo-Hartwig et Carrington, 1992). On peut donc imaginer un modèle dans lequel la protéine 6K2 aurait un rôle important sur l'activité de la NIa et sur sa concentration dans le cytoplasme. Quand le précurseur 6K2-NIa n'est pas clivé, il demeure dans le cytoplasme et peut ainsi cliver en

trans d'autres sites de clivage de type NIa. Lorsque la jonction 6K2-NIa est clivée de façon autocatalytique, la NIa peut alors être acheminée vers le noyau sous sa forme complète, ou se cliver en VPg et NIa-Pro.

L'activité de la NIa était également totalement inhibée lorsque l'un des acides aminés (His) des trois qui composent son site actif était remplacé par la glutamine (Article 4, Fig. 3). Ceci est donc une indication que la protéase NIa du TuMV est bien une protéase du même type que la protéase 3C des picornavirus, qui clive dans la grande majorité des cas après le résidu Gln, mais également dans certains cas, après le résidu Glu (Palmenberg, 1990).

Les sites de clivages NIa-NIb et NIb-CP ont également été démontrés chez *E.coli* (communication de J.-F. Laliberté). La protéine NIa est donc la première protéine clonée du TuMV dont l'activité biologique a été démontrée.

La 6K2 du TuMV est une petite protéine de 6 kDa qui a un domaine très hydrophobe (Ile¹⁸⁹⁰ à Val¹⁹⁰⁹), et qui est située en N-terminal de la VPg. Ces trois caractéristiques sont analogues à celles de la protéine 3A des picornavirus, qui a un rôle dans la régulation de la réplication (Giachetti et Semler, 1991), ce qui permet d'attribuer une fonction identique à la 6K2.

La protéine 6K1 de 6 kDa du TuMV présente plus de 82 % d'homologie avec celle des autres potyvirus, alors que les protéines adjacentes, la P3 et la CI, n'en présentent environ que 30 % et 64 % respectivement (Article 3, Fig. 7). Cette nette différence d'homologie permet de conclure que la 6K1 est une protéine extrêmement importante pour

la multiplication virale puisqu'elle est très conservée. On y trouve également un domaine très hydrophobe (Arg¹¹⁸⁶ à Phe¹¹⁹⁶) qui pourrait servir à ancrer une des protéines adjacentes (la P3 ou/et la CI) aux membranes cellulaires.

La protéine N1b du TuMV est constituée de 518 acides aminés et on y retrouve plusieurs des huit motifs communs à la grande majorité des ARN polymérase ARN-dépendantes virales, dont le tripeptide (Gly²⁷⁰⁹-Asp-Asp²⁷¹¹) (Bruenn, 1991; Koonin *et al.*, 1991; Poch *et al.*, 1989). C'est pour le moment la seule évidence qui permet d'attribuer le rôle de réplacase à la N1b du TuMV.

La protéine CI du TuMV est la plus grosse protéine avec 643 acides aminés. Elle contient les deux motifs A et B qui constituent un site de fixation de nucléosides triphosphates (« NTBM »), commun aux protéines cellulaires ainsi qu'à un grand nombre de protéines virales qui ont une fonction hélicase (Gorbalenya et Koonin, 1989; Habili et Symons, 1989; Laín *et al.*, 1989b). Les motifs du TuMV sont les suivants: motif A [¹³⁰⁹Ile-Leu-Leu-Met-Gly-Ala-Val-Gly-Ser-Gly-Lys-Ser¹³²⁰] et motif B [¹³⁹⁸Phe-Val-Ile-Phe-Asp-Glu¹⁴⁰³]. Une fonction hélicase a été démontrée pour la protéine CI du PPV (Laín *et al.*, 1990, Laín *et al.*, 1991) et il est fort probable que celle du TuMV ait la même fonction.

2. CONSTRUCTION D'UN CLONE COMPLET DE TUMV.

Notre deuxième objectif était de construire un clone complet de TuMV pouvant mener à la synthèse *in vitro* d'ARN le plus analogue possible à l'ARN viral, étape importante conduisant à l'obtention d'un clone infectieux. Plusieurs clones de virus autres que les potyvirus pouvant donner des transcrits infectieux ont déjà été obtenus grâce à

l'utilisation du PCR (Boccard et Baulcombe, 1992; Hayes et Buck, 1990; Young *et al.*, 1991). Nous avons également choisi l'amplification par PCR pour assembler la séquence complète du TuMV. En effet, cette méthode permet d'introduire des sites de restriction ainsi que de nouvelles séquences comme par exemple, le promoteur de l'ARN polymérase du phage T7, ce qui évite d'avoir recours à la mutagenèse dirigée pour éliminer les séquences « non-virales ». Il était impossible d'assembler les clones obtenus pour le clonage de TuMV (Articles 1, 2 et 3) mis à part ceux de 1 200 pb (pTUM1200) et de 3 300 pb (pTUM3300) (Article 2), qui ont pu être assemblés par ligation grâce à la présence d'un site de restriction *BclI* dans la séquence commune à ces deux clones (Figure 12).

Il faut noter cependant la possibilité que la *Taq* polymérase introduise une mutation létale pour la réplication du virus; ce risque n'est d'ailleurs pas spécifique à l'utilisation du PCR, puisque la population d'ARN purifiés à partir d'une plante infectée n'est pas homogène et beaucoup de ces ARN, s'ils pouvaient être isolés et clonés, ne seraient pas infectieux. La *Taq* polymérase peut également provoquer la délétion ou l'insertion d'un nucléotide au cours de la polymérisation, mais à une fréquence plus faible que celle de la substitution (Tindall et Kunkel, 1988). La seule façon fiable d'éviter cette possibilité, serait de séquencer au complet tous les clones obtenus par PCR qui conduisent à la synthèse des transcrits synthétiques, afin d'éliminer les clones ayant des mutations importantes.

Nous avons procédé à l'assemblage du génome du TuMV à l'aide de six clones provenant d'amplification PCR distinctes (pTUM900, pTUM1400, pKS2400, pKS2500,

pTUM1200, pTUM3300); deux de ces clones ont été séquencés au complet (pTUM1200 et pTUM3300). Plutôt que d'utiliser un promoteur présent dans un vecteur de transcription tel que le pBluescript^{MD} I SK, la séquence du promoteur de l'ARN polymérase du phage T7 a été introduite dans l'amorce JF_{saT7} (Annexe 1) directement suivie de la séquence du TuMV. L'ARN polymérase du T7 a été choisie car elle reconnaît spécifiquement son promoteur et permet également d'obtenir de bons rendements d'ARN lors de la synthèse *in vitro*. Son seul inconvénient est qu'elle commence toujours la synthèse des transcrits par une guanine, alors que le premier nucléotide de l'ARN viral est une adénine. La présence d'un résidu supplémentaire en 5' ne semble pas éliminer le caractère infectieux des transcrits de potyvirus. L'adjonction en 5' d'un analogue de coiffe méthylée est même nécessaire pour que les transcrits soient infectieux (Domier *et al.*, 1989; Gal-On *et al.*, 1991; Riechmann *et al.*, 1990). Les ARN provenant de la descendance de ces transcrits perdent ce résidu ainsi que toutes les séquences « non-virales », ce qui démontre la propriété des protéines virales à répliquer spécifiquement l'ARN viral (Riechmann *et al.*, 1990).

Un site de restriction *Bam*HI a été ajouté vers le milieu du génome, grâce à deux amorces spécifiques FT₅₀ et son complémentaire JF₅₃ (Annexe 1), sans que cela ne modifie la séquence en acides aminés de la polyprotéine. Ce site de restriction a permis d'assembler les clones pKS4900 et pTUM5000, qui représentaient chacun la moitié du génome ADN du TuMV. La construction plasmidique finale (pTUM9900), qui contient la séquence complète du TuMV suivie d'une queue polyadénylée d'environ 55 résidus, a été linéarisée par l'enzyme *Sma*I avant de procéder à sa transcription *in vitro*. Les transcrits synthétiques obtenus n'avaient donc que trois bases supplémentaires en 3', et une

en 5', ce qui est donc presque identique à l'ARN viral; par comparaison, Riechmann *et al.* (1990) ont obtenu des transcrits infectieux du PPV qui possédaient 13 bases « non-virales ».

La transcription *in vitro* du clone pTUM9900 linéarisé par *SmaI* a produit environ 20 μg d'ARN coiffés pour 10 μg d'ADN, ce qui est un bon rendement compte tenu de la longueur des ARN et de la présence de l'analogue de coiffe. Les transcrits ont été analysés par électrophorèse sur gel d'ARN afin d'estimer la qualité de chaque préparation (Figure 16). On remarque que les transcrits de 2 500 bases (Figure 16, puits 4) sont peu dégradés comparés à ceux de 4 900 bases (Figure 16, puits 3) et ceux de 9 900 bases (Figure 16, puits 2); ceci suggère que la traînée située sous la bande principale de chaque puits correspond à des ARN dont la synthèse est incomplète plutôt qu'à un produit de dégradation.

Les transcrits coiffés, qui comportaient seulement quatre bases de plus par rapport à l'ARN viral, ont été testés sur deux espèces dont l'infection par le TuMV est systémique, et qui présentent rapidement des symptômes de mosaïques: le navet (*Brassica campestris ssp. rapifera* cv. « Just Right ») et la moutarde tendre (*Brassica perviridis* var. tendergreen). L'inoculation mécanique d'environ 2 μg d'ARN synthétique coiffé par plante (50 plantes testées par espèce) a échoué, mais le témoin positif pour cette expérience, qui consistait en l'inoculation des plantes avec l'ARN viral (20 plantes testées par espèce) n'a pas également provoqué l'apparition de symptômes. Le virus purifié était par contre infectieux, même à une dose de 100 ng, ce qui correspond à environ 5 ng d'ARN, et qui signifie que le protocole d'infection n'était pas en cause dans l'échec

obtenu. Trois hypothèses peuvent être considérées pour expliquer notre résultat préliminaire:

1) Les espèces choisies sont peut-être extrêmement difficiles à infecter par de l'ARN non-encapsidé, du fait de la présence de nucléases qui le dégradent très rapidement. Pour confirmer cette hypothèse, il faudrait tester les transcrits ou l'ARN viral sur d'autres espèces sensibles au TuMV, ou procéder à la transfection de protoplastes suivie de l'analyse des protéines virales, comme l'ont fait Domier *et al.* (1989) pour les transcrits du TVMV.

2) Le résultat négatif obtenu avec l'ARN natif peut s'expliquer par le fait qu'il ne possède plus de protéine VPg à son extrémité 5', car il a été préparé par digestion à la protéinase K; or, il a été mentionné pour d'autres potyvirus que l'élimination de la VPg diminuait considérablement le caractère infectieux du génome viral (Domier *et al.*, 1989; Riechmann *et al.*, 1989). Nous devrions donc inoculer beaucoup plus de plantes par l'ARN purifié du TuMV, ou le préparer sans digestion à la protéinase K, c'est-à-dire par extraction au phénol en présence de SDS suivie d'une précipitation à l'éthanol. La VPg ainsi que l'analogue de coiffe méthylée, conféreraient une certaine résistance aux RNases de la plante, permettant ainsi à l'ARN d'être traduit avant d'être dégradé (Riechmann *et al.*, 1990).

3) Le clone que nous avons construit (pTUM9900) peut être défectueux, auquel cas les transcrits présentent une mutation létale pour la multiplication virale. Pour éliminer cette hypothèse, il faudrait procéder au séquençage complet du clone pTUM9900 afin de vérifier qu'il ne comporte aucune mutation qui puisse affecter le cadre de lecture sur

l'ARN. Nous pourrions également répéter l'expérience en utilisant une ADN polymérase thermostable ayant une activité 3'-5' exonucléase, comme par exemple, la « Vent^{MD} DNA polymérase » (New England Biolab). De plus, nous nous sommes basés sur l'apparition des symptômes de mosaïques pour déclarer que nos transcrits n'étaient pas infectieux; il faudrait également considérer la possibilité qu'ils puissent être infectieux, mais qu'une mutation affecte la formation des symptômes sans diminuer la production de virus. Cette hypothèse peut être vérifiée, soit par la transfection de protoplastes et l'analyse des protéines virales, soit par la détection de la protéine de la capsid dans les plants infectés par les transcrits; ceci nécessiterait la mise au point de la détection rapide et sensible de cette protéine par une méthode immunoenzymatique de type « ELISA ».

Cette deuxième partie de mon travail de doctorat a donc montré qu'il était possible d'assembler le génome du TuMV à l'aide de quelques clones provenant d'amplifications par PCR; la transcription *in vitro* de très longs ARN coiffés par l'introduction d'un promoteur de la T7 ARN polymérase n'a pas non-plus posé de problèmes majeurs.

Nous pourrions envisager de construire un clone complet de TuMV sans procéder par PCR, mais par synthèse « classique » d'un ADN homologue à l'ARN viral (Gubler et Hoffmann, 1983); la synthèse du premier brin d'ADNc serait amorcée par un oligonucléotide du type FT_{oxk} (Annexe 1) à partir d'une grande quantité d'ARN viral, et le produit de cette synthèse par une transcriptase inverse serait sélectionné par électrophorèse sur gel alcalin; le deuxième brin serait alors amorcé par un oligonucléotide spécifique du type JF_{sarr7} (Annexe 1) et synthétisé grâce au fragment « Klenow » de l'ADN polymérase I de *E. coli*, puis purifié par électrophorèse sur gel. Cette méthode, qui a été

utilisée de façon efficace pour le clonage d'un népovirus (Le Gall *et al.*, 1988), pourrait éliminer le risque inhérent à l'utilisation de la *Taq* polymérase. Cela permettrait également de tester plusieurs clones différents issus de cette synthèse, ce que nous ne pouvons pas faire dans notre cas, puisque le clone pTUM9900 est issu d'une seule combinaison de différents clones.

La construction d'un clone complet et infectieux de TuMV pourrait donc nécessiter la mise au point d'une nouvelle stratégie d'assemblage, comme cela a été décrit ci-dessus, mais toutes les expériences pourraient être testées tout d'abord sur les transcrits synthétiques, qui constituent une source inépuisable d'ARN, avant de les réaliser sur l'ARN viral.

La maîtrise d'un clone infectieux d'un virus tel que le TuMV serait un atout important pour sa caractérisation moléculaire approfondie. En effet, cela donnerait la possibilité de manipuler génétiquement l'ARN tout en conservant son pouvoir infectieux. Nous pourrions ainsi faire des délétions de petites régions ou même de protéines complètes, afin d'observer leur rôle dans la pathogénèse, la multiplication et le transport du virus. Nous pourrions également échanger des protéines complètes avec d'autres potyvirus afin de déterminer ce qui définit le spectre d'hôte ou la pathologie.

CONCLUSION

L'objectif fixé pour ce travail de doctorat était d'entamer l'étude moléculaire du virus de la mosaïque du navet (TuMV), qui était encore en 1989 non caractérisé. Pour ce faire, une souche canadienne de TuMV a été purifiée et analysée. L'ARN viral a ensuite été extrait et nous avons procédé à son clonage et à son séquençage complet. Les différentes protéines virales ont été localisées sur la polyprotéine codée par l'ARN. Nous avons ensuite procédé à la construction d'un clone complet de TuMV potentiellement infectieux.

Pour parvenir à atteindre ces objectifs, nous avons amplifié l'ARN viral par polymérisation en chaîne (PCR) grâce à l'ADN polymérase thermostable de *Thermus aquaticus*, technique qui était très prometteuse en 1989 mais encore inutilisée pour le clonage des potyvirus. Les conditions d'amplification et de clonage du génome du TuMV ont été mises au point (Article 1 et 2) et cette méthode est maintenant utilisée de façon routinière au laboratoire du Dr. J.-F. Laliberté pour le clonage et l'expression des protéines du TuMV (Article 4).

Le génome du TuMV a ainsi pu être cloné puis séquencé au complet et les régions codantes et non-codantes ont été localisées (Article 3). Une polyprotéine de 3 163 acides aminés a été déterminée, à partir de laquelle nous avons localisé les sites de clivage reconnus par les trois protéases virales, dont la protéase NIa qui a été démontrée avoir une activité chez *E. coli* (Article 4).

Un clone complet de TuMV a également été construit en aval du promoteur de l'ARN polymérase du phage T7, permettant la synthèse *in vitro* de transcrits équivalents à l'ARN viral. Des transcrits synthétiques possédant un analogue de coiffe ont été

synthétisés et inoculés de façon mécanique à deux espèces sensibles au TuMV. Ces transcrits n'étaient pas infectieux, sur la base de la formation de symptômes de mosaïques, mais l'ARN purifié n'était pas non plus infectieux dans les tests d'infection réalisés.

Ce projet de doctorat a apporté quelques informations sur la biologie moléculaire du TuMV, et des anticorps dirigés contre les protéines CP, N1b, N1a, CI et P1 ont pu être obtenus, après clonage et expression de ces protéines chez *E. coli*. Plusieurs projets ont ainsi pu être initiés au laboratoire du Dr. J.-F. Laliberté, dont certains sont mentionnés ci-dessous:

- étude de l'interaction de la protéine P1 avec l'ARN viral dans le but de trouver un rôle potentiel de cette protéine dans le transport du virus de cellule à cellule.
- étude du rôle des protéines CI et N1b sur la réplication de l'ARN viral.
- étude biochimique de la protéase N1a avec son expression et son analyse structurale par cristallographie aux rayons X.
- étude du rôle de la séquence 5' non-codante sur la traduction *in vitro* et sur l'expression *in situ* de la β -glucuronidase.

REMERCIEMENTS

Je tiens à remercier cordialement mon directeur de recherche, le Dr. Jean-François Laliberté qui s'est avéré être un excellent enseignant et qui a été très motivant durant toute la durée de mes études.

Je remercie le Dr. Michel Trudel, directeur du Centre de recherche en virologie, de m'avoir accepté dans son centre et d'avoir mis à notre disposition le matériel nécessaire à la réalisation de ce projet de doctorat.

Je remercie les professeurs du comité de programme de doctorat en virologie qui m'ont donné la possibilité de réaliser ce doctorat, et de m'avoir suivi tout au long de ces trois années d'études. Je remercie également tous les professeurs de l'Institut Armand-Frappier qui ont su me faire partager leur plaisir à travailler en virologie ou en biologie moléculaire.

Je remercie les techniciennes qui m'ont aidé dans mes travaux, principalement madame Henriette Chatel qui m'a appris à travailler en équipe, madame Johanne Roger pour avoir parfaitement séquencé plus de 40 000 bases, et madame Lise Trempe pour son expertise sur la synthèse d'oligonucléotides.

Je remercie mes parents et mes soeurs, très proches de coeur mais loin dans l'espace, qui m'ont soutenu depuis si longtemps et sans arrêt.

BIBLIOGRAPHIE

AHLQUIST, P., R. FRENCH, J. JANDA et L.S. LOESCH-FRIES. 1984. Multicomponent RNA plant virus infection derived from cloned viral cDNA. Proc. Natl. Acad. Sci. USA 81: 7066-7070.

ALLISON, R.F., W.G. DOUGHERTY, T.D. PARKS, L. WILLIS, R.E. JOHNSTON, M. KELLY et F.B. ARMSTRONG. 1985. Biochemical analysis of the capsid protein gene and capsid protein of tobacco etch virus: N-terminal amino acids are located on the virion's surface. Virology 147: 309-316.

ALLISON, R.F., R.E. JOHNSTON et W.G. DOUGHERTY. 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology 154: 9-20.

ANDERER, F.A., H. UHLIG, E. WEBER et G. SCHRAMM. 1960. Primary structure of the protein of tobacco mosaic virus. Nature 186: 922-925.

ARUMUGANATHAN, K. et E.D. EARLE. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Repo. 9: 208-218.

ATREYA, P.L., C.D. ATREYA et T.P. PIRONE. 1991. Amino acid substitution in the coat protein result in loss of insect transmissibility of a plant virus. Proc. Natl. Acad. Sci. USA 88: 7887-7891.

ATREYA, C.D., P.L. ATREYA, D.W. THORNBURY et T.P. PIRONE. 1992. Site-directed mutations in the potyvirus HC-Pro gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. Virology 191: 106-111.

ATREYA, C.D., RACCAH B. et T.P. PIRONE. 1990. A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. Virology 178: 161-165.

BARNETT, O.W. 1991. *Potyviridae*, a proposed family of plant viruses. Arch. Virol. 118: 139-141.

BAUNOCH, D.A., P. DAS, M.E. BROWNING et V. HARI. 1991. A temporal study of the expression of the capsid, cytoplasmic inclusion and nuclear inclusion proteins of tobacco etch potyvirus in infected plants. J. Gen. Virol. 72: 487-492.

BAUNOCH, D.A., P. DAS et V. HARI. 1990. Potato virus Y helper component protein is associated with amorphous inclusions. J. Gen. Virol. 71: 2479-2482.

BAWDEN, F.C., N.W. PIRIE, J.D. BERNAL ET I. FANKUCHEN. 1936. Liquid crystalline substances from virus-infected plants. Nature 138: 1051-1052.

BAZAN, J.F. et R.J. FLETTERICK. 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. Proc. Natl. Acad. Sci. USA 85: 7872-7876.

- BEACHY, R.N., S. LOESCH-FRIES et N.E. TUMER. 1990. Coat protein-mediated resistance against virus infection. *Annu. Rev. Phytopathol.* 28: 451-474.
- BERGER, P.H., A.G. HUNT, L.L. DOMIER, G.M. HELLMANN, Y. STRAM, D.W. THORNBURY et T.P. PIRONE. 1989. Expression in transgenic plants of a viral gene product that mediates insect transmission of potyviruses. *Proc. Natl. Acad. Sci. USA* 86: 8402-8406.
- BERGER, P.H. et T.P. PIRONE. 1986. The effect of helper component on the uptake and localization of potyviruses in *Myzus persicae*. *Virology* 153: 256-261.
- BLOCH, W. 1991. A biochemical perspective of the polymerase chain reaction. *Biochemistry* 30: 2735-2745.
- BOCCARD, F. et D.C. BAULCOMBE. 1992. Infectious *in vitro* transcripts from amplified cDNAs of the Y and the Kin strains of cucumber mosaic virus. *Gene* 114: 223-227.
- BOURDIN, D. et H. LECOQ. 1991. Evidence that heteroencapsidation between two potyviruses is involved in aphid transmission of a non-aphid-transmissible isolate from mixed infections. *Phytopathology* 81: 1459-1464.
- BRANDES, J et R. BERCKS. 1965. Gross morphology and serology as a basis for classification of elongated plant viruses. *Adv. Virus Res.* 11: 1-24.
- BRANDES, J. et C. WETTER. 1959. Classification of elongated plant viruses on the basis of particle morphology. *Virology* 8: 99-115.
- BRUENN, J.A.. 1991. Relationships among the positive strand and double-strand RNA viruses as viewed through their RNA-dependent RNA polymerases. *Nucleic Acids Res.* 19: 217-226.
- BRYAN, G.T., R.C. GARDNER et R.L.S. FORSTER. 1992. Nucleotide sequence of the coat protein gene of a strain of clover yellow vein virus from New Zealand: conservation of a stem-loop structure in the 3' region of potyviruses. *Arch. Virol.* 124: 133-146.
- BURGER, J.T., R.J. BRAND et E.P. RYBICKI. 1990. The molecular cloning and nucleotide sequencing of the 3'-terminal region of *Ornithogalum* mosaic virus. *J. Gen. Virol.* 71: 2527-2534.
- CARRINGTON, J.C., S.M. CARY et W.G. DOUGHERTY. 1988. Mutational analysis of tobacco etch virus polyprotein processing: cis and trans proteolytic activities of polyproteins containing the 49-kilodalton proteinase. *J. Virol.* 62: 2313-2320.
- CARRINGTON, J.C., S.M. CARY, T.D. PARKS et W.G. DOUGHERTY. 1989a. A second proteinase encoded by a plant potyvirus genome. *EMBO J.* 8: 365-370.

CARRINGTON, J.C. et W.G. DOUGHERTY. 1987a. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. *J. Virol.* 61: 2540-2548.

CARRINGTON, J.C. et W.G. DOUGHERTY. 1987b. Processing of the tobacco etch virus 49K protease requires autoproteolysis. *Virology* 160: 355-362.

CARRINGTON, J.C. et W.G. DOUGHERTY. 1988. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. USA* 85: 3391-3395.

CARRINGTON, J.C. et D.D. FREED. 1990. Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. *J. Virol.* 64: 1590-1597.

CARRINGTON, J.C., D.D. FREED et A.J. LEINICKE. 1991. Bipartite signal sequence mediates nuclear translocation of the plant potyviral N1a protein. *Plant Cell* 3: 953-962.

CARRINGTON, J.C., D.D. FREED et C.S. OH. 1990. Expression of potyviral polyproteins in transgenic plants reveals three proteolytic activities required for complete processing. *EMBO J.* 9: 1347-1353.

CARRINGTON, J.C., D.D. FREED et T.C. SANDERS. 1989b. Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and *in vitro*. *J. Virol.* 63: 4459-4463.

CARRINGTON, J.C. et K.L. HERNDON. 1992. Characterization of the potyviral HC-Pro autoproteolytic cleavage site. *Virology* 187: 308-315.

CHOI, G.H., D.M. PAWLYK et D.L. NUSS. 1991. The autocatalytic protease p29 encoded by a hypovirulence-associated virus of the chestnut blight fungus resembles the potyvirus-encoded protease HC-Pro. *Virology* 183: 747-752.

DASGUPTA, A., M.H. BARON et D. BALTIMORE. 1979. Poliovirus replicase: a soluble enzyme able to initiate copying of poliovirus RNA. *Proc. Natl. Acad. Sci. USA* 76: 2679-2683.

DAWSON, W.O., D.L. BECK, D.A. KNORR et G.L. GRANTHAM. 1986. cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. *Proc. Natl. Acad. Sci. USA* 83: 1832-1836.

DEOM, C.M., M. LAPIDOT et R.N. BEACHY. 1992. Plant virus movement proteins. *Cell* 69: 221-224.

DEOM, C.M., M.J. OLIVER et R.N. BEACHY. 1987. The 30-Kilodalton gene product of tobacco mosaic virus potentiates virus movement. *Science* 237: 389-394.

- DINANT, S., H. LOT, J. ALBOUY, C. KUZIYAK, M. MEYER et S. ASTIER-MANIFACIER. 1991. Nucleotide sequence of the 3' terminal region of lettuce mosaic potyvirus RNA shows a Gln/Val dipeptide at the cleavage site between the polymerase and the coat protein. *Arch. Virol.* 116: 235-252.
- DOLJA, V.V., V.P. BOYKO, A.A. AGRANOVSKY et E.V. KOONIN. 1991. Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. *Virology* 184: 79-86.
- DOMIER, L.L., K.M. FRANKLIN, A.G. HUNT, R.E. RHOADS et J.G. SHAW. 1989. Infectious *in vitro* transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. *Proc. Natl. Acad. Sci. USA* 86: 3509-3513.
- DOMIER, L.L., K.M. FRANKLIN, M. SHAHABUDDIN, G.M. HELLMANN, J.H. OVERMEYER, S.T. HIREMATH, M.F.E. SIAW, G.P. LOMONOSSOFF, J.G. SHAW et R.E. RHOADS. 1986. The nucleotide sequence of tobacco vein mottling virus RNA. *Nucleic Acids Res.* 14: 5417-5430.
- DOMIER, L.L., J.G. SHAW et R.E. RHOADS. 1987. Potyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. *Virology* 158: 20-27.
- DOUGHERTY, W.G. et J.C. CARRINGTON. 1988. Expression and function of potyviral gene products. *Ann. Rev. Phytopathol.* 26: 123-143.
- DOUGHERTY, W.G., J.C. CARRINGTON, S.M. CARY et T.D. PARKS. 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J.* 7: 1281-1287.
- DOUGHERTY, W.G., S.M. CARY et T.D. PARKS. 1989a. Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* 171: 356-364.
- DOUGHERTY, W.G. et E. HIEBERT. 1980a. Translation of potyvirus RNA in a rabbit reticulocyte lysate: reaction conditions and identification of capsid protein as one of the products of *in vitro* translation of tobacco etch and pepper mottle viral RNAs. *Virology* 101: 466-474.
- DOUGHERTY, W.G. et E. HIEBERT. 1980b. Translation of potyvirus RNA in a rabbit reticulocyte lysate: identification of nuclear inclusion proteins as products of tobacco etch virus RNA translation and cylindrical inclusion proteins as a product of the potyvirus genome. *Virology* 104: 174-182.
- DOUGHERTY, W.G. et E. HIEBERT. 1980c. Translation of potyvirus RNA in a rabbit reticulocyte lysate: cell-free translation strategy and a genetic map of the potyviral genome. *Virology* 104: 183-194.

DOUGHERTY, W.G. et T.D. PARKS. 1989. Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile at two tobacco etch virus cleavage sites in cell-free assays. *Virology* 172: 145-155.

DOUGHERTY, W.G. et T.D. PARKS. 1991. Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology* 183: 449-456.

DOUGHERTY, W.G., T.D. PARKS, S.M. CARY, J.F. BAZAN et R.J. FLETTERICK. 1989b. Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. *Virology* 172: 302-310.

EGGENBERGER, A.L., D.M. STARK et R.N. BEACHY. 1989. The nucleotide sequence of a soybean mosaic virus coat protein-coding region and its expression in *Escherichia coli*, *Agrobacterium tumefaciens* and tobacco callus. *J. Gen. Virol.* 70: 1853-1860.

EDWARDSON, J.R. 1992. Inclusion bodies. *Arch Virol* [Suppl. 5]: 25-30.

EDWARDSON, J.R. et R.G. CHRISTIE. 1986. Turnip mosaic virus dans "Virus infecting forage legumes" II, monographie no 14, 438-453.

FAUQUET, C., J. DEJARDIN et J.-C. THOUVENEL. 1986a. Evidence that the amino acid composition of the particle proteins of plant viruses is characteristic of the virus group. I. Multidimensional classification of plant viruses. *Intervirology* 25: 1-13.

FAUQUET, C., J. DEJARDIN et J.-C. THOUVENEL. 1986b. Evidence that the amino acid composition of the particle proteins of plant viruses is characteristic of the virus group. II. Discriminant analysis according to structural biological and classification properties of plant viruses. *Intervirology* 25: 190-200.

FOURNEY, R.M., J. MIYAKOSHI, R.S. DAY et M.C. PATERSON. 1988. Northern blotting: efficient RNA staining and transfer. *FOCUS* 10: 5-6.

FROHMAN, M.A., M.K. DUSH et G. R. MARTIN. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85: 8998-9002.

GADH, I.P.S. et V. HARI. 1986. Association of tobacco etch virus related RNA with chloroplasts in extracts of infected plants. *Virology* 150: 304-307.

GAL-ON, A., Y. ANTIGNUS, A. ROSNER et B. RACCAH. 1991. Infectious *in vitro* RNA transcripts derived from cloned cDNA of the cucurbit potyvirus, zucchini yellow mosaic virus. *J. Gen. Virol.* 72: 2639-2643.

- GARCÍA, J.A., M.T. MARTÍN., M.T. CERVERA et J.L. RIECHMANN. 1992. Proteolytic processing of the plum pox potyvirus polyprotein by the N1a protease at a novel cleavage site. *Virology* 188: 697-703.
- GARCÍA, J.A., J.L. RIECHMANN et S. LAÍN. 1989a. Artificial cleavage site recognized by plum pox potyvirus protease in *Escherichia coli*. *J. Virol.* 63: 2457-2460.
- GARCÍA, J.A., J.L. RIECHMANN et S. LAÍN. 1989b. Proteolytic activity of the plum pox potyvirus N1a-like protein in *Escherichia coli*. *Virology* 170: 362-369.
- GARCÍA, J.A., J.L. RIECHMANN, M.T. MARTÍN et S. LAÍN. 1989c. Proteolytic activity of the plum pox potyvirus N1a-protein on excess of natural and artificial substrates in *Escherichia coli*. *FEBS Letters* 257: 269-273.
- GARDNER, R. et J.B. KENDRICK. 1921. Soybean mosaic. *J. Agri. Res.* 22: 111-114
- GIACHETTI, C. et B.L. SEMLER. 1991. Role of a viral membrane polypeptide in strand-specific initiation of poliovirus RNA synthesis. *J. Virol.* 65: 2647-2654.
- GIERER, A. et G. SCHRAMM. 1956. Infectivity of ribonucleic acid from tobacco mosaic virus. *Nature* 177: 702-703.
- GOELET, P., G.P. LOMONOSSOFF, P.J.G. BUTLER, M.E. AKAM, M.J. GAIT et J. KARN. 1982. Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. USA* 79: 5818-5822.
- GOLDBACH, R. 1986. Molecular evolution of plant RNA viruses. *Ann. Rev. Phytopathol.* 24: 289-310.
- GOLDBACH, R. et J. WELLINK. 1988. Evolution of plus-strand RNA viruses. *Intervirology* 29: 260-267.
- GORBALENYA, A.E. et E.V. KOONIN. 1989. Viral protein containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* 17: 8413-8440.
- GORBALENYA, A.E., E.V. KOONIN, A.P. DONCHENKO et V.M. BLINOV. 1988. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Letters* 235: 16-24.
- GOVIER, D.A. et B. KASSANIS. 1974a. Evidence that a component other than the virus particle is needed for aphid transmission of potato virus Y. *Virology* 57: 285-286.
- GOVIER, D.A. et B. KASSANIS. 1974b. A virus-induced component of plant sap needed when aphids acquire potato virus Y from purified preparations. *Virology* 61: 420-426.

- GOVIER, D.A., B. KASSANIS et T.P. PIRONE. 1977. Partial purification and characterization of potato virus Y helper component. *Virology* 78: 306-314.
- GUBLER, U. et B.J. HOFFMAN. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* 25: 263-269.
- GUNASINGHE, U.B. et P.H. BERGER. 1991. Association of potato virus Y gene products with chloroplasts in tobacco. *Mol. Plant-Microbe Interactions* 4: 452-457.
- HABILI, N. et R.H. SYMONS. 1989. Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. *Nucleic Acids Res.* 17: 9543-9555.
- HANAHAH, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557-580.
- HARI, V., A. SIEGEL, C. ROZEK et W.E. TIMBERLAKE. 1979. The RNA of tobacco etch virus contains poly(A). *Virology* 92: 568-571.
- HARI, V. 1981. The RNA of tobacco etch virus: further characterization and detection of protein linked to RNA. *Virology* 112: 391-399.
- HAYES, R.J. et K.W. BUCK. 1990. Infectious cucumber mosaic virus RNA transcribed *in vitro* from clones obtained from cDNA amplified using the polymerase chain reaction. *J. Gen. Virol.* 71: 2503-2508.
- HELLMANN, G.M., J.G. SHAW, J.A. LESNAW, L.-Y. CHU, T.P. PIRONE et R.E. RHOADS. 1980. Cell-free translation of tobacco vein mottling virus RNA. *Virology* 106: 206-216.
- HELLMANN, G.M., J.G. SHAW et R.E. RHOADS. 1988. *In vitro* analysis of tobacco vein mottling virus NIa cistron: evidence for a virus-encoded protease. *Virology* 163: 554-562.
- HELLMANN, G.M., D.W. THORNBURY, E. HIEBERT, J.G. SHAW, T.P. PIRONE et R.E. RHOADS. 1983. Cell-free translation of tobacco vein mottling virus RNA. II. Immunoprecipitation of products by antisera to cylindrical inclusion, nuclear inclusion and helper component proteins. *Virology* 124: 434-444.
- HIEBERT, E. et J.G. Mc DONALD. 1976. Capsid protein heterogeneity in turnip mosaic virus. *Virology* 70: 144-150.
- HIEBERT, E, D.E. PURCIFULL et R.G. CHRISTIE. 1984. Purification and immunological analysis of plant viral inclusion bodies. *Meth. Virol.* 8: 225-280.

- HILL, J.H. et R.J. SHEPHERD. 1972. Biochemical properties of turnip mosaic virus. *Virology* 47: 807-816.
- HULL, R. 1989. The movement of viruses in plants. *Ann. Rev. Phytopathol.* 27: 213-240.
- JAGADISH, M.N., C.W. WARD, K.H. GOUGH, P.A. TULLOCH, L.A. WHITTAKER et D.D. SHUKLA. 1991. Expression of potyvirus coat protein in *Escherichia coli* and yeast and its assembly into virus-like particles. *J. Gen. Virol.* 72: 1543-1550.
- JAYARAM, C., J.H. HILL et W.A. MILLER. 1992. Complete nucleotide sequences of two soybean mosaic virus strains differentiated by response of soybean containing the Rsv resistance gene. *J. Gen. Virol.* 73: 2067-2077.
- JOHANSEN, E., O.F. RASMUSSEN, M. HEIDE, et B. BORKHARDT. 1991. The complete nucleotide sequence of pea seed-borne mosaic virus RNA. *J. Gen. Virol.* 72: 2625-2632.
- KASHIWAZAKI, S. Y. MINOBE et H. HIBINO. 1991. Nucleotide sequence of barley yellow mosaic virus RNA 2. *J. Gen. Virol.* 72: 995-999.
- KASHIWAZAKI, S. Y. MINOBE, T. OMURA et H. HIBINO. 1990. Nucleotide sequence of barley yellow mosaic virus RNA 1: a close evolutionary relationship with potyviruses. *J. Gen. Virol.* 71: 22781-2790.
- KASSANIS, B. et D.A. GOVIER. 1971a. New evidence on the mechanism of aphid transmission of potato C and potato aucuba mosaic viruses. *J. Gen. Virol.* 10: 99-101.
- KASSANIS, B. et D.A. GOVIER. 1971b. The role of the helper virus in aphid transmission of potato aucuba mosaic virus and potato virus C. *J. Gen. Virol.* 13: 221-228.
- KEOHAVONG, P. et W.G. THILLY. 1989. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* 86: 9253-9257.
- KNUHTSEN, H., E. HIEBERT et D.E. PURCIFULL. 1974. Partial purification and some properties of tobacco etch virus intranuclear inclusions. *Virology* 61: 200-209.
- KOHNEN, P.D., W.G. DOUGHERTY et R.O. HAMPTON. 1992. Detection of pea seedborne mosaic potyvirus by sequence specific enzymatic amplification. *J. Virol. Meth.* 37: 253-258.
- KOONIN, E.V., G.H. CHOI, D.L. NUSS, R. SHAPIRA et J.C. CARRINGTON. 1991. Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive-strand RNA plant viruses. *Proc. Natl. Acad. Sci. USA* 88: 10647-10651.

- LAI, M.M.C. 1992. RNA recombination in animal and plant viruses. *Microbiol. Rev.* 56: 61-79.
- LAÍN, S., M.T. MARTÍN, J.L. RIECHMANN et J.A. GARCÍA. 1991. Novel catalytic activity associated with positive-strand RNA virus infection: nucleic acid-stimulated ATPase activity of the plum pox potyvirus helicase-like protein. *J. Virol.* 65: 1-6.
- LAÍN, S., J.L. RIECHMANN et J.A. GARCÍA JA. 1989a. The complete nucleotide sequence of plum pox potyvirus RNA. *Virus Res.* 13: 157-172.
- LAÍN, S., J.L. RIECHMANN et J.A. GARCÍA. 1990. RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* 18: 7003-7006.
- LAÍN, S., J.L. RIECHMANN, M.T. MARTÍN et J.A. GARCÍA. 1989b. Homologous potyvirus and flavivirus proteins belonging to a superfamily of helicase-like proteins. *Gene* 82: 357-362.
- LAÍN, S., J.L. RIECHMANN, E. MENDEZ et J.A. GARCÍA. 1988. Nucleotide sequence of the 3' terminal region of plum pox potyvirus RNA. *Virus Res.* 10: 325-342.
- LANGVELD, S.A., J.-M. DORE, J. MEMELINK, A.F.L.M. DERKS, C.I.M. VAN DER VLUGT, C.J. ASJES et J.F. BOL. 1991. Identification of potyviruses using the polymerase chain reaction with degenerate primers. *J. Gen. Virol.* 72: 1531-1541.
- LAWSON, C., W. KANIEWSKI, L. HALEY, R. ROZMAN, C. NEWELL, P. SANDERS et N.E. TUMER. 1990. Engineering resistance to mixed virus infection in a commercial potato cultivar: resistance to potato virus X and potato virus Y in transgenic Russet Burbank. *Bio/technology* 8: 127-134.
- LE GALL, O., T. CANDRESSE, V. BRAULT, C. BRETOUT, L. HIBRAND et J. DUNEZ. 1988. Cloning full-length cDNA of grapevine chrome mosaic nepovirus. *Gene* 73: 67-75.
- LINDBO, J.A. et W.G. DOUGHERTY. 1992a. Pathogen-derived resistance to a potyvirus immune and resistant phenotypes in transgenic tobacco expressing altered forms of a potyvirus coat protein nucleotide sequence. *Mol. Plant-Microbe Interactions* 5: 144-153.
- LINDBO, J.A. et W.G. DOUGHERTY. 1992b. Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. *Virology* 189: 725-733.
- LING-JIE, K., F. RONG-XIANG, C. ZHENG-HUA et M. KE-QUIANG. 1990. Molecular cloning and nucleotide sequence of coat protein gene of turnip mosaic virus. *Nucleic Acids Res.* 18: 5555.

- LÜTCKE, H.A., K.C. CHOW, F.S. MICKEL, K.A. MOSS, H.F.KERN et G.A. SCHEELE. 1987. Selection of AUG initiation codons differs in plants and animals. *EMBO J.* 6: 43-48.
- MAISS, E., E. BREYEL, A. BRISSE et R. CASPER. 1988. Molecular cloning of DNA complementary to the RNA-genome of plum pox virus (PPV). *J. Phytopathol.* 122: 222-231.
- MAISS, E., U. TIMPE, A. BRISSE, W. JELKMANN, R. CASPER, G. HIMMLER, D. MATTANOVICH et H.W. KATINGER. 1989. The complete nucleotide sequence of plum pox virus RNA. *J. Gen. Virol.* 70: 513-524.
- MAISS, E., U. TIMPE., A. BRISSEKERODE., D.E. LESEMANN et R. CASPER. 1992. Infectious *in vivo* transcripts of a plum pox potyvirus full-length cDNA clone containing the cauliflower mosaic virus-35S RNA promoter. *J. Gen. Virol.* 73: 709-713.
- MARTÍN, M.T. et J.A. GARCÍA. 1991. Plum pox potyvirus RNA replication in a crude membrane fraction from infected *Nicotiana clevelandii* leaves. *J. Gen. Virol.* 72: 785-790.
- MARTÍN, M.T., C. LOPEZ OTÍN, S. LAÍN et J.A. GARCÍA. 1990. Determination of polyprotein processing sites by amino terminal sequencing of nonstructural proteins encoded by plum pox potyvirus. *Virus Res.* 15: 97-106.
- MATTHEWS, R.E.F. 1991. *Plant Virology*, third edition. Academic Press Inc., Santiago, Harcourt Brace Jovanovich, Publishers.
- MAVANKAL, G. et R.E. RHOADS. 1991. *In vitro* cleavage at or near the N-terminus of the helper component protein in the tobacco vein mottling virus polyprotein. *Virology* 185: 721-731.
- MAYO, M.A. et C.A. JOLLY. 1991. The 5'-terminal sequence of potato leafroll virus RNA: evidence of recombination between virus and host RNA. *J. Gen. Virol.* 72: 2591-2595.
- MCDONALD, J.G. et J.B. BANCROFT. 1977. Assembly studies on potato virus Y and its coat protein. *J. Gen. Virol.* 35: 251-263.
- MCDONALD, J.G. et E. HIEBERT. 1975. Characterization of the capsid and cylindrical inclusion proteins of three strains of turnip mosaic virus. *Virology* 63: 295-303.
- MURPHY, J.F., R.E. RHOADS, A.G. HUNT et J.G. SHAW. 1990. The VPg of tobacco etch virus RNA is the 49-kDa proteinase or the N-terminal 24-kDa part of the proteinase. *Virology* 178: 285-288.

MURPHY, J.F., W. RYCHLIK, R.E. RHOADS, A.G.HUNT et J.G. SHAW. 1991. A tyrosine residue in the small nuclear inclusion protein of tobacco vein mottling virus links the VPg to the viral RNA. *J. Virol.* 65: 511-513.

OH, C.-S. et J.C. CARRINGTON. 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. *Virology* 173: 692-699.

OHKI, S.T. 1992. A potyvirus in nature: indistinct populations. *Arch. Virol* [suppl. 5]: 217-219.

PALMENBERG, A.C. 1990. Proteolytic processing of picornaviral polyprotein. *Ann. Rev. Microbiol.* 44: 603-623.

POCH, O., I. SAUVAGET, M. DELARUE et N. TORDO. 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* 8: 3867-3874.

PONZ, F. et G. BRUENING. 1986. Mecanismos of resistance to plant viruses. *Ann. Rev. Phytopathol.* 24: 355-381.

POWELL, P.A., P.R. SANDERS, N. TUMER, R.T. FRALEY et R.N. BEACHY. 1990. Protection against tobacco mosaic virus infection in transgenic plants requires accumulation of coat protein rather than coat protein RNA sequences. *Virology* 175: 124-130.

RACANIELLO, V.R. et D. BALTIMORE. 1981a. Molecular cloning of poliovirus DNA and determination of the complete nucleotide sequence of the viral genome. *Proc. Natl. Acad. Sci. USA* 78: 4887-4891.

RACANIELLO, V.R. et D. BALTIMORE. 1981b. Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* 214: 916-919.

RESTREPO, M.A., D.D. FREED et J.C. CARRINGTON. 1990. Nuclear transport of plant potyviral proteins. *Plant Cell* 2: 987-998.

RESTREPO-HARTWIG, M.A. et J.C. CARRINGTON. 1992. Regulation of nuclear transport of a plant potyvirus protein by autoproteolysis. *J. Virol.* 66: 5662-5666.

RIECHMANN, J.L., S. LAÍN et J.A. GARCÍA. 1989. The genome-linked protein and 5' end RNA sequence of plum pox potyvirus. *J. Gen. Virol.* 70: 2785-2789.

RIECHMANN, J.L., S. LAÍN et J.A. GARCÍA. 1990. Infectious in vitro transcripts from a plum pox potyvirus cDNA clone. *Virology* 177: 710-716.

RIECHMANN, J.L., S. LAÍN et J.A. GARCÍA. 1991. Identification of the initiation codon of plum pox potyvirus genomic RNA. *Virology* 185: 544-552.

- SHIELDS, S.A. et T.M.A. WILSON. 1987. Cell-free translation of turnip mosaic virus RNA. *J. Gen. Virol.* 68: 169-180.
- SHUKLA, D.D., P.M. STRIKE, S.L. TRACY, K.H. GOUGH et C.W. WARD. 1988. The N and the C termini of the coat proteins of potyviruses are surface-located and the N terminus contains the major virus-specific epitopes. *J. Gen. Virol.* 69: 1497-1508.
- SHUKLA, D.D. et C.W. WARD. 1988. Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *J. Gen. Virol.* 69: 2703-2710.
- SHUKLA, D.D. et C.W. WARD. 1989a. Identification and classification of potyviruses on the basis of coat protein sequence data and serology. *Arch. Virol.* 106: 171-200.
- SHUKLA, D.D. et C.W. WARD. 1989b. Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. *Adv. Virus Res.* 36: 273-314.
- SIAW, M.F.E., M. SHAHABUDDIN, S. BALLARD, J.G. SHAW et R.E. RHOADS. 1985. Identification of a protein covalently linked to the 5' terminus of tobacco vein mottling virus RNA. *Virology* 142: 134-143.
- SMITH, K.M. 1972. Turnip mosaic virus dans "A text book of plant viruses diseases", Academic Press, New-York, 556-560.
- SOUTHERN, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.
- STANLEY, W.M. 1935. Isolation of a crystalline protein possessing the properties of tobacco mosaic virus. *Science* 81: 644-645.
- STARK, D.M. et R.N. BEACHY. 1989. Protection against potyvirus infection in transgenic plants: evidence for broad spectrum resistance. *Bio/technology* 7: 1257-1262.
- STEINHAEUER, D.A. et J.J. HOLLAND. 1987. Rapid evolution of RNA viruses, *Ann. Rev. Microbiol.* 41: 409-433.
- STOBBS, L.W. et I. SHATTUCK. 1989. Turnip mosaic virus strains in southern Ontario, Canada. *Plant Disease* 73: 208-212.
- STOBBS, L.W. et J.G. VAN SCHAGEN. 1987. Occurrence and characterization of a turnip mosaic isolate infecting *Alliaria petiolata* in Ontario, Canada. *Plant Disease* 71: 965-968.
- THORNBURY, D.W., G.M. HELLMANN, R.E. RHOADS et T.P. PIRONE. 1985. Purification and characterization of potyvirus helper component. *Virology* 144: 260-267.

- THORNBURY, D.W., C.A. PATTERSON, J.T. DESSENS et T.P. PIRONE. 1990. Comparative sequence of the helper component (HC) region of potato virus Y and a HC-defective strain, potato virus C. 1990. *Virology* 178: 573-578.
- TINDALL, R.K. et T.A. KUNKEL. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27: 6008-6013.
- TOBIN, J.G., D.C. YOUNG et J.B. FLANEGAN. 1989. Self-catalyzed linkage of poliovirus terminal protein VPg to poliovirus RNA. *Cell* 59: 511-519.
- TURPEN, T. 1989. Molecular cloning of potato virus Y genome: nucleotide sequence homology in non-coding regions of potyviruses. *J. Gen. Virol.* 70: 1951-1960.
- VAN DER VLUGT, R.A.A, R.K. RUITER et R. GOLDBACH. 1992. Evidence for sense RNA-mediated protection to PVY (N) in tobacco plants transformed with the viral coat protein cistron. *Plant Mol. Biol.* 20: 631-639.
- VERCHOT, J.M., E.V. KOONIN et J.C. CARRINGTON. 1991. The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* 185: 527-535.
- WANG, D. et A.J. MAULE. 1992. Early embryo invasion as a determinant in pea of the seed transmission of pea seed-borne mosaic virus. *J. Gen. Virol.* 73: 1615-1620.
- WARD, C.D. et J.B. FLANEGAN. 1992. Determination of the poliovirus RNA polymerase error frequency at eight sites in the viral genome. *J. Virol.* 66: 3784-3793.
- WARD, C.W. et D.D SHUKLA. 1991. Taxonomy of potyviruses: current problems and some solutions. *Intervirology* 32: 269-296.
- WETZEL, T., T. CANDRESSE, M. RAVELONANDRO, R.P. DELBOS, H. MAZYAD, A.E. ABOUL-ATA et J. DUNEZ. 1991a. Nucleotide sequence of the 3'-terminal region of the RNA of the El Amar strain of plum pox potyvirus. *J. Gen. Virol.* 72: 1741-1746.
- WETZEL, T., T. CANDRESSE, M. RAVELONANDRO et J. DUNEZ. 1991b. A polymerase chain reaction assay adapted to plum pox potyvirus detection. *J. Virol. Meth.* 33: 355-365.
- YOUNG, M.J., L. KELLY, P.J. LARKIN, P.M. WATERHOUSE et W.L. GERLACH. 1991. Infectious *in vitro* transcripts from a cloned cDNA of barley yellow dwarf virus. *Virology* 180: 372-379.

ANNEXES

ANNEXE 1

AMORCES OLIGONUCLÉOTIDIQUES: Les nucléotides qui forment un site de restriction sont soulignés. [T]₁₇ représente une queue de 17 résidus T. Les nucléotides dégénérés sont indiqués entre parenthèses. La position du premier et du dernier nucléotide de TuMV (si c'est le cas) sont numérotés en exposant. Les oligonucléotides sont orientés de 5' vers 3'.

FT₀ GGTCTAGAGCTCGAG[T]₁₇

FT_{0XX} GGTCGGTACCCGGG[T]₁₇

FT₂ ⁹¹⁸⁴CGTTCTCAATGCACCAGACC⁹¹⁷⁰

FT₈ ⁸⁰⁷⁶ATACACGAATTCGGTGTAC⁸⁰⁵⁸

FT₁₂ ⁴⁸⁵²CGTAGTTAAAGCTTGTCTTTG⁴⁸³²

FT₁₆ ³⁹⁰⁶AGCTGGAAGCTTCTGGAGCATT³⁸⁸⁵

FT₁₈ ²³⁰²CCTTAGCCTGGGGGGTACCCCTTAACATT²²⁷⁸

FT₂₀ ⁷²²¹CCACGGATCCTGCTAGTTTTG⁷²⁰¹

FT₃₀ ¹⁵¹ATGCAAATGTGACTGCTGC¹³³

FT₅₀ ⁴⁸⁷³CAAGTCTTTGGATCCGCTCC⁴⁸⁵⁴

FT₅₂ ²⁵⁰⁹GAACATGATATCCAGTCGAG²⁴⁹⁰

ON₉ ⁶⁰⁴²CCTGCCTTTGCTCTTCC⁶⁰²⁶

JF₁ ⁸⁷⁸⁵GA(AG)GA(AG)CA(AG)AA(AG)CA(AG)GC⁸⁸⁰¹

JF₃ ⁷⁹³³TA(TC)TG(TC)GA(TC)GC(TC)GATGG(TC)TC⁷⁹⁵²

ANNEXE 1 (Suite)

- JF₅** ⁸⁴⁸⁰GTGGATCCCA(TC)(AC)G(AG)(TC)T(AG)GA(AG)GCAAT
(TC)TG(TC)GC⁸⁵¹⁰
- JF₉** ⁴⁷²¹GCGGATCCAA(TC)AT(TCA)AT(TC)GA(GA)AATGG⁴⁷⁴⁵
- JF₁₃** ³⁷²⁸GACTCGAGAGTGA(TC)(TG)GCGTGT(AT)CAA(GA)A³⁷⁵⁴
- JF₁₅** ⁴A(AT)T(AG)(AT)(AC)AGAATTCAA(CT)ACAACAT²⁶
- JF₁₇** ²²³⁰GC(ACT)GGTACCGG(AT)TATTG(TC)TA(TC)AT²²⁵²
- JF₂₁** ⁵⁷¹⁰TATGGAGCCCATGGAGGCAGTT⁵⁷³⁰
- JF₂₃** ⁵⁸⁷⁸TCTGAACCCCATGGCTCATGAA⁵⁸⁹⁸
- JF₅₁** ²⁴⁹⁰CTCGACTGGATATCATGTTC²⁵⁰⁹
- JF₅₃** ⁴⁸⁵³GGGAGCGGATCCAAAGACT⁴⁸⁷¹
- JF_{SALT7}** GCGGGTCGACTAATACGACTCACTATAG-
¹AAAATATAAAAACTCAACACAACAT²⁶
- PR₇** ⁸⁷²⁸GGTGTGAGGCTTGGATCCGAACCATGGCAGG-
TGAAACG⁸⁷⁶⁶

ANNEXE 2

AUTRE ARTICLE: Cet article démontre une autre utilisation de l'amplification par PCR ainsi que la transformation génétique des plantes, partie à laquelle j'ai collaboré pour cet article.

Laliberté, J.-F., O. Nicolas, S. Durand et R. Morosoli

The xylanase introns from Cryptococcus albidus are accurately spliced in transgenic tobacco plants.

Plant Molecular Biology 1992. 18: 447-451.