THE UNIVERSITY OF

WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/62029

The Deployment and Mechanism of Broad-spectrum Resistance to *Turnip mosaic virus* in *Brassica rapa*

by

Charlotte Florence Nellist

Thesis

Submitted to the University of Warwick
for the degree of

Doctor of Philosophy

Supervisors : John A. Walsh[1], Guy C. Barker[1] and William H. Briggs[2]

[1] School of Life Sciences, University of Warwick, CV35 9EF, UK.
[2] Syngenta Seeds, 1601 BK Enkhuizen, The Netherlands.

School of Life Sciences

September 2013

THE UNIVERSITY OF
WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

First and foremost, I would like to express my gratitude to Dr. John Walsh for his guidance, advice, encouragement and support throughout the project.

I would like to thank Dr. Guy Barker and Dr. Bill Briggs for helpful comments and discussions and also Dr. Carol Jenner for her help at the beginning of this research. I would also like to express my gratitude to all the members of the plant-virus interactions group, past and present, especially Isabel Saunders, Adam Baker and Max Newbert for their help. In addition, I would also like to thank Dr. Julie Jones, Dr. Jay Moore, Dr. Karl Morris, Dr. Rebecca Allen and Gary Grant for their diverse help.

I would also like to thank all the PhD students and research staff based at Wellesbourne for their support and friendship over the last four years.

I would like to thank my family, especially my parents Jane and Dave and also my siblings Bethan, Clara and Joseph and my friend Liam, whose support has been invaluable. Finally, I would like to thank Stephan Page for his continued support and encouragement.

Declarations

This thesis is presented in accordance with the regulations for the degree of Doctor of Philosophy. It has been written by myself and has not been submitted in any previous application for any degree. The work in this thesis has been undertaken by myself except where otherwise stated.

Work has been mentioned in the following publications:

Jenner C.E., <u>**Nellist C.F.**</u>, **Barker G.C.** and **Walsh J.A.** (2010) *Turnip mosaic virus* (TuMV) is able to use alleles of both *eIF4E* and *eIF(iso)4E* from multiple loci of the diploid *Brassica rapa*. *Molecular Plant-Microbe Interactions*, **23**(11), 1198-1505.

WO2011/161466, International patent cooperation treaty (PCT) application published 29th December 2011.

<u>**Nellist C.F.**</u>, **Qian W.**, **Jenner C.E.**, **Moore J.D.**, **Zhang S.**, **Wang X.**, **Briggs W.H.**, **Barker G.C.**, **Sun R.** and **Walsh J.A.** (2014) Multiple copies of eukaryotic translation initiation factors in *Brassica rapa* facilitate redundancy, enabling diversification through variation in splicing and broad-spectrum virus resistance. *The Plant Journal*, **77**(2), 261-268.

<div align="right">Charlotte Florence Nellist</div>

Abstract

The potyvirus *Turnip mosaic virus* (TuMV) is a major constraint on the cultivation of a wide range of plant species worldwide. It causes significant economic losses in brassica species such as Chinese cabbage (*Brassica rapa*), which is one the most important vegetable crops in the world. The *B. rapa* line RLR22 has broad-spectrum resistance to TuMV, which is undefeated. Many recessive resistances against plant viruses in the *Potyvirus* genus are based on mutations in plant eukaryotic translation initiation factor 4E (*eIF4E*), or its isoform *eIF(iso)4E*. *B. rapa* has three *eIF4E* genes and three *eIF(iso)4E* genes. Segregation following a cross between RLR22 and the TuMV-susceptible R-o-18 line of the closely related *B. rapa* ssp. *trilocularis* revealed the resistance was due to a recessive gene, *retr01* that was epistatic to a dominant gene, *ConTR01*. My research revealed that *retr01* is *BraA.eIF(iso)4E.a* and that *ConTR01* is probably *BraA.eIF(iso)4E.c*. It also showed that the highly sought after broad-spectrum resistance to TuMV is due to a novel, recessive, natural mechanism, based on the mis-splicing of *BraA.eIF(iso)4E.a* in *B. rapa*. This results in a range of *eIF(iso)4E* splice variants, the most common of which retained the whole of intron 1 and appears to be non-functional for the virus. As the susceptible parent in the original cross, R-o-18, was a different sub-species to RLR22 (*B. rapa* var. *pekinensis*, Chinese cabbage), the genetic inheritance of resistance was also investigated in crosses with Chinese cabbage lines; F_2 segregation ratios were consistent with those predicted for the single recessive gene (*retr01*). Yeast two-hybrid interactions between the viral protein genome-linked (VPg) of TuMV and *eIF(iso)4E* from *B. rapa* seem to be TuMV isolate-specific. Aphid transmission experiments to investigate the complementation of an *eIF(iso)4E Arabidopsis thaliana* knock-out line with *B. rapa BraA.eIF(iso)4E.a* confirmed the earlier results from mechanical inoculation of these plants. The inability of TuMV to access multiple copies of eIF(iso)4E in Chinese cabbage and the broad-spectrum of the resistance, suggest it may prove to be durable.

Abbreviations

Virus Abbreviations

BaMMV *Barley mild mosaic virus*

BaYMV *Barley yellow mosaic virus*

BCMV *Bean common mosaic virus*

BYMV *Bean yellow mosaic virus*

ChiVMV *Chilli veinal mottle virus*

ClYVV *Clover yellow vein virus*

LMV *Lettuce mosaic virus*

MNSV *Melon necrotic spot virus*

PepMoV *Pepper mottle virus*

PSbMV *Pea seed-borne mosaic virus*

PVMV *Pepper veinal mottle virus*

PVY *Potato virus Y*

RYMV *Rice yellow mottle virus*

TEV *Tobacco etch virus*

TMV *Tobacco mosaic virus*

TuMV *Turnip mosaic virus*

Common Abbreviations

A_{405nm} absorbance at wavelength 405 nm

bp base pair

cDNA complementary deoxyribonucleic acid

CDS coding DNA sequence

dH_2O distilled water

DNA deoxyribonucleic acid

DNase deoxyribonuclease

dSpm defective Suppressor-mutator element

eIF4E eukaryotic translation initiation factor 4E

eIF(iso)4E eukaryotic translation initiation factor isoform 4E

ELISA enzyme-linked immunosorbent assay

FPKM fragments per kilobase of exon per million fragments mapped

gDNA genomic deoxyribonucleic acid

g gram

H_2O water

His histidine

Kb kilobase

L litre

Leu leucine

M molar (moles per litre)

ml millilitre

mM millimolar

mg	milligram
mm	millimetre
nm	nanometres
nt	nucleotide
ORF	open reading frame
PCR	polymerase chain reaction
poly(A) tail	polyadenylated tail
pH	$-\log_{10}[\text{H}^+]$
REML	residual maximum likelihood
RT-PCR	reverse transcriptase polymerase chain reaction
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
rRNA	ribosomal ribonucleic acid
SD	synthetic defined (base media for yeast)
SNP	single nucleotide polymorphism
Tris	tris(hydroxymethyl)aminomethane
Trp	tryptophan
UV	ultra violet
Ura	uracil
V	volts
VPg	viral protein genome-linked
Y2H	yeast two-hybrid
YPDA	yeast peptone dextrose adenine

Chapter 1

General Introduction

1.1 *Brassicaceae*

The family *Brassicaceae* (formerly known as *Cruciferae*) is an economically important family of angiosperm plants. The family consists of over 330 genera and approximately 3,700 species (Warwick *et al.*, 2006), including the genus *Brassica* and the model plant *Arabidopsis thaliana*. The family is considered diverse in distribution, but is mainly concentrated in the northern temperate regions along with a high concentration of diversity around the Mediterranean area. The family comprises mostly of herbaceous plants with annual, biennual, or perennial lifespans.

The *Brassica* genus includes three diploid crop species (*Brassica rapa*, *Brassica oleracea* and *Brassica nigra*) and three amphidiploid crop species (*Brassica napus*, *Brassica juncea* and *Brassica carinata*) (U, 1935), which are interrelated as described in Figure 1.1. Brassica crops worldwide supply the highest diversity of products used by the human race from a single genus (Dixon, 2007). Brassica crops collectively provide leaf, flower and root vegetables, they can be used for condiments as well as fodder and forage and provide sources of edible oil, industrial lubricants and fuel oil (Dixon, 2007).

(Chinese cabbage, Turnip)   (Indian mustard, Chinese mustard)   (Black mustard)

Brassica rapa   **Brassica juncea**   **Brassica nigra**
n=10   n=18   n=8
AA   AABB   BB

Brassica napus   **Brassica carinata**
(Oilseed rape, Swede)   n=19   n=17   (Ethiopian mustard)
AACC   BBCC

Brassica oleracea
n=9
CC
(Cabbage, Broccoli, Brussels sprouts, Cauliflower, Kale)
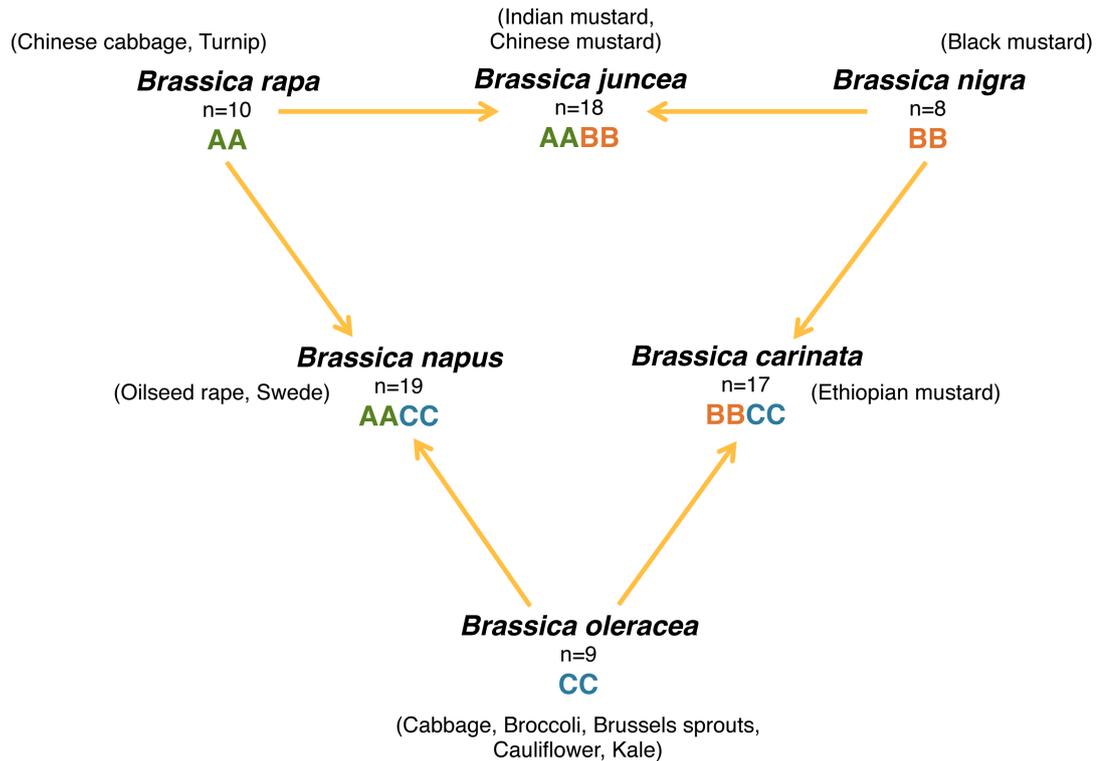
Figure 1.1 – Triangle of U, showing the genetic relationships between the six crop species of the genus *Brassica* and common crops from each species, adapted from U (1935).

1.1.1 *Brassica rapa*

The diploid species *B. rapa* (AA genome; Figure 1.1) has 10 chromosomes. Comparative physical mapping studies of Arabidopsis and *B. rapa* ssp. *pekinensis* confirmed genome triplication and rearrangements in Chinese cabbage (Park *et al.*, 2005; Wang *et al.*, 2011).

B. rapa is a diverse species, that includes seven *B. rapa* vegetable types; var. *campestris* (e.g. Field mustard), var. *chinensis* (e.g. Pak choi), var. *japonica* (e.g. Mizuna), var. *narinosa* (e.g. Tatsoi), var. *parachinenesis* (e.g. Yu choy sum), var. *pekinensis* (e.g. Chinese cabbage) and var. *rapa* (e.g. Turnip) (Dixon, 2007). In Japan and China, *B. rapa* crop plants are used as leaf or root vegetables, whereas in India they were developed as oil plants (Dixon, 2007).

Chinese Cabbage

Chinese cabbage (*B. rapa* var. *pekinensis*) is a leaf vegetable and is native of China. It forms distinct heads of leaves and is often used as a salad vegetable. It

is thought that it evolved from the natural crossing of Pak Choi (*B. rapa* var. *chinensis*) and turnip (*B. rapa* var. *rapa*), with much of the diversity within the varieties developing over the last 600 years (Dixon, 2007). It was introduced into Korea, Japan and other south-eastern countries and today is grown worldwide (Dixon, 2007), including in Europe and America (McNaughton, 1995). Chinese cabbage is a fundamental part of many Asian dishes.

1.2 Plant Viruses

Plant viruses are important plant pathogens, not only as model agents for studying interactions with their hosts, but also economically. It is estimated that plant viral pathogens cause US$60 billion loss in crop yields worldwide per annum (Klausner, 1987).

A virus is defined as an infective, intracellular, obligate parasite, consisting of two parts, the genome (nucleic acid) and protein which forms the protective shell (capsid). The capsid structure can take the form of two fundamental types of symmetry. The first type of symmetry is helical (elongated) and incorporates the rigid rod-shaped particles (e.g. *Tobacco mosaic virus*) and the filamentous flexuous rod-shaped particles (e.g. *Potato virus Y*). The second type of symmetry is icosahedral (spherical) and incorporates isometric shaped particles (e.g. *Tobacco necrosis virus*), the bacilliform (short round-ended rods) shaped particles (e.g. *Cocoa swollen shoot virus*) and geminate (twinned isometric) shaped particles (e.g. *Maize streak virus*). Plant viruses belong to three orders of viruses and there are 21 families which have plant virus members (King *et al.*, 2012).

Plant viral genomes are either circular or linear and can be monopartite (made up of a single component) or multipartite (made up of more than one component). Plant viruses can be classified by the nature of their genome; single-stranded DNA (ssDNA), reverse transcribing viruses, double-stranded RNA viruses (dsRNA), negative sense single-stranded RNA (ssRNA-) and positive sense single-stranded RNA (ssRNA+). Most plants viruses are transmitted via vectors as the plant cell wall offers robust protection, which most viruses cannot penetrate. Aphids (Aphididae) are considered the most important sap-feeding vectors, along with nematodes, fungi and other insects making up common vectors (Katis *et al.*, 2007).

1.2.1 *Potyviridae*

The *Potyviridae* is the largest family of plant viruses and is comprised of eight genera: *Brambyvirus*, *Bymovirus*, *Ipomovirus*, *Macluravirus*, *Poacevirus*, *Potyvirus*, *Rymovirus*, *Tritimovirus* (King *et al.*, 2012). All members of the *Potyviridae* family have flexuous filamentous rod-shaped particles with positive-sense single-stranded RNA genomes (Shukla *et al.*, 1994). All members possess monopartite genomes apart from the genus *Bymovirus*, which possesses a bipartite genome (King *et al.*, 2012). The members of the *Potyviridae* have a viral protein genome-linked (VPg) covalently attached at the 5' end of the RNA and a polyadenylated (poly(A)) tail at the 3' end. In addition, all members of the *Potyviridae* family induce the formation of cylindrical inclusion bodies, a characteristic of the family (López-Moya *et al.*, 2009).

1.3 *Turnip mosaic virus*

Turnip mosaic virus (TuMV) is a member of the *Potyvirus* genus, which is part of the family *Potyviridae* (Shukla *et al.*, 1994). The *Potyvirus* genus contains 36% of all known plant viruses and is responsible for enormous losses in a variety of crops (Ward and Shukla, 1991). TuMV has the widest host range of any of the potyviruses and is the only potyvirus known to infect brassicas (Tomlinson, 1987; Shukla *et al.*, 1994; Walsh and Jenner, 2002).

TuMV is a huge constraint on the cultivation of a wide range of plant species worldwide. It causes significant economic losses in *Brassica* species (Walsh *et al.*, 2002). Tomlinson (1987) described the virus as being the second most significant virus infecting field-grown vegetables worldwide. The virus is particularly damaging in brassicas in parts of Europe, Asia and North America (Walsh and Jenner, 2002). TuMV affects both the yield and quality of commercially grown brassica crops, is described as the most important virus affecting brassicas in many Asian countries (Yoon *et al.*, 1993) and considered more important than any fungal or bacterial pathogen in some regions.

1.3.1 Classification of Isolates

Several systems have been developed for classifying TuMV isolates into unique groups based on host range and symptoms produced in indicator plants. Ear-

lier systems, for example, six pathotypes (C1-C6) have been described based on interactions with Chinese cabbage (*B. rapa*); C1-C4 were described by Provvidenti (1980), C5 by Green and Deng (1985) and C6 by Stobbs and Shattuck (1989). Seven pathotypes (Tu1-Tu7) were described from nineteen isolates identified from ten regions of China (Liu *et al.*, 1990a,b). Whereas in *B. napus*, twelve distinctive pathotypes (1-12) have been characterised based on interactions with four *B. napus* differentials (Walsh, 1989; Jenner and Walsh, 1996; Table 1.1). The latter system is considered more comprehensive due to it describing a greater number of TuMV isolates from around the world and characterisation of the resistance genes in the differentials and the viral determinants of pathogenicity (Walsh and Jenner, 2002).

Table 1.1 – Interactions of *Turnip mosaic virus* pathotypes with differential lines of *Brassica napus* (oilseed rape and swede) (Jenner and Walsh, 1996).

Virus pathotype	Plant line			
	Rape S6	Rape R4	Swede 165	Swede S1
1	$+^a$	0^b	0	+
2	R^c	R	0	R
3	+	$+_N$	0	+
4	+	+	+	+
5	+	+	0	+
6	+	$+_N$	0	R
7	+	0	0	R
8	+	$+_N$	R_N	R
9	+	R_N	0	R
10	+	+	0	0
11	R	+	0	R
12	+	+	$+_N$	+

[a] +, systemic infection, plants were susceptible.
[b] 0, no infection, plants appeared to be immune.
[c] R, local infection, no systemic spread.
Local symptoms were chlorotic and systemic symptoms were mosaic unless indicated by $_N$ (necrotic).

1.3.2 Symptoms

Symptoms in infected plants include leaf curling, stunted growth, black necrotic lesions, severe chlorotic local lesions, systemic vein clearing, chlorotic mottling and the classical systemic mosaic patterning in the leaves. These symptoms result in the plants being visually undesirable and often unmarketable. Different pathotypes of the virus have varying effects on the diverse crop species of brassica and on each cultivar.

Studies on TuMV infecting winter oilseed rape (*B. napus* ssp. *oleifera*) showed that the seed pods of infected plants can be reduced in number as well as size, depending on the susceptibility to TuMV and some can be malformed and seedless (Walsh and Tomlinson, 1985). The size of individual seeds and seed yield are reduced and seed viability can also be affected in *B. napus* ssp. *oleifera* (Walsh and Tomlinson, 1985).

1.3.3 Controlling the Spread of TuMV

Controlling the spread of TuMV is very difficult as the virus is transmitted in a non-persistent manner by aphid vectors. Cultural control methods such as the removal of TuMV-infected plant material may help to reduce the virus inoculum and restrict the spread of the virus. The scheduling of planting dates to avoid peak aphid migration periods, allowing susceptible plants to develop as much as possible before being exposed to infection, may also help to reduce the impact of the virus (Shattuck, 1992). Unlike fungal and bacterial pathogens, presently there are no direct methods for controlling viruses. Chemicals do not protect the plants from TuMV infection. Chemical control of aphids is impractical and not sufficient to control the virus, as virulliferous aphids are able to migrate and spread the virus before insecticides are able to take effect. Tolerant plants may be used but are not preferred as they can still serve as reservoirs of TuMV. The use of immune, or highly resistant cultivars continues to be the cheapest and most successful mode of protecting plants against losses in productivity from TuMV infection (Shattuck, 1992).

1.3.4 Genome Structure and Organisation

TuMV has flexuous filamentous rod-shaped particles (Figure 1.2), typical of the family *Potyviridae*, with a positive-sense, single-stranded RNA genome (Shukla *et al.*, 1994).
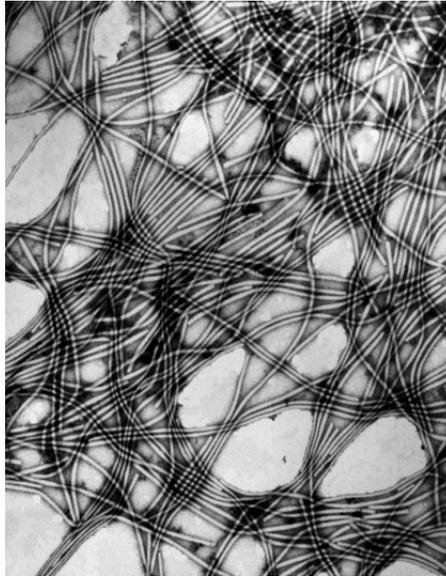
Figure 1.2 – Electron micrograph of the long (720 x 15-20 nm) flexuous filamentous rod-shaped particles of *Turnip mosaic virus* (Colin Clay).

Infection of plants requires the interaction between host factors and viral proteins, as well as RNA for replication and systemic spread (Hwang *et al.*, 2009). The genomic RNA has two roles, it provides the template for replication as well as being the messenger RNA (mRNA) for protein synthesis (Walsh and Jenner, 2002). The genome of TuMV is comprised of approximately 9830 nucleotides (nt), with a poly(A) tail at the 3' end and a 22 kDa VPg at the 5' end (Ohshima *et al.*, 1996; Walsh and Jenner, 2002; Figure 1.3). The large open reading frame (ORF) is flanked by two untranslated regions (UTR). The ORF is translated into a large polyprotein, which is post-translationally cleaved into ten mature proteins (Figure 1.3) by virally encoded proteases. The mature proteins are the 40 kDa P1, 52 kDa helper component protease (HC-Pro), 40 kDa P3, 6 kDa 6K1 (which may be attached to P3), 72 kDa cylindrical/cytoplasmic inclusion protein (CI), 6 kDa 6K2, 22 kDa VPg, 27 kDa nuclear inclusion protein a (NIa, which may be attached to VPg), 60 kDa nuclear inclusion protein b (NIb) and the 33 kDa coat protein (CP) (Walsh and Jenner, 2002) (Figure 1.3). An overview of each protein and its function is described in Table 1.2.

7
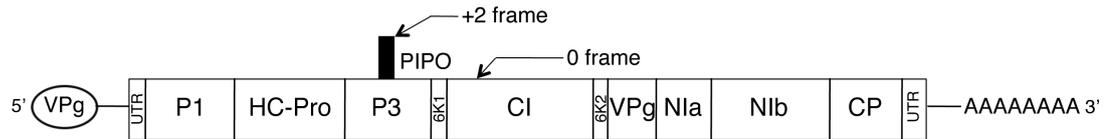
Figure 1.3 – Schematic representation of *Turnip mosaic virus* genome, showing the position of the overlapping coding region within the P3 cistron, the viral protein genome-linked (VPg) covalently attached to the 5' end of the genome and the polyadenylated tail at the 3' end (adapted from Chung *et al.*, 2008).

Table 1.2 – Functions of mature potyviral proteins.

Name	Function	References
P1	Proteinase	Carrington *et al.*, 1990
	Genome amplification	Verchot and Carrington, 1995
	Suppressor of gene silencing	Anandalakshmi *et al.*, 1998
HC-Pro	Aphid transmission	Pirone and Thornbury, 1984
	Proteinase	Carrington *et al.*, 1989
	Systemic movement	Rojas *et al.*, 1997
	Suppressor of gene silencing	Anandalakshmi *et al.*, 1998
P3	Genome amplification	Rodríguez-Cerezo *et al.*, 1993
	Avirulence gene	Jenner *et al.*, 2003
6K1	RNA replication	Riechmann *et al.*, 1992
CI	ATPase/RNA helicase	Laín *et al.*, 1990
	Cell-to-cell movement	Carrington *et al.*, 1998
	Avirulence gene	Jenner *et al.*, 2000
6K2	Virus replication	Restrepo-Hartwig and Carrington, 1994
	Long-distance movement	Rajamäki and Valkonen, 1999
VPg	Interaction with eIF4E and eIF(iso)4E	Léonard *et al.*, 2000
	Replication	Schaad *et al.*, 1997
	Cell-to-cell and systemic movement	Schaad *et al.*, 1997
NIa	Proteinase	Dougherty *et al.*, 1989
NIb	RNA-dependent RNA polymerase	Hong and Hunt, 1996
CP	Encapsidation	Shukla and Ward, 1989
	Aphid transmission	Atreya *et al.*, 1990
	Cell-to-cell and systemic movement	Dolja *et al.*, 1994
	Virus assembly	Dolja *et al.*, 1994
	Genome amplification	Mahajan *et al.*, 1996
PIPO	Virus movement	Wei *et al.*, 2010

It was believed up until recently that the virus consisted of a single ORF. However, a short overlapping coding sequence (CDS), *pipo* (Pretty Interesting *Potyviridae* ORF; Figure 1.3) encoding the PIPO protein which is present within the P3 cistron, was discovered recently (Chung *et al.*, 2008). Chung *et al.* (2008) observed that the PIPO protein is not expressed independently but rather as a fusion protein with the N-terminal of P3 (P3N-PIPO). Wei *et al.* (2010) demonstrated that the P3N-PIPO of TuMV is a plasmodesmata (PD)-located protein that physically interacts with the CI protein *in planta*. They suggest that the intercellular movement of potyviruses in infected plants is facilitated by PD-associated structures, which is coordinated by the CI protein and the P3N-PIPO complex. The recent discovery of *pipo* suggests that there may be other short CDSs present within the TuMV genome that have yet to be discovered (Chung *et al.*, 2008).

Viral Protein Genome-linked (VPg)

The VPg protein has an intrinsically disordered structure, enabling it to bind several proteins (Jiang and Laliberté, 2011). The VPg is described as a hub protein that plays a pivotal role in many phases of potyvirus infection. It is able to control many processes leading to virus production and spread, including playing a key role in viral RNA translation, replication, viral movement and possibly host gene regulation (Jiang and Laliberté, 2011). The potyviral VPg is versatile and has been linked with interacting with many proteins, most significantly, host factors involved in protein synthesis within viral replication factories, or within the nucleus. VPgs have been found to also be linked to the genomic RNA of several vertebrate viruses. However, they have been shown not to possess any homology in their amino acid sequence with the plant virus VPgs, although, they have been shown to share several properties (Jiang and Laliberté, 2011).

1.3.5   TuMV Life-Cycle

Viruses move through the plant by exploiting pre-existing pathways for macromolecular movement (Carrington *et al.*, 1996). Viruses move from cell-to-cell through plasmodesmata and utilise the plant's vasculature to travel long-distance (Carrington *et al.*, 1996). Once the virus is in a cell, the viral particle is uncoated. Genome replication occurs in the cytoplasm in close proximity with membrane surfaces (Carrington *et al.*, 1996).

Viruses are not capable of completing their biological cycle on their own and so they recruit the host's biochemical machinery to carry out their cycle. Potyviruses such as TuMV utilise the eukaryotic initiation factor complex to complete its infection cycle.

Acquisition and Transmission of Potyviruses by Aphids

TuMV is acquired from (and also introduced into) plant cells via the stylet of aphids when probing or feeding and is transmitted in a non-persistent stylet-borne manner (Walsh and Jenner, 2002). The helper component protein (HC, encoded by HC-Pro; Figure 1.3) aids transmission of potyviruses by acting as a 'bridge' between the viral coat protein (CP) and the aphid stylet, allowing the retention of the virions in the mouthpart ready for inoculation (Govier and Kassanis, 1974). Located within the N-terminal cysteine-rich domain of HC-Pro is a lysine motif (KITC), which along with another motif PTK, have been shown to be crucial for aphid transmission (Pirone and Blanc, 1996).

It is reported that at least 89 species of aphid can transmit TuMV (Edwardson and Christie, 1986). Aphids usually remain infective for a short period of time after acquiring TuMV; in *Myzus persicae*, the peach-potato aphid, which is notorious for spreading TuMV, this is typically 3-5 hours (Sylvester, 1954). Once the primary infections are established in the field, TuMV can be spread fairly quickly from plant to plant if aphids are not controlled (Shattuck, 1992). TuMV spread can be affected by aphid activity and migration patterns, which are influenced by temperature and weather conditions. Global warming is likely to have an effect on aphid reproduction and survival, which in turn could increase virus spread and epidemics.

Potyvirus Movement

Systemic infection of plants relies on the intercellular transport of the virus through cytoplasmic connections, PD. Potyviruses do not possess a single designated movement protein (MP), but rather several viral proteins have been reported to have MP-related roles. HC-Pro and CP have been shown to increase the size exclusion limit (SEL) of PDs (Rojas *et al.*, 1997). CP and CI are also required for virus intercellular movement and have been shown to be associated with PD (Carrington *et al.*, 1998; Roberts *et al.*, 1998). The recently discovered potyviral protein, P3N-PIPO has been shown to interact with the CI protein *in planta* and

is involved in the cell-to-cell movement of viral particles, by targeting CI to PD (Wei *et al.*, 2010). P3N-PIPO is targeted to the plasma membrane through an interaction with the host protein PCaP1 (Vijayapalani *et al.*, 2012). PCaP1 is a cation-binding protein that attaches to the plasma membrane via myristoylation.

1.4 Eukaryotic Translation Initiation

The process of protein synthesis in higher plants is a very complex procedure facilitated by numerous macromolecules (Browning, 1996). The first step in the process is the transcription of DNA into mRNA, this occurs in the nucleus of the cell. The mature mRNAs are then transported into the cytoplasm for translation into a polypeptide chain. In eukaryotes, translation initiation is thought to occur by one of two mechanisms, a cap-dependent mechanism, involving the eukaryotic translation initiation factors, or a cap-independent mechanism, involving ribosomes which are recruited to an internal ribosome entry site (IRES) in the cellular mRNA. Once translation is initiated, the polypeptide chain is elongated until a stop codon in the mRNA is reached, indicating termination of protein synthesis.

1.4.1 Cap-dependent Initiation

The cap-dependent method of initiation is the most common method. In this mechanism, translation is initiated through the interaction of the cap structure (7-methyl-guanosine-containing cap) at the 5' end of the mRNA and the eukaryotic initiation factor 4E (eIF4E), known as the cap binding protein. eIF4E is unique with a high content of tryptophan (Trp) residues; the conservation of the position and number of Trp residues is thought to be involved in the binding of the cap structure of mRNA (Browning, 1996). The complex consists of numerous proteins that facilitate eukaryotic translation (Browning, 1996; Figure 1.4). eIF4E binds to the eIF4G (forming the eIF4F complex), which provides scaffolding for other members of the initiation complex (Browning, 1996). The interaction of eIF4F and the mRNA facilitates binding to the 40S ribosomal subunit through the eIF3 complex and eIF4G scaffold. The poly(A) binding protein (PABP) binds the poly(A) tail of the mRNA enabling 'circularisation' of the transcript.
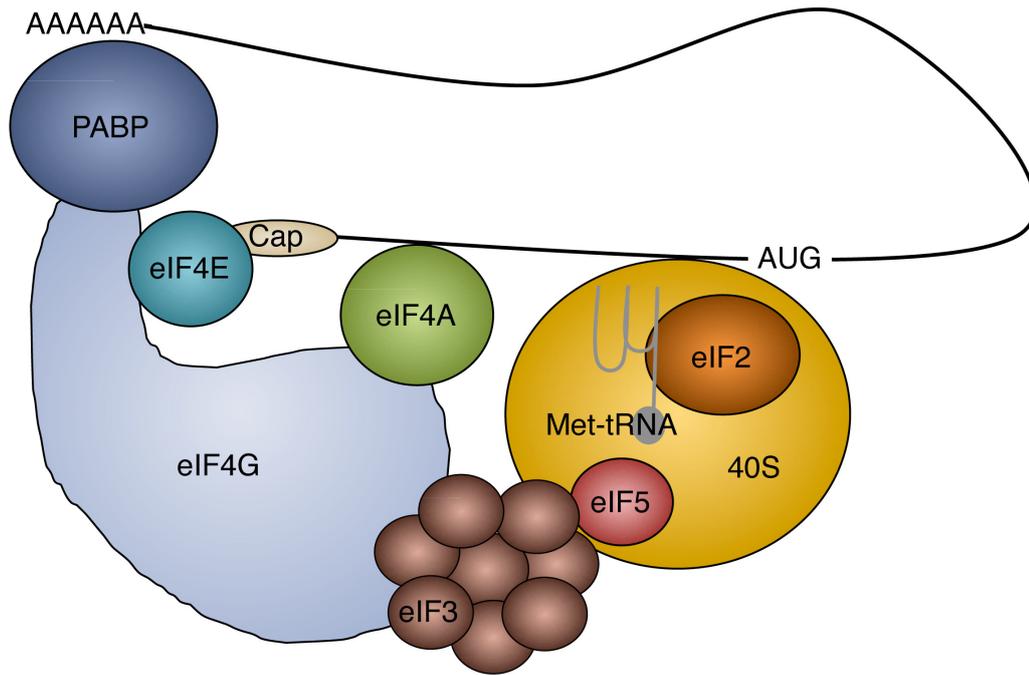
Figure 1.4 – Eukaryotic translation initiation complex, showing the interactions of the cap structure of mRNA and eukaryotic initiation factor 4E (eIF4E) and the polyadenylated tail (poly(A)) and the poly(A) binding protein (PABP) (adapted from Robaglia and Caranta, 2006).

Plants possess more than one family of eIF4E, this was first discovered by Browning *et al.* (1987) when looking at wheat germ. It was assumed that eIF4E was a single protein, but it was later revealed that there were two forms of eIF4E present; named eIF4E and eIF(iso)4E (Browning *et al.*, 1987). Over ten years later this was supported by research on Arabidopsis which confirmed that not only the family of isoforms were present in species other than wheat, but that Arabidopsis also contained a similar novel cap-binding protein (nCBP) (Ruud *et al.*, 1998). It is now recognised that nearly all eukaryotes express multiple *eIF4E* family members (Rhoads, 2009). Other examples of initiation factors with multiple family members are eIF4A and eIF4G. eIF(iso)4E binds with eIF(iso)4G to form eIF(iso)4F, which has a role comparable to eIF4F. eIF4E and eIF(iso)4E are similar in their amino acid sequence, in wheat and rice, showing approximately 50% similarity (Browning, 1996). Both proteins are approximately 24 kDa in molecular mass, compared to eIF4G and eIF(iso)4G which are very different in mass, with eIF4G being much larger, 180 kDa compared to 86 kDa for eIF(iso)4G (Gallie and Browning, 2001). Gallie and Browning (2001) found that unstructured mRNAs are preferentially translated by eIF(iso)4F, whereas mRNAs with a structured 5' leader that are either uncapped or contain multiple cistrons are translated by eIF4F.

Viruses hijack the host's machinery to complete their life cycle. The potyvirus VPg plays a role in viral translation and replication. The potyvirus VPg mimics the 5' cap of mRNA and interacts with eIF4E or eIF(iso)4E, utilising the host's initiation complex to initiate viral translation. This process may also inhibit host cell translation; Léonard *et al.* (2000) suggested that TuMV VPg and host cellular mRNAs compete for eIF(iso)4E in Arabidopsis. The role of eIF4E and eIF(iso)4E in the potyvirus life-cycle and recessive resistance has recently been reviewed by Robaglia and Caranta (2006).

1.4.2   Cap-independent Initiation

The second mechanism of translation initiation of mRNA and/or viruses, involves the recruitment of ribosomes to an IRES, which is a nucleotide sequence that allows translation initiation in the middle of a mRNA sequence. It is most common that IRESs are located in the 5'-UTR of the RNA and allow translation initiation in circumstances when cap-dependent translation is inhibited, or restricted (Reviewed by Filbin and Kieft, 2009).

1.5   Resistance to TuMV

Most types of resistance to plant pathogens are dominant and specific against a narrow range of isolates/strains. The first brassica resistance gene to be mapped was from the A genome of *B. napus*, *TuMV RESISTANCE IN BRASSICA 01* (*TuRB01*). It provides extreme resistance to pathotype 1 isolates (Walsh *et al.*, 1999), but is overcome by single nucleotide mutations in the CI protein of TuMV, A5056G and A5570G (Jenner *et al.*, 2000). In total, six *TuRB0* genes (including *TuRB01*) responsible for dominant strain-specific resistance have been discovered. *TuRB01b*, provides extreme resistance to pathotype 1 isolates in *B. rapa* (Rusholme, 2000) and is also overcome by the same single nucleotide mutation (A5770G) in the CI protein of TuMV as the *TuRB01* based resistance in *B. napus* (Walsh *et al.*, 2002). *TuRB02* was identified on the C-genome of *B. napus* and provides control of the degree of susceptibility to the TuMV pathotype 1 isolate CHN 1 in a quantitative manner (Walsh *et al.*, 1999). *TuRB03* was also identified in *B. napus* and provides resistance to some pathotype 3 and some pathotype 4 TuMV isolates (including isolate CDN 1)

(Hughes *et al.*, 2003), the avirulence determinant was found to be a single amino acid in the P3 protein of TuMV (Jenner *et al.*, 2003). *TuRB04* and *TuRB05* were also identified in *B. napus* and provide resistance to TuMV isolates from pathotypes 1 and 3, the avirulence determinants were found to be the P3 and CI proteins respectively (Jenner *et al.*, 2002a).

However, potentially durable, broad-spectrum resistance has recently been discovered and characterised in the *B. rapa* plant line RLR22. *B. rapa* ssp. *pekinensis* (Chinese cabbage) line RLR22 shows broad-spectrum resistance to TuMV (Walsh *et al.*, 2002). *B. rapa* L. ssp. *trilocularis* (Roxb.) Hanelt. (yellow sarson) line R-o-18 is extremely susceptible to a wide range of TuMV isolates (Walsh *et al.*, 2002). A summary of the phenotypes of these two plant lines to a selection of TuMV isolates is shown in Table 1.3. The TuMV-resistant line RLR22 was crossed with the TuMV-susceptible line R-o-18. Both dominant and recessive resistance genes have been reported in association with TuMV in *B. rapa* (Rusholme *et al.*, 2007). Rusholme *et al.* (2007) showed that the broad-spectrum resistance to TuMV in *B. rapa* RLR22 is controlled by two genes. The recessive gene, <u>recessive</u> <u>T</u>uMV <u>resistance</u> <u>01</u> (*retr01*) is located on the upper part of chromosome A4. The epistatic gene <u>Conditional</u> <u>T</u>uMV <u>resistance</u> <u>01</u> (*ConTR01*), which possesses a dominant allele is located on the upper part of chromosome A8.

Table 1.3 – Phenotypes of *Brassica rapa* lines following challenge with *Turnip mosaic virus* isolates (Walsh *et al.*, 2002).

B. rapa	Virus isolate (pathotype)						
	UK 1 (1)	CHN 5 (3)	CDN 1 (4)	JPN 1 (7)	DEU 7 (8)	GK 1 (9)	UK 4 (12)
R-o-18	$+^a$	+	+	+	+	+	+
RLR22	R^b	R	R	0^c	R	R	R

[a] Systemic spread.
[b] Infection of inoculated leaves, but no systemic spread.
[c] No infection.

1.6 Recessive Resistance

More than half of the resistance genes that have been discovered for potyviruses are recessive; this is considerably higher than recorded for other plant viruses, for

which only approximately 20% of the resistance genes are recessive (Provvidenti and Hampton, 1992). It is believed that recessive resistance to potyviruses is based on a passive mechanism (Diaz-Pendon *et al.*, 2004). Passive resistance can be defined as the lack of activity required from the plant (Fraser, 1986). Fraser postulated that resistant plants do not possess a specific host factor that is required by the virus to complete its life cycle. This is in contrast to dominant resistance, in which recognition of the virus by the plant triggers an active response, for example the hypersensitive response which leads to resistance. In recent years many plant species have been characterised with natural mutations in members of the translation initiation complex, which leads to resistance to particular RNA viruses (Robaglia and Caranta, 2006). A selection of these resistance genes are shown in Table 1.4.

1.6.1 Plant Resistance to Viruses Mediated by the Interaction of VPgs and Translational Initiation Factors

Recessive resistance to potyviruses has been reported to be associated with the lack of interaction of the VPg and the host's translation initiation complex (Robaglia and Caranta, 2006). Mutations in *eIF4E/eIF(iso)4E* have resulted in the VPg being unable to interact with host's translation machinery and ultimately unable to complete it's life-cycle. The first examples of resistances based on the *eIF4E* gene family were by Wittmann *et al.* (1997) and Léonard *et al.* (2000) on Arabidopsis and TuMV and work by Ruffel *et al.* (2002) on *Capsicum annum*, *Potato virus Y* (PVY) and *Tobacco etch virus* (TEV). Since the original discoveries, there have been numerous reports of recessive resistance to plant viruses, for example the allele *mo1* in lettuce which confers resistance to *Lettuce mosaic virus* (LMV) (Nicaise *et al.*, 2003) and the alleles *pvr1/pvr2* in *Capsicum* species which confer resistance to PVY (Kang *et al.*, 2005a). Further examples of natural recessive virus resistance genes associated with translation initiation factors are shown in Table 1.4. These genes control diverse resistance phenotypes from restriction of cell-to-cell movement, to complete restriction of virus accumulation (Robaglia and Caranta, 2006). This type of resistance is not exclusive to potyviruses and has been associated with the closely related *Bymovirus* genus (Kanyuka *et al.*, 2005) in the *Potyviridae* family and also to the more distantly related genera of *Carmovirus* (Nieto *et al.*, 2006) and *Sobemovirus* (Albar *et al.*, 2006). Caliciviruses also possess a VPg and some have been shown to utilise eIF4E to translate their genome in mammals (Goodfellow and Roberts,

Table 1.4 – Examples of natural recessive virus resistance genes associated with translation initiation factors (adapted from Le Gall et al., 2011).

Gene	Plant species	Virus	Translation factor	Reference(s)
nsv	Melon (*Cucumis melo*)	MNSV	eIF4E	Nieto et al. (2006)
mo1	Lettuce (*Lactuca sativa*)	LMV	eIF4E	Nicaise et al. (2003)
pvr1/pvr2	Pepper (*Capsicum* spp.)	PVY, TEV, PepMoV	eIF4E	Ruffel et al. (2002, 2006); Kang et al. (2005a); Charron et al. (2008)
sbm1/wlv/cyv-2	Pea (*Pisum sativum*)	PsBMV, BYMV, ClYVV	eIF4E	Andrade et al. (2009); Gao et al. (2004); Bruun-Rasmussen et al. (2007)
bc-3	Bean (*Phaseolus vulgaris*)	BCMV	eIF4E	Naderpour et al. (2010)
rym4/rym5/rym6	Barley (*Hordeum vulgare*)	BaMMV, BaYMV	eIF4E	Kanyuka et al. (2005)
pot1	Wild tomato (*Solanum habrochaites*)	PVY, TEV	eIF4E	Ruffel et al. (2005)
pvr6 (+pvr2)	Pepper (*Capsicum* spp.)	PVMV, ChiVMV	eIF(iso)4E	Ruffel et al. (2006)
sbm2	Pea (*Pisum sativum*)	PSbMV	eIF(iso)4E	Gao et al. (2004)
retr01	Chinese cabbage (*Brassica rapa*)	TuMV	eIF(iso)4E	Rusholme et al. (2007)
rymv1	Rice (*Oryza sativa*)	RYMV	eIF(iso)4G	Albar et al. (2006)

2008).

In most plant-potyvirus interactions described, the VPg is the avirulence determinant (for example Keller *et al.*, 1998; Ayme *et al.*, 2006). Although other viral genes may be involved in these recessive resistances (P3, Hjulsager *et al.*, 2006 and CI, Abdul-Razzak *et al.*, 2009). The precise role of the VPg is still unclear (Le Gall *et al.*, 2011). The maturation process leads to different VPg-containing polypeptides, including VPg-Pro and 6K-VPg-Pro (Léonard *et al.*, 2004; Beauchemin *et al.*, 2007; Thivierge *et al.*, 2008). The VPg-Pro is targeted to the nucleus, where it can interact with cellular factors including eIF4E, PABP and eIF1A (Léonard *et al.*, 2004; Thivierge *et al.*, 2008). The membrane-associated 6K-VPg-Pro is targeted to the endoplasmic reticulum, to the cytoplasmic vesicles, where it interacts with viral and cellular factors including eIF4E (Beauchemin *et al.*, 2007). The VPg appears to be able to interact with multiple host and viral proteins (Truniger and Aranda, 2009).

The majority of eIF4E-mediated potyvirus recessive resistances are a result of one to five non-conservative amino acid substitutions in the protein (Charron *et al.*, 2008). The amino acid changes are clustered in two neighbouring regions of eIF4E, located near the cap-binding pocket and at the surface of the protein (Robaglia and Caranta, 2006). Mutational analysis of eIF4E from lettuce suggested the role of eIF4E in the LMV cycle and its physiological role in host mRNA translation is possibly discrete (German-Retana *et al.*, 2008). Mutational analysis of eIF4E from pea in bimolecular fluorescence complementation and expression in an eIF4E-defective yeast strain supported the view that the binding site for the VPg of the potyvirus *Pea seed-borne mosaic virus* (PSbMV) overlapped with that of the mRNA cap (Ashby *et al.*, 2011).

Mutations in the central region of the VPg protein, which are thought to be exposed on the surface of the protein have been associated with eIF4E/eIF(iso)4E-resistance breaking (Roudet-Tavert *et al.*, 2007). This implies that this central region of the VPg could be involved in interacting with host eIF4E/eIF(iso)4E (Truniger and Aranda, 2009).

1.7 TuMV and Eukaryotic Initiation Factor Interactions

It was first discovered that plant viruses directly interacted with the eukaryotic translation initiation complex of plants in yeast two-hybrid assays (Wittmann *et al.*, 1997). Theses studies showed that the VPg of TuMV bound to the eIF(iso)4E of Arabidopsis (Wittmann *et al.*, 1997). A single amino acid mutation (D77N) in the TuMV VPg abolishes this interaction with Arabidopsis and results in the virus being unable to infect *B. rapa* var. *perviridis* (Léonard *et al.*, 2000). Léonard *et al.* (2000) showed that the VPg-eIF(iso)4E interaction is critical to preserve virus infectivity *in planta*. Further studies confirmed that VPg of TuMV does interact *in planta* with *B. rapa* var. *perviridis* eIF4E/eIF(iso)4E (Léonard *et al.*, 2004).

1.7.1 TuMV and Arabidopsis Interactions

Arabidopsis possesses three genes encoding eIF4E (eIF4E1: *At4g18040*; eIF4E2: *At1g29550* and eIF4E3: *At1g29590*), one encoding eIF(iso)4E (*At5g35620*) and another related gene encoding a novel cap-binding protein (*At5g18110*) (Ruud *et al.*, 1998). Arabidopsis possesses a single gene encoding eIF4G (*At3g60240*) and two genes encoding the eIF(iso)4G subfamily proteins (eIF(iso)4G1: *At5g57870*; eIF(iso)4G2: *At4g24050*) (Robaglia and Caranta, 2006). A genetic screen for ethyl methane sulfonate (EMS)-induced mutations in Arabidopsis identified eIF(iso)4E as a requirement for TuMV susceptibility (Lellis *et al.*, 2002) and a transposon knock-out of *eIF(iso)4E* (Col-0::*dSpm*) led to potyvirus resistance, particularly TuMV and LMV (Duprat *et al.*, 2002). In contrast, Arabidopsis lines with T-DNA insertions in *eIF4E* (Sato *et al.*, 2005), *eIF4G*, *eIF(iso)4G1* or *eIF(iso)4G2* (Nicaise *et al.*, 2007) remained susceptible to TuMV. However the plants with the double mutants, *eIF(iso)4G1* and *eIF(iso)4G2*, showed resistance to TuMV (Nicaise *et al.*, 2007). Single amino acid changes in the VPg of TuMV overcome the *eIF(iso)4E* and *eIF(iso)4G1* and *eIF(iso)4G2* resistances in Arabidopsis (Gallois *et al.*, 2010).

The eIF4E and eIF(iso)4E proteins are thought to have distinctive roles in plant development and metabolism, eIF4E is found in all tissues except the root specialisation zones and eIF(iso)4E is markedly abundant in floral tissues and young tissue (Rodriguez *et al.*, 1998). Studies in maize showed that eIF4E and

eIF(iso)4E have different regulatory control pathways (Dinkova *et al.*, 2000).

1.7.2   TuMV and *B. rapa* Interactions

To complicate matters, diploid brassicas (*B. rapa*) have three genes encoding eIF4E; *BraA.eIF4E.a*, *BraA.eIF4E.b* and *BraA.eIF4E.c* and three genes encoding eIF(iso)4E; *BraA.eIF(iso)4E.a*, *BraA.eIF(iso)4E.b* and *BraA.eIF(iso)4E.c* (Jenner *et al.*, 2010; Nellist *et al.*, 2014). Of the three copies of *eIF4E* identified, *BraA.eIF4E.b* appears to be a pseudogene, lacking Exons 2 and 3 (Jenner *et al.*, 2010). A summary of the introns and exons of *eIF4E/eIF(iso)4E* genes of *B. rapa* lines R-o-18 and RLR22 is shown in Figure 1.5. Prior to my research, very little was known about whether there are any differences in expression of the genes, whether some are only expressed in certain tissues and if the proteins have different binding properties.
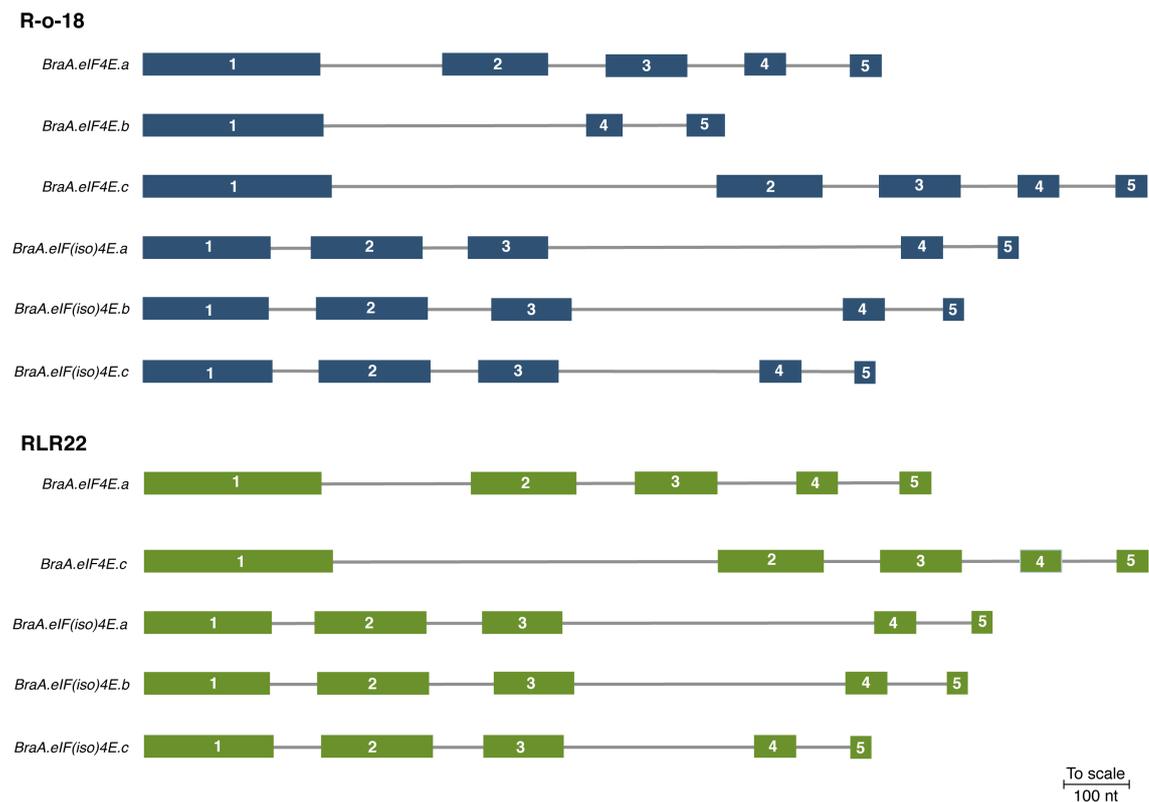
Figure 1.5 – The structure of *Brassica rapa eIF4E* and *eIF(iso)4E* genes from R-o-18 and RLR22 indicating positions of exons (coloured boxes) and introns (grey lines).

It was shown by complementation of an Arabidopsis eIF(iso)4E knock-out, that TuMV can use at least two copies of both *eIF4E* and *eIF(iso)4E* from TuMV-

susceptible *B. rapa* R-o-18, i.e. *BraA.eIF4E.a*, *BraA.eIF4E.c*, *BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c* when ectopically expressed in Arabidopsis (Jenner *et al.*, 2010). These findings might help to explain why recessive resistance to TuMV in *Brassica* species is less common than recessive resistance to *Potyvirus* species. in some other plant species. Interestingly and confusingly, complementation also showed that TuMV can use the same copies of *eIF4E* and *eIF(iso)4E* from the TuMV-resistant line RLR22 when ectopically expressed in Arabidopsis (Nellist *et al.*, 2014).

The interaction between the TuMV VPg protein and the *Brassica* eIF4E and eIF(iso)4E proteins is poorly understood. Recent unpublished work by our group revealed that the strong candidates for *ConTR01* were *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* and the strong candidate for *retr01* was *BraA.eIF(iso)4E.a*. It is not clearly understood which gene alleles are required for TuMV translation. Investigating this interaction in brassicas and determining which forms of the eIF4E and eIF(iso)4E proteins are required for the infection cycle of the virus would provide further understanding of this important interaction and enable potentially durable resistance to be introgressed into commercial *B. rapa* cultivars more rapidly than currently possible.

1.8 Aims and Objectives

The main aim of the study was to investigate the broad-spectrum resistance of the *B. rapa* line RLR22 to TuMV.

The specific objectives of the study were:

- To investigate if the yeast two-hybrid assay could be a useful tool for assessing TuMV VPg - *B. rapa* eIF4E/eIF(iso)4E interactions.

- To identify the two resistance genes, *retr01* and *ConTR01*, controlling the broad-spectrum resistance to TuMV in RLR22.

- To determine the mechanism of broad-spectrum resistance to TuMV in RLR22.

- To investigate the deployment of broad-spectrum resistance to TuMV into commercial Chinese cabbage lines.

- To investigate the complementation of an eIF(iso)4E Arabidopsis knock-out line with *B. rapa BraA.eIF(iso)4E.a* from RLR22.

Chapter 2

Materials and Methods

2.1 Plant Lines

2.1.1 *Brassica* Lines

Brassica rapa R-o-18 and RLR22

Brassica rapa L. ssp. *trilocularis* (Roxb.) Hanelt. (yellow sarson) line R-o-18 is extremely susceptible to a wide range of *Turnip mosaic virus* (TuMV) isolates (Walsh *et al.*, 2002). *B. rapa* ssp. *pekinensis* (Chinese cabbage) line RLR22 shows broad-spectrum resistance to TuMV (see Table 1.3).

Further *Brassica* Lines

Six further brassica lines, known to have broad-spectrum resistance, or derived from lines with broad-spectrum resistance to TuMV were investigated to see whether they have the same resistance mechanism as RLR22 and a line with a dominant resistance gene against one pathotype of TuMV was used as a control (summarised in Table 2.1). The plant lines were obtained from group stocks and from the Genetic Resource Unit (GRU) at Wellesbourne.

Two further brassica lines were used as controls in the study. *Brassica juncea* (L.) Czern. Cv. Tendergreen (Tendergreen Mustard; TGM) is susceptible to TuMV and was used as a TuMV-susceptible control for back-inoculations and also to maintain the TuMV isolates. The *Brassica napus* line R4 (var. *oleifera*) possesses the dominant resistance gene *TuRB01* (Walsh *et al.*, 1999), the line

shows resistance to TuMV isolate UK 1 and other pathotype 1 isolates, but the *TuRB01* resistance is overcome by the TuMV UK 1 mutant isolate, vVIR24. The *B. napus* L. cv. Mikado line was used to rear aphid populations on.

Table 2.1 – Details of further *Brassica rapa* lines resistant to *Turnip mosaic virus* that were investigated.

Plant Line	Plant Species	Broad-spectrum Resistance	Reference
B. rapa rapa	Turnip (*B. rapa* var. *rapa*)	+	Kassem and Walsh (2008)
BR03023	*B. rapa*	+	(Unpublished)
BR05060	*B. rapa*	+	(Unpublished)
BP058	*B. rapa*	+	Walsh *et al.* (2002)
Jin G55	*B. rapa*	+/-[a]	Hughes *et al.* (2002)
Jong Bai No. 2	*B. rapa*	+	Hughes *et al.* (2002)
S_3.RL2-43[b]	*B. rapa*	-	Walsh *et al.* (2002)

[a] Segregating for resistance.
[b] Resistant to Pathotype 1 isolates only.

Syngenta Chinese Cabbage Lines

Six inbred *B. rapa* var. *pekinensis* lines (CK 1 - CK 6) used for breeding F_1 hybrids in Syngenta's breeding programme in both Beijing, China and Enkhuizen, The Netherlands were received. These lines were crossed by Syngenta with a B_1F_4 plant (85-40), which originated from the initial crosses between RLR22 and R-o-18 and was fixed for *retr01* and *ConTR01*. F_2 material was sent from Syngenta to be genotyped and phenotyped, to inform which plants to take forward in the crossing process (Figure 2.1) and check on stability of resistance and the reliability of the marker for *retr01*.
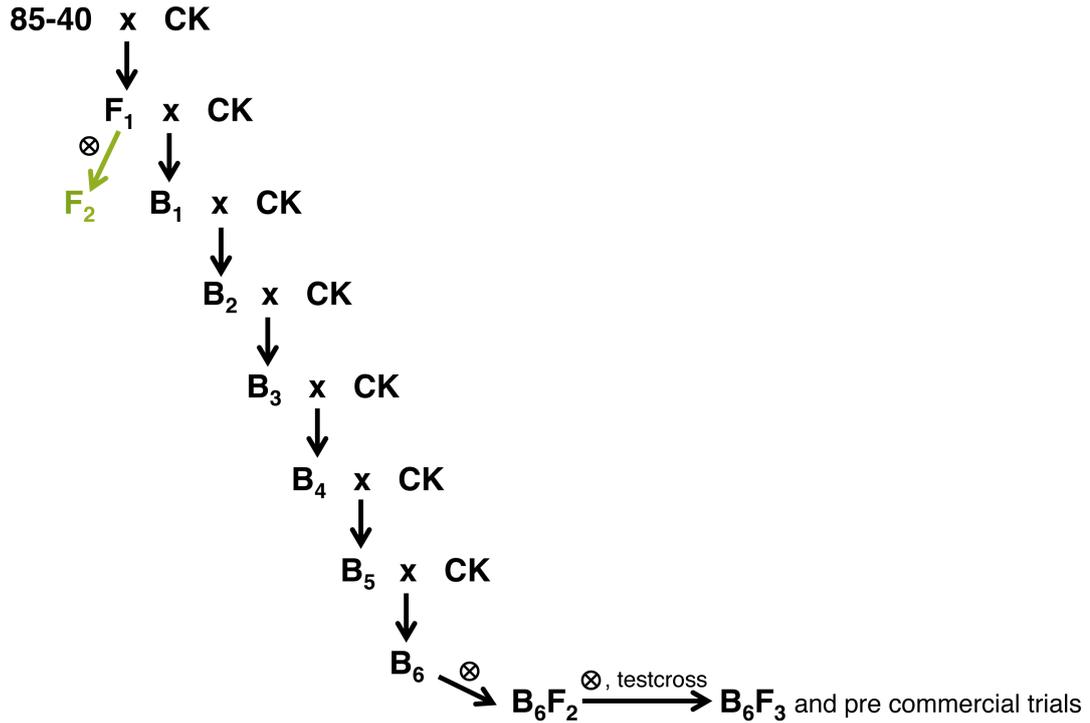
Figure 2.1 – Syngenta back-crossing strategy.
F_2 material phenotyped and genotyped at Warwick. 85-40, B_1F_4 fixed for *retr01* and *ConTR01*; CK, Syngenta Chinese cabbage susceptible parental lines (six in total); x, Cross pollination; ⊗, Self-pollination; B, Back-cross generation.

2.1.2 *Arabidopsis thaliana* Lines

Arabidopsis thaliana (L.) Heynh. Col-0/Col-4 are extremely susceptible to TuMV infection. An Arabidopsis line (Duprat *et al.*, 2002) with a defective Suppressor-mutator (*dSpm*) element inserted into exon 2 of *eIF(iso)4E* (At5g35620), Col-0::*dSpm* lacks both *eIF(iso)4E* mRNA and protein, resulting a line that is able to grow normally and is resistant to TuMV. Copies of *BraA.eIF(iso)4E.a* from R-o-18 and RLR22 were transformed into the lines Col-0/Col-4 and Col-0::*dSpm* by C. E. Jenner (Jenner *et al.*, 2010; Nellist *et al.*, 2014). Untransformed Col-0/Col-4 and Col-0::*dSpm* plants along with transformed Col-0/Col-4 were used as controls in the study.

Table 2.2 – Details of *Arabidopsis thaliana* lines used in transformation studies (Jenner *et al.*, 2010; Nellist *et al.*, 2014).

Line Code	Arabidopsis Line	Transposon[a]	*Brassica* Transgene
AT08002	Col-0	-	-
Col-4	Col-4	-	-
AT07001	Col-0	+	-
TP083-4	Col-0	-	R-o-18 *BraA.eIF(iso)4E.a*
TP388-1	Col-4	-	RLR22 *BraA.eIF(iso)4E.a*
AT23-4	Col-0	+	R-o-18 *BraA.eIF(iso)4E.a*
AT34-2	Col-0	+	RLR22 *BraA.eIF(iso)4E.a*

[a] *dSpm* transposon insertion is in *eIF(iso)4E* (*At5g35620*), inactivating the gene (Duprat *et al.*, 2002).

2.2 Plant Growth Methods

2.2.1 *Brassica* Growth Methods

Brassica plants were grown in Pot & Bedding - M2 compost (Levington; medium grade sphagnum moss peat 100% (pH 5.3 6.0; N = 200, P = 150, K = 200 mg/litre)) in the glasshouse under natural light conditions. During the winter months the day length was extended using halogen lamps. The temperature was maintained at 18°C $\pm$ 2°C in the glasshouse compartments. Seeds were sown into P7 pots (7 cm diameter) in a glasshouse compartment used for rearing healthy plants, the plants were grown there until ready to be challenged with TuMV (after approximately three weeks) at which point they were moved to another compartment to be challenged. The plants were watered daily.

Brassica Self-fertilisation

Brassica plants were selfed using flies (Family *Calliphoridae*) or by cutting off unopened buds, peeling back the petals and rubbing the pollen onto freshly opened buds. Once this had been performed, bread bags were used to cover the plants to stop cross-pollination from other plants.

2.2.2 Arabidopsis Growth Methods

Arabidopsis plants were grown in 6:1:1, Seed & Modular - F2S compost (Levington; fine grade sphagnum moss peat:sand; 24:1 (pH 5.3 6.0; N = 150, P = 200,

K = 200 mg/litre)):sand:vermiculite fine grade, in P40 trays in a growth room maintained at 18°C ± 2°C with a 9 hour day length (short-day, to prevent flowering). The trays were pre-watered and seeds were over-sown. The trays were then covered with tin foil and placed in the cold room (approximately 5°C) for three days prior to being taken to the growth room. Multiple P40 trays were placed in large trays and the plants were watered three times a week in the trays. Extra seedlings were removed when big enough to handle, leaving a single plant per pot. Plants were mechanically inoculated with TuMV/challenged with aphids 33 days after sowing.

2.3 *Turnip mosaic virus*

2.3.1 TuMV Isolates

Multiple TuMV isolates were used during the study. The main TuMV isolate used, was CDN 1 from Canada, which is a pathotype 4 isolate (Walsh, 1989) and has been shown to overcome a number of dominant brassica resistance genes (Jenner *et al.*, 2002b). TuMV isolate UK 1, is a pathotype 1 isolate which is unable to overcome the resistance gene *TuRB01* (Walsh *et al.*, 1999). TuMV UK 1 mutant isolate, vVIR24 contains a single nucleotide change in the cylindrical inclusion (CI) protein, A5056G (capable of overcoming *TuRB01*; Jenner *et al.*, 2000) and was used to test lines to indicate the presence of the *TuRB01* resistance gene. All isolates were maintained in *B. juncea* TGM by mechanical inoculation, as described by Walsh (1989).

2.3.2 Mechanical Transmission

Mechanical transmission of TuMV in brassica plants was performed following the method described by Walsh (1989); infected leaf material was mixed with inoculation buffer (1 g K_2HPO_4; 0.1 g $Na_2SO_3.7H_2O$; 100 ml dH_2O) and the abrasive carborundum (which enabled transmission of the virus when rubbed onto uninfected leaves using muslin) and ground up using a pestle and mortar. Healthy leaves were processed the same way and used for mock-inoculating plants.

Mechanical transmission of TuMV to Arabidopsis plants was performed in a similar manner, but using a cotton bud to rub the infected leaf material and

inoculation buffer onto uninfected leaves.

2.3.3   Aphid Transmission

Aphid transmission of TuMV was performed as described by Jenner *et al.*
(2010). The challenge was performed using the peach-potato aphid, *Myzus per-
sicae*. Aphid stocks were cultured on *B. napus* L. cv. Mikado. Leaves with a
good aphid population were chosen and placed underside-up on a sheet of white
paper. Aphids were exposed to bright light for a few minutes under a Tung-
sten filament lamp until they moved. The leaf was then tapped to remove the
aphids. The aphids were then transferred to a petri dish, where they were starved
for approximately two hours in the dark, before being moved, a couple at a time
with a fine brush, to feed on TuMV-infected *B. juncea* TGM leaves for a few
minutes. After a couple of minutes, the fine brush was used to 'tickle' the aphid
until it began to move. Five aphids were transferred in this way from the infected
leaves onto test plants, where they were left for two days. Non-viruliferous aphids
were used for mock-challenged plants. The aphids were then killed by spraying
with Aphox (Syngenta).

2.4   Detection of TuMV

2.4.1   Visual Assessments

Visual assessments of the plants were performed 1 week, 2 weeks and 3 weeks
post-inoculation, assessing both inoculated and uninoculated leaves. Plants were
scored for severity of symptoms, descriptions were recorded of whether the symp-
toms were fully systemic, or if they were limited, whether the chlorotic spots had
necrotic centres, or anything unusual about the symptoms/plants.

2.4.2   Enzyme-linked Immunosorbent Assay

For brassica plants, single young leaves were harvested 3 weeks post-inoculation
and a macerator (Pollähne) was used to extract the sap. For Arabidopsis
plants, five central leaves were harvested and plant sap was extracted using
Bioreba extraction bags (Bioreba AG) with a homogeniser hand model (Bioreba
AG). Plants were tested for the presence of TuMV and virus quantities by

plate-trapped antigen enzyme-linked immunosorbent assay (PTA-ELISA) using paired wells in microtitre plates (96-well; Nunc Maxisorp) based on a method described by Walsh *et al.* (1999). To confirm the symptoms were due to viral infection, the first antibody was a mouse monoclonal antibody EMA67 (Jenner *et al.*, 1999) and the second antibody was a goat anti-mouse IgG conjugated to alkaline phosphatase (Sigma). For each ELISA plate, 30 test plant samples were randomly assigned to paired wells, with a positive and negative control on each plate, surrounded by blank wells. The absorbance was measured using a Biochem Anthos 2010 microplate reader at 405 nm (Biochem Ltd.).

ELISA data was transferred to Microsoft Excel (Microsoft Corporation, v2011 Mac). Infection was determined by taking the average of the two healthy (Negative) control wells and subtracting value from the average of the two values for each test sample. Samples with negative corrected values were rounded up to 0.

2.4.3   Back-inoculation

To confirm questionable symptoms and borderline ELISA results, systemic lesions on leaves were ground up and mechanically inoculated to the TuMV-susceptible *B. juncea* TGM plants (see Section 2.3.2).

2.5   Nucleic Acid Techniques

2.5.1   Nucleic Acid Extraction

Genomic DNA Extraction

Leaf samples were collected and stored in 2 ml microcentrifuge tubes (Eppendorf) at -70°C prior to DNA extraction. Frozen leaves were crushed to a fine powder using a pre-cooled modified drill-piece, that fits flush with the microcentrifuge tubes. The Qiagen DNasey Kit (Qiagen) was used to extract DNA using silica-membrane purification. The kit was used according to the manufacturer's guidelines. The concentration and contamination of the DNA was assessed using a NanoDrop® ND-100 spectrophotometer (Thermo Scientific). DNA was stored at -20°C.

Total RNA Extraction

Leaf samples were collected in 2 ml microcentrifuge tubes (Eppendorf), flash frozen at -170°C in liquid nitrogen and stored at -70°C prior to RNA extraction. Frozen leaves were crushed to a fine powder using a pre-cooled modified drill-piece, that fits flush with the microcentrifuge tubes. The Ambion RNA Extraction Kit (Life Technologies) was used to extract total RNA using non-toxic guanidine-isothiocyanate lysis and silica-membrane purification. The kit was used according to the manufacturer's guidelines. The concentration and contamination of the RNA was assessed using a NanoDrop® ND-100 spectrophotometer (Thermo Scientific). RNA was stored at -70°C.

2.5.2   Primers

Primers were designed for the isolation of the plant genes, eukaryotic translation initiation factor 4E (*eIF4E*) and its isoform, *eIF(iso)4E* and for the viral protein genome-linked (VPg), cylindrical inclusion (CI) and 6K2 proteins of TuMV (Table 2.3) for yeast two-hybrid (Y2H) experiments. These primers were designed to include the whole gene, including a start and stop codon where applicable. The primers contained *att*B adapters for use with the Gateway Technology cloning method to provide an efficient and rapid way of moving DNA sequences into vectors. Internal primers were also designed for gene sequencing and plant genotyping (Table 2.4). Primers were also used from kits and based on previously published primers (Table 2.5).

2.5.3   DNase Treatment

The DNase I, RNase-free protocol (Fermentas) for the treatment of RNA by digesting single- and double-stranded DNA, was followed as described. The prepared RNA was then used as a template for the reverse transcription reaction.

2.5.4   Reverse Transcriptase Reaction

Complementary DNA (cDNA) was synthesised from the DNase treated RNA, following the Superscript^{TM} II Reverse Transcriptase kit (Invitrogen). Gene-specific reverse primers were used along with the recombinant ribonuclease inhibitor RNaseOUT^{TM} (Invitrogen). Reactions were performed in a thermocycler

Table 2.3 – Details of primers designed for whole gene amplification.

Primer Name	Sequence (5'-3')[a]	Orientation	Target Region
CN1	**AAAAAGCAGGCT**CAAAAGGAAAGAGGCAAAGACAG	Forward	5'-end of TuMV VPg
CN2	**AGAAAGCTGGGT**CTCACTCGTCGTGGTCCACTGGGACGA	Reverse	3'-end of TuMV VPg
CN3	**AAAAAGCAGGCT**CGATGGCGACAGAGGATG	Forward	5'-end of *BraA.eIF(iso)4E.a*
CN4	**AGAAAGCTGGGT**TCAGACACAGTGAACCTAGTTCTTC	Reverse	3'-end of R-o-18 *BraA.eIF(iso)4E.a*
CN5	**AGAAAGCTGGGT**TCAGACACAGTGAACCGAGTTCTTC	Reverse	3'-end of Chinese cabbage *BraA.eIF(iso)4E.a*
CN6	**AAAAAGCAGGCT**CAATGGCGACCGATGATGTGAACG	Forward	5'-end of Arabidopsis *eIF(iso)4E*
CN7	**AGAAAGCTGGGT**TCAGACACAGTGAACCGGCTTCTTC	Reverse	3'-end of Arabidopsis *eIF(iso)4E*
CN8	**AAAAAGCAGGCT**CAATGGCGGGTAGAAGACACTCCC	Forward	5'-end of Arabidopsis *eIF4E*
CN9	**AGAAAGCTGGGT**TCAAGCGGGTGTAAGCGTTCT	Reverse	3'-end of Arabidopsis *eIF4E*
CN10	**AAAAAGCAGGCT**CAATGGCGGGTAGAAGACACTT	Forward	5'-end of *BraA.eIF4E.c*
CN11	**AAGAAAGCTGGGT**TCAAGCGGTGTAAGCGCTCTTC	Reverse	3'-end of *BraA.eIF4E.c*
CN12	**AAAAAGCAGGCT**CTATGGCGACAGAGGATGTGA	Forward	5'-end of *BraA.eIF(iso)4E.c*
CN13	**AAGAAAGCTGGGT**TTAGACACTAAAATCGACTTCTT	Reverse	3'-end of R-o-18 *BraA.eIF(iso)4E.c*
CN14	**AAGAAAGCTGGGT**TCAGACACTAAATCGACTTCTT	Reverse	3'-end of RLR22 *BraA.eIF(iso)4E.c*
CN15	**AAAAAGCAGGCT**CGACACTCAATGAAATAGAGGACG	Forward	5'-end of TuMV CI
CN16	**AAGAAAGCTGGGT**CTATTGGTGATGAACTGCC	Reverse	3'-end of TuMV CI
CN17	**AAAAAGCAGGCT**CAAGCACCAACGAAATGAGCAAGT	Forward	5'-end of TuMV 6K2
CN18	**AAGAAAGCTGGGT**CTATTCATGAGTTACGGGTTCGGA	Reverse	3'-end of TuMV 6K2
CN23	**AAAAAGCAGGCT**CGATGGCGGTAGAAGACACA	Forward	5'-end of R-o-18 *BraA.eIF4E.a*
CN24	**AGAAAGCTGGGT**TCAGGCAGTGTAAGCGCTCTTTGCG	Reverse	3'-end of R-o-18 *BraA.eIF4E.a*
CN25	**AAAAAGCAGGCT**CGATGGCGGGTAGAAGACACTCT	Forward	5'-end of RLR22 *BraA.eIF4E.a*
CN26	**AGAAAGCTGGGT**TCAGGCAGTGTAAGCGCTCTTTGCGCC	Reverse	3'-end of RLR22 *BraA.eIF4E.a*

[a] attB adapter in bold.

Table 2.4 – Details of primers designed for genotyping and sequencing.

Primer Name	Sequence (5'-3')	Orientation	Target Region
CN19	AGAAGACACATTATCATTCCAAG	Forward	TuMV CDN 1 CI (4482-4504)
CN20	TTTGAACAAACTGGCCATACCCA	Forward	TuMV CDN 1 CI (5181-5203)
CN21	TGATCTTGCCAGAGTAGTCACAC	Reverse	TuMV CDN 1 CI (4412-4389)
CN22	ACCTCACTACTTGGCTTGTGTAG	Reverse	TuMV CDN 1 CI (5111-5088)
CN44	TTTCTTGTTGGGTTAAGTGAAG	Forward	RLR22 *BraA.eIF4E.c* Middle of intron 1
CN45	CAAGCAACTACACATGGAAAAAAC	Reverse	RLR22 *BraA.eIF4E.c* Beginning of exon 3
CN55	TCTTTGTTGGTGGGTTAGATTCCG	Forward	R-o-18 *BraA.eIF4E.c* Beginning of intron 1
CN56	ATCAACGCAAGCAACTACACATCGAG	Reverse	R-o-18 *BraA.eIF(iso)4E.a* Intron 2/exon 3 splice junction
CN57	GACTTCTGGGGTTTGCACGAGACT	Forward	RLR22 *BraA.eIF(iso)4E.a* Exon 1/exon 2 splice junction
CN59	CCTCATCAAATTGCTCTCCGACA	Reverse	RLR22 *BraA.eIF(iso)4E.a* Beginning of exon 3

Table 2.5 – Details of kit and published primers used.

Primer Name	Sequence (5'-3')	Orientation	Target Region	Reference
KO1	TTGACCCAATAGAGTCCAGAAAT	Forward	5'-end of *At.eIF(iso)4E* cDNA	Duprat *et al.* (2002)
DSPM1	CTTATTTCAGTAAGAGTGTGGGGTTTTGG	Reverse	3'-end of *dSpm*	Duprat *et al.* (2002)
attB1	GGGGGACAAGTTTGTACAAAAAAGCAGGCT	Forward	attB adapter	Invitrogen
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGT	Reverse	attB adapter	Invitrogen
M13 F	GTAAAACGACGGCCAG	Forward	N-terminal sequence of *lacZ*	Invitrgoen
M13 R	CAGGAAACAGCTATGAC	Reverse	N-terminal sequence of *lacZ*	Invitrgoen
DEST32 F	AACCGAAGTGCGCCAAGTGTCTG	Forward	Bait	Invitrgoen
DEST22 F	AGCCGACAACCTTGATTGGAGAC	Forward	Prey	Invitrogen
DEST R	TATAACGCGTTTGGAATCACT	Reverse	Bait and Prey	Invitrogen
BR14	TAGACAAGGCTTGGCTTGAAACTG	Froward	*B. rapa eIF(iso)4E* 3'-end of exon 2	Nellist *et al.* (2014)
BR2	TCTCCTCCACTTCTTCCCAATAC	Reverse	*B. rapa eIF(iso)4E* 5'-end of exon 4	Nellist *et al.* (2014)

32

(Bio-Rad, MyCycler^{TM}). A total volume of 20 μl was prepared and stored at -20°C.

2.5.5 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed on cDNA and genomic DNA (gDNA); 5 μl of cDNA from the RT-reaction was added to a 45 μl PCR reaction and gDNA was used at various volumes depending on concentration (total reaction volume of 50 μl used). The PCR reaction mixture comprised of 5 μl 10x PCR Buffer (Invitrogen), 1.5 mM magnesium chloride (2 mM - 2.5 mM, also used), 0.2 mM dNTPs (Invitrogen), 2 U Taq-DNA Polymerase (Invitrogen) and 5 μM of each specific primer. The PCR reaction mixture was amplified under the following conditions; 95°C for 5 mins, 30 cycles at 95°C for 30 secs, between 50 and 62°C for 30 secs depending on the specific primer pair and 72°C for 1 min (2 mins for larger products, whole genes) and one cycle at 72°C for 10 mins. Reactions were performed in a thermocycler (Bio-Rad, MyCycler^{TM}). PCR products were stored at -20°C.

PCR reactions for the isolation of genes for Y2H were performed with KOD Hot Start DNA Polymerase (Novagen). The PCR mixture comprised 5 μl 10x Buffer for KOD Hot Start DNA Polyermase (Novagen), 1.5 mM magnesium sulphate, 0.2 mM dNTPs (Novagen), 0.02 U/μl KOD Hot Start DNA Polymerase, 10 μM of each specific primer, template DNA, in a total volume of 50 μl.

Bacterial PCR

Individual bacterial (*Escherichia coli*) colonies were stabbed with a 20 μl pipette tip and the cells were transferred into a 0.2 ml PCR tube (Thermo Scientific) with 10 μl dH$_2$O. The pipette tip was agitated to dislodge the cells and the 10 μl of re-suspended cells were used in standard PCR.

Yeast PCR

DNA was extracted from yeast (*Saccharomyces cerevisiae*) following a user developed protocol (Isolation of plasmid DNA from yeast (PR04.doc Oct-01, www.qiagen.com)) adapted from the QIAprep$^{®}$ Spin Miniprep Kit (Qiagen).

2.5.6 Agarose Gel Electrophoresis of DNA

DNA fragments were separated and quantified using electrophoresis. UltraPure^{TM} Agarose (Invitrogen) powder was dissolved in 1x Tris-Borate-EDTA (TBE) buffer (made from 9 parts dH$_2$O and 1 part 10x TBE; National Diagnostics) to make 1% (1.5% and 2% for distinguishing between similar sized PCR products) agarose gels, with 2 μg/50 ml GelRed (Biotium Inc.). Gels with a thickness of 7-10 mm were prepared in gel trays and run in the corresponding tanks.

DNA samples were loaded into the gel after the addition of 1/10 volume of 6x DNA Loading Buffer IV (0.25 % Bromophenol blue, 40 % (w/v) sucrose in dH$_2$O). The 1 Kb Plus DNA Ladder (Invitrogen) was used at a concentration of 0.5 μg/μl to estimate molecular mass of double-stranded DNA bands and give an indication of concentration. Gels were run at room temperature between 70-120 V for anything between 60-240 mins for diagnostic tests or 30-70 V for isolation of products of similar size. After electrophoresis, separated DNA fragments were viewed on a Syngene G Box transilluminator, using GeneSnap 7.07 (Syngene, a division of Synoptics Ltd.).

A clean scalpel was used to cut out bands to be extracted. Bands were purified following the QIAquick Gel Extraction Kit Protocol (Qiagen) using a microcentrifuge. EB elution buffer (10 mM Tris Cl, pH 8.5) at a volume of 30 μl was used to elute the DNA.

2.5.7 Cloning and Transformation of Plasmids into *E. coli*

TA Cloning

TA cloning was used to sequence mixed PCR products and was performed following the manufacturer's instructions (TOPO TA Cloning® Kit for Sequencing; Invitrogen) for transforming chemically competent One Shot® (Invitrogen) *E. coli* cells.

BP and LR Recombination Reactions

Gateway cloning was used in preparation for Y2H and was performed following the manufacturer's instructions (Gateway® Technology with Clonase^{TM} II; Invitrogen). RNA was extracted from *B. rapa* plants as described in Section 2.5.1. *E. coli*

with plasmids containing Arabidopsis *eIF4E1* (U12635) and *eIF(iso)4E* (U16070) were obtained from The Arabidopsis Information Resource (TAIR). TuMV RNA was extracted from *B. rapa* plants. RT-PCR was performed to obtain cDNA (see Section 2.5.4). The genes of interest were then amplified by PCR using the appropriate primers (Table 2.3). The primers contained *att*B adapters to aid the introgression of the target genes into entry clones (Tables 2.3 and 2.5). A BP recombination reaction was performed to create the entry clone with the *att*B-PCR products and pDONR221, which was used to transform the host *E. coli*. The expression clone was generated using the LR recombination reaction, using the previously mentioned entry clones to transfer the gene into the Gateway destination vectors, pDEST32 and pDEST22. Constructs were sequenced after each step to ensure the correct sequence was present. Each gene was transformed into both the pDEST32 and pDEST22 vectors.

TempliPhi Reaction

The TempliPhi Amplification Kit (GE Healthcare) was used to amplify plasmid DNA in preparation for Y2H, following the manufacturers instructions.

2.5.8   Sequencing

DNA Sequencing

Two organisations were used for sequencing, some sequencing was performed in-house (School of Life Sciences Genomics Centre) and some sequencing was performed by an external company, GATC Biotech.

For in-house sequencing, PCR reactions had a final volume of 10 μl; 2 μl BigDye (Applied Biosystems), 2 μl sequencing buffer, 1 μl of 5 μM appropriate primer (Tables 2.3 - 2.5), up to 5 μl of 10-50 ng template DNA and remaining volume of dH$_2$O. Reactions were performed in a thermal-cycler (Bio-Rad MyCycler^{TM}), using the temperature profile published by Applied Biosystems; 96°C for 1 min, 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min. The products were then sequenced using an Applied Biosystems 3130xl Genetic Analyser (School of Life Sciences Genomics Centre).

For GATC Biotech sequencing, a total volume of 10 μl was prepared; 5 μl

template DNA (80-100 ng/μl purified plasmid DNA or 20-80 ng/μl purified PCR product) and 5 μl of 5 μM appropriate primer (Tables 2.3 - 2.5).

Transcriptome Sequencing

Transcriptome sequencing was performed on R-o-18 and RLR22 plants to profile the expression of the *eIF4E* and *eIF(iso)4E* genes. Total RNA was extracted from young leaves of three R-o-18 plants and three RLR22 plants (see Section 2.5.1). The RNA was sent to SeqWright (SeqWright, Inc.), ribosomal RNA depletion was performed using the RiboMinus Eukaryotic Kit for RNA-Seq (Life Technologies, A10837-08). Libraries were constructed using the TruSeq RNA Sample Preparation Kit (Illumina Inc.) and sequences of mRNA were determined using an Illumina HiSeq 2000 system (Illumina Inc.). Sequence reads (100-base paired-end) were analysed using Illumina CASAVA version 1.8.

The transgenic Arabidopsis Col-0::*dSpm* line with the RLR22 *BraA.eIF(iso)4E.a* transgene was also transcriptome sequenced in-house to profile the expression of the transgene. Total RNA was extracted from a young leaf (see Section 2.5.1) and Oligo(dT) selection was performed twice using Dynal magnetic beads (Invitrgoen). Illumina library preparation was performed according to the manufacturer's protocol (15018818 revA), using mRNA-TruSeq Sample Prep Kit Version 5 (Illumina Inc.). Illumina's GAIIx sequencing system at School of Life Sciences Genomics Centre was used to sequence the library. Sequence reads (70 base paired-end) were base-called and scored for read quality, using the CASAVA pipeline.

2.6   Yeast Two-hybrid System

The Y2H method was followed from the ProQuest Two-Hybrid System (Invitrogen). The yeast strain (*S. cerevisiae*) used, MaV203, contains single copies of the following three reporter genes, *HIS3*, *URA3* and *lacZ*. The Y2H system utilises the fact that transcription factors consist of two domains, a DNA binding domain (DBD) and an activation domain (AD). Two separate hybrid proteins are constructed, referred to as bait and prey (Figure 2.2). If the bait and prey proteins investigated do not interact, there is no reporter gene expression (Figure 2.2 (a)). If they do interact, the reporter gene is expressed (Figure 2.2 (b)).
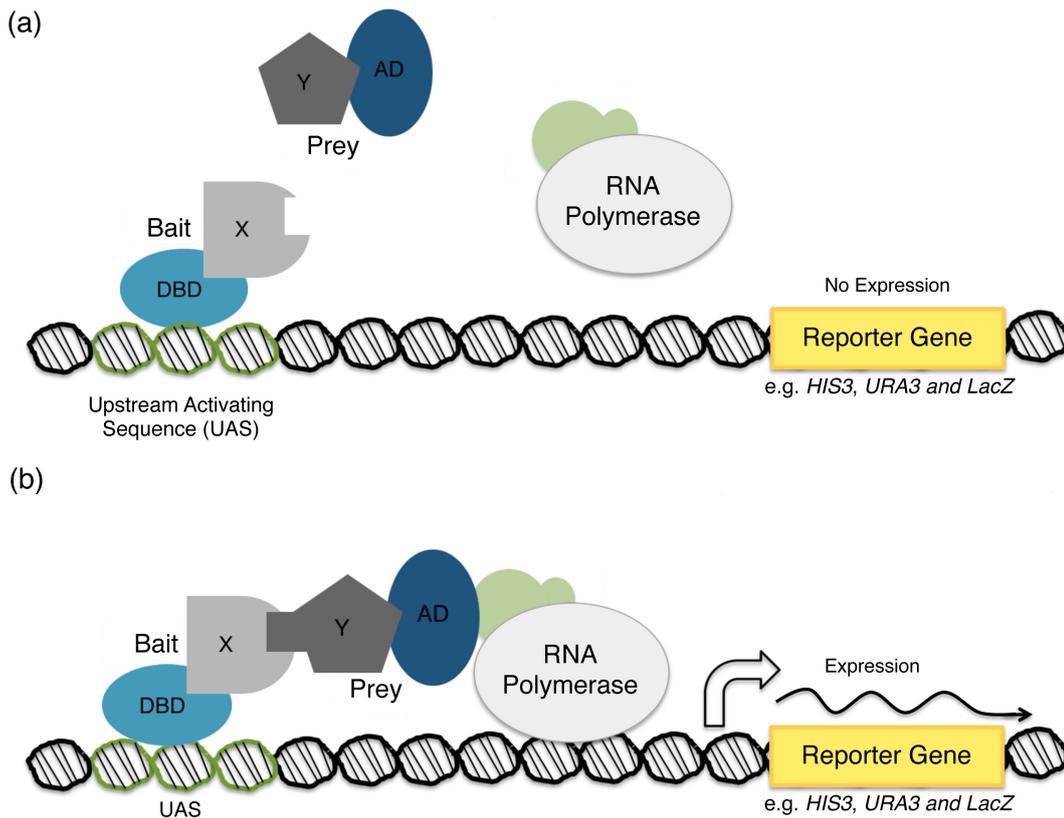
Figure 2.2 – Model of protein interactions in the yeast two-hybrid assay (adapted from Invitrogen Gateway Technology with Clonase II manual). The bait construct is comprised of protein X and the DNA binding domain (DBD). The prey construct is comprised of protein Y and the activation domain (AD).
(a) If proteins X and Y do not interact, there is no reporter gene expression. (b) If proteins X and Y do interact then the reporter gene is expressed.

2.6.1 Preparation for the Y2H System

In preparation for Y2H, the 'Gateway Technology with Clonase II' manual (Invitrogen) was followed. N-terminal fusion proteins were designed for the assay. The two sets of expression vectors (pDEST32, known as the "bait" and possessing the DBD and pDEST22, known as the "prey" and possessing the AD) were then used to transform yeast following the ProQuest Two Hybrid System manual.

BraA.eIF(iso)4E.a from both R-o-18 and RLR22 was investigated along with BraA.eIF4E.c and BraA.eIF(iso)4E.c from R-o-18 to test for an interaction with three TuMV proteins, VPg, 6K2 and CI (Table 2.6). The TuMV VPg was also paired with Arabidopsis eIF(iso)4E and eIF4E as controls, positive and negative respectively. Four or five Invitrogen kit controls, varying in interaction strength were also used: A, no interaction; B, weak interaction; C, moderately strong

interaction; D, strong interaction; E, very strong interaction (Invitrogen). Plate plans were then designed to include the relevant controls and the target gene combinations in both bait and prey form (Table 2.6 and Figure 2.3). One plate plan includes the negative interaction control TuMV CDN 1 VPg + Arabidopsis eIF4E and not Invitrogen kit control E (Figure 2.3 (a)) and the other includes the Invitrogen kit control E and not the negative interaction control TuMV CDN 1 VPg + Arabidopsis eIF4E (Figure 2.3 (b)).

Once the sets of expression vectors were in the transformed yeast, they were grown on selective media to characterise the transformants; SD-Leu-Trp-Ura (Clontech), SD-Leu-Trp-His+3-amino-1,2,4-triazole (3AT) (Clontech) and YPAD with a filter for X-gal assay. The characterisation of transformants was followed from the ProQuest Two-Hybrid System.
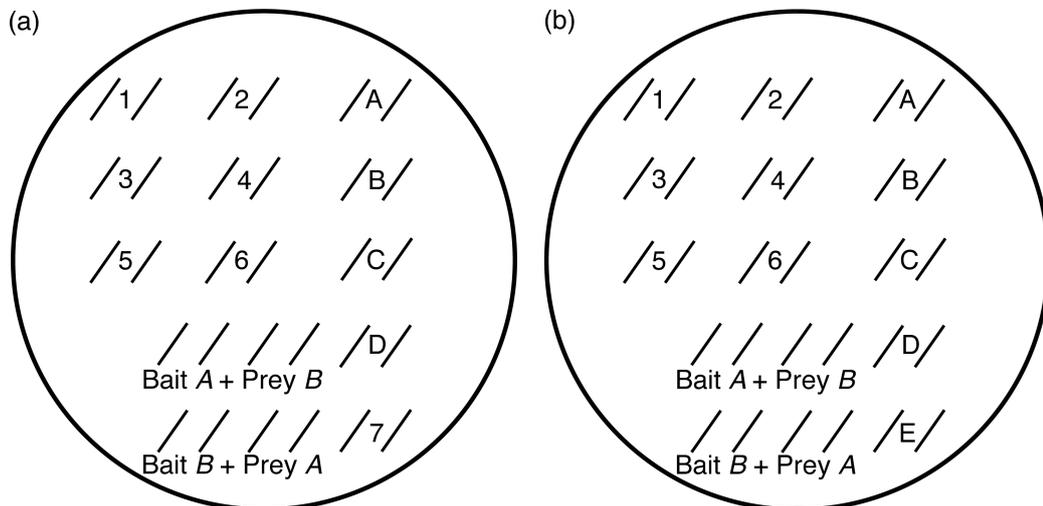
Figure 2.3 – Yeast two-hybrid plate plans. (a) with the negative interaction control *Turnip mosaic virus* CDN 1 VPg + *Arabidopsis thaliana* eIF4E and (b) with the Invitrogen kit control E.
A, Negative control; B, Weak interaction control; C, Moderately strong interaction control; D, Strong interaction control; E, Very strong interaction control; 1, Yeast MaV203 (Negative control); 2, pDEST32 + pDEST22 (Negative control); 3, TuMV VPg + Arabidopsis eIF(iso)4E (Positive control); 4, pDEST32 + Arabidopsis eIF(iso)4E (Negative control); 5, Bait + pDEST22 (Negative control); 6, pDEST32 + Prey (Negative control); 7, TuMV CDN 1 VPg + Arabidopsis eIF4E (Negative control).

Table 2.6 – Combinations of plant eukaryotic translation initiation factors and *Turnip mosaic virus* CDN 1 viral proteins tested in the yeast two-hybrid system.

TuMV protein	R-o-18			RLR22		Arabidopsis Controls	
	BraA.eIF(iso)4E.a	BraA.eIF(iso)4E.c	BraA.eIF4E.c	BraA.eIF(iso)4E.a[a] (Whole of intron 1 retained)	BraA.eIF(iso)4E.a[b] (15 nt intron 1 retained)	At.eIF(iso)4E	At.eIF4E
VPg	✓	✓	✓	✓	✓	✓	✓
CI	✓	✓	✓	NT[c]	NT	✓	✓
CI-6K2	✓	✓	✓	NT	NT	✓	✓
6K2	✓	✓	✓	NT	NT	✓	✓
6K2-VPg	✓	✓	✓	NT	NT	✓	✓

[a] Construct of RLR22 BraA.eIF(iso)4E.a retaining the extra G (indel) and the whole of intron 1.
[b] Construct of RLR22 BraA.eIF(iso)4E.a retaining the last 15 nt of intron 1.
[c] Not tested.

2.7 Data Analysis

Bioinformatics

Primers were designed in SeqBuilder (DNASTAR, Lasergene v10.0; Burland, 1999). The ABI trace files of forward and reverse sequences were aligned and edited in SeqMan (DNASTAR Lasergene v10.0) and the consensus was exported in fasta format. Sequences were compared to previously published results using the Basic Local Alignment Search Tool (BLAST; Altschul *et al.*, 1997) on the National Centre for Biotechnology Information (NCBI), GenBank (http://www.ncbi.nlm.nih.gov/; Benson *et al.*, 1999). SeqMan was also used to infer amino acid sequences from DNA sequences. MegAlign (DNASTAR Lasergene v10.0) was used to compare amino acid sequences using the ClustalW algorithm (Thompson *et al.*, 1994). MEGA 5 (Tamura *et al.*, 2011) was also used to compare amino acid sequences using the MUSCLE algorithm (Edgar, 2004a,b). Some amino acid sequences were manually adjusted due to unnecessary gaps created by the MUSCLE algorithm.

The R-o-18 and RLR22 transcriptome sequence reads were aligned to the published *B. rapa* (Wang *et al.*, 2011) genome assembly using Bowtie algorithm (Langmead *et al.*, 2009). The cufflinks algorithm (Trapnell *et al.*, 2010) was also used to calculate and compare fragments per kilobase of exon per million mapped fragments (FPKM), to estimate relative transcript abundances. The Arabidopsis transcriptome sequence reads were aligned to the *B. rapa* RLR22 *BraA.eIF(iso)4E.a* sequence using Tophat (Trapnell *et al.*, 2009) and sequences were viewed using the Integrative Genome Viewer (IGV; Thorvaldsdóttir *et al.*, 2012).

Statistics

Statistical analyses were performed using both Microsoft Excel (Microsoft Corporation, v2011 Mac) and GenStat (VSN International Ltd., v13). ELISA readings of transformed and untransformed Arabidopsis plants were subjected to a variance stabilising transformation (angular/arcsine) prior to analysis using a residual maximum likelihood method (REML). Differences between treatment means were compared using an approximate least significant difference (LSD) test. A maximum LSD value was used in place of individual LSDs for each treatment combination.

To test for the segregation of resistance and susceptibility in F_2 populations from the RLR22 and Syngenta CK crosses, a chi-squared test was used to test whether the resistance segregated 3:1 Susceptible:Resistant, as expected for a single recessive gene. Plants showing any symptoms were classed as susceptible. The resistance and susceptibility was also tested for a 1:2:1 relationship, for Fully susceptible:Limited infection:Resistant. T-tests were also performed on ELISA values from the same plant lines, to compare values for the plants heterozygous for *retr01* with values for plants homozygous for the absence of *retr01* in order to determine whether there were any differences in the accumulation of TuMV.

Chapter 3

Interaction of *Brassica rapa* BraA.eIF(iso)4E.a and *Turnip mosaic virus* CDN 1 VPg

3.1 Background

3.1.1 Use of Yeast Two-hybrid in Plant-Virus Studies

The yeast two-hybrid (Y2H) assay is a powerful tool for assessing protein-protein interactions by testing for physical interactions such as binding. It was first described by Fields and Song (1989) and enables the observation of the interaction of individual proteins. The Y2H system provided the first evidence that plant viruses directly interacted with the eukaryotic translation initiation complex of plants, by showing the virus-encoded genome-linked protein (VPg) of *Turnip mosaic virus* (TuMV) bound to *Arabidopsis thaliana* eukaryotic translation initiation factor isoform 4E (eIF(iso)4E) (Wittmann *et al.*, 1997). Since then, the system has revealed an interaction between eIF4E from susceptible *Capsicum annuum* and *Tobacco etch virus* (TEV) VPg, but not eIF4E from the *pvr1* and *pvr2*[1] (also known as *pvr1*[2]) alleles (Kang *et al.*, 2005a). The system has also been used to show an interaction between TEV VPg and eIF4E from tomato and tobacco (Schaad *et al.*, 2000), *Lettuce mosaic virus* (LMV) and lettuce (*Lactuca sativa*) eIF4E and others (Reviewed by Le Gall *et al.*, 2011). Additionally, it has been used to show an interaction between the VPg of TuMV and a copy of eIF(iso)4E from *Brassica rapa*, but the authors did not indicate which loci of eIF(iso)4E (Hwang *et al.*, 2009).

42

The main aim of the work described in this chapter was to test whether the Y2H system would provide a simple assay that could be used to assess which of the eIF4E and eIF(iso)4E proteins of *B. rapa* might be functional for TuMV isolates. The overall aim being to identify alleles of the genes that would confer susceptibility and resistance to TuMV. After the first efforts to detect an interaction between BraA.eIF(iso)4E.a from a TuMV-susceptible plant and the VPg of TuMV were unsuccessful, further viral genes were investigated to test whether they play a role in the interaction between TuMV and *B. rapa* eIF4E and eIF(iso)4E as has been shown for LMV and lettuce eIF4E (Abdul-Razzak *et al.*, 2009).

3.2 Results

3.2.1 Investigation of the Interaction of *B. rapa* BraA.eIF(iso)4E.a with TuMV CDN 1 VPg

The assessment of interaction was performed using three independent reporter genes, *HIS3*, *URA3* and *lacZ*, to reduce the detection of false positives. BraA.eIF(iso)4E.a from TuMV-susceptible R-o-18 and TuMV-resistant RLR22 *B. rapa* were tested against the VPg of TuMV isolate CDN 1 (see plate design, Figure 2.3). When amplifying cDNA of *BraA.eIF(iso)4E.a* from R-o-18 and RLR22, in preparation for the Y2H assay, it was observed that the RLR22 product was larger than the R-o-18 product (see Chapter 4, Figure 4.3). Sequencing of the RLR22 cDNA larger product revealed that *BraA.eIF(iso)4E.a* was mis-spliced and the variant retained an extra G (indel) and the whole of intron 1. A less common minor variant was also detected in RLR22, retaining the last 15 nucleotides (nt) of intron 1. Both of these variants were included in the Y2H assay to see if the VPg of TuMV CDN 1 directly interacted with them. No mis-splicing was detected in R-o-18 and so only the correctly spliced gene was studied (see Chapter 4 for further explanation). Arabidopsis eIF(iso)4E was used as a positive interaction control with the VPg of TuMV CDN 1 and Arabidopsis eIF4E was used as a negative interaction control with the VPg of TuMV CDN 1, as TuMV has been shown to only use eIF(iso)4E in Arabidopsis (Duprat *et al.*, 2002; Sato *et al.*, 2005). Nine other controls were used on each plate to check for false positives. These included four/five kit controls (see Section 2.6.1)

varying in interaction strength, the untransformed yeast strain MaV203 and combinations of empty vectors and constructs (see plate design, Figure 2.3). To check negative results, a second transformation of yeast was performed with R-o-18 BraA.eIF(iso)4E.a, both RLR22 BraA.eIF(iso)4E.a variants and the VPg of TuMV CDN 1 constructs. The second Y2H experiments produced the same results.

Confirmation of Genes in Yeast

DNA was extracted from the yeast colonies and PCR was performed to check the correct genes were present. An example agarose gel shows the gene for the TuMV VPg to be present in the yeast (Figure 3.1; lanes 1-9). PCR was also performed to confirm the presence of R-o-18 and RLR22 *BraA.eIF(iso)4E.a* and Arabidopsis *eIF(iso)4E* and *eIF4E* in the corresponding colonies (data not shown). The correct genes were present in the respective yeast colonies.
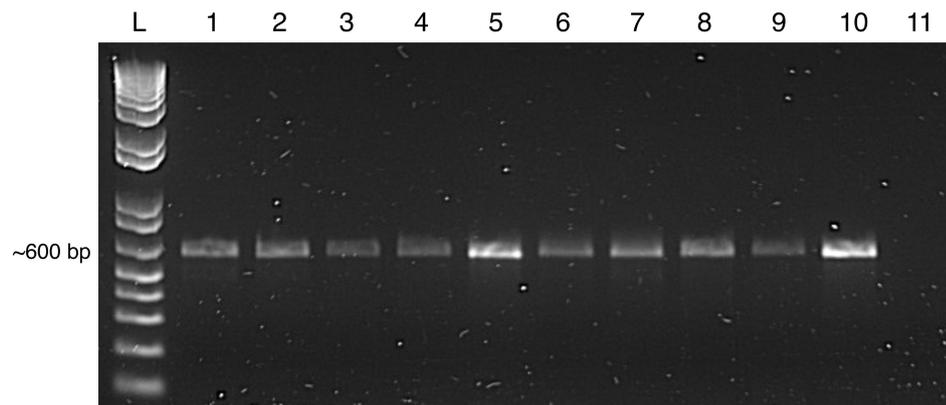
Figure 3.1 – PCR amplification of the *Turnip mosaic virus* isolate CDN 1 VPg (∼600 bp) from yeast, using specific primers CN1 and CN2.
L, 1 Kb Plus DNA Ladder; 1, yeast colony transformed with the VPg of TuMV CDN 1 and R-o-18 BraA.eIF(iso)4E.a; 2, yeast colony transformed with the VPg of TuMV CDN 1 and R-o-18 BraA.eIF(iso)4E.a; 3, yeast colony transformed with the VPg of TuMV CDN 1 and RLR22 BraA.eIF(iso)4E.a (extra G and whole of intron 1 retained); 4, yeast colony transformed with the VPg of TuMV CDN 1 and RLR22 BraA.eIF(iso)4E.a (extra G and whole of intron 1 retained); 5, yeast colony transformed with the VPg of TuMV CDN 1 and RLR22 BraA.eIF(iso)4E.a (last 15 nt of intron 1 retained); 6, yeast colony transformed with the VPg of TuMV CDN 1 and RLR22 BraA.eIF(iso)4E.a (last 15 nt of intron 1 retained); 7, yeast colony transformed with the VPg of TuMV CDN 1 and Arabidopsis eIF(iso)4E; 8, yeast colony transformed with the VPg of TuMV CDN 1 and Arabidopsis eIF(iso)4E; 9, yeast colony transformed with the VPg of TuMV CDN 1 and Arabidopsis eIF4E; 10, TuMV isolate CDN 1 VPg cDNA (positive control); 11, dH$_2$O (negative control).

HIS3 Reporter Gene

In both experiments, the Arabidopsis eIF(iso)4E with TuMV CDN 1 VPg control was positive on all *HIS3* plates and the Arabidopsis eIF4E with TuMV CDN 1 VPg was negative on all plates (examples, Figures 3.2 and 3.3). All empty vectors and construct combinations tested negative (examples, Figures 3.2 and 3.3). The untransformed MaV203 strain and kit negative control also tested negative (examples, Figures 3.2 and 3.3). No interaction was detected between either the correctly spliced R-o-18 BraA.eIF(iso)4E.a and the VPg of TuMV CDN 1 (Figure 3.2), or for either of the RLR22 BraA.eIF(iso)4E.a variants and the VPg of TuMV CDN 1 (example, Figure 3.3).
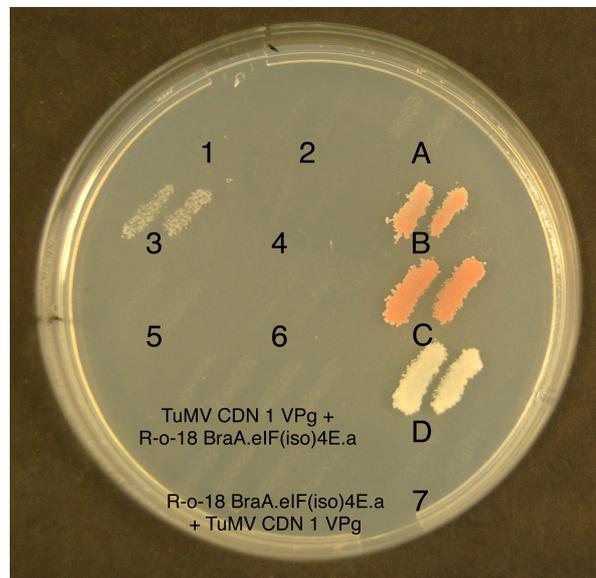
Figure 3.2 – Interaction of R-o-18 BraA.eIF(iso)4E.a + *Turnip mosaic virus* CDN 1 VPg, using the *HIS3* reporter gene where growth indicates an interaction.
A, Negative control; B, Weak interaction control; C, Moderately strong interaction control; D, Strong interaction control; 1, MaV203 (Negative control); 2, Empty pDEST32 + pDEST22 vectors (Negative control); 3, TuMV CDN 1 VPg + *Arabidopsis thaliana* eIF(iso)4E (Positive control); 4, Empty pDEST32 vector + Arabidopsis eIF(iso)4E (Negative control); 5, R-o-18 BraA.eIF(iso)4E.a + Empty pDEST22 vector (Negative control); 6, Empty pDEST32 vector + R-o-18 BraA.eIF(iso)4E.a (Negative control); 7, TuMV CDN 1 VPg + Arabidopsis eIF4E (Negative control).

URA3 Reporter Gene

In both experiments, the *URA3* reporter gene produced the same results as the *HIS3* reporter gene. The positive interaction control Arabidopsis eIF(iso)4E with
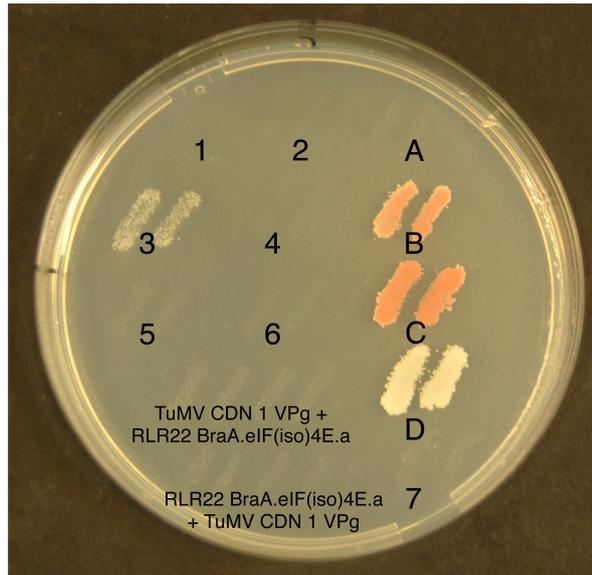
Figure 3.3 – Interaction of RLR22 BraA.eIF(iso)4E.a (last 15 nt of intron 1 retained) + *Turnip mosaic virus* CDN 1 VPg, using the *HIS3* reporter gene where growth indicates an interaction.
A, Negative control; B, Weak interaction control; C, Moderately strong interaction control; D, Strong interaction control; 1, MaV203 (Negative control); 2, Empty pDEST32 + pDEST22 vectors (Negative control); 3, TuMV CDN 1 VPg + *Arabidopsis thaliana* eIF(iso)4E (Positive control); 4, Empty pDEST32 vector + Arabidopsis eIF(iso)4E (Negative control); 5, RLR22 BraA.eIF(iso)4E.a (last 15 nt of intron 1 present) + Empty pDEST22 vector (Negative control); 6, Empty pDEST32 vector + RLR22 BraA.eIF(iso)4E.a (last 15 nt of intron 1 present) (Negative control); 7, TuMV CDN 1 VPg + Arabidopsis eIF4E (Negative control).

TuMV CDN 1 VPg showed growth along with the kit controls B, C and D. No interaction was observed between correctly spliced R-o-18 BraA.eIF(iso)4E.a and the VPg of TuMV CDN 1, or for either of the RLR22 BraA.eIF(iso)4E.a splice variants and the VPg of TuMV CDN 1 (plates not shown).

LacZ Reporter Gene

In both experiments, the *lacZ* reporter gene did not appear to be as sensitive as the other two selections and only the stronger kit controls showed an interaction, indicated by the blue colour (C, D and E; Figure 3.4). Control B had a slight blue tinge in all plates, indicating a weak interaction. No interaction was detected for Arabidopsis eIF(iso)4E with TuMV CDN 1 VPg (example in Figure 3.4). No interaction was detected for correctly spliced R-o-18 BraA.eIF(iso)4E.a with TuMV CDN 1 VPg (Figure 3.4), or for either of the RLR22 BraA.eIF(iso)4E.a splice variants with TuMV CDN 1 VPg.
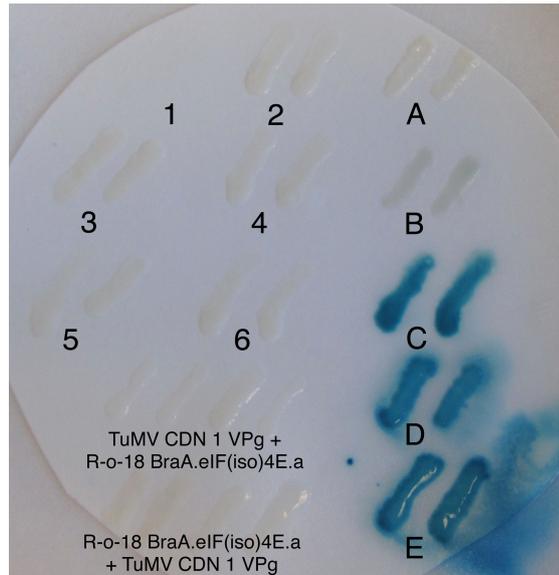
Figure 3.4 – Interaction of R-o-18 BraA.eIF(iso)4E.a + *Turnip mosaic virus* CDN 1 VPg, using the *lacZ* reporter gene where blue coloured colonies indicate an interaction.
A, Negative control; B, Weak interaction control; C, Moderately strong interaction control; D, Strong interaction control; E, Very strong interaction control; 1, MaV203 (Negative control); 2, Empty pDEST32 + pDEST22 vectors (Negative control); 3, TuMV CDN 1 VPg + *Arabidopsis thaliana* eIF(iso)4E (Positive control); 4, Empty pDEST32 vector + Arabidopsis eIF(iso)4E (Negative control); 5, R-o-18 BraA.eIF(iso)4E.a + Empty pDEST22 vector (Negative control); 6, Empty pDEST32 vector + R-o-18 BraA.eIF(iso)4E.a (Negative control).

3.2.2 Investigation of Interaction with Further *B. rapa* and TuMV Viral Proteins

Two further TuMV viral proteins were investigated, the cylindrical inclusion (CI) protein (1932 nt) and the 6K2 protein (159 nt; see Section 1.3.4 and Figure 1.3). Two further *B. rapa* proteins were included, BraA.eIF(iso)4E.c and BraA.eIF4E.c, the candidates for the dominant resistance gene *ConTR01*. Efforts were focused on investigating interactions between the TuMV-susceptible plant line, R-o-18 proteins and the TuMV CDN 1 viral proteins and combinations of viral proteins (6K2-VPg and CI-6K2). Unfortunately it was not possible to transform the CI-6K2-VPg construct into the vectors due to the large size of the construct (2,667 nt in total).

No interaction could be detected between any of the R-o-18 eIF4E/eIF(iso)4E proteins and any of the TuMV proteins, or protein combinations, using all three reporter genes (Summary of data in Table 3.1).

Table 3.1 – Interactions between plant eukaryotic translation initiation factors and *Turnip mosaic virus* CDN 1 viral proteins using the yeast two-hybrid system.

TuMV	R-o-18			RLR22		Arabidopsis Controls	
	BraA.eIF(iso)4E.a	BraA.eIF(iso)4E.c	BraA.eIF4E.c	BraA.eIF(iso)4E.a[a] (Whole of intron 1 retained)	BraA.eIF(iso)4E.a[b] (15 nt of intron 1 retained)	At.eIF(iso)4E	At.eIF4E
VPg	-[c]	-	-	-	-	+[d]	-
CI	-	-	-	NT[e]	NT	-	-
CI-6K2	-	-	-	NT	NT	-	-
6K2	-	-	-	NT	NT	-	-
6K2-VPg	-	-	-	NT	NT	-	-

[a] Construct of RLR22 BraA.eIF(iso)4E.a retaining an extra G (indel) and the whole of intron 1.
[b] Construct of RLR22 BraA.eIF(iso)4E.a retaining the last 15 nt of intron 1.
[c] No interaction detected with *HIS3*, *URA3* or *lacZ* reporter genes.
[d] Interaction detected with *HIS3* and *URA3* reporter genes.
[e] Not tested.

3.3 Discussion

3.3.1 Overview of the Y2H System

It was hoped that the Y2H system would provide a simple and effective way to screen *B. rapa* eIF4E/eIF(iso)4E proteins for interaction with the VPgs of different TuMV isolates, but no interactions were detected between the VPg of TuMV isolate CDN 1 and these proteins from the susceptible plant line and a TuMV isolate able to infect this line. The Arabidopsis eIF(iso)4E with TuMV CDN 1 VPg interaction control tested positive with two of the reporter genes (*HIS3* and *URA3*). This was a useful control as it showed that the system worked. No interaction was observed for Arabidopsis eIF(iso)4E with TuMV CDN 1 VPg using the *lacZ* reporter gene. Control B, described as a weak interaction control, showed a weak interaction as indicated by the slight blue tinge (for example, Figure 3.4), this indicated that the Arabidopsis eIF(iso)4E with TuMV isolate CDN 1 VPg interaction was weaker than control B in the Y2H system, this does not necessarily reflect on the strength of the interaction *in planta*. It also indicated that the *lacZ* reporter was less sensitive than both *HIS3* and *URA3* reporter genes. The positive and negative kit controls verified the plate selections were working for each of the reporter genes. Although the controls verified the results, it is not possible to say that there is no interaction between the brassica and viral proteins. It is possible that the interaction between *B. rapa* BraA.eIF(iso)4E.a and the VPg of TuMV CDN 1 was very weak (much weaker than for Arabidopsis eIF(iso)4E) and that the tests were not sensitive enough to detect an interaction between these proteins. It is also possible that the system is not suitable for detecting an interaction between proteins from *B. rapa* and this TuMV isolate. The presence of the target genes in the transformants was checked by PCR, however, it would have been more informative to check for complementary DNA (cDNA) and confirm the expression of the genes in the yeast.

Abdul-Razzak *et al.* (2009) reported the first example of the CI protein's involvement in overcoming eIF4E-based resistance. Their work in lettuce showed that the C-terminus of the CI protein of LMV was involved in overcoming both *mo1*[1] and *mo1*[2] resistances. My work was predominantly on the VPg, but also investigated the CI and 6K2 proteins of TuMV isolate CDN 1 to see whether they interacted with *B. rapa* eIF4E/eIF(iso)4E. No interaction was observed between any of the TuMV viral proteins and brassica proteins. Experiments using the whole of the CI protein (644 amino acids) may have been over ambitious, as

larger proteins are more complex, although there are reports of proteins as large as 750 amino acids in length being successfully studied using Y2H (Wang *et al.*, 2004).

3.3.2 Recent Advances

Subsequent to my Y2H experiments, colleagues in China managed to confirm an interaction between BraA.eIF(iso)4E.a from a susceptible *B. rapa* line (Ji Zao Chun) and the VPg of TuMV isolate C4 (Nellist *et al.*, 2014). The same group were unable to detect an interaction between BraA.eIF(iso)4E.a from a TuMV-resistant plant line, BP8407, with the same resistance mechanism as RLR22 and the VPg of TuMV isolate C4. They were also unable to detect an interaction between the VPg of TuMV UK 1 and the TuMV-susceptible Ji Zao Chun BraA.eIF(iso)4E.a (W. Qian, Personal Communication). It is not clear why no interaction was observed between the VPg of TuMV UK 1 and BraA.eIF(iso)4E.a from the TuMV-susceptible *B. rapa* line Ji Zao Chun, when the system worked for the VPg of TuMV isolate C4. Possible explanations are discussed below.

3.3.3 Appropriateness of Experiments

The Y2H assay is subject to a significant degree of false negatives, which have been described as genuine interactions that have been tested experimentally but failed to be detected in the Y2H assay (Edwards *et al.*, 2002). All the Y2H work performed in this chapter was based on N-terminal fusion proteins. Stellberger *et al.* (2010) recommended using both C- and N-terminal fusion proteins and combinations of the two for both bait and prey proteins to improve the screen. Colleagues from China used C-terminal eIF(iso)4E and N-terminal VPg constructs (W. Qian, Personal Communication), this may possibly explain their success with the system. Although, this still does not explain the lack of interaction observed for the VPg of TuMV isolate UK 1 and the BraA.eIF(iso)4E.a from the TuMV-susceptible *B. rapa* line Ji Zao Chun. Despite the VPg being the avirulence determinant (Keller *et al.*, 1998), Gao *et al.* (2004) could not detect an interaction between pea eIF4E and the VPg of *Pea seed-borne mosaic virus* (PSbMV) using the Y2H system. The findings in this chapter show that the Y2H assay does not work for all potyvirus VPg-eIF4E/eIF(iso)4E interactions and is not suitable for all viruses and in the case of TuMV appears to be virus isolate-specific. This may explain why the system did not work for TuMV and

B. rapa. In some cases other viral genes (P3 and CI) are involved in overcoming the recessive resistance (Hjulsager *et al.*, 2006; Abdul-Razzak *et al.*, 2009), which complicates the situation. It is possible that the VPg is the sole determinant interacting with the *B. rapa* eIF(iso)4E, as a single amino acid substitution in the VPg of TuMV (that abolishes the interaction with Arabidopsis eIF(iso)4E in Y2H experiments), results in the virus unable to infect *B. rapa* var. *perviridis* (Léonard *et al.*, 2000). Further studies failed to detect an interaction between the VPg of TuMV and eIF(iso)4E/eIF4E from *B. rapa* var. *perviridis in planta* (Léonard *et al.*, 2004).

Alternative techniques that have been utilised in detecting eIF4E-VPg interactions include co-immunoprecipitation (for example, Plante *et al.*, 2004) and firefly luciferase complementation (for example, Truniger *et al.*, 2008). Work by Ashby *et al.* (2011) highlighted the usefulness of the bimolecular fluorescence complementation (BiFC) technique when studying the eIF4E-VPg interaction between pea and PSbMV. Ashby *et al.* (2011) observed an interaction *in vivo* between an eIF4E from a susceptible plant and PSbMV VPg using BiFC. The use of other techniques, such as BiFC may lead to the detection of a direct interaction between R-o-18 BraA.eIF(iso)4E.a and the VPg of TuMV isolate CDN 1. The *dSpm* Arabidopsis complementation demonstrated by Jenner *et al.* (2010) and Nellist *et al.* (2014) indicated that TuMV could use two copies of *eIF4E* (*BraA.eIF4E.a* and *BraA.eIF4E.c*) and two copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c*) from both the TuMV-susceptible line R-o-18 and the TuMV-resistant line RLR22, which it clearly could not use in *B. rapa*. Hence any positive Y2H interaction results may be mis-leading in terms of what happens *in vivo* in *B. rapa*.

To further this particular area of research, a more informative approach to determining which copies of eIF4E/eIF(iso)4E TuMV can access and use would be to knock-out the genes in *B. rapa*. This would however, be a complicated task as *B. rapa* has three copies of *eIF4E* (although one, *BraA.eIF4E.b*, is a pseudogene) and three copies of *eIF(iso)4E* and so could involve numerous combinations of knock-outs. It is also not known which copies are critical for the plant's survival, so knocking out particular copies of *eIF4E/eIF(iso)4E* is likely to have adverse effects on the plant, or be lethal.

3.4 Conclusions

The main aim of the work in this chapter was to test whether the Y2H system would provide a simple assay that could be used to assess which of the eIF4E and eIF(iso)4E proteins of *B. rapa* might be functional for TuMV isolates. The variability between the TuMV isolates used indicates that the Y2H assay is not a good assay for assessing interaction as it seems to be isolate-specific.

Chapter 4

Expression and Mis-splicing of *retr01*

4.1 Background

4.1.1 Recessive Resistance Mechanisms

The involvement of eukaryotic translation initiation factor 4E (eIF4E), its isoform eIF(iso)4E (Robaglia and Caranta, 2006) and to a lesser extent eIF4G (Le Gall *et al.*, 2011) has been identified in the investigation of recessive resistances to plant viruses. This was predominantly evident for recessive resistances to members of the *Potyviridae* family. Mutations in *eIF4E* and *eIF(iso)4E* in a range of plant species have been shown to confer resistance to a range of potyviruses (Robaglia and Caranta, 2006). Robaglia and Caranta (2006) concluded the diversity of the resistance phenotypes was a result of a few amino acid changes in the eIF4E proteins encoded by the recessive resistance alleles. The amino acid changes associated with resistance were clustered near the cap-binding pocket and at the surface of the protein (Robaglia and Caranta, 2006).

4.1.2 Intron Splicing in Eukaryotes

Eukaryotic genes are comprised of exons, the protein-coding sequence and introns, the non-coding sequence. Transcription takes place in the nucleus and involves copying the whole of the gene from the transcription start site to the transcription stop site to produce a pre-messenger RNA (pre-mRNA) that pos-

sess the transcribed intron sequences. Intron splicing is the modification of the pre-mRNA transcript, by removing the introns and joining the exons to produce a translatable mature mRNA that has a poly(A) tail attached and exported from the nucleus (Figure 4.1). The splicing process in eukaryotes is performed by the spliceosome. The spliceosome is an extremely complex ribonucleoprotein megaparticle that assembles around splice sites at each intron. The majority of spliceosomal introns contain |GU at the donor splice site (5' splice site; splice site is shown by vertical line) and AG| at the acceptor splice site (3' splice site) (Rogozin *et al.*, 2012).
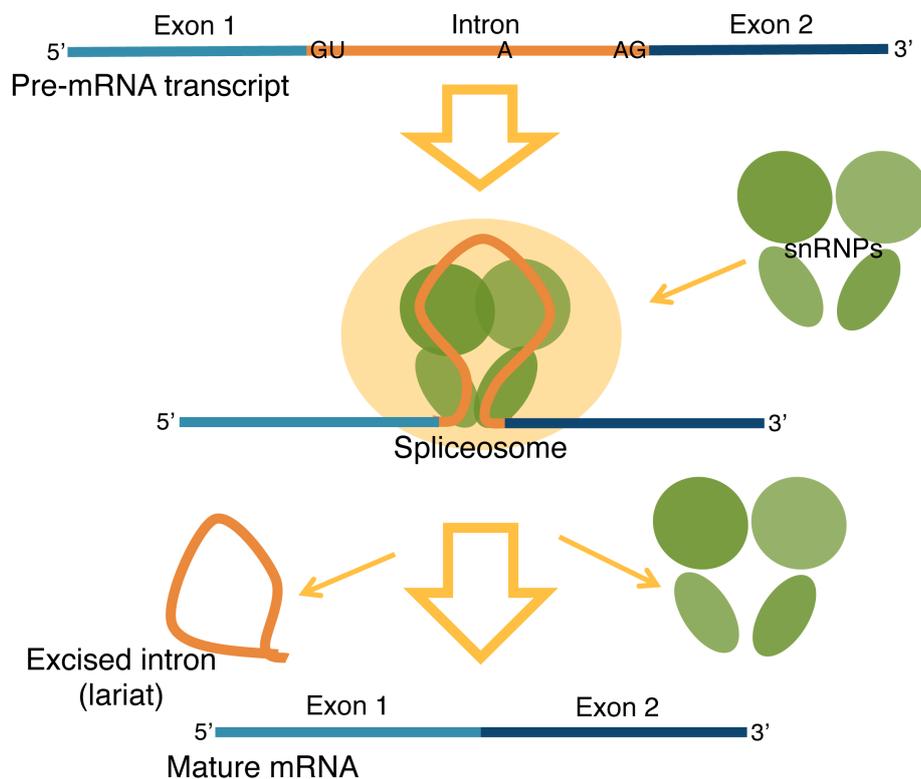
Figure 4.1 – Simplified diagram of pre-mRNA splicing. Introns are recognised at their donor splice site (GU) and their acceptor splice site (AG) and are removed by small nuclear ribonucleoproteins (snRNPs), which bind to form the spliceosome. The exons are spliced together to produce translatable messenger RNA (mature mRNA).

Alternative Splicing in Plants

Alternative splicing is an important mechanism, which increases the diversity of the transcriptome and the proteome (Keren *et al.*, 2010). Alternative splicing is the process by which the transcription of a gene can be altered slightly by the combination of exons and/or inclusion of introns to produce variants of the

complete protein. There are four basic alternative splicing events: alternative 5' splice-site, alternative 3' splice-site, cassette-exon inclusion, or skipping and intron retention (Nilsen and Graveley, 2010). It has been reported in plants, that intron retention is the most common alternative event with 30-50% of *Arabidopsis thaliana* alternative splicing events and 54% of rice alternative splicing events being intron retention compared to approximately 10% in humans (Ner-Gaon *et al.*, 2004; Wang and Brendel, 2006).

Alternative splicing has previously been associated with resistance to a plant virus (Dinesh-Kumar and Baker, 2000). The *N* resistance gene in tobacco confers resistance to *Tobacco mosaic virus* (TMV) (Whitham *et al.*, 1994). The *N* gene was alternatively spliced into two transcripts, N_S encoding the full length N protein and N_L encoding a truncated N protein (N^{tr}), lacking 13 of the 14 repeats of the leucine-rich region (LRR). Transgenic tobacco plants that possessed both of these transcripts had complete resistance to TMV. Transgenic plants possessing only the full length N protein did not show complete resistance to TMV (Dinesh-Kumar and Baker, 2000).

4.1.3 Previous Work on RLR22

Brassica rapa var. *pekinensis* (Chinese cabbage) line RLR22, has broad-spectrum resistance to *Turnip mosaic virus* (TuMV) (Walsh *et al.*, 2002) and was derived from an accession identified in a screen of >3,000 lines (Liu *et al.*, 1996). To map the resistance gene(s), the TuMV-resistant line RLR22 was crossed with the TuMV-susceptible line, R-o-18, of the closely related *B. rapa* ssp. *trilocularis* (Roxb.) Hanlet. (yellow sarson). Three copies of *eIF4E* (*BraA.eIF4E.a*, *BraA.eIF4E.b* and *BraA.eIF4E.c*) and three copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.a*, *BraA.eIF(iso)4E.b* and *BraA.eIF(iso)4E.c*) have been identified and sequenced from the *B. rapa* lines R-o-18 (Jenner *et al.*, 2010) and RLR22 (Nellist *et al.*, 2014). Segregation of resistance/susceptibility in the offspring revealed that the resistance was due to a recessive gene, *retr01* (coincident with the copy of *eIF(iso)4E* on chromosome A4 (*BraA.eIF(iso)4E.a*)) that was epistatic to a dominant gene, *ConTR01* (coincident with one of the other copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.c*), or one of the copies of *eIF4E* (*BraA.eIF4E.c*), both on chromosome A8) (Rusholme *et al.*, 2007; Nellist *et al.*, 2014).

This chapter describes work on a number of *B. rapa* lines with broad-spectrum resistance to TuMV. The main aim of the study was to confirm the identity of the recessive resistance gene *retr01* and to explain the mechanism of the *retr01*-based resistance.

4.2 Results

4.2.1 Discovery of the Mis-splicing of *retr01*

BraA.eIF(iso)4E.a is located on chromosome A4 of *B. rapa* (Rusholme *et al.*, 2007; Nellist *et al.*, 2014). Comparison of the genomic DNA (gDNA) sequences of *BraA.eIF(iso)4E.a* from R-o-18 and RLR22 revealed that there was an extra G present at the exon 1-intron 1 splice junction, at position 201 nucleotides (nt) in RLR22 (Figure 4.2).

Figure 4.2 – DNA sequence around the intron 1 splice site junction in *Brassica rapa* R-o-18 and RLR22. Exons are show in green boxes and any difference between the two sequences in this region are in red text.
(a) *B. rapa* R-o-18 DNA segment. (b) *B. rapa* RLR22 DNA segment, showing the indel (extra G present at position 201 nt).

RT-PCR analysis of *BraA.eIF(iso)4E.a* from R-o-18 and RLR22 in preparation for yeast two-hybrid assay revealed that the complementary DNA (cDNA) product of RLR22 *BraA.eIF(iso)4E.a* was larger (Figure 4.3; lane 2) than the product of R-o-18 *BraA.eIF(iso)4E.a* (Figure 4.3; lane 1). It was expected that the cDNA products would be the same size (both predicted to be 600 nt). Sequencing of the RLR22 *BraA.eIF(iso)4E.a* cDNA product revealed that it had retained the extra G from the gDNA (indel) and also the whole of intron 1 (63 nt). The retention of the extra G and the whole of intron 1 resulted in a premature stop codon at position 234 nt in *BraA.eIF(iso)4E.a* in RLR22 (Figure 4.9, (b)).

Figure 4.3 – RT-PCR amplification of *Brassica rapa* and *Arabidopsis thaliana* *eIF4E* and *eIF(iso)4E* cDNA products using specific primers CN3 and CN4/CN5 for *B. rapa* and CN6/8 and CN7/9 for Arabidopsis. Arabidopsis *eIF4E* and *eIF(iso)4E* were 708 bp and 597 bp respectively, whilst R-o-18 and RLR22 *BraA.eIF(iso)4E.a* were expected to be the same size, approximately 600 bp.
Lanes: L, 1 Kb Plus DNA Ladder; 1, R-o-18 *BraA.eIF(iso)4E.a*; 2, RLR22 *BraA.eIF(iso)4E.a*; 3, Arabidopsis *eIF4E*; 4, Arabidopsis *eIF(iso)4E*; 5, dH$_2$O (negative control).

4.2.2 Investigation of Other *B. rapa* Lines

Verification of Resistance/Susceptibility

Several TuMV-resistant lines were investigated to see whether their resistances to TuMV were related to the presence of the indel and mis-splicing of *BraA.eIF(iso)4E.a*. Plants to be genotyped were challenged with TuMV isolate CDN 1 to verify resistance/susceptibility. The lines R-o-18 (Rusholme *et al.*, 2007) (Figure 4.4), S$_3$.RL2-43 (Walsh *et al.*, 2002), CK 1 and CK 2 (see Chapter 6 for details) were uniformly susceptible to TuMV isolate CDN 1, all displaying systemic mosaic symptoms in uninoculated leaves following TuMV CDN 1 challenge. Infection was verified by ELISA. The line Jin G55 (Hughes *et al.*, 2002) segregated for resistance, two plants showed no systemic spread of the virus (for example, Figure 4.6, (a)), two plants showed severe systemic mosaic symptoms (for example, Figure 4.6, (d)) and six plants showed varying limited systemic symptoms (for example, Figure 4.6, (b) and (c)). Infection was confirmed by ELISA, the plants with milder symptoms, had lower ELISA values. The lines RLR22 (Walsh

57

et al., 2002) (Figure 4.5), BP058 (Walsh *et al.*, 2002) and Jong Bai No. 2 (Hughes *et al.*, 2002) were resistant to TuMV isolate CDN 1, showing chlorotic spots on inoculated leaves, but no systemic spread of the virus. No systemic spread of the virus could be detected by ELISA on uninoculated leaves. The lines *B. rapa rapa* (Kassem and Walsh, 2008), BR03023 (Unpublished) (Figure 4.7) and BR05060 (Unpublished) were immune to TuMV, showing no symptoms in inoculated leaves and no systemic spread of the virus. No systemic spread of the virus could be detected by ELISA on uninoculated leaves.
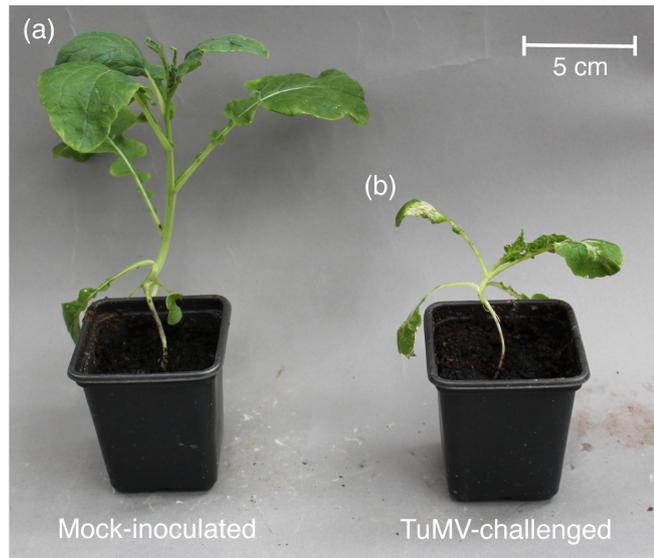
Figure 4.4 – R-o-18 symptoms following mechanical challenge with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
(a) Mock-inoculated R-o-18 control plant. (b) R-o-18 plant challenged with TuMV CDN 1 displaying systemic mosaic symptoms.
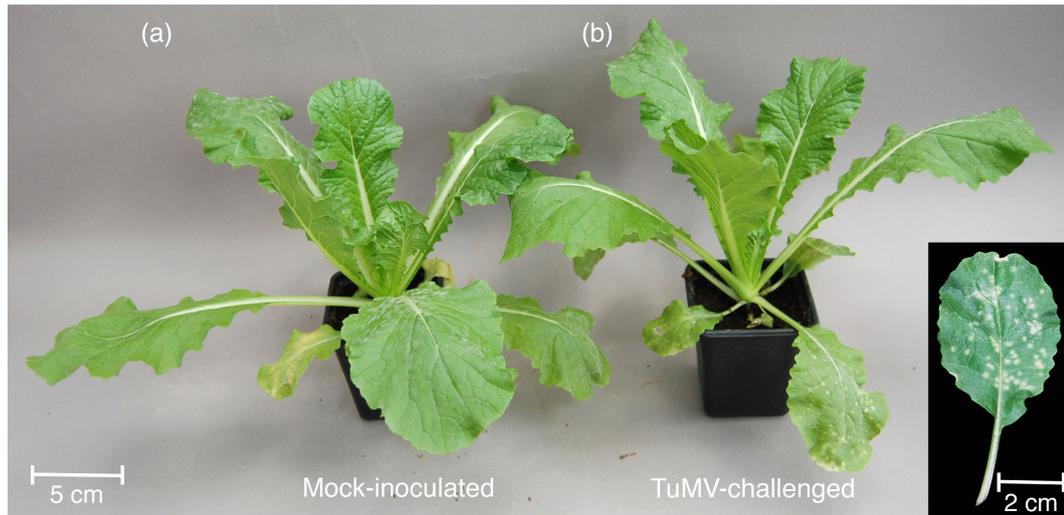
Figure 4.5 – RLR22 symptoms following mechanical challenge with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
(a) Mock-inoculated RLR22 control plant. (b) RLR22 plant challenged with TuMV isolate CDN 1 displaying chlorotic spots on inoculated leaves but no systemic symptoms. Close up of chlorotic spots on an inoculated leaf (inset).

Expression of BraA.eIF(iso)4E.a

To study the expression of *BraA.eIF(iso)4E.a*, RT-PCR analysis was performed on all the lines and the products were sequenced. All CDN 1 TuMV-susceptible lines (lacking the indel) produced a single product approximately 600 nt in size (Figure 4.8; R-o-18, CK 1, CK 2, Jin G55 completely susceptible plants and S$_3$RL2-43), corresponding to the correctly spliced version of the gene. All the lines possessing the indel were resistant to TuMV and had a larger major product of approximately 664 nt, plus a smaller minor product of approximately 600 nt (Figure 4.8; RLR22, BP058, Jong Bai No. 2 and Jin G55 resistant plants). Sequencing of larger RT-PCR products revealed that they retained the extra G and the whole of intron 1. The smaller products appeared to be of a similar size to the correctly spliced version of the gene. Three lines with broad-spectrum resistance that did not possess the indel, did not mis-splice *BraA.eIF(iso)4E.a* (Figure 4.8; BR03023, BR05060 and *B. rapa rapa*). Sequencing of RT-PCR products of *BraA.eIF(iso)4E.a* showed that introns 2-4 were correctly spliced in all resistant and susceptible plants. The retention of intron 1 by all lines possessing the indel, resulted in the introduction of a premature stop codon at position 234 nt (Figure 4.9, (b)). All lines possessing the indel, mis-spliced *BraA.eIF(iso)4E.a* (*retr01*) and had broad-spectrum resistance to TuMV (Table 4.1).
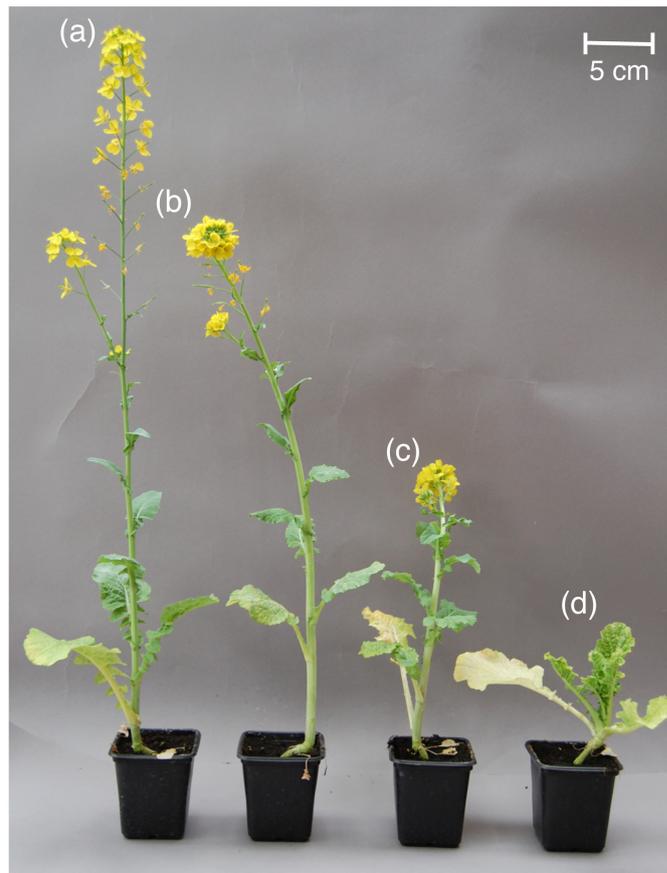
Phenotype	0	(+)	(+)	+
ELISA optical density	0.012	0.195	0.315	0.619

~600 bp

Figure 4.6 – Jin G55 symptoms, ELISA results and *BraA.eIF(iso)4E.a* cDNA following mechanical challenge with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation. Phenotype 0, no systemic spread of TuMV; (+), limited systemic spread of TuMV; +, full systemic spread of TuMV. ELISA absorbance measured at 405 nm. RT-PCR amplification of *BraA.eIF(iso)4E.a* in Jin G55 using specific primers CN3 and CN5. Correctly spliced versions of *BraA.eIF(iso)4E.a* are approximately 600 nt.

(a) Jin G55 plant showing no systemic spread of TuMV, low ELISA reading and mis-splicing of *BraA.eIF(iso)4E.a* as indicated by larger major band (~664 bp) and smaller minor band (~600 bp). (b) Jin G55 plant displaying limited systemic symptoms with mild chlorotic spots on uninoculated leaves, with equal strength products for the two bands (~664 bp and ~600 bp). (c) Jin G55 plant displaying slightly more severe limited systemic mosaic symptoms with a brighter lower band (~600 bp). (d) Jin G55 plant displaying systemic mosaic symptoms with a relatively high ELISA reading and a single product of correctly spliced size (~600 bp) detected.
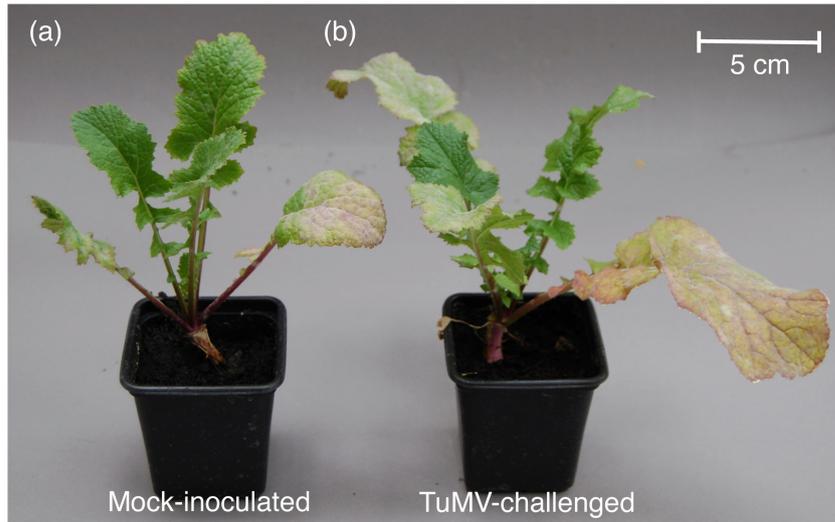
Figure 4.7 – BR03023 symptoms following mechanical challenge with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
(a) Mock-inoculated BR03023 control plant. (b) BR03023 plant challenged with TuMV isolate CDN 1 displaying no infection.
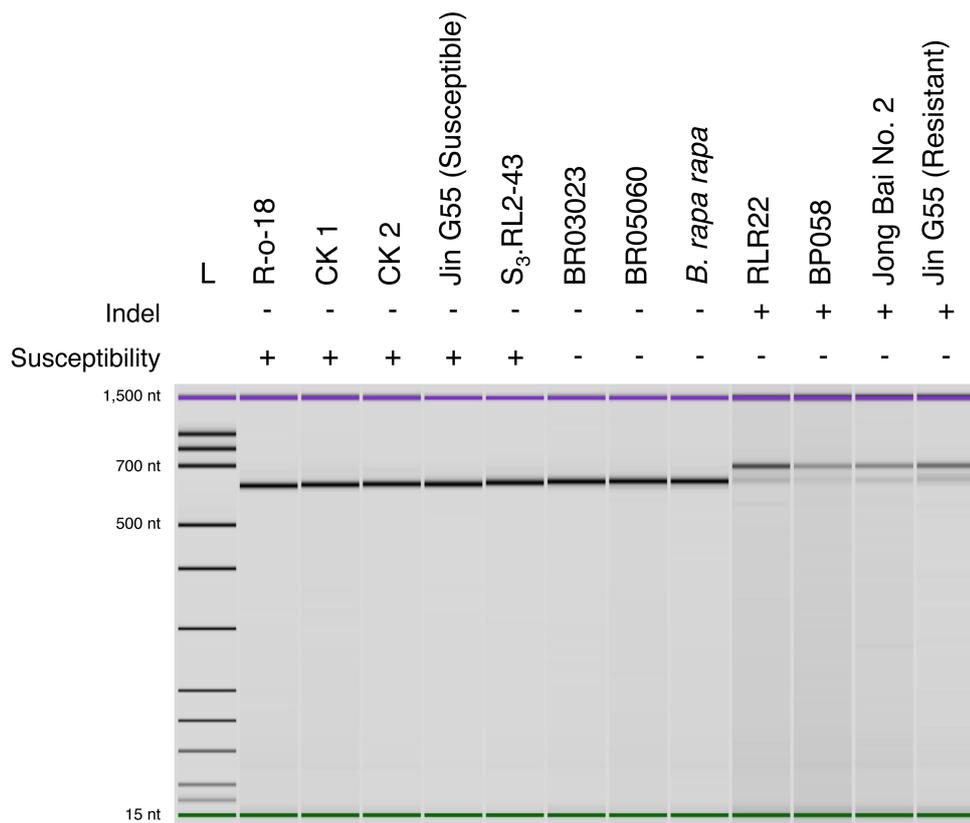
Figure 4.8 – Detection of a single transcript in *Brassica rapa* plants lacking the indel, corresponding to correctly spliced *BraA.eIF(iso)4E.a* and splice variants in *Turnip mosaic virus*-resistant *B. rapa* plants possessing the indel, separated using an Agilent 2100 Bioanalyser.

Table 4.1 – Phenotypes and *BraA.eIF(iso)4E.a* genotypes of *Brassica* lines following challenge with *Turnip mosaic virus* isolate CDN 1.

Plant Line	Broad-spectrum Resistance	Indel	Phenotype[a]	Mis-spliced
B. rapa rapa	+	-	0[b]	-
BR03023	+	-	0	-
BR05060	+	-	0	-
BP058	+	+	R[c]	+
Jong Bai No. 2	+	+	R	+
Jin G55[d]	+/-	+/-	R/+[e]	+/-
S$_3$.RL2-43[f]	-	-	+	-
CK 1	-	-	+	-
CK 2	-	-	+	-
R-o-18	-	-	+	-
RLR22	+	+	R	+

[a] Plants challenged with TuMV isolate CDN 1.

[b] No infection.

[c] Infection of inoculated leaves but no systemic spread, no virus detected by ELISA in uninoculated leaves.

[d] Line segregating for the presence of the indel.

[e] Systemic mosaic symptoms, infection confirmed by ELISA.

[f] Line possesses the dominant resistance gene *TuRB01* which confers resistance to TuMV pathotype 1 isolates.

Further analysis of the Jin G55 plants segregating for resistance and susceptibility to TuMV, found that the resistant plants mis-spliced *BraA.eIF(iso)4E.a* (Figure 4.6, (a)) and the susceptible plants did not (Figure 4.6, (d)). Interestingly plants displaying limited systemic symptoms appeared to have equal strength bands for mis-spliced and correctly spliced products (Figure 4.6, (b)). It is likely these plants were heterozygous for the presence of the indel, although this was not determined.

4.2.3 RT-PCR Analysis of the Expression of *BraA.eIF(iso)4E.a* in RLR22

Cloning of the less abundant RT-PCR products from RLR22 revealed four, less common variants, as well as the major mis-spliced variant retaining the extra G and the whole of intron 1 (Figure 4.9, (b)). The less common variants included a variant retaining the extra G (indel) and the last 14 nt of intron 1 (or alternatively, if the G comprises part of intron 1, then the variant retained the last 15 nt of intron 1), resulting in a slightly elongated, in-frame mRNA sequence (Figure 4.9, (c)), one with intron 1 excised along with the last 3 nt of exon 1, resulting in a slightly shortened, in-frame mRNA with a substitution (Figure 4.9, (d)), one with an extra G at the end of exon 1, resulting in a premature stop codon at position 228 nt (Figure 4.9, (e)) and one that was correctly spliced (Figure 4.9, (a)).
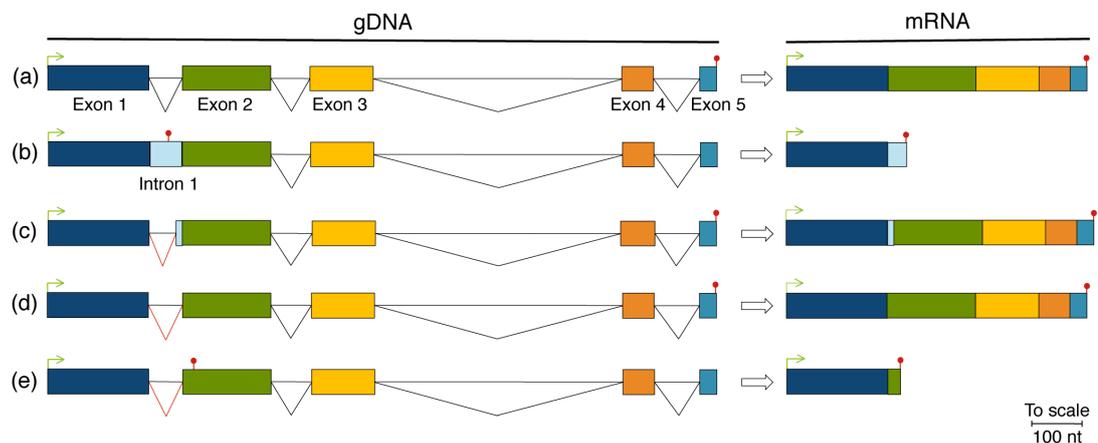
Figure 4.9 – *BraA.eIF(iso)4E.a* splice variants detected in *Brassica rapa* RLR22 by RT-PCR.
(a) Correctly spliced RLR22 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Most common mis-spliced variant retaining the extra G (indel) and whole of intron 1 resulting in a premature stop codon at position 234 nt (664 nt in length). (c) Variant retaining the extra G (indel) and the last 14 nt of intron 1/retaining the last 15 nt of intron 1, resulting in a slightly elongated, in-frame mRNA (615 nt). (d) Variant lacking the last 3 nt of exon 1, resulting in a slightly shortened, in-frame mRNA with a substitution (597 nt). (e) Variant retaining the extra G (indel) resulting in a premature stop codon at position 228 nt.

4.2.4 Analysis of *BraA.eIF(iso)4E.a* Splice Variants Detected in RLR22

In total there were four *BraA.eIF(iso)4E.a* splice variants detected in RLR22, of these, two introduced a premature stop codon, severely truncating the protein and two were in-frame, one with a substitution and missing an amino acid and the other with an additional five amino acids. Interestingly, a correctly spliced version of *BraA.eIF(iso)4E.a* was also detected. Comparing the splice variants with the correctly spliced *BraA.eIF(iso)4E.a* and the published *eIF4E* and *eIF(iso)4E* sequences in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008), indicated that both the variants that are in-frame could be functional for TuMV. The variant possessing an additional five amino acids had the correct amino acids at important sites (Figure 4.10, (b)). The variant with a substitution and lacking one amino acid had the correct amino acids at important sites, apart from missing a tryptophan residue at position 67 (Figure 4.10, (c)). The other two variants, which comprised the majority of those transcripts sequenced, were unlikely to be functional for TuMV, as they were missing some of the important amino acids (Figure 4.10, (d) and (e)).

```
MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG-----LHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG
--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+
        10        20        30        40        50        60        70        80        90        100       110
--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+
```

(a) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG-----LHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG 104
(b) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGIFFCSLHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG 109
(c) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDF------CLHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG 103
(d) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGVCLSSPFTYC.   78
(e) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGFARDYIHP.   76
(f) MATDDVNEPLPAAAELPATE-AEKQP-HKLERKWSFWFDNQS--KKGAAWGASLRKAYTFDTVEDFWG-----LHETIFQTSKLTANAEIHLFKAGVEPKWEDPECANGG 101
(g)   W   W   W   E WE C

```
KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV-
--------+---------+---------+---------+---------+---------+---------+---------+---------+--------
        120       130       140       150       160       170       180       190       200
--------+---------+---------+---------+---------+---------+---------+---------+---------+--------
```

(a) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV. 200
(b) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV. 205
(c) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV. 199
(d)   78
(e)   76
(f) KWTWVVTANRKEALDKGWLETLMALIGEQFDEADEICGVVASVRPQSKQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFNNHDDSRRSRFTV. 199
(g)   GE   C   R   K W

Figure 4.10 – Multiple alignment of amino acid sequences of correctly and mis-spliced *BraA.eIF(iso)4E.a* detected in RLR22, based on homology with *eIF4E* and *eIF(iso)4E* from other plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008). Amino acids that are most directly involved in binding the mRNA cap-structure are highlighted in green. Amino acids that are involved in binding eIF4G are highlighted in blue. Cys residues involved in forming a di-sulphide bridge in wheat and implicated in mRNA cap binding are highlighted in yellow.
(a) RLR22 BraA.eIF(iso)4E.a, correctly spliced (Figure 4.9, (a)). (b) Variant retaining extra G (indel) and last 14 nt of intron 1/last 15 nt of intron 1 (Figure 4.9, (c)). (c) Variant lacking the last 3 nt of exon 1 (Figure 4.9, (d)). (d) Variant retaining extra G (indel) and the whole of intron 1 (Figure 4.9, (b)). (e) Variant retaining extra G (indel) (Figure 4.9, (e)). (f) *Arabidopsis thaliana* eIF(iso)4E. (g) Important amino acids.

4.3   Discussion

4.3.1   Identification of *retr01*

The link between plants with mutations in their eukaryotic translation initiation factors and the inability of potyviruses to infect these plants has been well documented (Robaglia and Caranta, 2006; Gómez *et al.*, 2009; Truniger and Aranda, 2009; Le Gall *et al.*, 2011; Wang and Krishnaswamy, 2012). However, most of these incompatibilities have been found to be strain-specific. Unlike many previous examples of translation factor-based resistance, the *retr01*-based mechanism provides broad-spectrum resistance. This study is the first example of natural mis-splicing of a eukaryotic initiation factor gene resulting in virus resistance. The study also confirmed the identity of *retr01* as *BraA.eIF(iso)4E.a* and clearly demonstrated that the presence of the indel caused the mis-splicing

of *BraA.eIF(iso)4E.a* and was responsible for the broad-spectrum resistance observed in the *B. rapa* plants. It is unknown if the different *retr01*-based resistant lines identified in my work have a common origin, or evolved independently. It is probable that they have a common origin, although the lineage of each line is unknown.

4.3.2 Implications *in planta*

The most-common mis-splicing of the *BraA.eIF(iso)4E.a* (resulting in the variant retaining the extra G and the whole of intron 1) introduced a stop codon at position 234 nt; it was predicted that this would result in a truncated protein, that is non-functional. Colleagues in China studied the plant line BP8407, that also mis-splices *BraA.eIF(iso)4E.a*. They expressed the most common mis-spliced *BraA.eIF(iso)4E.a* variant (retaining the extra G and the whole of intron 1) in *Escherichia coli* Rosetta(DE3)pLysS with a 6x Histidine tag. The deduced sizes of the protein compared to *BraA.eIF(iso)4E.a* from the TuMV-susceptible line Ji Zao Chun was consistent with the predicted sizes (8.67 kDa and 22.50 kDa, respectively) (Nellist *et al.*, 2014). This confirmed the mis-splicing of *BraA.eIF(iso)4E.a* resulted in a truncated protein, which is unlikely to be functional for either the plant or TuMV (Nellist *et al.*, 2014). Resistance to potyviruses mediated by eIF4E has principally been associated with a small number of changes in the amino acid sequence of the eIF4E protein (Robaglia and Caranta, 2006). In contrast, the *retr01*-based resistance is a result of an insertion of an extra G at position 201 nt, which leads to the majority of *BraA.eIF(iso)4E.a* transcripts being mis-spliced and the associated proteins being truncated. The *Capsicum annuum pvr6* resistance locus was also found to result in a defective *eIF(iso)4E* (Ruffel *et al.*, 2006), where an 82 nt deletion in gDNA resulted in a truncated protein. As three copies of *eIF4E* and three copies of *eIF(iso)4E* had been confirmed in *B. rapa* (Jenner *et al.*, 2010; Nellist *et al.*, 2014), it facilitates that a copy/copies that TuMV would normally use in *B. rapa*, could be non-functional for both plant and virus, without disadvantaging the plant. This supports the suggestions of Charron *et al.* (2008) that a trade-off between potyvirus resistance and eIF4E host functions can result in defective eIF4E proteins.

4.3.3 Induced Mis-splicing

The technique of TILLING (targeting induced local lesions in genomes) artificially induced mis-splicing of *eIF4E* in tomato (Piron *et al.*, 2010). The induced mutation led to a substitution in the splice site of *eIF4E* (G1485A), which resulted in the deletion of exons 2 and 3 in the mRNA, the truncation of the eIF4E protein and resistance to the potyvirus *Pepper mottle virus* (PepMoV) and strain-specific resistance to *Potato virus Y* (PVY) (Piron *et al.*, 2010). It was reported that at least one isolate of PVY was able to overcome the resistance and was suggested that the resistance-breaking PVY isolate was able to use a different copy of *eIF4E* (Piron *et al.*, 2010). Segregation ratios and mapping showed that TuMV was not able to use other copies of eIF4E, or eIF(iso)4E to spread systemically in RLR22 *B. rapa* (Rusholme *et al.*, 2007).

4.3.4 Reports of Recessive Broad-spectrum TuMV Resistance in Chinese Cabbage

Genetic analysis of broad-spectrum resistance to TuMV observed in Chinese cabbage line BP8407, indicated that the resistance was controlled by a single recessive gene (Qian *et al.*, 2013). The resistance was mapped to chromosome A4 (same as *retr01*), however, they concluded that a single nucleotide polymorphism (SNP), which resulted in an amino acid change G152N and correlated with the susceptibility/resistance (Qian *et al.*, 2013). The resistance gene was named *retr02*. Having obtained the sequence of the reported *retr02* gene, it was found to be identical to *retr01*. This along with the fact the gene is mis-spliced giving the same major variant shows that *retr02* is in fact *retr01* (Nellist *et al.*, 2014).

Broad-spectrum resistance to TuMV in Chinese cabbage has also been described in two lines, SB18 and SB22, by Kim *et al.* (2013). Genetic analysis indicated that the resistance was also controlled by a single recessive gene, *trs* (TuMV resistance discovered at Seoul National University), located on chromosome A4. They identified two alleles of *BraA.eIF(iso)4E.a*, *Braiso4Ea-1* and *Braiso4Ea-2* (missing exons 4 and 5, compared to *Braiso4Ea-1*). However, they ruled out *BraA.eIF(iso)4E.a* as the resistance gene, as the presence/absence of the *BraA.eIF(iso)4E.a* sequence deletion was not consistent between resistant and susceptible lines (Kim *et al.*, 2013). They suggest the gene may be tightly linked to *retr02* or another allele. They also suggested that the gene may not

belong to the eIF4E family (Kim *et al.*, 2013).

4.3.5   Durability of the *retr01*-based Resistance

The inability of a wide range of TuMV isolates from around the world to over-come the *retr01*-based resistance in RLR22 (Walsh *et al.*, 2002), indicates that the broad-spectrum resistance mechanism could provide durable resistance to TuMV. This however, will be dependant upon TuMV not mutating and being able to utilise/access other copies of eIF4E/eIF(iso)4E *in planta*. Selection pressures could also lead to TuMV evolving to be capable of effective cap-independent translation as described by Basso *et al.* (1994). Only by inoculating excessive amounts of TuMV to RLR22 resistant plants was it possible to induce the limited infection (chlorotic spots) in inoculated leaves, which was not able to spread systemically. This indicates that there was very little eIF(iso)4E protein present that was functional for the virus. It is also possible that the virus was less efficient than mRNA cap in competing for the protein, as evidence to date suggests that the viral protein genome-linked (VPg) of potyviruses competes with host plant mRNA cap for eIF4E binding (Gao *et al.*, 2004) and eIF(iso)4E binding (Plante *et al.*, 2004).

4.3.6   Broad-spectrum TuMV Resistance in Plants Lacking the Indel

The discovery of the resistance mechanism of *retr01* also facilitated the identification of lines with broad-spectrum resistance to TuMV (*B. rapa rapa* (Kassem and Walsh, 2008), BR03023 and BR05060) that did not possess the indel associated with mis-splicing of *BraA.eIF(iso)4E.a*. These lines must evade TuMV infection by a different mechanism to the mis-splicing of *BraA.eIF(iso)4E.a*. This provides the opportunity to develop commercial plant varieties possessing these alternative resistances. Such varieties could be exploited by alternating their plantings with lines possessing *retr01*, in order to reduce the selection pressure for resistance-breaking viral mutants and thereby extend the durability of *retr01*-based resistance and these other resistance sources in the field.

4.4 Conclusions

The main aim of this chapter was to confirm the identity of the recessive resistance gene *retr01* and explain the mechanism of the *retr01*-based resistance. *BraA.eIF(iso)4E.a* was confirmed as *retr01* and the mechanism of resistance was identified as the novel mis-splicing of *BraA.eIF(iso)4E.a*, caused by the presence of the indel (extra G, position 201 nt), resulting in non-functional Bra.A.eIF(iso)4E.a, which TuMV is unable to use.

Chapter 5

Identification of *ConTR01*

5.1   Background

5.1.1   Previous Work on the RLR22 and R-o-18 Cross

In previous work, Rusholme *et al.* (2007) crossed the TuMV-resistant plant line RLR22 with the TuMV-susceptible plant line R-o-18 to map the broad-spectrum resistance to TuMV in *Brassica rapa* (Figure 5.1).   The F_1 plants were all susceptible to *Turnip mosaic virus* (TuMV), indicating that the resistance is recessive.   RLR22 was selfed to produce more seed ($RLR22S_1$), which was grown on and was back-crossed with the F_1 plants.   One hundred and twenty B_1 plants were grown on and selfed to produce 120 B_1S_1 families.   The B_1S_1 population segregated for resistance to TuMV and it was determined that there were two loci controlling resistance to TuMV, *retr01* and *ConTR01* (Rusholme *et al.*, 2007).

In the previous chapter, the recessive resistance gene *retr01* was identified and the mechanism of resistance was determined (see Chapter 4).   *ConTR01* was mapped to the upper portion of the A8 chromosome and was coincident with one of the copies of eukaryotic translation initiation factor 4E (*eIF4E*), or one of the copies of its isoform (*eIF(iso)4E*) (Rusholme *et al.*, 2007).   Three copies of *eIF4E* (*BraA.eIF4E.a*, *BraA.eIF4E.b* and *BraA.eIF4E.c*) and three copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.a*, *BraA.eIF(iso)4E.b* and *BraA.eIF(iso)4E.c*) have been identified and sequenced from the *B. rapa* lines R-o-18 (Jenner *et al.*, 2010) and RLR22 (Nellist *et al.*, 2014).   One of the copies of *eIF4E* (*BraA.eIF4E.b*) lacked exons 2 and 3 and appeared to be a pseudogene (Jenner *et al.*, 2010).   *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* were mapped to chromosome

A8 and were coincident with *ConTR01* (Rusholme *et al.*, 2007). Sequencing of *BraA.eIF4E.c* revealed three amino acid differences between R-o-18 and RLR22, A35V, G45T and R105K (C. E. Jenner, Personal communication). Sequencing of *BraA.eIF(iso)4E.c* revealed four amino acid differences between R-o-18 and RLR22, L36F, V52A, T80I and Q150P (C. E. Jenner, Personal communication).
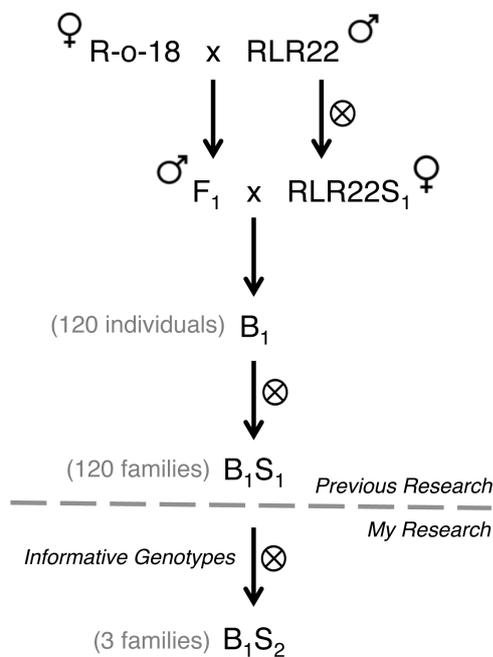
Figure 5.1 – Crossing strategy used to develop the segregating lines. $\otimes$, Self-pollination; $\male$, male parent; $\female$, female parent. RLR22 is resistant to *Turnip mosaic virus* isolate CDN 1 and R-o-18 is susceptible to isolate CDN 1.

5.1.2 Previous Unpublished Work on the Identification of *ConTR01*

Previous work on the RLR22 and R-o-18 cross indicated that *BraA.eIF(iso)4E.c* was a candidate for *ConTR01* (Rusholme *et al.*, 2007). The experiments showed that the plants homozygous for the RLR22 alleles of *BraA.eIF(iso)4E.a* (*retr01*) and *BraA.eIF4E.c* and the R-o-18 allele of *BraA.eIF(iso)4E.c* were susceptible, whilst plants that were homozygous for *retr01*, homozygous for the RLR22 allele of *BraA.eIF4E.c* and homozygous for the RLR22 allele of *BraA.eIF(iso)4E.c* were resistant. However, plants that were homozygous for *retr01*, homozygous for the RLR22 allele of *BraA.eIF4E.c* but heterozygous for *BraA.eIF(iso)4E.c* segregated for resistance (C. E. Jenner and J. Bambridge, Personal communication; Table 5.1). It should be noted that this was based on a small number of plants

of some genotypes.

Table 5.1 – Previous work on the involvement of *BraA.eIF(iso)4E.c* in broad-spectrum resistance to *Turnip mosaic virus* in RLR22.

Individual plant Genotype[b]	A$_{405}$ from ELISA on uninoculated leaf	Phenotype[a]
rr Ii CC	0.08	(+)[c]
rr Ii CC	0.02	(+)
rr Ii CC	0.02	R[d]
rr Ii CC	0.02	R
rr Ii CC	0.025	R
rr Ii CC	0.016	R
rr Ii CC	0.022	R
rr ii CC	0.3	(+)
rr II CC	0.023	R

[a] Plants challenged with TuMV isolate CDN 1.
[b] *r*, *retr01*; *I*, RLR22 allele of *BraA.eIF(iso)4E.c*; *i*, R-o-18 allele of *BraA.eIF(iso)4E.c*; *C*, RLR22 allele of *BraA.eIF4E.c*.
[c] Limited systemic spread of virus, infection confirmed by ELISA.
[d] Infection of inoculated leaves but no systemic spread.

This chapter describes work on the identification of the second resistance gene *ConTR01* in RLR22. It describes work on the transcriptome sequencing of R-o-18 and RLR22 as well as the genotyping and phenotyping of plants from three B$_1$S$_2$ families to identity *ConTR01*.

5.2 Results

5.2.1 Transcriptome Sequencing of *B. rapa* R-o-18 and RLR22

To identify the second resistance gene (*ConTR01*) involved in the resistance in RLR22 (Rusholme *et al.*, 2007), the expression of the *ConTR01* candidates *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* in the *B. rapa* lines R-o-18 and RLR22 were investigated to see if there were any differences between the two lines. Transcriptome sequencing was performed on three individual plants from each line. There was no significant difference in the expression of *BraA.eIF4E.c*, or *BraA.eIF(iso)4E.c* between the lines (Table 5.2). Also, there was no significant difference in the expression of *BraA.eIF4E.a*, or *BraA.eIF(iso)4E.b* between the

two lines (Table 5.2). The only significant difference observed between the two lines was for $BraA.eIF(iso)4E.a$ ($retr01$) ($Q \leq 0.005$; Table 5.2).

5.2.2 Investigation of Segregating Families

As no difference in the expression of the two $ConTR01$ candidates was observed, families segregating for the candidate genes were investigated. Seed of the segregating lines was obtained from colleagues in Canada, where the crossing had been done and the seed had been stored. Unfortunately, the seed had not been stored under optimum conditions and the germination rate of the seed was very poor. Whilst every effort was made to optimise germination, the relatively low frequency of informative genotypes meant that the plants needed to be selfed to bulk up the seed to produce B_1S_2 families (Figure 5.1). Three B_1S_1 families derived from the original cross between RLR22 and R-o-18 were identified, plants grown and selfed to produce B_1S_2 families homozygous for $retr01$ ($BraA.eIF(iso)4E.a$) and segregating for either $BraA.eIF4E.c$, or $BraA.eIF(iso)4E.c$, or both genes. The B_1S_2 families were grown, phenotyped for resistance/susceptibility to TuMV isolate CDN 1 and genotyped at the $BraA.eIF4E.c$ and $BraA.eIF(iso)4E.c$ loci.

B_1S_2 Family Segregating for *BraA.eIF4E.c*

The B_1S_2 family homozygous for $retr01$ and the RLR22 allele of $BraA.eIF(iso)4E.c$, but segregating for $BraA.eIF4E.c$, showed no symptoms when challenged with TuMV isolate CDN 1 (example plant, Figure 5.2). The lack of systemic infection was verified by ELISA and also by back-inoculation to susceptible *Brassica juncea* plants.

Genotyping of a sample of 12 B_1S_2 plants confirmed that all plants tested were homozygous for $retr01$ (RLR22 allele of $BraA.eIF(iso)4E.a$) and the RLR22 allele of $BraA.eIF(iso)4E.c$ (Figure 5.3) and that they segregated for $BraA.eIF4E.c$ (Table 5.3). All three genotypes were detected (Table 5.4), six plants were homozygous for RLR22 $BraA.eIF4E.c$ (plant nos. CN9c.001, CN9c.014, CN9c.021, CN9c.023, CN9c.031 and CN9c.033), two plants were homozygous for R-o-18 $BraA.eIF4E.c$ (plant nos. CN9c.010 and CN9c.029) and four plants were heterozygous for $BraA.eIF4E.c$ (plant nos. CN9c.009, CN9c.019, CN9c.025 and CN9c.038) (Table 5.3).

Table 5.2 – Expression levels of copies of *eIF4E* and *eIF(iso)4E* from *Turnip mosaic virus*-susceptible *Brassica rapa* R-o-18 and TuMV-resistant RLR22 plants measured in fragments per million fragments mapped (FPKM) from transcriptome analysis.

	Gene					
Plant line	*BraA.eIF(iso)4E.a*	*BraA.eIF(iso)4E.b*	*BraA.eIF(iso)4E.c*	*BraA.eIF4E.a*	*BraA.eIF4E.c*
R-o-18	7.7	0.05	10.5	3.7	13.5
RLR22	2.0	0.07	12.6	2.9	11.8
Q value	0.0047*[a]	1	0.7810	0.8040	0.8573

[a] Statistically significant ($Q \leq 0.005$).

BraA.eIF4E.b is a pseudogene lacking exons 2 and 3 (Jenner *et al.*, 2010).
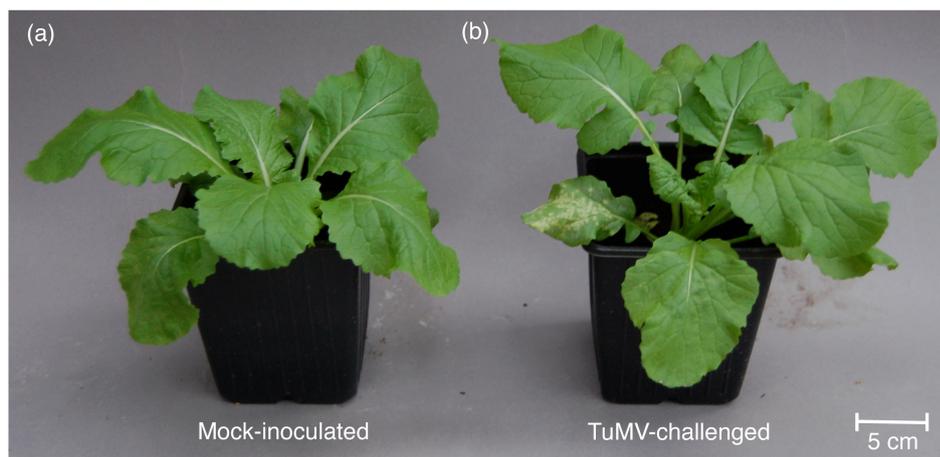
Figure 5.2 – Phenotype of a B_1S_2 plant from the family segregating for *BraA.eIF4E.c*, following mechanical inoculation with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
(a) Mock-inoculated RLR22 control plant. (b) B_1S_2 plant homozygous for *retr01*, the RLR22 allele of *BraA.eIF(iso)4E.c* and the R-o-18 allele of *BraA.eIF4E.c*, challenged with TuMV isolate CDN 1 displaying chlorotic spots on inoculated leaves but no systemic symptoms.
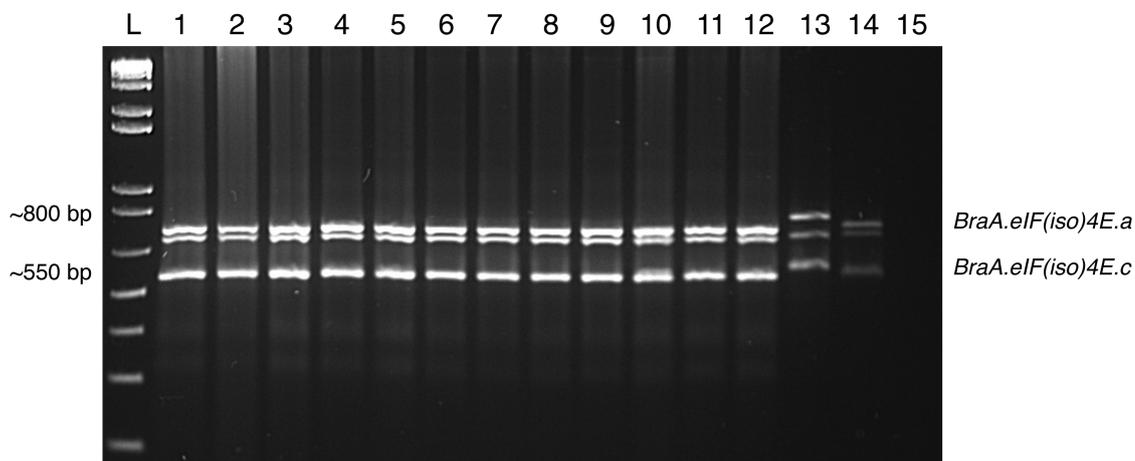
Figure 5.3 – PCR amplification of *eIF(iso)4E* from a B_1S_2 family segregating for *BraA.eIF4E.c* using primers BR14 and BR2. R-o-18 sizes: *BraA.eIF(iso)4E.a* is ~799 bp, *BraA.eIF(iso)4E.b* is ~697 bp and *BraA.eIF(iso)4E.c* is ~566 bp. RLR22 sizes: *BraA.eIF(iso)4E.a* is ~747 bp, *BraA.eIF(iso)4E.b* is ~697 bp and *BraA.eIF(iso)4E.c* is ~546 bp. The R-o-18 allele of *BraA.eIF(iso)4E.b* is identical to RLR22.
L, 1 Kb Plus DNA Ladder; 1, Plant CN9c.001; 2, Plant CN9c.009; 3, Plant CN9c.010; 4, Plant CN9c.014; 5, Plant CN9c.019; 6, Plant CN9c.021; 7, Plant CN9c.023; 8, Plant CN9c.025; 9, Plant CN9c.029; 10, Plant CN9c.031; 11, Plant CN9c.033; 12, Plant CN9c.038; 13, R-o-18 gDNA (positive control); 14, RLR22 gDNA (positive control); 15, dH$_2$O (negative control).

Table 5.3 – *BraA.eIF4E.c* genotypes of plants from the B_1S_2 family segregating for *BraA.eIF4E.c*.

Plant	R-o-18 *BraA.eIF4E.c* (CN55 and CN56)[a]	RLR22 *BraA.eIF4E.c* (CN44 and CN45)[b]	Genotype
CN9c.001	-	+	CC
CN9c.009	+	+	Cc
CN9c.010	+	-	cc
CN9c.014	-	+	CC
CN9c.019	+	+	Cc
CN9c.021	-	+	CC
CN9c.023	-	+	CC
CN9c.025	+	+	Cc
CN9c.029	+	-	cc
CN9c.031	-	+	CC
CN9c.033	-	+	CC
CN9c.038	+	+	Cc

[a] Primers specific to R-o-18 *BraA.eIF4E.c*.
[b] Primers specific to RLR22 *BraA.eIF4E.c*.

Table 5.4 – Phenotypes and genotypes of plants from the B_1S_2 family segregating for *BraA.eIF4E.c* when challenged with *Turnip mosaic virus* isolate CDN 1.

Genotype[b]	Phenotype[a]	
	R[c]	+[d]
rr II CC	6	0
rr II Cc	4	0
rr II cc	2	0

[a] Plants challenged with TuMV isolate CDN 1.
[b] r, *retr01*; I, RLR22 allele of *BraA.eIF(iso)4E.c*; C, RLR22 allele of *BraA.eIF4E.c*; c, R-o-18 allele of *BraA.eIF4E.c*.
[c] Infection of inoculated leaves but no systemic spread, no virus detected by ELISA in uninoculated leaves.
[d] Systemic infection.

The detection of all three *BraA.eIF4E.c* genotypes and the lack of infection, indicated that TuMV can not use BraA.eIF4E.c from either plant line and that *BraA.eIF4E.c* is not *ConTR01*.

B$_1$S$_2$ Family Segregating for *BraA.eIF(iso)4E.c*

The B$_1$S$_2$ family homozygous for *retr01* and the RLR22 allele of *BraA.eIF4E.c*, but segregating for *BraA.eIF(iso)4E.c*, segregated for resistance to TuMV isolate CDN 1. The symptoms observed in the uninoculated leaves of susceptible plants were very mild and limited (Figure 5.4) but were confirmed by positive ELISA values.
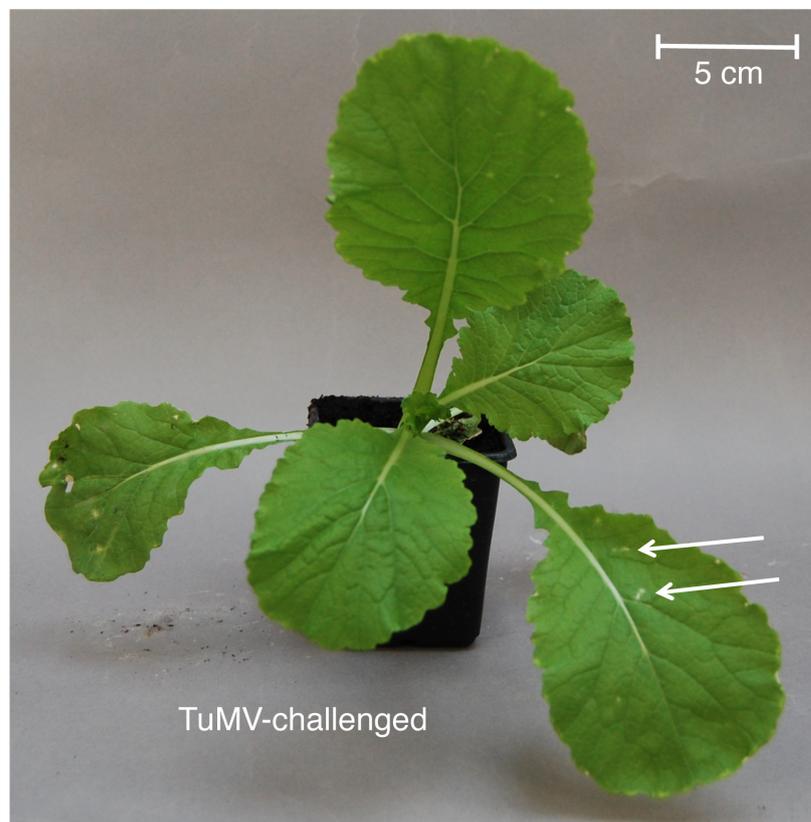
Figure 5.4 – Phenotype of a B$_1$S$_2$ plant from the family segregating for *BraA.eIF(iso)4E.c*, following mechanical challenge with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
B$_1$S$_2$ plant homozygous for *retr01*, the RLR22 allele of *BraA.eIF4E.c* and homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c*, challenged with TuMV isolate CDN 1 displaying limited small chlorotic spots on uninoculated leaves (at arrows).

Genotyping of the plants confirmed that all plants were homozygous for *retr01* (RLR22 allele of *BraA.eIF(iso)4E.a*; Figure 5.5) and the RLR22 allele of *BraA.eIF4E.c* (data not shown). The family segregated for *BraA.eIF(iso)4E.c*, with all three genotypes detected (homozygous for the RLR22 allele of *BraA.eIF(iso)4E.c*, homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c* and heterozygous for *BraA.eIF(iso)4E.c*; Figure 5.5).
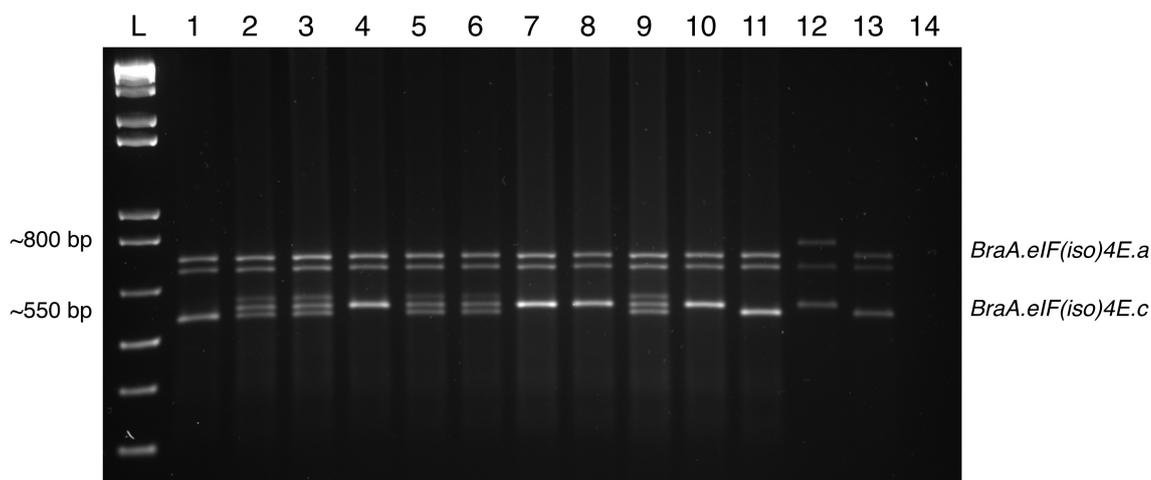
Figure 5.5 – PCR amplification of *eIF(iso)4E* from B_1S_2 plants segregating for *BraA.eIF(iso)4E.c* using primers BR14 and BR2. R-o-18 sizes: *BraA.eIF(iso)4E.a* is ∼799 bp, *BraA.eIF(iso)4E.b* is ∼697 bp and *BraA.eIF(iso)4E.c* is ∼566 bp. RLR22 sizes: *BraA.eIF(iso)4E.a* is ∼747 bp, *BraA.eIF(iso)4E.b* is ∼697 bp and *BraA.eIF(iso)4E.c* is ∼546 bp. The R-o-18 allele of *BraA.eIF(iso)4E.b* is identical to RLR22. Plants heterozygous for *BraA.eIF(iso)4E.c* produce three bands.
L, 1 Kb Plus DNA Ladder; 1, Homozygous for the RLR22 allele of *BraA.eIF(iso)4E.c*; 2, Heterozygous for *BraA.eIF(iso)4E.c*; 3, Heterozygous for *BraA.eIF(iso)4E.c*; 4, Homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c*; 5, Heterozygous for *BraA.eIF(iso)4E.c*; 6, Heterozygous for *BraA.eIF(iso)4E.c*; 7, Homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c*; 8, Homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c*; 9, Heterozygous for *BraA.eIF(iso)4E.c*; 10, Homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c*; 11, Homozygous for the RLR22 allele of *BraA.eIF(iso)4E.c*; 12, R-o-18 gDNA (positive control); 13, RLR22 gDNA (positive control); 14, dH_2O (negative control).

Comparing the genotypic and phenotypic data revealed that all plants homozygous for the RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* were resistant to TuMV isolate CDN 1, (Table 5.5). The plants heterozygous for *BraA.eIF(iso)4E.c* segregated for resistance, with seven plants resistant to TuMV isolate CDN 1 and three plants showing systemic chlorotic spots (Table 5.5). The plants homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c* also segregated for resistance, with four

plants showing no systemic symptoms and two plants showing very mild limited symptoms (Figure 5.4).

Table 5.5 – Phenotypes and genotypes of plants from the B_1S_2 family segregating for *BraA.eIF(iso)4E.c* following challenge with *Turnip mosaic virus* isolate CDN 1.

Genotype[b]	Phenotype[a]	
	R[c]	(+)[d]
rr II CC	8	0
rr Ii CC	7	3
rr ii CC	4	2

[a] Plants challenged with TuMV isolate CDN 1.
[b] r, *retr01*; I, RLR22 allele of *BraA.eIF(iso)4E.c*; i, R-o-18 allele of *BraA.eIF(iso)4E.c*; C, RLR22 allele of *BraA.eIF4E.c*.
[c] Infection of inoculated leaves but no systemic spread, no virus detected by ELISA in uninoculated leaves.
[d] Limited systemic chlorotic spots, infection confirmed by ELISA.

B_1S_2 Family Segregating for Both *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*

The B_1S_2 family homozygous for *retr01*, but segregating for both *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*, segregated for resistance to TuMV isolate CDN 1. Genotyping of the plants confirmed that the plants in this family segregated for both genes (*BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*; gels not shown). Not all potential genotypes were detected (Table 5.6). The systemic symptoms where observed, were mild and limited (Figures 5.6 and 5.7).

Figure 5.6 – Phenotype of a B_1S_2 plant homozygous for *retr01* and heterozygous for *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c*, following mechanical inoculation with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation. Plant displaying limited systemic spread of TuMV, with many small chlorotic spots in three leaves.

All plants homozygous for *retr01* and homozygous for the R-o-18 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* were susceptible to TuMV isolate CDN 1. The vast majority of plants heterozygous at both loci (*BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*) were susceptible to TuMV isolate CDN 1, showing limited systemic spread of TuMV (Table 5.6 and example plant, Figure 5.6). Interestingly in this experiment, the plants homozygous for *retr01* and the RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* segregated for resistance to TuMV isolate CDN 1. The symptoms observed were limited and mild in comparison to symptoms observed in R-o-18 (see Figure 4.4, (b)). Plant CN9d.055 (homozygous for *retr01* and the RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*) displayed the most severe symptoms observed for this genotype, with chlorotic spots present on four uninoculated leaves (Figure 5.7 and Table 5.7).
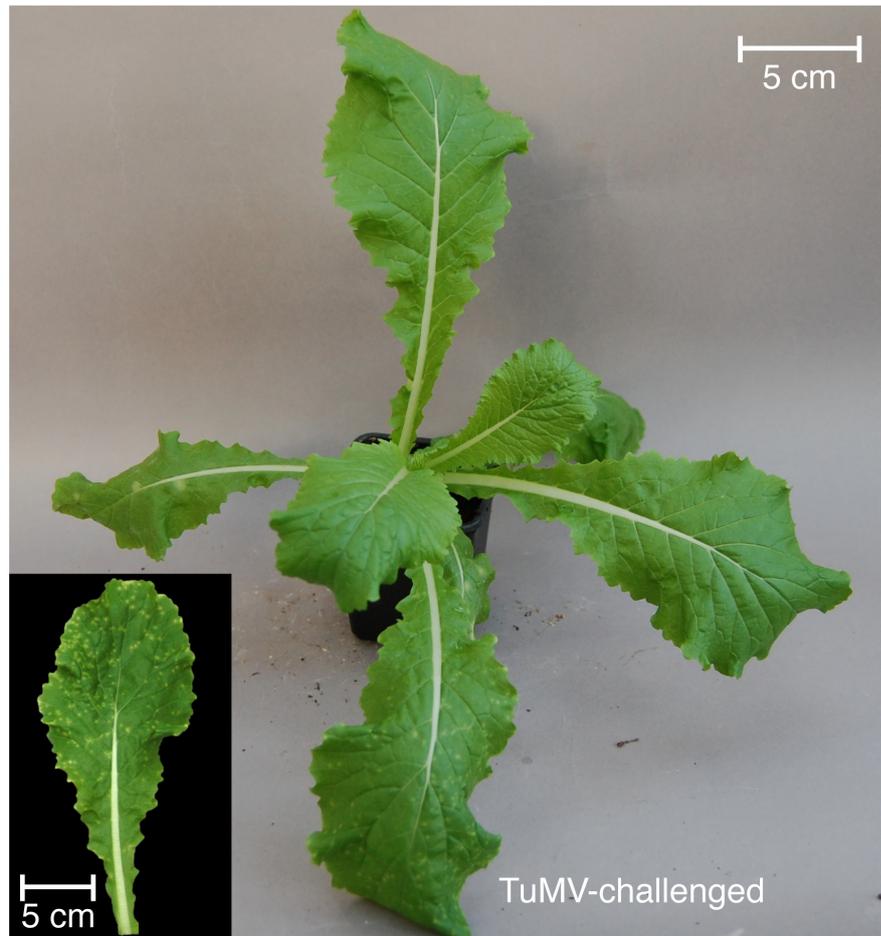
Figure 5.7 – Phenotype of a B_1S_2 plant (CN9d.055) homozygous for *retr01* and the RLR22 alleles of *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c*, following mechanical inoculation with *Turnip mosaic virus* isolate CDN 1; three weeks post-inoculation.
CN9d.055, displaying larger chlorotic spots on inoculated leaves and limited small chlorotic spots on four leaves. Close up of chlorotic spots on an uninoculated leaf (inset).

The symptoms of the eight plants homozygous for all three copies of RLR22 genes, showed varied systemic spread of TuMV isolate CDN 1 (Table 5.7). CN9d.029 showed the mildest symptoms, with only one chlorotic spot present on one leaf and CN9d.055 showed the most severe limited systemic spread with the highest ELISA value (Figure 5.7 and Table 5.7). The lesion from CN9d.029 and lesions from CN9d.055 were back-inoculated onto susceptible *Brassica juncea* plants, which confirmed the TuMV infection (Table 5.7).

Table 5.6 – Phenotypes and genotypes of the B_1S_2 family segregating for both *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* following challenge with *Turnip mosaic virus* isolate CDN 1.

Genotype[b]	Phenotype[a]	
	R[c]	(+)[d]
rr II CC	3	8
rr II Cc	-[e]	-
rr II cc	-	-
rr Ii CC	0	2
rr Ii Cc	1	27
rr Ii cc	-	-
rr ii CC	-	-
rr ii Cc	0	2
rr ii cc	0	18

[a] Plants challenged with TuMV isolate CDN 1.

[b] *r*, *retr01*; *I*, RLR22 allele of *BraA.eIF(iso)4E.c*; *i*, R-o-18 allele of *BraA.eIF(iso)4E.c*; *C*, RLR22 allele of *BraA.eIF4E.c*; *c*, R-o-18 allele of *BraA.eIF4E.c*.

[c] Infection of inoculated leaves but no systemic spread, no virus detected by ELISA in uninoculated leaves.

[d] Limited systemic chlorotic spots, infection confirmed by ELISA.

[e] No plants with this genotype detected.

Table 5.7 – Further investigation of plants homozygous for *retr01* and RLR22 alleles of *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c*, but displaying limited systemic symptoms to *Turnip mosaic virus*.

Plant Number	Phenotype (symptoms in uninoculated leaves)	ELISA Value[a]	No. of *B. juncea* plants infected/no. inoculated[b]
CN9d.018	Chlorotic spots on 4 leaves	1.18	NT[c]
CN9d.019	Chlorotic spots on 3 leaves	0.237	NT
CN9d.029	One chlorotic spot on 1 leaf	0.033	2/3
CN9d.041	Chlorotic spots on 2 leaves	0.163	NT
CN9d.047	Chlorotic spots on 3 leaves	>3.3[d]	NT
CN9d.050	Chlorotic spots on 3 leaves	0.239	NT
CN9d.051	Chlorotic spots on 2 leaves	0.443	NT
CN9d.055	Chlorotic spots on 4 leaves	>3.3	3/3

[a] ELISA absorbance measured at 405 nm.

[b] To check further for TuMV infection, lesions from B_1S_2 plants were ground up and mechanically back-inoculated onto susceptible *Brassica juncea* plants.

[c] Not tested.

[d] Optical density readings greater than 3.3, overflow.

5.2.3 RT-PCR Analysis of the Expression of RLR22 *BraA.eIF(iso)4E.a* in the Susceptible Plants that were Homozygous for *retr01*, *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c*

To explain the limited symptoms observed in the eight plants homozygous for *retr01*, and the *ConTR01* candidates (RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*), RT-PCR analysis was performed on *BraA.eIF(iso)4E.a*. A larger major band approximately 664 nucleotides (nt) in size was detected along with smaller minor bands (Figure 5.8). Sequencing of the larger band confirmed the mis-spliced variant retaining the extra G (indel, at position 201 nt) and the whole of intron 1 (Figure 5.9, (b)). Cloning of the less abundant RT-PCR products from CN9d.055 did not reveal a correctly spliced version of *BraA.eIF(iso)4E.a*, but did reveal two splice variants. A variant retaining the extra G (indel) at the end of exon 1, resulting in a premature stop codon at position 234 nt was detected (Figure 5.9, (c)) and a variant retaining the extra G (indel) and the last 14 nt of intron 1 (or alternatively, if the G comprises part of intron 1, then the variant retained the last 15 nt of intron 1), resulting in a slightly elongated, in-frame mRNA sequence (Figure 5.9, (d)). Both of these variants were detected previously in RLR22, see Section 4.2.3. They were also compared to published *eIF4E* and *eIF(iso)4E* sequences in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008). The variant with an additional five amino acids possessed the correct amino acids at important sites (as discussed in the previous chapter, see Section 4.2.4, Figure 4.10).
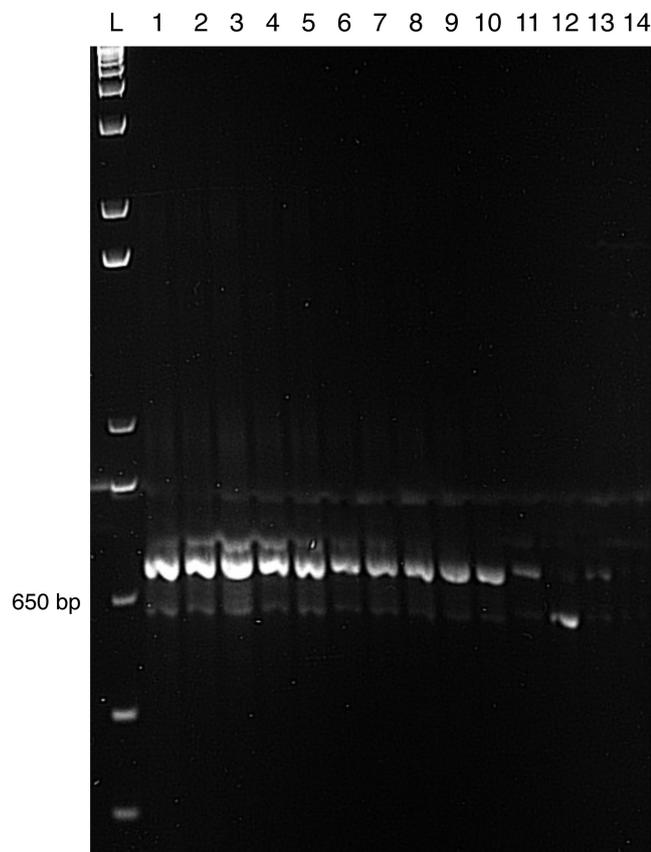
Figure 5.8 – RT-PCR amplification of *BraA.eIF(iso)4E.a* from a B_1S_2 family segregating for both *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* using primers CN3 and CN5. All plants were homozygous for *retr01* and mis-spliced *BraA.eIF(iso)4E.a*.

L, 1 Kb Plus DNA Ladder; 1, Plant CN9d.018; 2, Plant CN9d.019; 3, Plant CN9d.029; 4, Plant CN9d.041; 5, Plant CN9d.047; 6, Plant CN9d.050; 7, Plant CN9d.051; 8, Plant CN9d.055; 9, Plant homozygous for *retr01*, RLR22 allele of *BraA.eIF4E.c* and R-o-18 allele of *BraA.eIF(iso)4E.c*; 10, Plant homozygous for *retr01*, RLR22 allele of *BraA.eIF4E.c* and heterozygous for *BraA.eIF(iso)4E.c*; 11, Plant homozygous for *retr01*, R-o-18 allele of *BraA.eIF4E.c* and RLR22 allele of *BraA.eIF(iso)4E.c*; 12, R-o-18 cDNA (positive control); 13, RLR22 cDNA (positive control); 14, dH_2O (negative control).
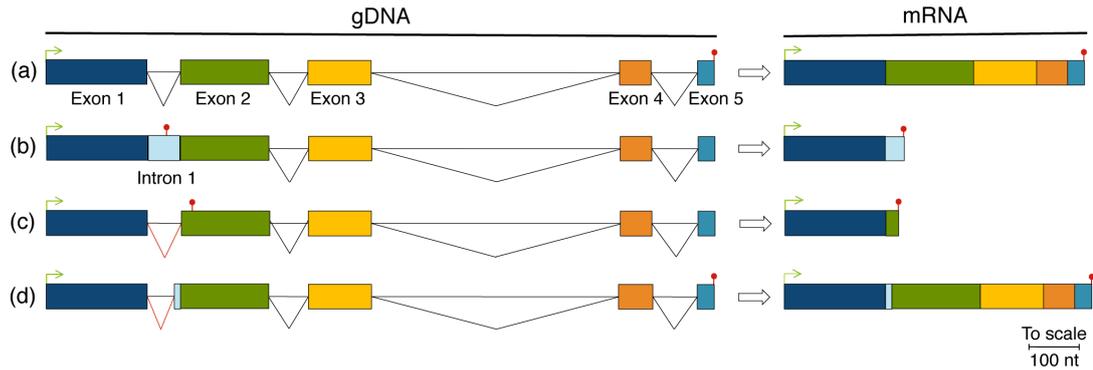
Figure 5.9 – *BraA.eIF(iso)4E.a* splice variants detected in B$_1$S$_2$ plant CN9d.055 by RT-PCR.
(a) For comparison, correctly spliced RLR22 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Most common mis-spliced variant retaining the extra G (indel) and whole of intron 1 resulting in a premature stop codon at position 234 nt (664 nt in length). (c) Variant retaining the extra G (indel) resulting in a premature stop codon at position 228 nt. (d) Variant retaining the extra G (indel) and the last 14 nt of intron 1/retaining the last 15 nt of intron 1, resulting in a slightly elongated, in-frame mRNA (615 nt).

5.3 Discussion

5.3.1 *BraA.eIF(iso)4E.c* is the Only Candidate for *ConTR01*

It is clear from the segregation data that TuMV can not use *BraA.eIF4E.c* from either parent. This rules out *BraA.eIF4E.c* as a candidate for *ConTR01*. It appears that TuMV can use *BraA.eIF(iso)4E.c* from R-o-18, but only to a limited extent as indicated by the limited systemic spread of TuMV in plants possessing *retr01* and the R-o-18 allele of *BraA.eIF(iso)4E.c*. The work discussed in this chapter does not confirm that *BraA.eIF(iso)4E.c* is *ConTR01*, further work is needed if we are to be completely sure of this.

In the two experiments looking at the segregation of the individual genes (*BraA.eIF4E.c*, section 5.2.2 and *BraA.eIF(iso)4E.c*, section 5.2.2), all the plants homozygous for *retr01* and the RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* were resistant to TuMV, with plants showing no systemic symptoms. The results of these two families seem to correlate with each other and the hypothesis that TuMV can not use BraA.eIF4E.c from either parent

and can only use BraA.eIF(iso)4E.c to a limited degree. These results also support previous research by C. E. Jenner and J. Bambridge. In the experiment investigating the B_1S_2 family segregating for both genes (Section 5.2.2), the genotype, homozygous for *retr01* and the RLR22 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c*, segregated for resistance and susceptibility. This could be the result of differing viral loads as the experiments were carried out separately. The experiment on the B_1S_2 family segregating for *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* may have used a higher viral load resulting in the limited systemic spread. Not all potential genotypes were detected in this family, which indicated that that the genes might be tightly linked.

It is evident from the genotyping and phenotyping of the B_1S_2 families, that *retr01* is the main gene controlling the resistance in plants derived from the cross between RLR22 and R-o-18. Plants that were homozygous for *retr01* but homozygous for the R-o-18 alleles of *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* only displayed limited symptoms and nothing as severe as observed in R-o-18.

The results show that *BraA.eIF(iso)4E.c* is the only remaining candidate for *ConTR01*. It appears that TuMV isolate CDN 1 is limited in its ability to use the R-o-18 BraA.eIF(iso)4E.c protein as the plants homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c* segregated for resistance. Alternatively, it could be that the viral RNA is less efficient at interacting with this protein than host mRNA.

5.3.2 Explaining Susceptibility of Some Plants Homozygous for *retr01* and *ConTR01* Candidates from the B_1S_2 Family Segregating for *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c*

The detection of the potentially functional *BraA.eIF(iso)4E.a* splice variant (possessing an additional five amino acids, relative to the correctly spliced version) might explain why the B_1S_2 plants homozygous for *retr01* and the RLR22 allele of *BraA.eIF(iso)4E.c* were susceptible to TuMV isolate CDN 1. This variant possesses all of the correct amino acids at important sites (Figure 4.10) that have been identified from comparing it with published *eIF4E* and *eIF(iso)4E* sequences in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008) (Section 4.2.4). This variant (possessing an additional five amino

acids) was sent to colleagues in China, who had previously shown interactions between the viral protein genome-linked (VPg) of TuMV and *BraA.eIF(iso)4E.a* from TuMV-susceptible *B. rapa* using the yeast two-hybrid (Y2H) assay. They found that this variant interacted with the VPg of TuMV isolate C4 in the Y2H assay (Nellist *et al.*, 2014). This strengthens my hypothesis that this variant is functional for TuMV in the B_1S_2 plants and is leading to the limited systemic symptoms observed in some plants. As we only saw susceptibility in plants with RLR22 copies of *retr01* and *ConTR01* candidates in this family, there may be some genetic factor in this family affecting the splicing of *retr01*, which could explain the potential functional variants.

Only a proportion of plants possessing both *retr01* and the RLR22 allele of *BraA.eIF(iso)4E.c* displayed limited systemic spread when mechanically inoculated with TuMV. The identification of a potentially functional mis-spliced *BraA.eIF(iso)4E.a* variant could explain the mild susceptibility, but mechanical inoculation is an artefact and is not a true representation of what happens in the field. Further work could also include investigating the B_1S_2 families to see if using aphids to challenge the plants, still resulted in limited systemic spread in plants homozygous for both *retr01* and the RLR22 allele of *BraA.eIF(iso)4E.c*.

5.3.3 Explaining the Susceptibility of Heterozygotes in the B_1S_2 Family Segregating for *BraA.eIF(iso)4E.c*, or *BraA.eIF4E.c*

The family segregating for both *BraA.eIF(iso)4E.c* and *BraA.eIF4E.c* genes seemed to contradict the results in the other two families, with the majority of plants that were heterozygous for both genes showing limited systemic spread. If *ConTR01* was dominant, it would be expected that these plants would be mostly resistant. However, as TuMV appears to only be able to use *BraA.eIF(iso)4E.c* to a limited degree, this may explain why it does not behave completely dominantly.

5.3.4 Discrepancies Between Transcriptome Data and Ectopic Expression of Genes in *Arabidopsis thaliana* Col-0::*dSpm eIF(iso)4E*

The transcriptome analysis of the *eIF4E* and *eIF(iso)4E* genes in RLR22 and R-o-18 showed that both *eIF4E* genes and two *eIF(iso)4E* genes (*BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c*) were expressed in R-o-18 and in RLR22. However, the expression of *BraA.eIF(iso)4E.a* was significantly lower in RLR22 than R-o-18, perhaps consistent with its lack of function resulting from mis-splicing. Complementation of an *Arabidopsis thaliana* knock-out with *B. rapa* R-o-18 and RLR22 alleles of *eIF4E* and *eIF(iso)4E* showed that TuMV could use both copies of *eIF4E* (*BraA.eIF4E.a* and *BraA.eIF4E.c*) and both copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c*) from both plant lines, when ectopically expressed in Arabidopsis (Jenner *et al.*, 2010; Nellist *et al.*, 2014). This was misleading as it indicated that TuMV could use both copies of eIF4E in R-o-18 and RLR22. This is clearly not the case in *B. rapa* as earlier segregation data (Rusholme *et al.*, 2007) and my results (Nellist *et al.*, 2014) showed no linkage between these genes and susceptibility.

5.4 Conclusions

The main aim of this chapter was to identify the second resistance gene *ConTR01*. Although apparent contradictions in the data prevent unambiguous confirmation of *ConTR01*, it was possible to rule out one of the candidates, leaving only *BraA.eIF(iso)4E.c*. Further work is needed to confirm this as *ConTR01*.

Chapter 6

Introgression of the *retr01*-based Resistance into Chinese Cabbage Varieties

6.1 Background

6.1.1 The Market

Chinese cabbage (*Brassica rapa* var. *pekinensis*) is an important crop, grown worldwide. It is a fundamental part of many Asian dishes. *Turnip mosaic virus* (TuMV) is considered the most important disease affecting crucifers, including Chinese cabbage (Sako, 1981). As discussed in Section 1.3, TuMV infection can result in severe symptoms in crops, which may lead to severe yield losses. Breeding for resistance is considered the most efficient and reliable control measure against TuMV. Currently there are no uniformly broad-spectrum TuMV-resistant Chinese cabbage varieties available on the market.

6.1.2 Methods of Introgressing Resistance to Viruses

One of the most efficient and simplest approaches for virus control remains conventional breeding of natural sources of resistance (Lecoq *et al.*, 2004). Both dominant and recessive resistance genes have been identified and widely used to confer resistance to viruses. Marker-assisted selection can greatly increase the speed of introgression of natural resistances.

The other major strategy for crop protection against viruses is engineered virus resistance. The first demonstration that the expression of a viral coat protein gene in transgenic plants could confer resistance to that virus was provided by Powell-Abel *et al.* (1986), in the demonstration of resistance to *Tobacco mosaic virus* (TMV) in transgenic tobacco plants. Since then, this pathogen-derived resistance approach has resulted in the creation of numerous virus-resistant transgenic plants (Ritzenthaler, 2005). Engineered resistance has also broken down the species barrier, enabling resistance genes from unrelated species to be expressed in transgenic plants, for example the expression of lectin from snowdrop (*Galanthus nivalis*) in transgenic wheat to protect plants against cereal aphids (Stoger *et al.*, 1999).

Commercial F_1 hybrids are produced by crossing two inbred lines to produce uniform progeny with the desired attributes from both parents and added vigour. As the broad-spectrum resistance to TuMV is controlled by a recessive gene, it is necessary to introgress the resistance into both parents of all hybrids. This chapter describes work on the introgression of the *retr01*-based resistance into commercial Chinese cabbage varieties by conventional breeding and the testing of F_2 populations to evaluate the transfer of resistance to progeny.

6.2   Results

6.2.1   Susceptibility of Parental Lines

Susceptibility to TuMV Isolates CDN 1 and UK 1

Prior to introgressing the broad-spectrum resistance to TuMV into commercial Chinese cabbage varieties, it was important to know the resistance status of the recipient plant lines. Six inbred Syngenta parental lines of F_1 hybrid varieties (CK 1, CK 2, CK 3, CK 4, CK 5 and CK 6; *B. rapa* var. *pekinensis*) were mechanically inoculated with TuMV isolates CDN 1 (pathotype 4) and UK 1 (pathotype 1).

All lines showed uniform symptoms when challenged with the virus isolates. All six lines were susceptible to mechanical challenge of TuMV isolate CDN 1, resulting in systemic mosaic symptoms. Infection was confirmed by
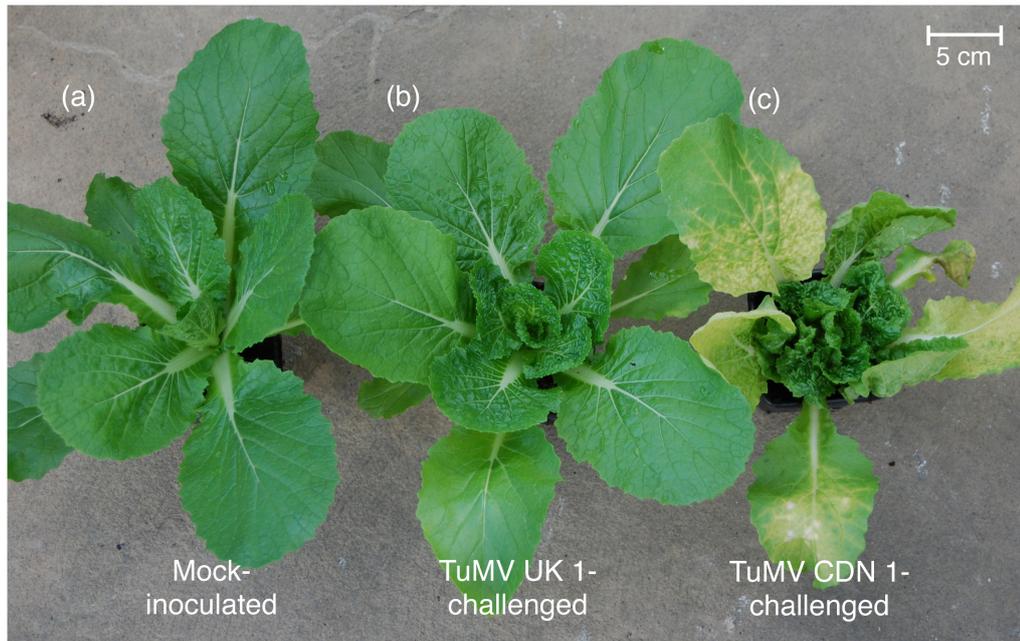
Figure 6.1 – *Brassica rapa* line CK 1 plants following mechanical inoculation with *Turnip mosaic virus* isolates UK 1 and CDN 1; three weeks post-inoculation.
(a) Mock-inoculated CK 1 control plant. (b) CK 1 plant challenged with TuMV isolate UK 1 displaying no symptoms. (c) CK 1 plant challenged with TuMV isolate CDN 1 displaying systemic mosaic symptoms.
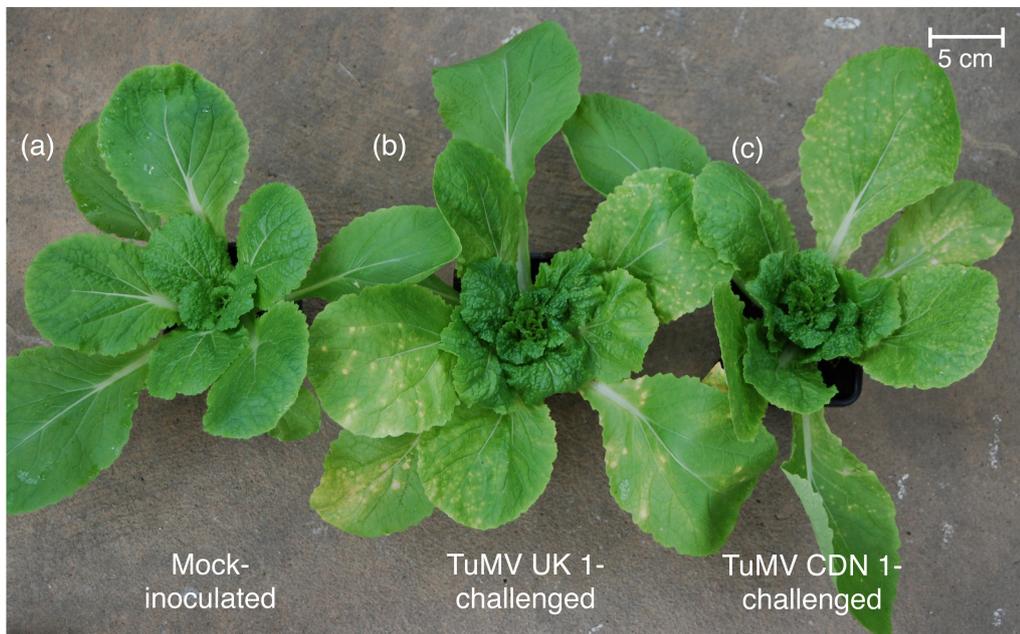
Figure 6.2 – *Brassica rapa* line CK 2 plants following mechanical inoculation with *Turnip mosaic virus* isolates UK 1 and CDN 1; three weeks post-inoculation.
(a), Mock-inoculated CK 2 control plant. (b), CK 2 plant challenged with TuMV isolate UK 1 displaying systemic mosaic symptoms. (c), CK 2 plant challenged with TuMV isolate CDN 1 displaying systemic mosaic symptoms.

91

ELISA (CK 1, Figure 6.1; CK 2, Figure 6.2 and Table 6.1). The lines CK 2 (Figure 6.2), CK 3 and CK 5 were also susceptible to mechanical challenge of TuMV isolate UK 1, infection confirmed by ELISA (Table 6.1). Lines CK 1 (Figure 6.1), CK 4 and CK 6 were resistant to mechanical challenge of TuMV isolate UK 1, with no symptoms of TuMV infection observed and no virus detected by ELISA (Table 6.1).

Testing for the Presence of the *TuRB01* Resistance Gene

Subsequently, three of the lines (CK 1, CK 4 and CK 6), which showed resistance to TuMV isolate UK 1, were mechanically challenged with a TuMV UK 1 mutant isolate derived from the vVIR24 construct (Jenner *et al.*, 2000), to test whether the dominant resistance gene *TuRB01* (Walsh *et al.*, 1999) was present. The presence of *TuRB01* could possibly interfere with the assessment of susceptibility in subsequent lines. The plant line R4 (*Brassica napus* var. *oleifera*), possesses *TuRB01* (Walsh *et al.*, 1999) and was used as a control, as the TuMV UK 1 mutant isolate vVIR24 overcomes the *TuRB01* resistance gene.
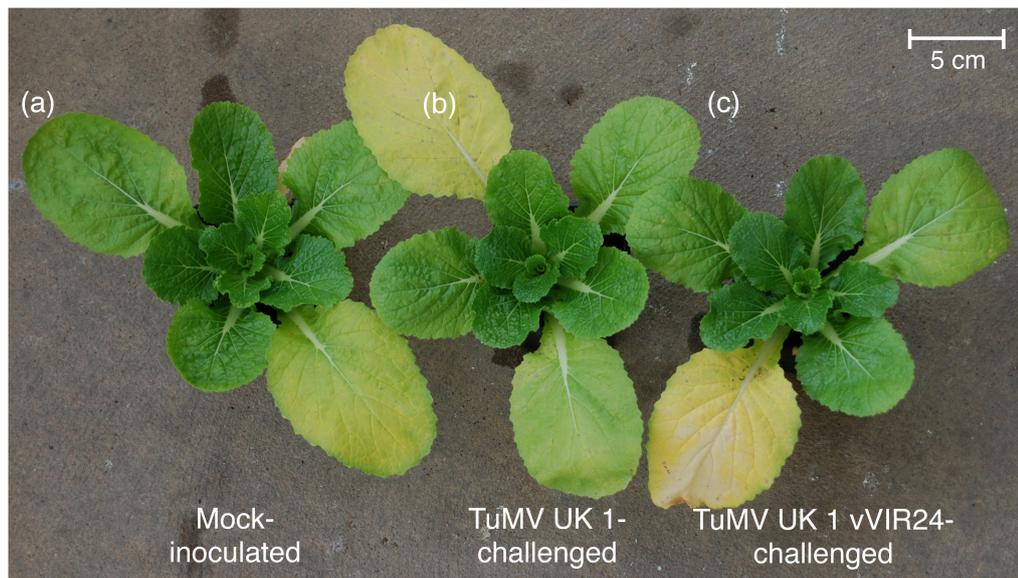
Figure 6.3 – *Brassica rapa* line CK 1 plants following mechanical inoculation with *Turnip mosaic virus* isolates UK 1 and UK 1 mutant derived from the vVIR24 construct; three weeks post-inoculation.
(a) Mock-inoculated CK 1 control plant. (b), CK 1 plant challenged with TuMV isolate UK 1 displaying no infection. (c), CK 1 plant challenged with TuMV UK 1 mutant isolate derived from the vVIR24 construct displaying no systemic symptoms.

All three lines, CK 1 (Figure 6.3), CK 4 and CK 6, were resistant to the TuMV UK 1 mutant isolate derived from the vVIR24 construct (Table 6.1), indicating that the resistance gene present in these lines is not *TuRB01*.

Table 6.1 – Phenotypes of Syngenta *Brassica rapa* parental lines and *Brassica* control plants mechanically challenged with *Turnip mosaic virus* isolates CDN 1, UK 1 and UK 1 mutant derived from the vVIR24 construct.

Plant Line	TuMV Inoculum		
	CDN 1	UK 1	UK 1 vVIR24[a]
CK 1	+[b]	0[c]	0
CK 2	+	+	-[d]
CK 3	+	+	-
CK 4	+	0	0
CK 5	+	+	-
CK 6	+	0	0
R-o-18	+	+	-
RLR22	R[e]	R	-
R4[f]	-	0	(+)[g]
TGM[h]	-	+	+

[a] TuMV UK 1 mutant capable of overcoming and infecting plants possessing the *TuRB01* resistance gene, no ELISA was carried out on these plants.

[b] Systemic mosaic symptoms, infection confirmed by ELISA.

[c] No infection, no virus detected by ELISA in uninoculated leaves or by grinding up uninoculated leaves and back-inoculating to susceptible *Brassica juncea*.

[d] Not tested.

[e] Infection of inoculated leaves but no systemic spread, no virus detected by ELISA in uninoculated leaves.

[f] Experimental control plant line possessing the resistance gene *TuRB01*, to check the TuMV vVIR24 mutant isolate.

[g] Limited systemic spread of TuMV.

[h] TuMV-susceptible *B. juncea* line.

Sequencing of *BraA.eIF(iso)4E.a* from CK Parental Lines

Sequencing of the genomic DNA (gDNA) of *BraA.eIF(iso)4E.a* (eukaryotic translation initiation factor isoform 4E) from five of the parental CK lines revealed that none of the lines possessed the extra G (indel, at position 201 nucleotides (nt)). *BraA.eIF(iso)4E.a* from CK 2 and CK 4 were identical and differed from RLR22 by lacking the indel and 1 nt in intron 2. *BraA.eIF(iso)4E.a* from CK 1 and CK 6 were identical and differed from RLR22 by lacking the indel and 1 nt in intron 2, as well as having 1 nt substitution in exon 3. *BraA.eIF(iso)4E.a* from CK 5 was different from all others and was different from RLR22 by lacking the indel and 1 nt in intron 2 as well as having a different 1 nt substitution in exon 3.

Sequencing revealed the amino acid sequence of BraA.eIF(iso)4E.a from CK 2 and CK 4 was identical to correctly spliced RLR22. It also revealed that BraA.eIF(iso)4E.a from CK 1 and CK 6 had a single non-synonymous mutation relative to correctly spliced RLR22 and that BraA.eIF(iso)4E.a from CK 5 had a different non-synoymous mutation relative to correctly spliced RLR22 allele.

The discovery that the Syngenta commercial lines do not possess the indel in *BraA.eIF(iso)4E.a*, facilitated the design of a single nucleotide polymorphism (SNP) marker to develop a simple assay to test for the presence of *retr01* in the crosses. Due to the SNP's codominant nature, the plants were also tested for homozygosity.

6.2.2   Testing for TuMV Resistance in F_2 Generations

Six F_2 populations (CH0001, CH0011, CH0017, CH0021, CH0030 and CH0033) were produced from crosses between a self of RLR22 (85-40) and the six Syngenta parental lines (CK 1, CK 2, CK 3, CK 4, CK 5 and CK 6, respectively) (Figure 2.1) by Syngenta. A SNP marker in *BraA.eIF(iso)4E.a* was developed by Syngenta based on the indel, to determine whether the plants were homozygous for the presence of *retr01* (possessing the extra G (indel) at position 201 nt), heterozygous for *retr01*, or homozygous for the absence of *retr01* (presence of the CK allele of *BraA.eIF(iso)4E.a*; Table 6.2). Fluorophores were conjugated to the probes targeting the resistant and susceptible alleles for *retr01*. The resistant probe was conjoined to the fluorophore FAM and the susceptible probe was conjoined to the fluorophore VIC. Plants homozygous for *retr01* gave a high value at the FAM wavelength and a low value at the VIC wavelength (Figure 6.4). Plants homozygous for the absence of *retr01* gave a high value at the VIC wavelength and a low value at the FAM wavelength (Figure 6.4). Plants heterozygous for *retr01* gave an intermediate emission at both wavelengths (Figure 6.4). Plants that did not cluster for any of the genotypes, or did not give a detectable result were re-tested, however some remained unscored (Figure 6.4 and Table 6.2).
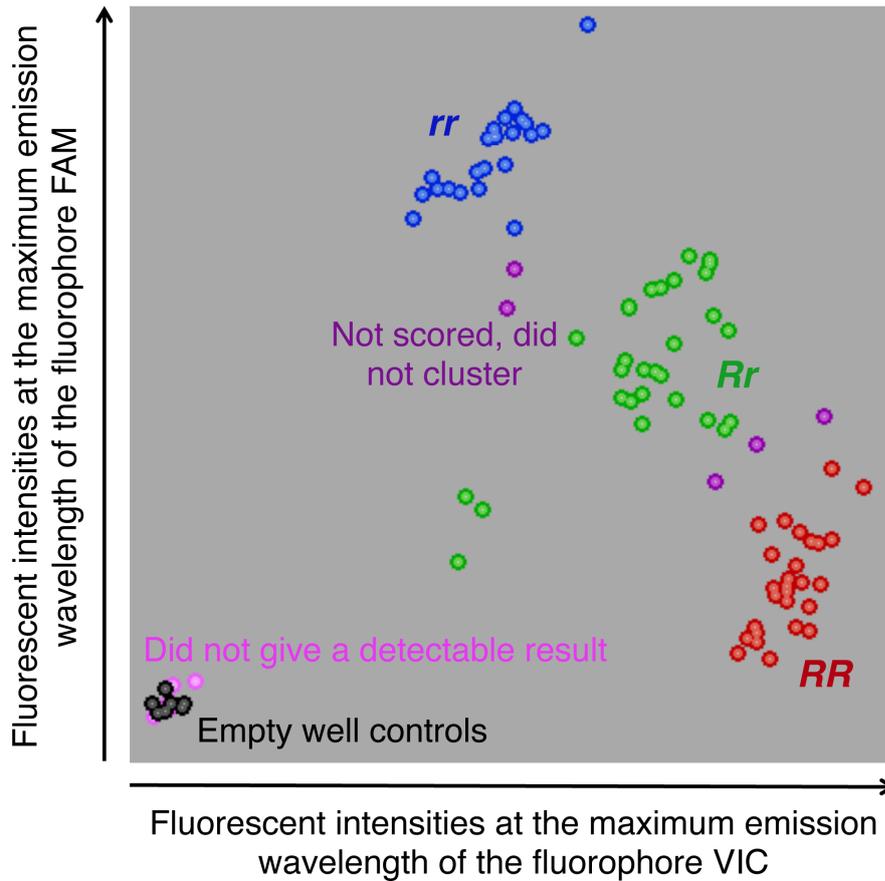
Figure 6.4 – *retr01* allelic discrimination plot for *Brassica rapa* CH0001 plants using the SNP marker (performed by Syngenta).
Blue dots - homozygous for *retr01* (*rr*); Green dots - heterozygous for *retr01* (*Rr*); Red dots - homozygous for the absence of *retr01* (*RR*); Purple dots - not scored because they did not cluster; Pink dots - did not give a detectable result and black dots were empty well controls.

Table 6.2 – Genotypes of the six *Brassica rapa* F$_2$ populations using the SNP marker developed by Syngenta.

| Plant Line | Genotype | | | | Total No. |
	RR^a	Rr^b	rr^c	-^d	of plants
CH0001	25	47	34	2	108
CH0011	33	55	18	2	108
CH0017	36	45	19	6	106
CH0021	27	33	31	13	104
CH0030	28	53	27	0	108
CH0033	14	18	20	0	52

[a] Plants homozygous for the CK allele of *BraA.eIF(iso)4E.a*.
[b] Plants heterozygous for *retr01*.
[c] Plants homozygous for *retr01*.
[d] Unscored plants.

Plants from the six F_2 populations were mechanically inoculated with TuMV isolate CDN 1 and the plants for all six lines segregated for resistance to TuMV. Research focused on further analysis of three of the populations, CH0001, CH0011 and CH0030. Plants homozygous for *retr01* in the other three lines (CH0017, CH0021 and CH0033) were also investigated.

Plants homozygous for the CK allele of *BraA.eIF(iso)4E.a* always displayed systemic mosaic symptoms when mechanically challenged with TuMV isolate CDN 1; infection was confirmed by ELISA (Figure 6.5, (b) and Table 6.3). Plants heterozygous for *BraA.eIF(iso)4E.a* segregated for resistance to TuMV isolate CDN 1, with the majority of plants showing limited systemic spread of TuMV (infection confirmed by ELISA; Figure 6.5, (a)) and some showing no systemic spread (no virus detected by ELISA) (Table 6.3). Plants homozygous for *retr01*, mostly showed resistance to TuMV isolate CDN 1, no virus detected by ELISA (Figure 6.5, (c) and Table 6.3). However, at least one plant, homozygous for *retr01*, in each line, except CH0033 showed limited systemic spread of TuMV; infection was confirmed by ELISA (Figure 6.6 and Table 6.3). The symptoms in these plants were extremely limited and only present in a maximum of three uninoculated leaves in the worst affected plants (Figure 6.6).
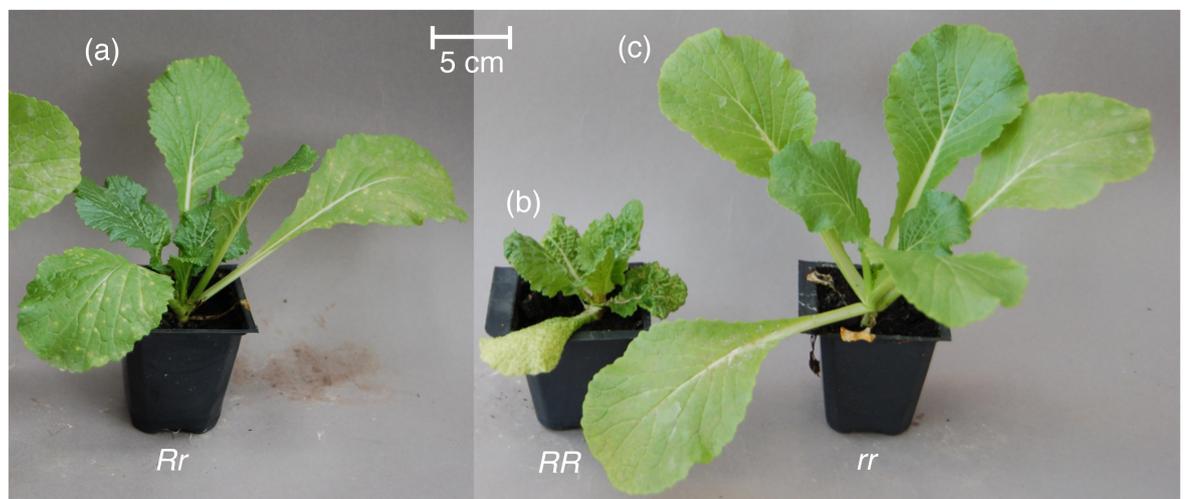
Figure 6.5 – Susceptibility of *Brassica rapa* CH0001 F_2 plants to *Turnip mosaic virus* isolate CDN 1, following mechanical inoculation; three weeks post-inoculation.
(a) Plant heterozygous for *retr01* (*BraA.eIF(iso)4E.a*) with limited systemic infection. (b) Plant homozygous for the CK 1 allele of *BraA.eIF(iso)4E.a* with severe systemic infection. (c) Plant homozygous for *retr01*, resistant to TuMV.

Table 6.3 – Genotypes and phenotypes of *Brassica rapa* F$_2$ populations CH0001, CH0011 and CH0030, following mechanical inoculation with *Turnip mosaic virus* isolate CDN 1.

F$_2$ family	Genotype	Phenotype		
		R[a] No symptoms and no systemic spread	(+)[b] Mild symptoms and limited systemic spread	+[b] Severe symptoms and full systemic infection
CH0001	*RR*[c]	0	0	25
	Rr[d]	1	46	0
	rr[e]	28	6	0
CH0011	*RR*	0	0	33
	Rr	10	45	0
	rr	17	1	0
CH0030	*RR*	0	0	28
	Rr	11	42	0
	rr	26	1	0
CH0017	*rr*	18	1	0
CH0021	*rr*	30	1	0
CH0033	*rr*	20	0	0

[a] No virus detected in uninoculated leaves by ELISA.
[b] Infection confirmed by ELISA.
[c] *RR*, plants homozygous for the respective CK allele of *BraA.eIF(iso)4E.a.*
[d] *Rr*, plants heterozygous for *retr01*.
[e] *rr*, plants homozygous for *retr01*.
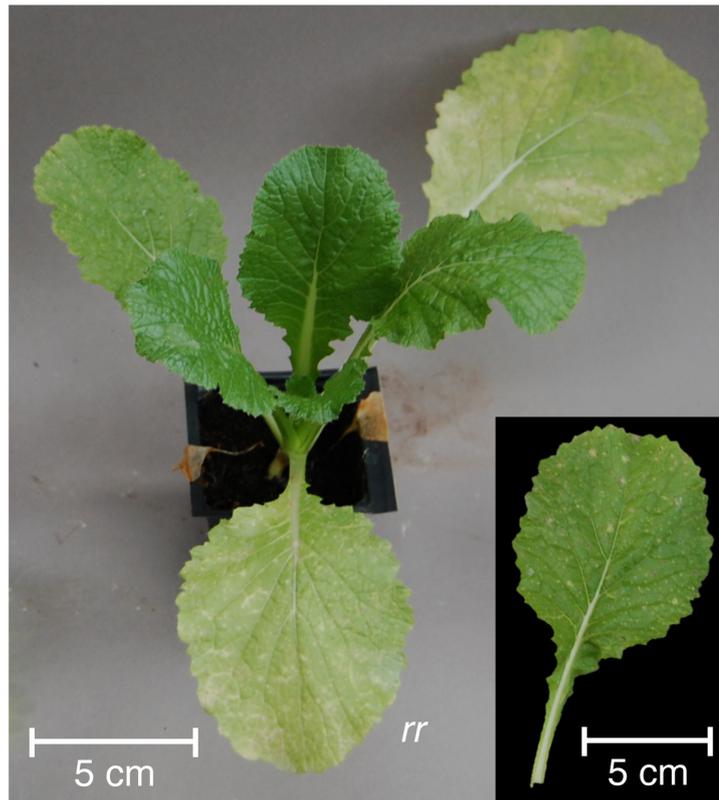
Figure 6.6 – Phenotype of a *Turnip mosaic virus*-susceptible *Brassica rapa* F_2 CH0001 plant homozygous for *retr01*, following mechanical inoculation with TuMV isolate CDN 1; three weeks post-inoculation.
Plant homozygous for *retr01* with limited systemic spread of TuMV in three uninoculated leaves. Close up of chlorotic spots on an uninoculated leaf (inset).

χ^2 Tests for Segregation of Resistance

Segregation ratios of phenotypes in the F_2 families CH0001 and CH0011 (Table 6.3 and Section 2.7), were consistent with those predicted for a single recessive gene (3:1, Susceptible:Resistant) ($\chi^2 = 0.62$ and 0.11 respectively, all df=1, $P>0.05$). Segregation ratios of phenotypes in the F_2 families CH0001, CH0011 and CH0030 (Table 6.3 and Section 2.7), were consistent with 1:2:1 (Fully susceptible:Limited infection:Resistant) ($\chi^2 = 0.58$, 1.87 and 5.98 respectively, all df=2, $P>0.05$).

T-tests for Accumulation of TuMV

Interestingly, from the F_2 populations (CH0001, CH0011 and CH0030), plants heterozygous at the *retr01* locus were less susceptible to TuMV than those that were homozygous for the allele from the susceptible CK parent (e.g. Figure 6.5). ELISA

confirmed the heterozygotes (mean A_{405} from ELISA for CH0001 heterozygotes 0.81 ($\pm$ 0.074), CH0011 heterozygotes 0.28 ($\pm$ 0.040) and CH0030 heterozygotes 0.24 ($\pm$ 0.028)) accumulated significantly less virus than plants homozygous for the CK alleles of *BraA.eIF(iso)4E.a* (mean A_{405} for CH0001 homozygotes 2.09 ($\pm$ 0.132), CH0011 homozygotes 1.47 ($\pm$ 0.159) and CH0030 homozygotes 1.34 ($\pm$ 0.098)) (all $P<0.001$).

6.2.3 Sequencing of *BraA.eIF(iso)4E.c* from CK Parental Lines

Since the resistance segregated as a single gene in the RLR22 and CK crosses, *BraA.eIF(iso)4E.c* was sequenced from the gDNA of the six parental Syngenta lines (CK 1 - CK 6). Sequencing revealed that they were identical to RLR22 *BraA.eIF(iso)4E.c*, showing they already possessed the *ConTR01* candidate, explaining why the resistance was inherited as a single recessive gene.

6.2.4 Was Susceptibility to TuMV in Plants Homozygous for *retr01* Due to a Mutation in the Virus?

It has been shown that mutations in the viral protein genome-linked (VPg) of TuMV isolate CDN 1 resulted in infection of Arabidopsis Col-0::*dSpm* (Gallois *et al.*, 2010). To test whether a mutation in the VPg of TuMV caused the limited infection in plants homozygous for *retr01*, the VPg from TuMV isolate CDN 1 was sequenced from a selection of F_2 plants, including plants that were homozygous for *retr01* and showed limited systemic TuMV infection. Only a single synonymous polymorphism at position 216 nt of the VPg was detected, relative to the published CDN 1 sequence (GenBank: AB093610; Tomimura *et al.*, 2003). This showed that mutations in the VPg were not responsible for the infections observed in the plants homozygous for *retr01*. Sequencing of the current TuMV isolate CDN 1, maintained in *Brassica juncea* was also checked and no mutations were detected.

6.2.5 RT-PCR Analysis of the Expression of *BraA.eIF(iso)4E.a* in the F_2 Populations

Since no mutations could be detected in the TuMV VPg, RT-PCR was performed on a selection of plants to investigate the expression of *BraA.eIF(iso)4E.a* to

see if there were any potentially functional variants present that may explain the limited infection in some plants homozygous for *retr01*. RT-PCR analysis of plants homozygous for the CK alleles of *BraA.eIF(iso)4E.a* detected a single band approximately 600 nt in size (Figure 6.7, lanes 1 and 2). Sequencing of this band revealed the correctly spliced version of the respective CK allele of *BraA.eIF(iso)4E.a*. RT-PCR analysis of plants heterozygous for *retr01* detected two equal strength bands, one approximately 664 nt in size and the other approximately 600 nt in size, in plants that showed susceptibility and resistance to TuMV (Figure 6.7, lanes 3-6). RT-PCR analysis of plants homozygous for *retr01* detected two bands, a larger major band approximately 664 nt in size along with a smaller minor band (Figure 6.7, lanes 7-15). Sequencing of the larger band confirmed a mis-spliced variant retaining the extra G (indel, at position 201 nt) and the whole of intron 1 (Figure 6.8, (b)). To investigate why a proportion of plants homozygous for *retr01* showed limited systemic spread of TuMV, the minor product of one plant from the CH0001 population (Figure 6.7, lane 7) and one from CH0011 population (Figure 6.7, lane 10) were sequenced, to see if there were any potentially functional variants of *BraA.eIF(iso)4E.a*. Cloning of the less abundant RT-PCR products did not reveal a correctly spliced version of *BraA.eIF(iso)4E.a*, but did reveal two splice variants. A variant retaining the extra G (indel) at the end of exon 1, resulting in a premature stop codon at position 234 nt was detected (Figure 6.8, (c)) and a variant lacking the last 3 nt of exon 1, resulting in a shortened in-frame mRNA sequence with a substitution (Figure 6.8, (d)). Both of these variants were detected previously in RLR22, see Section 4.2.3.

Genotype		RR		Rr				rr												
Phenotype		+	+	(+)	(+)	R	R	(+)	(+)	(+)	(+)	(+)	(+)	R	R	R				
	L	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

650 bp
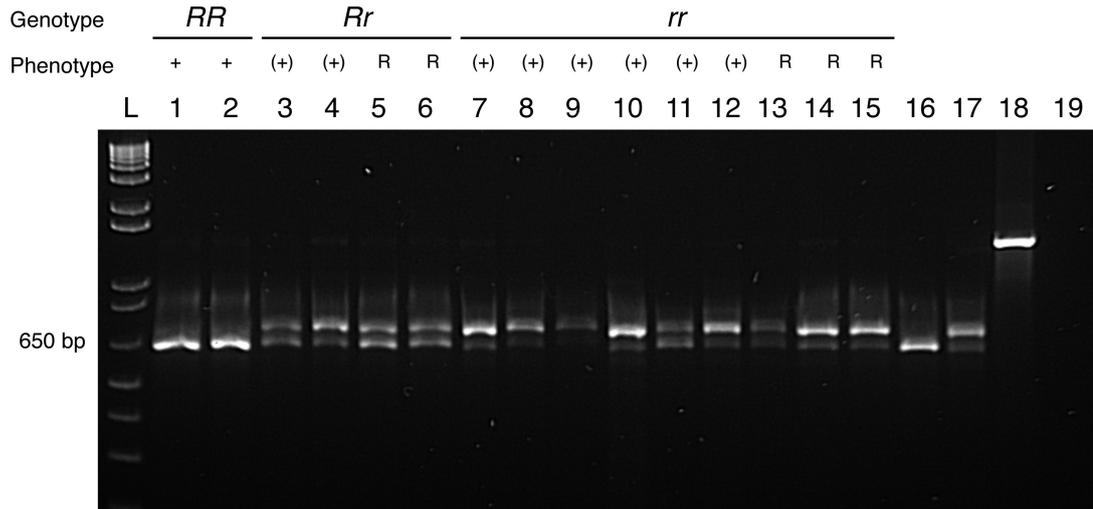
Figure 6.7 – Mis-splicing of *BraA.eIF(iso)4E.a* from a selection of F$_2$ plants from the CK and RLR22 crosses. *RR*, Plants homozygous for the CK allele of *BraA.eIF(iso)4E.a*; *Rr*, plants heterozygous for *retr01*; *rr*, plants homozygous for *retr01*; +, systemic mosaic symptoms; (+), limited systemic spread; R, no systemic spread. RT-PCR using primers CN3 and CN5.
L, 1 Kb Plus DNA Ladder; 1, CH0001 plant; 2, CH0011 plant; 3, CH0001 plant; 4, CH0030 plant; 5, CH0011 plant; 6, CH0011 plant; 7, CH0001 plant; 8, CH0001 plant; 9, CH0001 plant; 10, CH0011 plant; 11, CH0030 plant; 12, CH0030 plant; 13, CH0001 plant; 14, CH0011 plant; 15, CH0030 plant; 16, R-o-18 cDNA (positive control); 17, RLR22 cDNA (positive control); 18, RLR22 gDNA (positive control); 19, dH$_2$O (negative control).
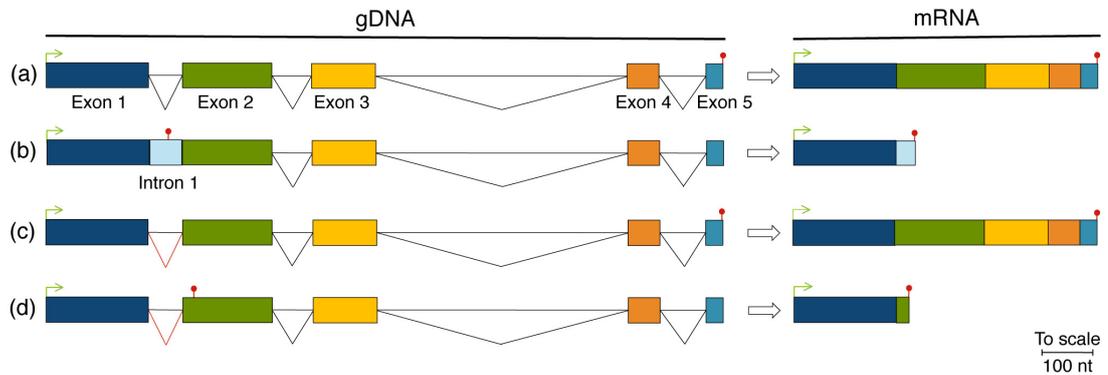
Figure 6.8 – *BraA.eIF(iso)4E.a* splice variants detected in two F$_2$ plants from the CK and RLR22 crosses, homozygous for *retr01* by RT-PCR.
(a) For comparison, correctly spliced RLR22 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Most common mis-spliced variant retaining the extra G (indel) and whole of intron 1 resulting in a premature stop codon at position 234 nt (664 nt in length). (c) Variant lacking the last 3 nt of exon 1, resulting in a slightly shortened, in-frame mRNA with a substitution (597 nt). (d) Variant retaining the extra G (indel) resulting in a premature stop codon at position 228 nt.

6.3 Discussion

6.3.1 Segregation of Resistance as a Single Recessive Gene

As the sequence of *BraA.eIF(iso)4E.c* from all six of the Syngenta parental lines was identical to that in RLR22, all the susceptible CK Chinese cabbage parents already had *ConTR01*, explaining why resistance segregated monogenically in the F_2 generations of crosses with RLR22. This will simplify the introgression process for Syngenta, as only one gene will need to be introgressed, ultimately speeding up the process.

Contradictory Results from Other TuMV Resistance Sources

Resistance to TuMV strains C4 and C5 has been described by Yoon *et al.* (1993) and they found that when they crossed the TuMV-resistant line 0-2 with susceptible Chinese cabbage lines, the resistance was controlled by two recessive genes. However, Suh *et al.* (1995) found that when an 0-2 resistant plant was crossed with different susceptible Chinese cabbage lines, the resistance was inherited as a single dominant gene, or a double dominant gene depending on the TuMV strain and the cross. The difference in inheritance of TuMV resistance in the offspring of 0-2 resistant plant and susceptible Chinese cabbage crosses, along with the R-o-18 and RLR22 cross and the RLR22 and Syngenta CK crosses, highlights the need to be aware of the genetic background of the parental plants intended for crossing to the resistant line.

6.3.2 Explaining the Limited Susceptibility to TuMV of F_2 Plants Homozygous for *retr01*

Mutations in the VPg of TuMV isolate CDN 1 were shown to result in the virus being able to infect Arabidopsis Col-0::*dSpm eIF(iso)4E* knock-out plants, which had previously been resistant (Gallois *et al.*, 2010). As the only polymorphism that was identified in the VPg of TuMV isolate CDN 1, in this study was synonymous, it is clear mutations in the VPg were not responsible for these infections.

As seen in previous chapters (Sections 4.2.4 and 5.2.3), a potentially functional variant of *BraA.eIF(iso)4E.a* was detected, which might explain the limited systemic spread of TuMV observed in some plants homozygous for

retr01. The variant that was most likely to be functional was detected in the F$_2$ plants, lacked the last 3 nt of exon 1, which resulted in the loss of one amino acid and had a substitution (W66C). As discussed in Section 4.2.4, this variant could potentially be functional for TuMV as it possesses all but one of the important amino acids when compared to published *eIF4E* and *eIF(iso)4E* sequences in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008). As discussed in Chapter 5 (Section 5.3.2), this variant (with the loss of one amino acid and a substitution) was also sent to colleagues in China, who had previously shown interactions between the VPg of TuMV and *B. rapa* using the Y2H assay. They also found that this variant interacted with the VPg of TuMV isolate C4 in the Y2H assay (Nellist *et al.*, 2014). This strengthens the hypothesis that this variant is functional for TuMV in the F$_2$ plants and is leading to the limited systemic symptoms observed in some plants. It would be interesting to investigate whether the plants homozygous for *retr01* displaying limited systemic spread of TuMV were an artefact of mechanical inoculation, by challenging the plants using aphids.

6.3.3 Exploitation of *retr01*-based Resistance in Other *Brassica* Species

The *retr01*-based resistance was identified in the diploid species *B. rapa* (AA genome). Introgression of the resistance into the diploid *Brassica oleracea* (CC genome) is desirable, as it represents a high diversity of crops which would benefit from the broad-spectrum TuMV-resistance. Successful examples of the use of interspecific crosses to introgress resistance from the A genome of *B. rapa* into the C genome of *B. oleracea* include the introgression of clubroot resistance from the A genome of the amphidiploid (AACC genome) rutabaga line (swede; *B. napus* L. ssp. *rapifera* (Metzg., Sinsk) cv. Wilhelmsburger), which shows resistance to race 2 of *Plasmodiophora brassicae* (causal agent of clubroot), into *B. oleracea*. The rutabaga line was crossed with a tetraploid (CCCC genome) cabbage (*B. oleracea* L. ssp. *capitata* cv. Châteauguay) (Chiang *et al.*, 1980; Landry *et al.*, 1992). Another successful example, again, resistance to clubroot, was performed by Syngenta. It involved the introgression of the clubroot resistance from *B. rapa* cv. Parkin into *B. oleracea* crops, using an embryo-rescue technique (Harberd, 1969) and repeated back-crosses.

Syngenta are attempting to introgress the *retr01*-based resistance into *B. oler-*

acea. An investigation of which copies of *eIF4E/eIF(iso)4E* TuMV can use in *B. oleracea* would be useful, as the resistance is recessive and based on a passive mechanism.

6.4 Conclusions

The main aim of this chapter was to investigate the introgression of broad-spectrum TuMV-resistance into commercial Chinese cabbage lines. The resistance segregated as a single recessive gene in the cross between RLR22 and the closely related Syngenta parental lines (*B. rapa* var. *pekinensis*). This chapter highlights the need to be aware of the genetic background of the plants which the resistant plant line is to be crossed with, as this can affect the segregation of resistance in subsequent generations.

Chapter 7

Complementation of the *Arabidopsis thaliana dSpm* Mutant with *Brassica rapa BraA.eIF(iso)4E.a*

7.1 Background

7.1.1 Use of *Arabidopsis thaliana* in Eukaryotic Translation Related Plant-Virus Studies

The model plant *Arabidopsis thaliana* has been widely used to study plant-virus interactions. Arabidopsis is appropriate for such studies because of its short life-cycle, small, well characterised and sequenced genome with vast genomic resources and the availability of abundant mutants. It was first discovered that plant viruses directly interacted with the eukaryotic translation initiation complex of plants in yeast two-hybrid (Y2H) binding assays (Wittmann *et al.*, 1997). These studies showed that the virus-encoded genome-linked protein (VPg) of *Turnip mosaic virus* (TuMV) bound to Arabidopsis eukaryotic translation initiation factor iso-form 4E (eIF(iso)4E; Wittmann *et al.*, 1997). Duprat *et al.* (2002) showed that a transposon knock-out of eIF(iso)4E (Col-0::*dSpm*) led to potyvirus resistance, particularly TuMV and *Lettuce mosaic virus* (LMV). The mutant line was found to completely lack both eIF(iso)4E mRNA and protein, but was identical to the wild-type plants, under standard laboratory conditions (Duprat *et al.*, 2002).

7.1.2 Previous Work with Arabidopsis Col-0::*dSpm* and *Brassica rapa eIF4E/eIF(iso)4E*

Previous work used the Arabidopsis Col-0::*dSpm* line, a transposon knock-out of *eIF(iso)4E* (Duprat *et al.*, 2002) to test *Brassica rapa* R-o-18 (Jenner *et al.*, 2010) and RLR22 (Nellist *et al.*, 2014) copies of *eIF4E* and *eIF(iso)4E* from multiple loci for their ability to complement the *dSpm* knock-out. It was found that TuMV could use two copies of *eIF4E* (*BraA.eIF4E.a* and *BraA.eIF4E.c*) and two copies of *eIF(iso)4E* (*BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c*) from both plant lines to complete its infection-cycle when transformed into Arabidopsis Col-0::*dSpm*. This was interesting as complementation with RLR22 *BraA.eIF(iso)4E.a* was unexpected. The experiments were performed using mechanical inoculation, which is not the natural mechanism by which TuMV is transmitted and could have resulted in susceptibility due to the plant being overloaded by viral inoculum. Work described by Marco *et al.* (2003) is an example of when the inoculation method can have an effect on susceptibility. Also, Rusholme *et al.* (2007) showed that although it was possible to induce local infection of RLR22 following mechanical inoculation of TuMV, aphid challenges failed to induce any local TuMV infection. To test whether the amount of viral inoculum used resulted in the susceptibility phenotype in the Arabidopsis Col-0::*dSpm* plants complemented with RLR22 *BraA.eIF(iso)4E.a* and to replicate what would happen in the field, an experiment directly comparing mechanical inoculation and aphid transmission was conducted. Efforts focused on *BraA.eIF(iso)4E.a*, as this was the candidate for *retr01*.

This chapter describes work using the Arabidopsis Col-0::*dSpm* eIF(iso)4E knock-out line transformed with *BraA.eIF(iso)4E.a* transgenes to try to determine if there were any differences between the utilisation of *BraA.eIF(iso)4E.a* from R-o-18 and RLR22 by TuMV during the infection process. The aim of this work was to try and explain the susceptibility phenotype observed in Arabidopsis *dSpm* plants complemented with the RLR22 *BraA.eIF(iso)4E.a* (*retr01*).

7.2 Results

7.2.1 Complementation of Arabidopsis with *B. rapa* *BraA.eIF(iso)4E.a*

When challenged with TuMV isolate CDN 1, all Arabidopsis control Col-0 and Col-4 plants (without brassica transgene) developed severe symptoms after mechanical inoculation of TuMV (Figure 7.1). Of the aphid-challenged control Col-0 plants 2/5 did not develop any symptoms and were ELISA negative, indicating the transmission was not successful, the remaining three were severely stunted and ELISA positive (Figure 7.1 and Table 7.1). Arabidopsis Col-0::*dSpm* plants (without brassica transgene; (Duprat *et al.*, 2002)) did not develop any symptoms and showed resistance to TuMV, both by mechanical inoculation and aphid challenge (Figure 7.1; Table 7.1). All mock-inoculated plants and plants colonised with non-viruliferous aphids remained symptomless. The brassica *BraA.eIF(iso)4E.a* transgenes from R-o-18 and RLR22 complemented the Arabidopsis *eIF(iso)4E* knock-out and were able to be used by TuMV as indicated by susceptible phenotypes, both by mechanical inoculation and aphid challenge (Figure 7.1). Arabidopsis Col-0::*dSpm* plants with brassica transgenes produced symptoms as severe as the untransformed Col-0 line. The visual symptoms were consistent between the plants within the lines, except for plants challenged with aphids that did not develop symptoms. Visually, there were no differences between the symptoms presented by the Col-0::*dSpm* plants complemented with R-o-18 *BraA.eIF(iso)4E.a*, or RLR22 *BraA.eIF(iso)4E.a* (Figure 7.1). This was confirmed by ELISA (Table 7.1), there was no significant difference (P=0.05) in the amount of TuMV accumulating in the two lines when subjected to residual maximum likelihood (REML) analysis. Overall, plants challenged with TuMV using aphids were less efficiently infected, not all potential susceptible controls were infected. The virus titer for aphid challenged plants as measured by ELISA was also significantly reduced for each brassica gene compared to mechanically inoculated plants (P<0.05; Table 7.1) when analysed by REML.
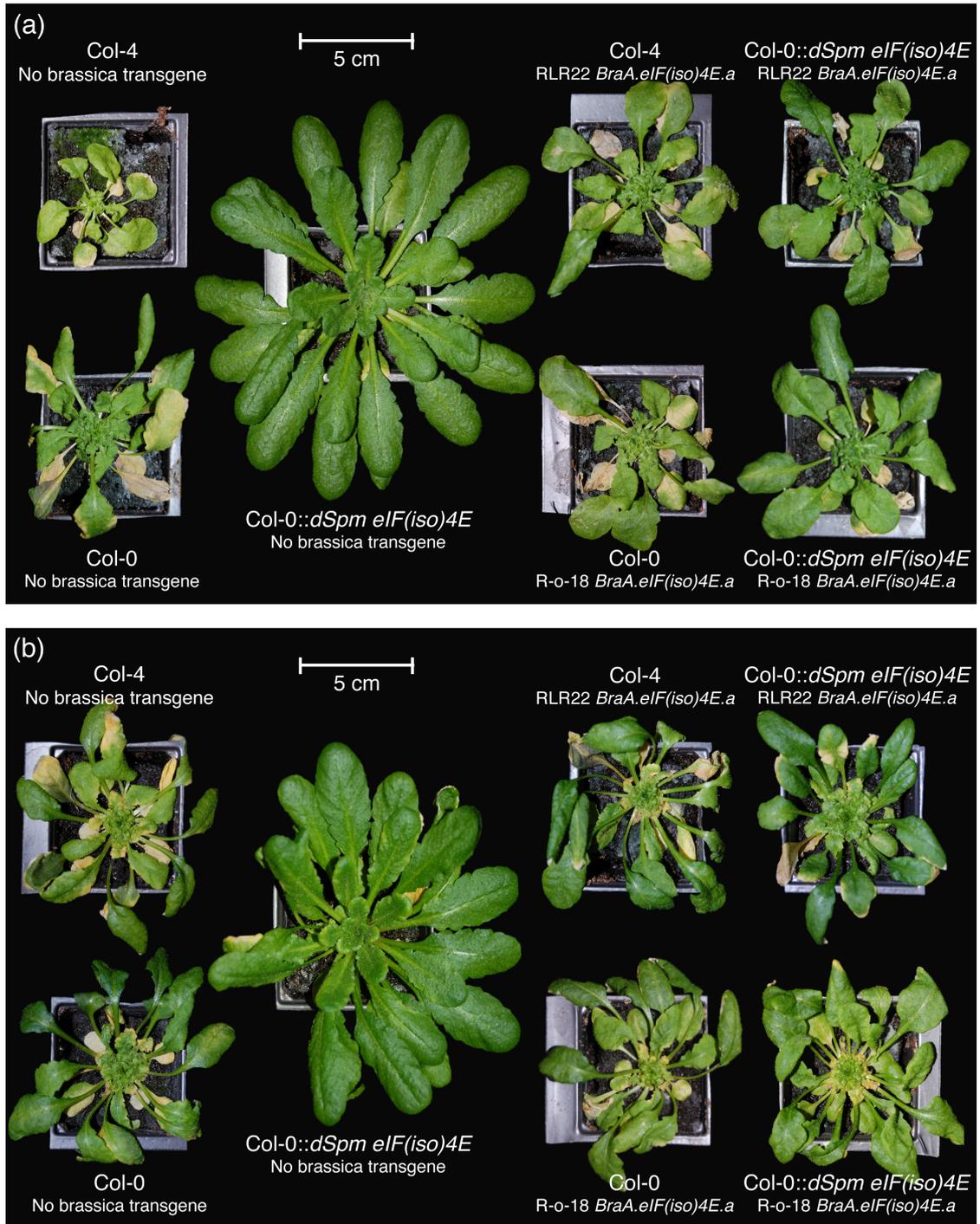
Figure 7.1 – Complementation of *Arabidopsis thaliana* Col-0::*dSpm* plants with R-o-18 and RLR22 *BraA.eIF(iso)4E.a*, following *Turnip mosaic virus* isolate CDN 1 challenge by (a) mechanical inoculation and (b) aphid challenge, both 33 days post inoculation.

Table 7.1 – Phenotypic and ELISA data of *Arabidopsis thaliana* plants complemented with *Brassica rapa BraA.eIF(iso)4E.a*.

| Brassica transgene | Inoculation method | No. of plants with phenotypes[a] | | | | Mean ELISA optical density[b] | | | | | |
| | | Col-0 | | Col-0::dSpm[c] | | Col-0 | | | Col-0::dSpm | | |
		0[d]	SM[e]	0	SM	0	SM	Mock[f]	0	SM	Mock
None	Mechanical	0	5	5	0	-[g]	0.26	0.04	0.05	-	0.05
	Aphid	2	3	5	0	0.04	0.19	0.04	0.04	-	0.04
R-o-18 *BraA.eIF(iso)4E.a*	Mechanical	0	12	4	8	-	0.22	0.05	0.04	0.20	0.05
	Aphid	6	6	7	5	0.04	0.16	0.04	0.04	0.13*[h]	0.04
RLR22 *BraA.eIF(iso)4E.a*	Mechanical	0	12	1	11	-	0.25	0.04	0.05	0.20	0.04
	Aphid	6	6	6	6	0.04	0.18	0.04	0.04	0.13*[i]	0.04

[a] Twelve plants inoculated per T_2 family. Families had not been selected for transgene homozygosity and were segregating for the presence of the transgene.

[b] ELISA absorbance measured at 405 nm.

[c] *dSpm* transposon insertion inactivating *At.eIF(iso)4E* (*At5g35620*).

[d] No symptoms observed.

[e] Systemic mosaic symptoms.

[f] Mock-inoculated control plants.

[g] Not applicable.

[h] A_{405} values from ELISA of aphid challenged R-o-18 *BraA.eIF(iso)4E.a* plants were statistically significant different (following residual maximum likelihood analysis) compared with R-o-18 *BraA.eIF(iso)4E.a* mechanically inoculated plants, indicated by * ($P<0.05$).

[i] A_{405} values from ELISA of aphid challenged RLR22 *BraA.eIF(iso)4E.a* plants were statistically significant different (following residual maximum likelihood analysis) compared with RLR22 *BraA.eIF(iso)4E.a* mechanically inoculated plants, indicated by * ($P<0.05$).

PCR tests were conducted to check for the presence/absence of the transposon and brassica transgenes in a sample of plants. PCR confirmed the presence of the *dSpm* element in the untransformed Col-0::*dSpm* line, the Col-0::*dSpm* line transformed with R-o-18 *BraA.eIF(iso)4E.a* and the Col-0::*dSpm* line transformed with RLR22 *BraA.eIF(iso)4E.a* (Figure 7.2; lanes 6-16). The absence of the *dSpm* element was also confirmed by PCR of the transformed and untransformed Col-0/Col-4 control lines (Figure 7.2; lanes 1-5). PCR also confirmed the presence of the R-o-18 transgene in the susceptible Col-0 and Col-0::*dSpm* plants transformed with R-o-18 *BraA.eIF(iso)4E.a* (Figure 7.3; lanes 3-6) and the absence of the transgene in the untransformed Col-0, Col-0::*dSpm* controls and four TuMV-resistant T_2 Col-0::*dSpm* plants (CN1.131, CN1.134, CN1.139 and CN1.141) (Figure 7.3; lanes 1, 2 and 7-10). The presence of the RLR22 transgene in the susceptible Col-0 and Col-0::*dSpm* plants transformed with RLR22 *BraA.eIF(iso)4E.a* (Figure 7.4; lanes 3-6) and the absence of the transgene in the untransformed Col-0, Col-0::*dSpm* controls and a TuMV-resistant T_2 Col-0::*dSpm* plant (CN1.154) (Figure 7.4; lanes 1, 2 and 7) was also confirmed by PCR.
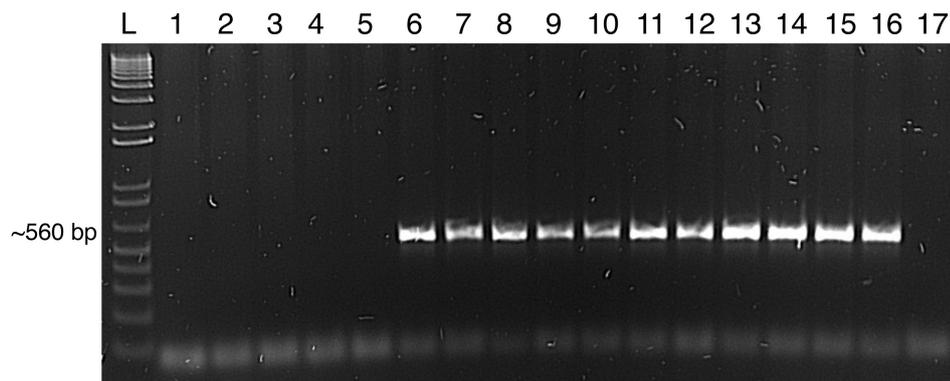
Figure 7.2 – PCR amplification of *dSpm* insertion (∼560 bp) in *Arabidopsis thaliana* using specific primers KO1 and DSPM1.
Lanes: L, 1 Kb Plus DNA Ladder; 1, Untransformed Col-0 ; 2-3, Col-4 transformed with RLR22 *BraA.eIF(iso)4E.a*; 4-5, Col-4 transformed with R-o-18 *BraA.eIF(iso)4E.a*; 6-7, Untransformed Col-0::*dSpm*; 8-13, Col-0::*dSpm* transformed with R-o-18 *BraA.eIF(iso)4E.a* (10, CN1.131; 11, CN1.134; 12, CN1.139; 13, CN1.141); 14-16, Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a* (16, CN1.154); 17, dH$_2$O (negative control).
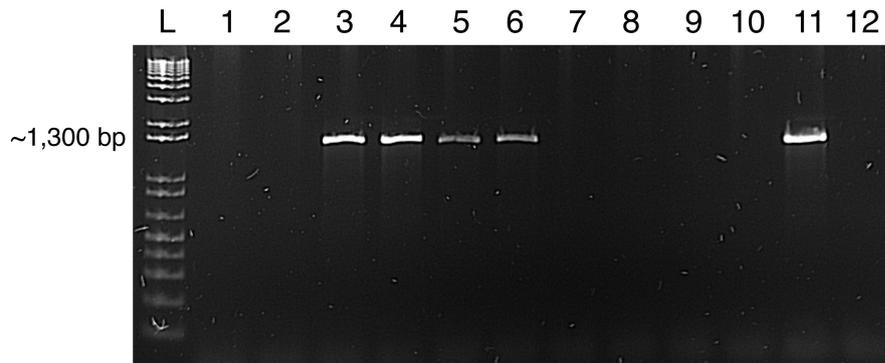
Figure 7.3 – PCR amplification of *Brassica rapa* R-o-18 *BraA.eIF(iso)4E.a* transgene (∼1300 bp) in *Arabidopsis thaliana* using specific primers CN3 and CN4.
Lanes: L, 1 Kb Plus DNA Ladder; 1, Untransformed Col-0; 2, Untransformed Col-0::*dSpm*; 3-4, Col-0 transformed with R-o-18 *BraA.eIF(iso)4E.a*; 5-10, T_2 Col-0::*dSpm* plants from a T_1 plant transformed with R-o-18 *BraA.eIF(iso)4E.a* (7, CN1.131; 8, CN1.134; 9, CN1.139; 10, CN1.141); 11, R-o-18 gDNA (positive control); 12, dH$_2$O (negative control).
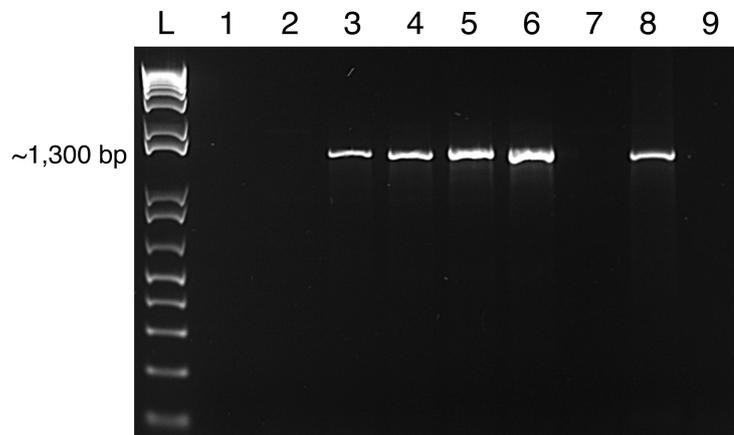
Figure 7.4 – PCR amplification of *Brassica rapa* RLR22 *BraA.eIF(iso)4E.a* transgene (∼1300 bp) in *Arabidopsis thaliana* using specific primers CN3 and CN5.
Lanes: L, 1 Kb Plus DNA Ladder; 1, Untransformed Col-0; 2, Untransformed Col-0::*dSpm*; 3-4, Col-4 transformed with RLR22 *BraA.eIF(iso)4E.a*; 5-7, T_2 Col-0::*dSpm* plants from a T_1 plant transformed with RLR22 *BraA.eIF(iso)4E.a* (7, CN1.154); 8, RLR22 gDNA (positive control); 9, dH$_2$O (negative control).

7.2.2 Transcriptome Analysis of Arabidopsis Col-0::*dSpm* Complemented with RLR22 *BraA.eIF(iso)4E.a*

As the mechanism of resistance to TuMV in the RLR22 plants had been identified as the mis-splicing of *BraA.eIF(iso)4E.a*, the transcriptome expression profile of an Arabidopsis Col-0::*dSpm* plant complemented with *BraA.eIF(iso)4E.a* from RLR22 was determined to try and explain why the plants were susceptible to TuMV when transformed with the *retr01*-resistance gene. A total of 3,252,900, 70 base paired-end sequence reads were recorded within RLR22 *BraA.eIF(iso)4E.a*, with good coverage of reads across the whole gene (Figure 7.5).

A correctly spliced version of the gene (600 nucleotides (nt) in length) was not detected in the expression profile, the majority of transcripts observed had the mis-spliced variant that retained the extra G (indel) and the whole of intron 1 (Figure 7.5 and Figure 7.6, (b)), which introduced a premature stop codon at position 234 nt. Four further mis-spliced variants were detected, one variant that retained the extra G (indel) and lacked the first 65 nt of exon 2 (or alternatively, if the G comprises part of exon 2, then the variant lacks the first 66 nt of exon 2; Figure 7.5, (a) and Figure 7.6, (c)), resulting in a premature stop codon at 288 nt, one variant retaining the extra G (indel) and the first 4 nt of intron 1 and lacking the first 66 nt of exon 2 (Figure 7.5, (b) and Figure 7.6, (d)), resulting in an in-frame shortened sequence (540 nt in length), one variant retaining the extra G (indel) and 28 nt of intron 1 and lacking the first 66 nt of exon 2 (Figure 7.6, (e)), resulting in an in-frame shortened sequence (564 nt in length) and one variant lacking the last 93 nt of exon 1 (Figure 7.6, (f)), resulting in an in-frame shortened sequence (507 nt in length).
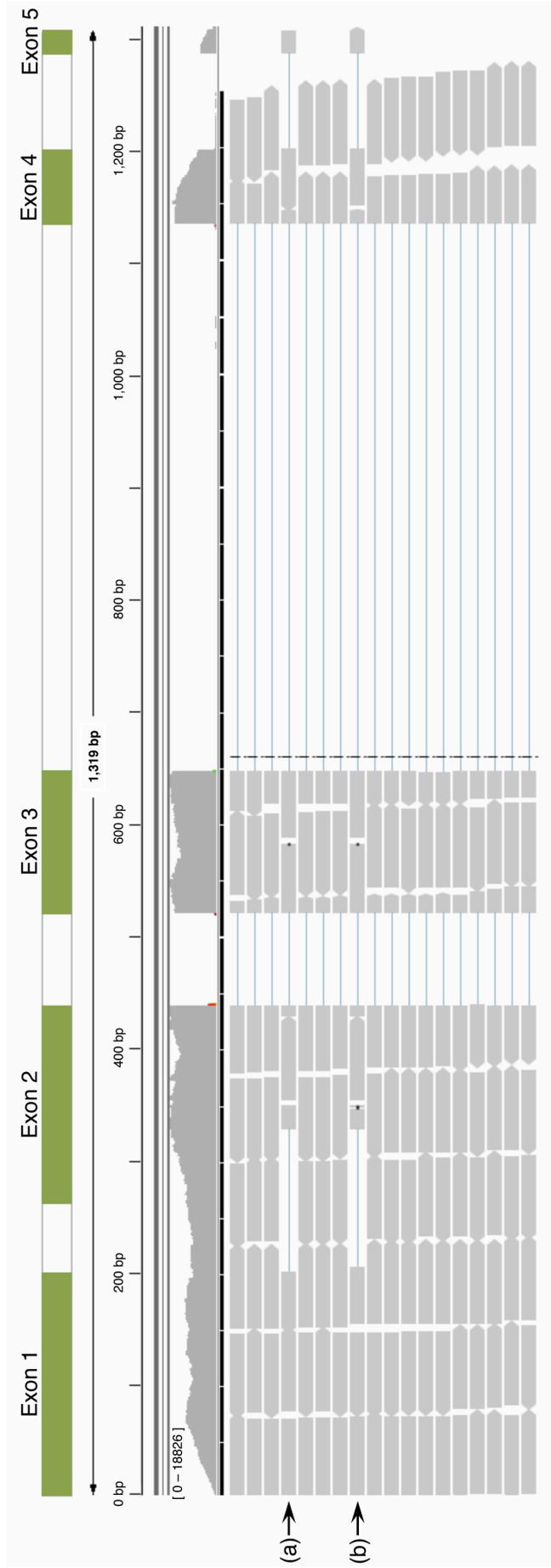
Figure 7.5 – A section of the coverage report from the Integrative Genome Viewer (IGV), of transcript reads of RLR22 *BraA.eIF(iso)4E.a* in *Arabidopsis thaliana* Col-0::*dSpm* from transcriptome sequencing. Exons (green boxes) and introns (white boxes) overlaid at the top of the diagram. The majority of transcripts possessed an extra G (indel) and the whole of intron 1. (a) Variant retaining extra G (indel) and lacking the first 65 nt of exon 2/lacking first 66 nt of exon 2. (b) Variant retaining extra G (indel) and the first 4 nt of intron 1 and lacking the first 65 nt of exon 2.

113
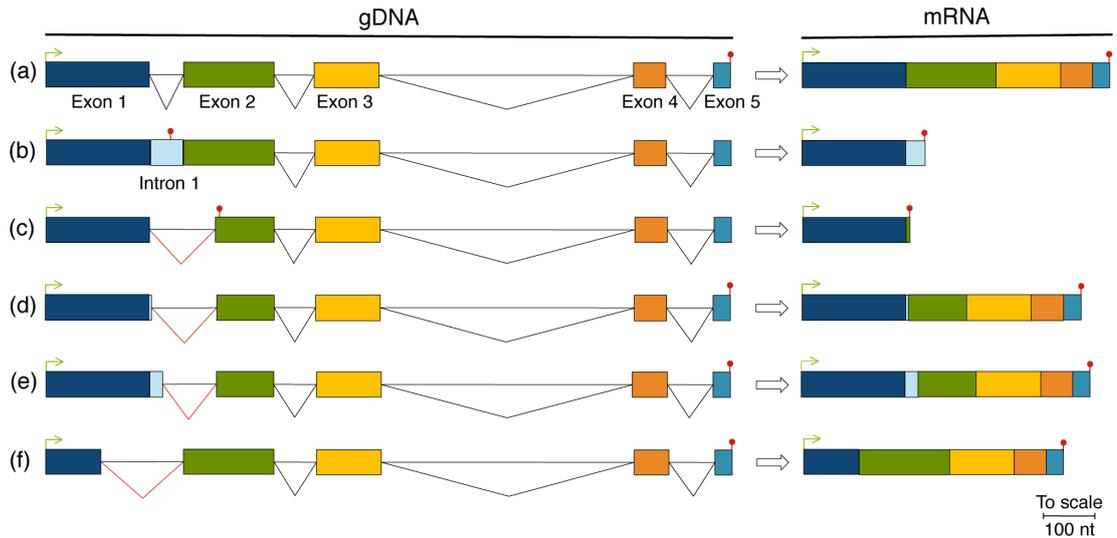
Figure 7.6 – *BraA.eIF(iso)4E.a* splice variants detected in *Arabidopsis thaliana* Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a* by transcriptome sequencing, compared to correctly spliced *BraA.eIF(iso)4E.a*.
(a) For comparison, correctly spliced RLR22 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Most common mis-spliced variant retaining the extra G (indel) and whole of intron 1 resulting in a premature stop codon at position 234 nt. (c) Variant retaining the extra G (indel) and lacking first 65 nt of exon 2/lacking the first 66 nt of exon 2 resulting in a premature stop codon at position 207 nt. (d) Variant retaining the extra G and first 4 nt of intron 1 and lacking the first 66 nt of exon 2 (in-frame, 540 nt in length). (e) Variant retaining the extra G (indel) and 28 nt of intron 1 and lacking the first 66 nt of exon 2 (in-frame, 564 nt in length). (f) Variant lacking the last 93 nt of exon 1 (in-frame, 507 nt in length).

7.2.3 RT-PCR Analysis of the Expression of RLR22 *BraA.eIF(iso)4E.a* in Arabidopsis Col-0::*dSpm*

The transcriptome sequencing did not reveal any correctly spliced versions of *BraA.eIF(iso)4E.a*. RT-PCR analysis was performed to see if there were any less common variants that were correctly spliced, or other mis-spliced variants that could be functional for TuMV, to explain the susceptible phenotype. Primers for RLR22 *BraA.eIF(iso)4E.a* were used (CN3 and CN 5, see Section 2.5.2). The initial gel showed multiple products, a major larger product of ∼650 bp and then smaller minor products between ∼500-650 bp (Figure 7.7; lane 2). A correctly spliced version of the gene would be 600 bp in size. Sequencing revealed that the major product contained a mis-spliced variant with an extra G (indel) and the whole of intron 1 present (Figures 7.5 and 7.8, (b)), introducing a premature stop codon at position 234 nt. This is the same as was revealed by the transcriptome

sequencing of Arabidopsis Col-0::*dSpm* transformed with *BraA.eIF(iso)4E.a* and also in *B. rapa* RLR22 (Section 4.2.3).
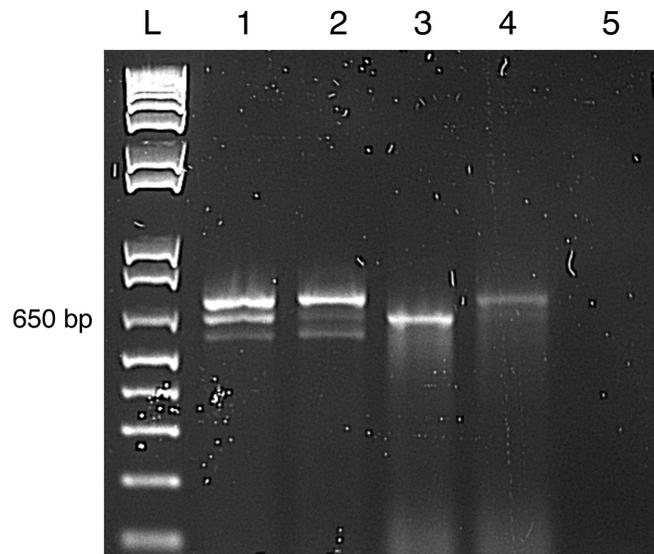
Figure 7.7 – Mis-splicing of R-o-18 and RLR22 *BraA.eIF(iso)4E.a* in *Arabidopsis thaliana* Col-0::*dSpm* transformed with R-o-18 and RLR22 *BraA.eIF(iso)4E.a* using specific primers CN3 and CN4/5.
Lanes: L, 1 Kb Plus DNA Ladder; 1, Col-0::*dSpm* transformed with R-o-18 *BraA.eIF(iso)4E.a*; 2, Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a*; 3, *BraA.eIF(iso)4E.a* cDNA from R-o-18; 4, *BraA.eIF(iso)4E.a* cDNA from RLR22; 5, dH$_2$O (negative control).

Cloning and sequencing of the minor products did not reveal any correctly spliced transcripts of the gene. Two further variants that were detected with the transcriptome sequencing were also detected by RT-PCR, the variant that retained the extra G (indel) and the whole of intron 1 (Figure 7.8, (b)) as well as the variant with the extra G (indel) and lacking the first 65 nt of exon 2/lacking the first 66 nt of exon 2 (Figure 7.8, (c)). A further three mis-spliced variants were also detected, one variant with an extra G at the end of exon 1 (Figure 7.8, (d)), introducing a premature stop codon at position 228 nt, one variant with the whole of intron 1 retained (but not the indel; Figure 7.8, (e)) introducing a premature stop codon at position 243 nt and one variant retaining the extra G (indel) and the last 14 nt of intron 1/retaining the last 15 nt of intron 1 (Figure 7.8, (f)), resulting in a slightly elongated mRNA (615 nt in length). The latter variant could potentially be functional for TuMV, as it would only result in an additional five amino acids relative to the correctly spliced version.
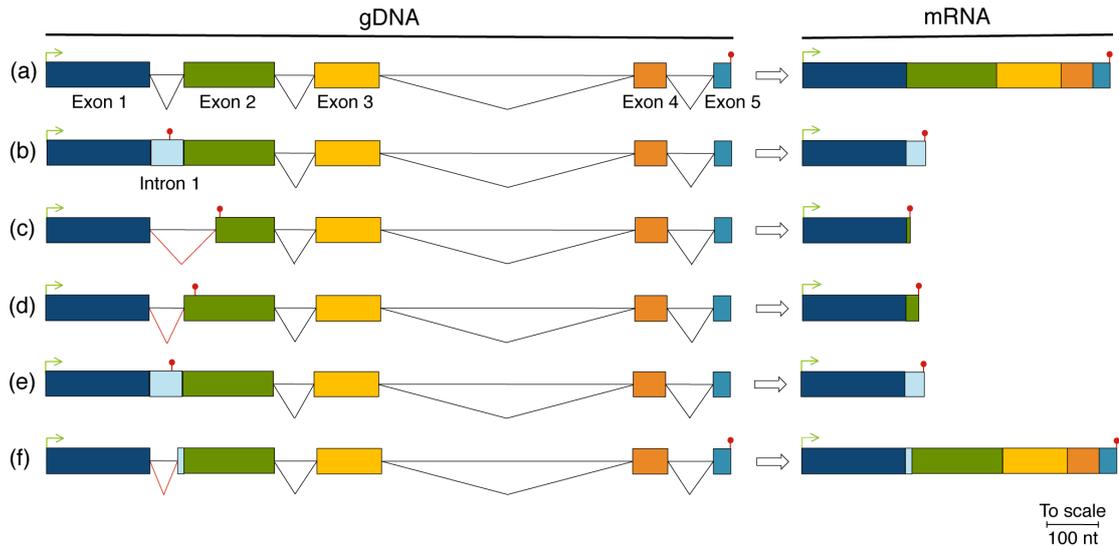
Figure 7.8 – *BraA.eIF(iso)4E.a* splice variants detected in *Arabidopsis thaliana* Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a* by RT-PCR.

(a) For comparison, correctly spliced RLR22 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Most common mis-spliced variant retaining the extra G (indel) and whole of intron 1 resulting in a premature stop codon at position 234 nt. (c) Variant retaining the extra G (indel) and lacking first 65 nt of exon 2/lacking the first 66 nt of exon 2 resulting in a premature stop codon at position 207 nt. (d) Variant retaining extra G (indel) resulting in a premature stop codon at position 228 nt. (e) Variant retaining the whole of intron 1 (but not indel) introducing a premature stop codon at position 243 nt. (f) Variant retaining the extra G (indel) and the last 14 nt of intron 1/retaining the last 15 nt of intron 1, resulting in a slightly elongated mRNA (615 nt in length).

7.2.4 RT-PCR Analysis of the Expression of R-o-18 *BraA.eIF(iso)4E.a* in Arabidopsis Col-0::*dSpm*

RT-PCR analysis was also performed on the Arabidopsis Col-0::*dSpm* line transformed with R-o-18 *BraA.eIF(iso)4E.a*. Multiple bands were observed on an agarose gel (Figure 7.7, lane 1). When these products were sequenced, a correctly spliced version of the gene was detected (600 nt in length; Figure 7.7, lane 1 middle band and Figure 7.9, (a)), a variant retaining the whole of intron 1 was also detected (Figure 7.7, lane 1, upper band and Figure 7.9, (b)). This introduced a premature stop codon at 243 nt. A smaller variant was also detected lacking the first 65 nt of exon 2 (Figure 7.7, lane 1 lower band and Figure 7.9 (c)). This introduced a premature stop codon at 210 nt.
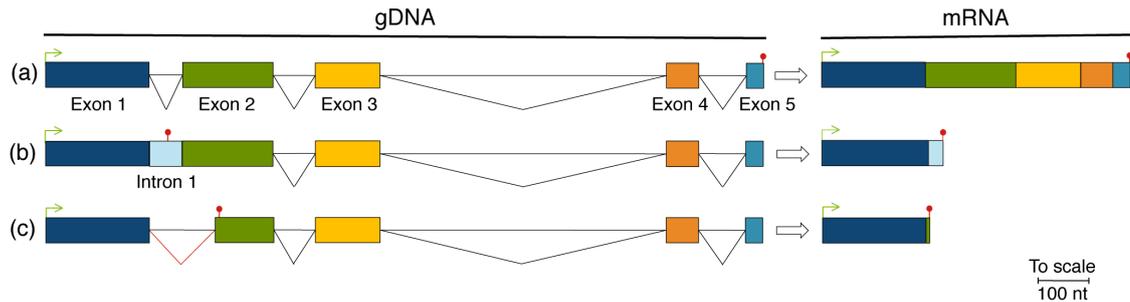
Figure 7.9 – *BraA.eIF(iso)4E.a* transcripts detected in *Arabidopsis thaliana* Col-0::*dSpm* transformed with R-o-18 *BraA.eIF(iso)4E.a* by RT-PCR. (a) Correctly spliced R-o-18 *BraA.eIF(iso)4E.a* (600 nt in length). (b) Variant retaining whole of intron 1, resulting in a premature stop codon at position 234 nt. (c) Variant lacking the first 65 nt of exon 2, resulting in a premature stop codon at position 210 nt.

7.2.5 Analysis of RLR22 *BraA.eIF(iso)4E.a* Splice Variants Detected in Transgenic Arabidopsis Col-0::*dSpm*

There were eight RLR22 *BraA.eIF(iso)4E.a* splice variants detected in Arabidopsis Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a* either by transcriptome analysis, or RT-PCR. Of these, four introduced a premature stop codon, severely truncating the protein, three were in-frame but missing a considerable number of amino acids (between 12-31) and one had an additional five amino acids. Comparing the sequences with the correctly spliced BraA.eIF(iso)4E.a and the published eIF4E and eIF(iso)4E sequences in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008) indicated that the variant with an additional five amino acids had the correct amino acids at important sites (Figure 7.10, (b)). Another two variants possess all the correct amino acids at important sites but were less similar to the correctly spliced version of *BraA.eIF(iso)4E.a*; the variant retaining an extra G (indel) and the first 4 nt of intron 1 but lacking the first 66 nt of exon 2 (Figure 7.10, (c)) and the variant retaining an extra G (indel) and first 28 nt of intron 1 but lacking the first 66 nt of exon 2 (Figure 7.10, (d)). The rest of the variants were unlikely to be functional for TuMV as they were missing some of the important amino acids (Figure 7.10).

```
MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG-----------------------GVEPKWEDPECANGG
---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+
         10        20        30        40        50        60        70        80        90        100       110
---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+

(a) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG-----LHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG   104
(b) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGIFFCSLHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG   109
(c) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG----------------------VSGVEPKWEDPECANGG      84
(d) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWG--------VCLSSPFT-------YSGVEPKWEDPECANGG      92
(e) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSF----------------------------C-----LHETIFIPSKLTPNAEIHMFKAGVEPKWEDPECANGG      73
(f) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGVCLSSPFTYC.                                  78
(g) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGFVCLLLLLIVSDL.                                81
(h) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGFARDYIHP.                                    76
(i) MATEDVNEAL-AAAEVPATETTEKQPADKLERKWSFWFDNQSKPKQGAAWGASLRKAYTFDTVQDFWGLVLSLSGKILSVLMAESGLLLLPPTASLL.              97
(j) MATDDVNEPLPAAAELPATE-AEKQP-HKLERKWSFWFDNQS--KKGAAWGASLRKAYTFDTVEDFWG-----LHETIFQTSKLTANAEIHLFKAGVEPKWEDPECANGG   101
(k)                                   W                  W              W                      E   WE    C


KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV-
---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
         120       130       140       150       160       170       180       190       200
---------+---------+---------+---------+---------+---------+---------+---------+---------+--------

(a) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV.   200
(b) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV.   205
(c) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV.   180
(d) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV.   188
(e) KWTFVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRP--KQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV.   169
(f)                                                                                                      78
(g)                                                                                                      81
(h)                                                                                                      76
(i)                                                                                                      97
(j) KWTWVVTANRKEALDKGWLETLMALIGEQFDEADEICGVVASVRPQSKQDKLSLWTRTKSNEAVLMGIGKKWKEILDVTDKITFNNHDDSRRSRFTV.   199
(k)                          GE       C       R   K   W
```

Figure 7.10 – Multiple alignment of amino acid sequences of correctly and mis-spliced RLR22 *BraA.eIF(iso)4E.a* detected in *Arabidopsis thaliana* Col-0::*dSpm*, based on homology with *eIF4E* and *eIF(iso)4E* from other plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008). Amino acids that are most directly involved in binding the mRNA cap-structure are highlighted in green. Amino acids that are involved in binding eIF4G are highlighted in blue. Cys residues involved in forming a di-sulphide bridge in wheat and implicated in mRNA cap binding are highlighted in yellow.
(a) RLR22 BraA.eIF(iso)4E.a, correctly spliced. (b) Variant retaining extra G (indel) and last 14 nt of intron 1/last 15 nt of intron 1 (Figure 7.8, (f)). (c) Variant retaining extra G (indel) and first 4 nt of intron 1 and lacking first 66 nt of exon 2 (Figure 7.6, (d)). (d) Variant retaining extra G (indel) and first 28 nt of intron 1 and lacking first 66 nt of exon 2 (Figure 7.6, (e)). (e) Variant lacking last 93 nt of exon 1 (Figure 7.6, (f). (f) Variant retaining the extra G (indel) and whole of intron 1 (Figure 7.6, (b)). (g) Variant lacking extra G (indel) and retaining the whole of intron 1 (Figure 7.8, (e)). (h) Variant retaining extra G (indel) (Figure 7.8, (d)). (i) Variant retaining extra G (indel) and lacking first 65 nt of exon 2/lacking first 66 nt of exon 2 (Figure 7.8, (c)). (j) Arabidopsis eIF(iso)4E. (k) Important amino acids.

7.3   Discussion

7.3.1   Complementation of the Arabidopsis *eIF(iso)4E* Knock-out with *BraA.eIF(iso)4E.a*

The main aim of the chapter was to investigate the unexpected susceptibility phenotype observed in Arabidopsis Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a*. Both R-o-18 (Jenner *et al.*, 2010) and RLR22 (Nellist *et al.*, 2014) copies of *BraA.eIF(iso)4E.a* complemented Arabidopsis *eIF(iso)4E* in the Col-0::*dSpm* line, whether TuMV was transmitted by mechanical inoculation, or aphid challenge. The susceptibility phenotype observed in the aphid-challenged plants transformed with *BraA.eIF(iso)4E.a* indicated that the mechanically inoculated plants susceptibility was not due to virus overloading. The virus titre was significantly reduced in the aphid challenged plants compared to mechanically inoculated plants. This could be a result of the lower viral load delivered by the aphids. Although the accumulation of the virus was significantly reduced, the plants were still susceptible to TuMV when aphid challenged, visually there were no differences between symptoms observed in plants where TuMV was mechanically inoculated and plants where TuMV was aphid transmitted, indicating there were functional transcripts of *BraA.eIF(iso)4E.a* in the Arabidopsis plants that TuMV was able to utilise to infect the plants.

The Arabidopsis plants were transformed with genomic DNA (gDNA), but having discovered the mechanism of the *retr01*-based resistance after the Arabidopsis plants were transformed, a more informative experiment would have been to transform the Arabidopsis plants with complementary DNA (cDNA) rather than gDNA. Transforming Arabidopsis with the subsequently discovered individual variants revealed by RT-PCR analysis of RLR22 would have provided more insight into whether the particular variants could be functional for TuMV.

7.3.2   Analysis of *BraA.eIF(iso)4E.a* variants in Transgenic Arabidopsis

Subsequent to determining that the transmission method was not responsible for the susceptibility phenotype observed in transformed plants, the question of why the line transformed with RLR22 *BraA.eIF(iso)4E.a* was susceptible to TuMV still remained unanswered. Investigation of the expression of *BraA.eIF(iso)4E.a*

in this line revealed multiple mis-spliced variants. It is reported that intron retention is the most common alternative splicing event; in plants with 30-50% of Arabidopsis alternative splicing events and 54% of rice alternative splicing events being intron retention compared to approximately 10% in humans (Ner-Gaon *et al.*, 2004; Wang and Brendel, 2006). The most frequently detected mis-spliced variant in the plants transformed with RLR22 *BraA.eIF(iso)4E.a* was the variant retaining the extra G (indel) and the whole of intron 1, however this produces a severely truncated mRNA/amino acid sequence which would most likely be non-functional for TuMV. The most notable variant was the one retaining the extra G and the last 14 nt of intron 1/retaining the last 15 nt of intron 1 resulting in additional five amino acids, this variant was the most similar to a correctly spliced transcript. This variant had the correct amino acids at important sites (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008) indicating it could be functional for TuMV, possibly explaining why the knock-out line was complemented by the RLR22 allele of *BraA.eIF(iso)4E.a*. Another two variants could potentially be functional for TuMV but were less similar to the correctly spliced version of *BraA.eIF(iso)4E.a*; the variant retaining an extra G (indel) and the first 4 nt of intron 1 but lacking the first 66 nt of exon 2 and the variant retaining an extra G (indel) and first 28 nt of intron 1 but lacking the first 66 nt of exon 2.

As discussed in Chapter 5 (Section 5.3.2), the variant possessing an additional five amino acids, was sent to colleagues in China, who had previously shown interactions between the VPg of TuMV and BraA.eIF(iso)4E.a from TuMV-susceptible *B. rapa* using the Y2H assay. They found that this variant interacted with the VPg of TuMV isolate C4 in the Y2H assay (Nellist *et al.*, 2014). This strengthens the hypothesis that this variant is functional for TuMV in the transformed Arabidopsis plants and is leading to the systemic symptoms observed in the plants.

When trying to explain the susceptibility phenotype observed in Arabidopsis Col-0::*dSpm* transformed with RLR22 *BraA.eIF(iso)4E.a*, two techniques were used to look for correctly spliced, or functional variants of *BraA.eIF(iso)4E.a*, transcriptome sequencing and RT-PCR analysis. Transcriptome sequencing used random primers to amplify the RNA, which leads to 'contamination' with ribosomal RNA (rRNA) (Wall *et al.*, 2009), which could have resulted in a reduction of the amplification of less common RNAs. This may explain why more variants were detected after RT-PCR analysis. RT-PCR used specific

primers for RLR22 *BraA.eIF(iso)4E.a*, which helped to remove the distortion of rRNAs by only amplifying *BraA.eIF(iso)4E.a*. However, as the majority of transcripts of *BraA.eIF(iso)4E.a* retained the extra G (indel) and the whole of intron 1, it was difficult to detect less common variants as the primers used (CN3 and CN5, see Section 2.5.2) were located at the beginning and the end of the gene. In an attempt to detect correctly spliced *BraA.eIF(iso)4E.a*, primers were designed (CN57 and CN59; Table 2.4) to span the exon 1 - exon 2 splice site. However, despite using different temperatures, touchdown and different magnesium concentrations, the primers did not amplify correctly spliced *BraA.eIF(iso)4E.a* from cDNA of the genes from R-o-18 and RLR22. There could potentially be more mis-spliced variants, or even a correctly spliced version of the gene present in the Arabidopsis plants, as was the case in RLR22 (see Section 4.2.3), but in very low numbers.

The brassica genes were constitutively over-expressed in the Arabidopsis plants as they had the CaMV 35S promoter driving transcription. This along with intron retention being the frequent alternative splicing event in Arabidopsis (Ner-Gaon *et al.*, 2004) and splicing at that site being error prone, may explain why a variant retaining the whole of intron 1 and also one lacking the first 65 nt of exon 2 were detected in Arabidopsis Col-0::*dSpm* plants transformed with the R-o-18 *BraA.eIF(iso)4E.a*.

7.4   Conclusions

The main aim of this chapter was to explain the susceptibility phenotype observed in Arabidopsis *dSpm* plants complemented with RLR22 *BraA.eIF(iso)4E.a* (*retr01*). Aphid transmission experiments showed that earlier results from mechanical inoculations were not an artefact and that ectopic expression of *eIF4E* and *eIF(iso)4E* was confusing as it indicated that TuMV could use both copies of *eIF4E* and *eIF(iso)4E* which is clearly not the case in *B. rapa*. Potentially functional mis-spliced variants of *BraA.eIF(iso)4E.a* were discovered, providing an explanation for the susceptible phenotype observed in the Arabidopsis *dSpm* plants transformed with RLR22 *BraA.eIF(iso)4E.a*.

Chapter 8

General Discussion

8.1 Summary of Findings

The exploitation of natural durable resistance to plant pathogens is the best approach to disease control. Plant varieties with resistance to viruses are considered the most cost-effective and reliable approach to protection (Kang *et al.*, 2005b). *Turnip mosaic virus* (TuMV) is a huge constraint on the cultivation of a wide range of plant species worldwide (Walsh *et al.*, 2002). An important finding in my study was the discovery of the novel mechanism of resistance to TuMV in *Brassica rapa*, resulting from the mis-splicing of the isoform of eukaryotic translation factor 4E (*eIF(iso)4E*). This is the first example of natural translation factor-based resistance, that unlike virtually all other examples which are strain-specific, is broad-spectrum. The University of Warwick and Syngenta filed a joint UK patent application on the 25th June 2010, on the novel mechanism of resistance. This was subsequently used as priority data for an international patent cooperation treaty (PCT) application published as WO2011/161466 on 29th December 2011. The University of Warwick assigned its interest in the intellectual property to Syngenta in November 2012. Syngenta are pursuing the application in Japan, the European community, Korea, Australia and the USA.

The study has led to the following conclusions:

1. *retr01* has been identified as *BraA.eIF(iso)4E.a*.

2. The broad-spectrum resistance to TuMV in *B. rapa* RLR22 is novel and based on the mis-splicing of *BraA.eIF(iso)4E.a*, caused by the presence of the indel (extra G, position 201 nucleotides (nt)).

3. Yeast two-hybrid (Y2H) interactions between the viral protein genome-linked (VPg) of TuMV and *eIF(iso)4E* from *B. rapa* seem to be TuMV isolate-specific.

4. Aphid transmission experiments to investigate the complementation of an eIF(iso)4E *Arabidopsis thaliana* knock-out line with *B. rapa BraA.eIF(iso)4E.a* showed that earlier results from mechanical inoculation of these plants were not an artefact. Hence, ectopic expression of *eIF4E* and *eIF(iso)4E* was confusing as it indicated TuMV could use both copies of *eIF4E/eIF(iso)4E* which is clearly not the case in *B. rapa*.

5. *BraA.eIF(iso)4E.c* is the only candidate for *ConTR01*.

6. In some crosses, resistance segregated as a single recessive gene (*retr01*), whereas in the cross with a plant of the related sub-species *B. rapa* spp. *trilocularis*, it segregated as the recessive gene and a dominant gene (*ConTR01*). It is important to be aware of the genetic background of plants which the resistant plant line is to be crossed with, as this can affect the segregation of resistance in subsequent generations.

8.2 Identifying *retr01* and *ConTR01*

One aim of the study was to identify the two resistance genes controlling the broad-spectrum resistance to TuMV in the RLR22 and R-o-18 cross. The involvement of eIF4E/eIF(iso)4E has been identified in the investigation of recessive resistances to plant viruses (Robaglia and Caranta, 2006). Whilst work is still needed to confirm the identity of *ConTR01* as *BraA.eIF(iso)4E.c*, it is clear that *ConTR01* is not the original favoured candidate, *BraA.eIF4E.c* (Rusholme *et al.*, 2007). The data on *retr01* shows very clearly that it is *BraA.eIF(iso)4E.a*.

8.2.1 Investigating the Interaction of TuMV VPg and *B. rapa* eIF4E/eIF(iso)4E

A further aim of the study was to investigate if the Y2H assay could be a useful tool for assessing TuMV VPg - *B. rapa* eIF4E/eIF(iso)4E interactions. In this study, no interaction was detected between the VPg of TuMV isolate CDN 1 and any of the eIF4E/eIF(iso)4E genes from the TuMV-susceptible line R-o-18. Colleagues in China managed to show an interaction between the VPg of

123

TuMV isolate C4 and *BraA.eIF(iso)4E.a* from the TuMV-susceptible line Ji Zao Chun and also the VPg of TuMV isolate C4 and the mis-spliced variants very similar to the correctly spliced version (variant lacking the last 3 nt of exon 1 and the variant retaining the last 15 nt of intron 1) (Nellist *et al.*, 2014). However, they were unable to detect an interaction between the VPg from TuMV isolate UK 1 and the *BraA.eIF(iso)4E.a* allele from the TuMV-susceptible *B. rapa* line, indicating that these reactions are TuMV strain-specific and hence not a reliable indicator of *B. rapa* - TuMV interactions.

The complementation of the Arabidopsis eIF(iso)4E knock-out with the *B. rapa* RLR22 allele of *BraA.eIF(iso)4E.a* was confusing. It was hypothesised that the susceptible phenotype observed in the plants transformed with the RLR22 *BraA.eIF(iso)4E.a* transgene could have been an artefact of overloading the plant with TuMV by mechanical inoculation. However, the plants challenged with aphids were also susceptible to TuMV, indicating that this was not the case. Although no correctly spliced versions of *BraA.eIF(iso)4E.a* were detected in the Arabidopsis plants transformed with RLR22 *BraA.eIF(iso)4E.a*, the discovery of variants that were potentially functional for TuMV, may explain the susceptibility phenotype observed, particularly as two of the variants (one retaining the last 15 nt of intron 1 and one lacking the last 3 nt of exon 1) have been shown to interact with the VPg of TuMV isolate C4 in Y2H experiments (W. Qian, Personal communication). The complementation of the Arabidopsis *eIF(iso)4E* knock-out line with *B. rapa BraA.eIF4E.a*, *BraA.eIF4E.c* and *BraA.eIF(iso)4E.c* from RLR22 was also misleading, as it indicated that TuMV could use all four copies of *eIF4E*/*eIF(iso)4E* from RLR22. Transcriptome data in this study showed that both copies of *eIF4E* and *BraA.eIF(iso)4E.c* were expressed in RLR22. It is clear from my results on plants segregating for copies of *eIF4E* and *eIF(iso)4E* and earlier results (Rusholme *et al.*, 2007) that TuMV can not use/access these genes to sustain a full systemic infection in the RLR22 line of *B. rapa*.

8.2.2   Inheritance of Resistance

In crosses between the susceptible parent in the original cross, R-o-18 (*B. rapa* ssp. *trilocularis*, a different sub-species to RLR22, *B. rapa* var. *pekinensis*, Chinese cabbage) and RLR22, the resistance was found to be controlled by two genes, *retr01* and *ConTR01* (Rusholme *et al.*, 2007). In the crosses between

RLR22 and the same sub-species (CK Chinese cabbage lines), the resistance was found to be controlled by a single gene (*retr01*), as plants already possessed *ConTR01*. The difference in inheritance patterns highlights the need to be aware of the genetic background of the parental plants intended for crossing with the resistant RLR22 line. There can be differences in segregation patterns in offspring depending upon which copies of eIF(iso)4E TuMV can use, or access.

TILLING has also resulted in the artificial induction of mis-splicing of *eIF4E* in tomato, providing resistance to *Pepper mottle virus* and strain-specific resistance to *Potato virus Y* (Piron *et al.*, 2010). EcoTILLING could also be used to screen natural populations for such polymorphisms (Comai *et al.*, 2004).

8.2.3   Introgression of *retr01*-based Resistance

Syngenta are currently introgressing the *retr01*-based resistance into six commercial inbred Chinese cabbage lines. The results from the cross between RLR22 and Syngenta CK lines showed a clear linkage between *retr01* and the TuMV-resistance in the F_2 families. The SNP marker developed by Syngenta based on the indel we discovered is within *retr01* and is enabling Syngenta to efficiently track *retr01* through the back-crossing generations thereby dramatically speeding up the breeding process. Pre-commercial trials are required before marketing, to evaluate the Chinese cabbage varieties. Syngenta are also currently attempting to move the resistance into the C genome of *Brassica oleracea*.

8.2.4   Durability

The inability of a wide range of virus isolates from around the world representing different genotypes, pathotypes and serotypes to overcome the resistance in RLR22 (Walsh *et al.*, 2002) indicates that the resistance mechanism may provide durable potyvirus resistance in a wide range of plant species. The generic applicability is dependent upon the existence of multiple copies of eIF4E/eIF(iso)4E being present and some tolerance of functional redundancy within the host. Arabidopsis is a good example of this, as the knock-out of *eIF(iso)4E*, which led to resistance to TuMV and *Lettuce mosaic virus* did not appear to have an adverse affects on the plant (Duprat *et al.*, 2002).

The Infidelity of the Mis-splicing

The discovery of potentially functional variants of *B. rapa* in the R-o-18 and RLR22 cross and the RLR22 and Syngenta CK crosses might explain why limited systemic spread of TuMV was detected in some plants. The existence and low frequency of these potentially functional variants might be the reason why, following mechanical inoculation of large amounts of TuMV, the occasional unexpected very limited systemic infection of plants homozygous for *retr01* were seen. Mechanical inoculation is an artefact and overloads the plant with huge amounts of virus that the plant would not normally experience in the field. When challenged with aphids, RLR22 did not show any symptoms in inoculated leaves, which it did when mechanically inoculated with TuMV (Rusholme *et al.*, 2007). So, it is possible that if the offspring of the R-o-18 and RLR22 cross and the RLR22 and Syngenta CK crosses were challenged by aphid transmission, we might not see limited spread of TuMV due to a much decreased viral load. The difficulty experienced in detecting correctly spliced variants of *BraA.eIF(iso)4E.a* and variants with minor amino acid changes in RLR22, suggests that there was not sufficient quantities of these variants to facilitate large-scale TuMV replication and hence completely defeat the resistance. There is the possibility that in the future, the infidelity of the mis-splicing may result in a higher quantity of functional variants being produced. However, all evidence to date suggests this is unlikely, as the potentially functional copies are very difficult to detect, indicating that are only present in very small quantities.

The Possibility of TuMV Mutating to Overcome *retr01*

In many plant-potyvirus pathosystems, the ability of the virus to overcome eIF4E-mediated resistance has been shown to be due to amino acid changes in the VPg (Kang *et al.*, 2005a; Ayme *et al.*, 2006; Charron *et al.*, 2008). It may be possible that TuMV could evolve to become capable of effective cap-independent translation (Basso *et al.*, 1994) and thereby overcome *retr01*-based resistance. Gallois *et al.* (2010) reported that two independent mutations in the VPg were sufficient to restore TuMV virulence in Arabidopsis plants with the *eIF(iso)4E* knock-out and the double knock-out of *eIF(iso)4G1* and *eIF(iso)4G2*. They suggested that the virulent TuMV variants may use an eIF4F-independent pathway. The durability of the resistance will also be dependant on TuMV not mutating to be able to utilise/access other copies of eIF4E/eIF(iso)4E *in planta*.

Ultimately the durability of the resistance remains to be seen, however, the data so far suggests it could be durable. The durability will also depend on selection pressures in the field. The discovery of three other brassica lines (*B. rapa rapa*, BR03023 and BR05060) with broad-spectrum resistance to TuMV that is not based on *retr01* is promising. These other resistances could be exploited by introgression into commercial *B. rapa* types and alternating plantings with lines possessing *retr01*, in order to reduce the selection pressure for *retr01* resistance-breaking viral mutants, thereby extending the durability of *retr01*-based resistance and these other resistance sources in the field.

8.3   Future Work

Suggested future work could encompass testing how the *retr01*-based resistance would perform in the field. Aphids could be used to transmit TuMV; this would also test whether any plants homozygous for *retr01* displayed limited systemic TuMV spread and would indicate how the resistance would behave in the field.

The discovery of three brassica lines (*B. rapa rapa*, BR03023 and BR05060) with broad-spectrum resistance that is not based on *retr01* is interesting and further investigation of these lines is required to identify the plant genes involved and to identify the resistance mechanism. Standard mapping procedures could be followed to identify these plant genes.

Further studies are required to investigate the involvement of *ConTR01/BraA.eIF(iso)4E.c* in TuMV resistance in the RLR22 and R-o-18 cross. One experiment could look at knocking out *BraA.eIF(iso)4E.a* from R-o-18 and producing a homozygous line and separately knocking out *BraA.eIF(iso)4E.c* from R-o-18 and producing a further homozygous line. These two knock-out lines could then be crossed to produce a double homozygous knock-out line ($\Delta BraA.eIF(iso)4E.a$ and $\Delta BraA.eIF(iso)4E.c$). This line could then be challenged with TuMV and if the plants were resistant, it would provide evidence for *BraA.eIF(iso)4E.c* being *ConTR01*. Another experiment could involve transiently/constitutively expressing R-o-18 *BraA.eIF(iso)4E.c* in RLR22 to see if it can induce susceptibility when challenged with TuMV.

To speed up the introgression of the *retr01*-based resistance into *B. oleracea*, work would need to be done to assess which copies of eIF4E/eIF(iso)4E

TuMV can use in *B. oleracea*. This could possibly be achieved by knocking them out, such as through the use of targeted TALEN (transcription activator-like effector nuclease) or CRISPR (clustered regularly interspaced short palindromic repeats) mutagenesis.

Another area of further research could be to investigate the TuMV isolate UK 1 resistance observed in three of the CK parental lines. It is not commercially relevant to Syngenta as the plants were still susceptible to TuMV isolate CDN 1, but it could be scientifically interesting to try and identify the resistance gene(s) involved.

Bibliography

Abdul-Razzak A., **Guiraud T.**, **Peypelut M.**, **Walter J.**, **Houvenaghel M.C.**, **Candresse T.**, **Gall O.L.** and **German-Retana S.** (2009) Involvement of the cylindrical inclusion (CI) protein in the overcoming of an eIF4E-mediated resistance against *Lettuce mosaic potyvirus*. *Molecular Plant Pathology*, **10**(1), 109–113.

Albar L., **Bangratz-Reyser M.**, **Hébrard E.**, **Ndjiondjop M.N.**, **Jones M.** and **Ghesquière A.** (2006) Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to *Rice yellow mottle virus*. *The Plant Journal*, **47**(3), 417–426.

Altschul S.F., **Madden T.L.**, **Schäffer A.A.**, **Zhang J.**, **Zhang Z.**, **Miller W.** and **Lipman D.J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**(17), 3389–3402.

Anandalakshmi R., **Pruss G.J.**, **Ge X.**, **Marathe R.**, **Mallory A.C.**, **Smith T.H.** and **Vance V.B.** (1998) A viral suppressor of gene silencing in plants. *Proceedings of the National Academy of Sciences*, **95**, 13,079–13,084.

Andrade M., **Abe Y.**, **Nakahara K.S.** and **Uyeda I.** (2009) The *cyv-2* resistance to *Clover yellow vein virus* in pea is controlled by the eukaryotic initiation factor 4E. *Journal of General Plant Pathology*, **75**(3), 241–249.

Ashby J.A., **Stevenson C.E.M.**, **Jarvis G.E.**, **Lawson D.M.** and **Maule A.J.** (2011) Structure-based mutational analysis of eIF4E in relation to *sbm1* resistance to *Pea Seed-Borne Mosaic Virus* in Pea. *PLoS ONE*, **6**(1), e15,873.

Atreya C.D., **Raccah B.** and **Pirone T.P.** (1990) A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. *Virology*, **178**, 161–165.

Ayme V., **Souche S.**, **Caranta C.**, **Jacquemond M.**, **Chadœuf J.**, **Palloix A.** and **Moury B.** (2006) Different mutations in the Genome-Linked Protein

VPg of Potato virus Y confer virulence on the *pvr2³* resistance in pepper. *Molecular Plant-Microbe Interactions*, **19**(5), 557–563.

Basso J., **Dallaire P.**, **Charest P.J.**, **Devantier Y.** and **Laliberté J.F.** (1994) Evidence for an internal ribosome entry site within the 5' non-translated region of turnip mosaic potyvirus RNA. *Journal of General Virology*, **75**, 3157–3165.

Beauchemin C., **Boutet N.** and **Laliberté J.F.** (2007) Visualization of the interaction between the precursors of VPg, the viral protein linked to the genome of turnip mosaic virus, and the translation eukaryotic initiation factor iso 4E in planta. *Journal of Virology*, **81**(2), 775–782.

Benson D.A., **Boguski M.S.**, **Lipman D.J.**, **Ostell J.**, **Ouellette B.F.F.**, **Rapp B.A.R.** and **Wheeler D.L.** (1999) GenBank. *Nucleic Acids Research*, **27**(1), 12–17.

Browning K.S. (1996) The plant translational apparatus. *Plant Molecular Biology*, **32**, 107–144.

Browning K.S., **Lax S.R.** and **Ravel J.M.** (1987) Identification of two messenger RNA cap binding proteins in Wheat Germ. *The Journal of Biological Chemistry*, **262**(23), 11,228–11,232.

Bruun-Rasmussen M., **Møller I.S.**, **Tulinius G.**, **Hansen J.K.R.**, **Lund O.S.** and **Johansen I.E.** (2007) The same allele of translation initiation factor 4E mediates resistance against two *Potyvirus* spp. in *Pisum sativum*. *Molecular Plant-Microbe Interactions*, **20**(9), 1075–1094.

Burland T.G. (1999) DNASTAR's Lasergene Sequence Analysis Software, *in:* **S. Misener** and **S.A. Krawetz** (Editors), Bioinformatics Methods and Protocols (Humana Press, New Jersey), pages 71–91.

Carrington J.C., **Cary S.M.**, **Parks D.T.** and **Dougherty W.G.** (1989) A second proteinase encoded by a plant potyvirus genome. *The EMBO Journal*, **8**(2), 365–370.

Carrington J.C., **Freed D.D.** and **Oh C.S.** (1990) Expression of potyviral polyproteins in transgenic plants reveals three proteolytic activities required for complete processing. *The EMBO Journal*, **9**(5), 1347–1353.

Carrington J.C., **Jensen P.E.** and **Schaad M.C.** (1998) Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. *The Plant Journal*, **14**(4), 393–400.

Carrington J.C., **Kasschau K.D.**, **Mahajan S.K.** and **Schaad M.C.** (1996) Cell-to-cell and long-distance transport of viruses in plants. *The Plant Cell*, **8**, 1669–1681.

Charron C., **Nicolaï M.**, **Gallois J.L.**, **Robaglia C.**, **Moury B.**, **Palloix A.** and **Caranta C.** (2008) Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. *The Plant Journal*, **54**, 56–68.

Chiang B.Y., **Chiang M.S.**, **Grant W.F.** and **Crete R.** (1980) Transfer of resistance to race 2 of *Plasmodiophora brassicae* from *Brassica napus* to cabbage (*B. oleracea* spp. *capitata*). IV. A resistant 18-chromosome B_1 plant and its B_2 progenies. *Euphytica*, **29**, 47–55.

Chung B.Y.W., **Miller W.A.**, **Atkins J.F.** and **Firth A.E.** (2008) An overlapping essential gene in the Potyviridae. *Proceedings of National Academy of Sciences USA*, **105**(15), 5897–5902.

Comai L., **Young K.**, **Till B.J.**, **Reynolds S.H.**, **Greene E.A.**, **Codomo C.A.**, **Enns L.C.**, **Johnson J.E.**, **Burtner C.**, **Odden A.R.** and **Henikoff S.** (2004) Effcient discovery of DNA polymorphisms in natural populations by Ecotilling. *The Plant Journal*, **37**, 778–786.

Diaz-Pendon J., **Truniger V.**, **Nieto C.**, **Garcia-Mas J.**, **Bendahmane A.** and **Aranda M.A.** (2004) Advances in understanding recessive resistance to plant viruses. *Molecular Plant Pathology*, **5**(3), 223–233.

Dinesh-Kumar S.P. and **Baker B.J.** (2000) Alternatively spliced *N* resistance gene transcripts: Their possible role in tobacco mosaic virus resistance. *Proceedings of the National Academy of Sciences*, **97**(4), 1908–1913.

Dinkova T.D., **Aguilar R.** and **de Jiménez E.S.** (2000) Expression of maize eukaryotic initiation factor (eIF) iso4E is regulated at the translational level. *Biochemical Journal*, **351**, 825–831.

Dixon G.R. (2007) Vegetable brassicas and related crucifers, Crop Production Science in Horticulture (CAB International, Oxfordshire, U.K.).

Dolja V.V., **Haldeman R.**, **Robertson N.L.**, **Dougherty W.G.** and **Carrington J.C.** (1994) Distinct functions of capsid protein in assembly of tobacco etch potyvirus in plants. *The EMBO Journal*, **13**(6), 1482–1491.

Dougherty W.G., **Parks D.**, **Cary S.M.**, **Bazan J.F.** and **Fletterick R.J.** (1989) Characterization of the catalytic residues of the Tobacco etch virus 49-kDa proteinase. *Virology*, **172**, 302–310.

Duprat A., **Caranta C.**, **Revers F.**, **Menand B.**, **Browning K.S.** and **Robaglia C.** (2002) The *Arabidopsis* eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *The Plant Journal*, **32**, 927–934.

Edgar R.C. (2004a) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5), 1792–1797.

Edgar R.C. (2004b) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**(113).

Edwards A.M., **Kus B.**, **Jansen R.**, **Greenbaum D.**, **Greenblatt J.** and **Gerstein M.** (2002) Bridging structural biology genomics: assessing protein interaction data with known complexes. *Trends in Genetics*, **18**(10), 529–536.

Edwardson J.R. and **Christie R.G.** (1986) Turnip mosaic virus, *in:* Viruses Infecting Forage Legumes Volume 2 (Agricultural Experiment Stations, Institute of Food and Agricultural Sciences, University of Florida, Gainesville), pages 438–453.

Fields S. and **Song O.K.** (1989) A novel genetic system to detect protein protein interactions. *Nature*, **340**, 245–246.

Filbin M.E. and **Kieft J.S.** (2009) Toward a structural understanding of IRES RNA function. *Current Opinion in Structural Biology*, **19**(3), 267–276.

Fraser R.S.S. (1986) Genes for resistance to plant viruses. *Critical Reviews in Plant Sciences*, **3**(3), 257–294.

Gallie D.R. and **Browning K.S.** (2001) eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-indepedent translation, and translation of structured mRNAs. *Journal of Biological Chemistry*, **276**(40), 36,951–36,960.

Gallois J.L., **Charron C.**, **Sánchez F.**, **Pagny G.**, **Houvenaghel M.C.**, **Moretti A.**, **Ponz F.**, **Revers F.**, **Caranta C.** and **German-Retana S.** (2010) Single amino acid changes in the turnip mosaic virus viral genome-linked protein (VPg) confer virulence towards *Arabidopsis thaliana* mutants knocked

out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. *Journal of General Virology*, **91**(1), 288–293.

Gao Z., **Johansen E.**, **Eyers S.**, **Thomas C.L.**, **Ellis T.H.N.** and **Maule A.J.** (2004) The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *The Plant Journal*, **40**, 376–385.

German-Retana S., **Walter J.**, **Doublet B.**, **Roudet-Tavert G.**, **Nicaise V.**, **Lecampion C.**, **Houvenaghel M.C.**, **Robaglia C.**, **Michon T.** and **Le Gall O.** (2008) Mutational Analysis of Plant Cap-Binding Protein eIF4E Reveals Key Amino Acids Involved in Biochemical Functions and Potyvirus Infection. *Journal of Virology*, **82**(15), 7601–7612.

Gómez P., **Rodríguez-Hernández A.M.**, **Moury B.** and **Aranda M.A.** (2009) Genetic resistance for the sustainable control of plant virus diseases: breeding, mechanisms and durability. *European Journal of Plant Pathology*, **125**(1), 1–22.

Goodfellow I.G. and **Roberts L.O.** (2008) Eukaryotic initiation factor 4E. *The International Journal of Biochemistry and Cell Biology*, **40**, 2675–2680.

Govier D.A. and **Kassanis B.** (1974) A virus-induced component of plant sap needed when aphids acquire Potato Virus Y from purified preparations. *Virology*, **61**, 420–426.

Green S.K. and **Deng T.C.** (1985) Turnip mosaic virus strains in cruciferous hosts in Taiwan. *Plant Disease*, **69**(1), 28–31.

Harberd D.J. (1969) A simple effective embryo culture technique for brassica. *Euphytica*, **18**, 425–429.

Hjulsager C.K., **Olsen B.S.**, **Jensen D.M.K.**, **Cordea M.I.**, **Krath B.N.**, **Johansen I.E.** and **Lund O.S.** (2006) Multiple determinants in the coding region of *Pea seed-borne mosaic virus* P3 are involved in virulence against *sbm-2* resistance. *Virology*, **355**(1), 52–61.

Hong Y. and **Hunt A.G.** (1996) RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. *Virology*, **226**, 146–151.

Hughes S.L., **Green S.K.**, **Lydiate D.J.** and **Walsh J.A.** (2002) Resistance to *Turnip mosaic virus* in *Brassica rapa* and *B. napus* and the analysis of genetic inheritance in selected lines. *Plant Pathology*, **51**, 567–573.

Hughes S.L., **Hunter P.J.**, **Sharpe A.G.**, **Kearsey M.J.**, **Lydiate D.J.** and **Walsh J.A.** (2003) Genetic mapping of the novel *Turnip mosaic virus* resistance gene *TuRB03* in *Brassica napus*. *Theoretical and Applied Genetics*, **107**(7), 1169–1173.

Hwang J., **Li J.**, **Liu W.Y.**, **An S.J.**, **Cho H.**, **Her N.H.**, **Yeam I.**, **Kim D.** and **Kang B.C.** (2009) Double mutations in eIF4E and eIFiso4E confer recessive resistance to *Chilli veinal mottle virus* in pepper. *Molecules and Cells*, **27**(3), 329–336.

Jenner C.E., **Keane G.J.**, **Jones J.E.** and **Walsh J.A.** (1999) Serotypic variation in turnip mosaic virus. *Plant Pathology*, **48**, 101–108.

Jenner C.E., **Nellist C.F.**, **Barker G.C.** and **Walsh J.A.** (2010) *Turnip mosaic virus* (TuMV) is able to use alleles of both *eIF4E* and *eIF(iso)4E* from multiple loci of the diploid *Brassica rapa*. *Molecular Plant-Microbe Interactions*, **23**(11), 1198–1505.

Jenner C.E., **Sánchez F.**, **Nettleship S.B.**, **Foster G.D.**, **Ponz F.** and **Walsh J.A.** (2000) The cylindrical inclusion gene of *Turnip mosaic virus* encodes a pathogenic determinant to the *Brassica* resistance gene *TuRB01*. *Molecular Plant-Microbe Interactions*, **13**(10), 1102–1108.

Jenner C.E., **Tomimura K.**, **Ohshima K.**, **Hughes S.L.** and **Walsh J.A.** (2002a) Mutations in *Turnip mosaic virus* P3 and Cylindrical Inclusion Proteins Are Separately Required to Overcome Two *Brassica napus* Resistance Genes. *Virology*, **300**(1), 50–59.

Jenner C.E. and **Walsh J.A.** (1996) Pathotypic variation in turnip mosaic virus with special reference to Eutopian isolates. *Plant Pathology*, **45**, 848–856.

Jenner C.E., **Wang X.**, **Ponz F.** and **Walsh J.A.** (2002b) A fitness cost for *Turnip mosaic Virus* to overcome host resistance. *Virus Research*, **86**, 1–6.

Jenner C.E., **Wang X.**, **Tomimura K.**, **Ohshima K.**, **Ponz F.** and **Walsh J.A.** (2003) The dual role of the potyvirus P3 protein of turnip mosaic virus as a symptom and avirulence determinant in brassica. *Molecular Plant-Microbe Interactions*, **16**(9), 777–784.

Jiang J. and **Laliberté J.F.** (2011) The genome-linked protein VPg of plant viruses – a protein with many partners. *Current Opinion in Virology*, **1**(5), 347–354.

Kang B.C., **Yeam I.**, **Frantz J.D.**, **Murphy J.F.** and **Jahn M.M.** (2005a) The *pvr1* locus in *Capsicum* encodes a translation initiation factor eIF4E that interacts with *Tobacco etch virus* VPg. *The Plant Journal*, **42**(3), 392–405.

Kang B.C., **Yeam I.** and **Jahn M.M.** (2005b) Genetics of plant virus resistance. *Annual Review of Phytopathology*, **43**, 581–621.

Kanyuka K., **Druka A.**, **Caldwell D.G.**, **Tymon A.**, **McCallum N.**, **Waugh R.** and **Adams M.J.** (2005) Evidence that the recessive bymovirus resistance locus *rym4* in barley corresponds to the eukaryotic translation initiation factor 4E gene. *Molecular Plant Pathology*, **6**(4), 449–458.

Kassem A.A.H. and **Walsh J.A.** (2008) Characterising resistance to turnip mosaic virus (TuMV) in turnip (*Brassica rapa*). *Arab Journal of Plant Protection*, **26**(2), 168–172.

Katis N.I., **Tsitsipis J.A.**, **Stevens M.** and **Powell G.** (2007) 14 Transmission of Plant Viruses, *in:* **H.F. Van Emden** and **R. Harrington** (Editors), Aphids as Crop Pests (CAB International, Oxfordshire, U.K.), pages 353–390.

Keller K.E., **Johansen E.**, **Martin R.R.** and **Hampton R.O.** (1998) Potyvirus Genome-Linked Protein (VPg) Determines Pea Seed-Borne Mosaic Virus Pathotype-Specific Virulence in *Pisum sativum*. *Molecular Plant-Microbe Interactions*, **11**(2), 124–130.

Keren H., **Lev-Maor G.** and **Ast G.** (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nature Reviews Genetics*, **11**(5), 345–355.

Kim J., **Kang W.H.**, **Yang H.B.**, **Park S.**, **Jang C.s.**, **Yu H.J.** and **Kang B.C.** (2013) Identification of a broad-spectrum recessive gene in *Brassica rapa* and molecular analysis of the eIF4E gene family to develop molecular markers. *Molecular Breeding*.

King A.M.Q., **Adams M.J.**, **Lefkowitz E.J.** and **Carstens E.B.** (2012) Virus Taxonomy, Classification and Nomenclature of Viruses : Ninth Report of the International Committee on Taxonomy of Viruses (Elsevier, London, U.K.).

Klausner A. (1987) Immunoassays flourish in new markets. *Nature Biotechnology*, **5**(6), 551–556.

Laín S., **Riechmann J.L.** and **García J.A.** (1990) RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Research*, **18**(23), 7003–7006.

Landry B.S., **Hubert N.**, **Crete R.**, **Chiang M.S.**, **Lincoln S.E.** and **Etoh T.** (1992) A genetic map for *Brassica oleracea* based on RFLP markers detected with expresses DNA sequences and mapping of resistance genes to race 2 of *Plasmodiophora brassicae* (Woronin). *Genome*, **35**, 409–420.

Langmead B., **Trapnell C.**, **Pop M.** and **Salzberg S.L.** (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**(3), R25.

Le Gall O., **Aranda M.A.** and **Caranta C.** (2011) Plant Resistance to Viruses Mediated by Translation Initiation Factors, *in:* **C. Caranta**, **M.A. Aranda**, **M. Tepfer** and **J.J. López-Moya** (Editors), Recent Advances in Plant Virology (Caister Academic Press, Norfolk, UK), pages 177–194.

Lecoq H., **Moury B.**, **Desbiez C.**, **Palloix A.** and **Pitrat M.** (2004) Durable virus resistance in plants through conventional approaches: a challenge. *Virus Research*, **100**(1), 31–39.

Lellis A.D., **Kasschau K.D.**, **Whitham S.A.** and **Carrington J.C.** (2002) Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus infection. *Current Biology*, **12**, 1046–1051.

Léonard S., **Plante D.**, **Wittmann S.**, **Daigneault N.**, **Fortin M.G.** and **Laliberté J.F.** (2000) Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. *Journal of Virology*, **74**(17), 7730–7737.

Léonard S., **Viel C.**, **Beauchemin C.**, **Daigneault N.**, **Fortin M.G.** and **Laliberté J.F.** (2004) Interaction of VPg-Pro of *Turnip mosaic virus* with the translation initiation factor 4E and the poly(A)-binding protein *in planta*. *Journal of General Virology*, **85**, 1055–1063.

Liu X.P., **Lu W.C.**, **Liu B.X.**, **Li S.Y.**, **Li J.L.**, **Zhao Z.Y.**, **Wang H.J.** and **CH W.** (1990a) A study on TuMV strain differentiation on cruciferous vegetables from ten regions of China: identification results with Green's methods. *Virologica Sinica*, **1**, 82–87.

Liu X.P., **Lu W.C.**, **Liu Y.K.** and **Li J.L.** (1990b) A study on TuMV strain differentiation of crucierfous vegetables from ten provinces in China: new host differentiator screening and strain classification. *Chinese Science Bulletin*, **35**, 1734–1739.

Liu X.P., **Lu W.C.**, **Lui Y.K.**, **Wei S.Q.**, **Xu J.B.**, **Liu Z.R.**, **Zhang H.J.**, **Li J.L.**, **Ke G.L.**, **Yao W.Y.**, **Cai Y.S.**, **Wu F.Y.**, **Cao S.C.**, **Li Y.H.**, **Xie S.D.**, **Lin B.X.** and **Zhang C.L.** (1996) Occurrence and strain differentiation of Turnip mosaic potyvirus and sources of resistance in Chinese cabbage in China. *Acta Horticulturae*, **407**, 431–440.

López-Moya J.J., **Valli A.** and **García J.A.** (2009) Potyviridae, *in:* Encyclopedia of Life Sciences (ELS) (John Wiley & Sons, Ltd, Chichester, U.K.) (doi:10.1002/9780470015902.a0000755.pub2).

Mahajan S.K., **Dolja V.V.** and **Carrington J.C.** (1996) Roles of the sequence encoding tobacco etch virus capsid protein in genome amplification: requirements for the translation process and a *cis*-active element. *Journal of Virology*, **70**(7), 4370–4379.

Marco C.F., **Aguilar J.M.**, **Abad J.**, **Gómez-Guillamón M.L.** and **Aranda M.A.** (2003) Melon resistance to *Cucurbit yellow stunting disorder virus* is characterized by reduced virus accumulation. *Phytopathology*, **93**, 844–852.

McNaughton I.H. (1995) Turnip and relatives, *in:* **J. Smartt** and **N.W. Simmonds** (Editors), Evolution of Crop Plants (Longman Group UK Limited, Essex, UK), pages 62–68.

Monzingo A., **Dhaliwal S.**, **Dutt-Chaudhuri A.**, **Lyon A.**, **Sadow J.**, **Hoffman D.**, **Robertus J.** and **Browning K.S.** (2007) The Structure of Eukaryotic Translation Initiation Factor-4E from Wheat Reveals a Novel Disulfide Bond. *Plant Physiology*, **143**(4), 1504–1518.

Naderpour M., **Lund O.S.**, **Larsen R.** and **Johansen E.** (2010) Potyviral resistance derived from cultivars of *Phaseolus vulgarise* carrying *bc-3* is associated with the homozygotic presence of a mutated *eIF4E* allele. *Molecular Plant Pathology*, **11**(2), 255–263.

Nellist C.F., **Qian W.**, **Jenner C.E.**, **Moore J.D.**, **Zhang S.**, **Wang X.**, **Briggs W.H.**, **Barker G.C.**, **Sun R.** and **Walsh J.A.** (2014) Multiple copies

of eukaryotic translation initiation factors in *Brassica rapa* facilitate redundancy, enabling diversification through variation in splicing and broad-spectrum virus resistance. *The Plant Journal*, **77**(2), 261–268.

Ner-Gaon H., **Halachmi R.**, **Savaldi-Goldstein S.**, **Rubin E.**, **Ophir R.** and **Fluhr R.** (2004) Intron retention is a major phenomenon in alternative splicing in *Arabidopsis*. *The Plant Journal*, **39**(6), 877–885.

Nicaise V., **Gallois J.L.**, **Chafiai F.**, **Allen L.M.**, **Schurdi-Levraud V.**, **Browning K.S.**, **Candresse T.**, **Caranta C.**, **Gall O.L.** and **German-Retana S.** (2007) Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in *Arabidopsis thaliana*. *FEBS Letters*, **581**, 1041–1046.

Nicaise V., **German-Retana S.**, **Sanjuàn R.**, **Dubrana M.P.**, **Mazier M.**, **Maisonneuve B.**, **Candresse T.**, **Caranta C.** and **Le Gall O.** (2003) The Eukaryotic Translation Initiation Factor 4E Controls Lettuce Susceptibility to the Potyvirus Lettuce mosaic virus. *Plant Physiology*, **132**(3), 1272–1282.

Nieto C., **Morales M.**, **Orjeda G.**, **Clepet C.**, **Monfort A.**, **Sturbois B.**, **Puigdomènech P.**, **Pitrat M.**, **Caboche M.**, **Dogimont C.**, **Garcia-Mas J.**, **Aranda M.A.** and **Bendahmane A.** (2006) An *eIF4E* allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. *The Plant Journal*, **48**(3), 452–462.

Nilsen T.W. and **Graveley B.R.** (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**(7280), 457–463.

Ohshima K., **Tanaka M.** and **Sako N.** (1996) The complete nucleotide sequence of turnip mosaic virus RNA japanese strain. *Archives of Virology*, **141**, 1991–1997.

Park J.Y., **Koo D.H.**, **Hong C.P.**, **Lee S.J.**, **Jeon J.W.**, **Lee S.H.**, **Yun P.Y.**, **Park B.S.**, **Kim H.R.**, **Bang J.W.**, **Plaha P.**, **Bancroft I.** and **Lim Y.P.** (2005) Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kbp gene-rich region of *Arabidopsis* chromosome 4 and partially duplicated on chromosome 5. *Molecular Genetics and Genomics*, **274**(6), 579–588.

Piron F., **Nicolaï M.**, **Minoïa S.**, **Piednoir E.**, **Moretti A.**, **Salgues A.**, **Zamir D.**, **Caranta C.** and **Bendahmane A.** (2010) An induced mutation in Tomato eIF4E leads to immunity to two Potyviruses. *PloS one*, **5**(6), 1–10.

Pirone T.P. and **Blanc S.** (1996) Helper-dependent vector transmission of plant viruses. *Annual Review of Phytopathology*, **34**, 227–247.

Pirone T.P. and **Thornbury D.W.** (1984) The involvement of a helper component in nonpersistent transmission of plant viruses by aphids. *Microbiological sciences*, **1**(8), 191–193.

Plante D., **Viel C.**, **Léonard S.**, **Tampo H.**, **Laliberté J.F.** and **Fortin M.G.** (2004) Turnip mosaic virus VPg does not disrupt the translation initiation complex but interferes with cap binding. *Physiological and Molecular Plant Pathology*, **64**, 219–226.

Powell-Abel P., **Nelson R.S.**, **De B.**, **Hoffmann N.** and **Rogers S.G.** (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*, **232**(4751), 738–743.

Provvidenti R. (1980) Evaluation of Chinese cabbage cultivars from Japan and the People's Republic of China for resistance to turnip mosaic virus and cauliflower mosaic virus. *Journal of the American Society for Horticultural Science*, **105**, 571–573.

Provvidenti R. and **Hampton R.O.** (1992) Sources of resistance to viruses in the Potyviridae. *Archives of virology, Supplementum*, **5**, 189–211.

Qian W., **Zhang S.**, **Zhang S.**, **Li F.**, **Zhang H.**, **Wu J.**, **Wang X.**, **Walsh J.A.**, **Sun R.** and **Sun R.** (2013) Mapping and candidate-gene screening of the novel Turnip mosaic virus resistance gene *retr02* in Chinese cabbage (*Brassica rapa* L.). *Theoretical and Applied Genetics*, **126**(1), 179–188.

Rajamäki M.L. and **Valkonen J.P.T.** (1999) The 6K2 protein and the VPg of Potato virus A are determinants of systemic infection in *Nicandra physaloides*. *Molecular Plant-Microbe Interactions*, **12**(12), 1074–1081.

Rhoads R.E. (2009) eIF4E: New family members, new binding partners, new roles. *Journal of Biological Chemistry*, **284**(25), 16,711–16,715.

Riechmann J.L., **Laín S.** and **García J.A.** (1992) Highlights and prospects of potyvirus molecular biology. *Journal of General Virology*, **73**, 1–16.

Ritzenthaler C. (2005) Resistance to plant viruses: old issue, news answers? *Current Opinion in Biotechnology*, **16**(2), 118–122.

Robaglia C. and **Caranta C.** (2006) Translation initiation factors: a weak link in plant RNA virus infection. *TRENDS in Plant Science*, **11**(1), 1–6.

Roberts I.M., **Wang D.**, **Findlay K.** and **Maule A.J.** (1998) Ultrastructural and temporal observations of the Potyvirus Cylindrical Inclusions (CIs) show that the CI protein acts transiently in aiding virus movement. *Virology*, **245**, 173–181.

Rodriguez C.M., **Freire M.A.**, **Camilleri C.** and **Robaglia C.** (1998) The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. *The Plant Journal*, **13**(4), 465–473.

Rodríguez-Cerezo E., **Ammar E.D.**, **Pirone T.P.** and **Shaw J.G.** (1993) Association of the non-structural P3 viral protein with cylindrical inclusions in potyvirus-infected cells. *Journal of General Virology*, **74**, 1945–1949.

Rogozin I.B., **Carmel L.**, **Csuros M.** and **Koonin E.V.** (2012) Origin and evolution of spliceosomal introns. *Biology Direct*, **7**, 11.

Rojas M.R., **Zerbini F.M.**, **Allison R.F.**, **Gilbertson R.L.** and **Lucas W.J.** (1997) Capsid protein and helper component-proteinase function as potyvirus cell-to-cell movement. *Virology*, **237**, 283–295.

Roudet-Tavert G., **Michon T.**, **Walter J.**, **Delaunay T.**, **Redondo E.** and **Le Gall O.** (2007) Central domain of a potyvirus VPg is involved in the interaction with the host translation initiation factor eIF4E and the viral protein HcPro. *Journal of General Virology*, **88**(3), 1029–1033.

Ruffel S., **Dussault M.H.**, **Palloix A.**, **Moury B.**, **Bendahmane A.**, **Robaglia C.** and **Caranta C.** (2002) A natural recessive resistance gene against potato virus y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). *The Plant Journal*, **32**, 1067–1075.

Ruffel S., **Gallois J.L.**, **Lesage M.L.** and **Caranta C.** (2005) The recessive potyvirus resistance gene *pot-1* is the tomato orthologue of the pepper *pvr2-eIF4E* gene. *Molecular Genetics and Genomics*, **274**(4), 346–353.

Ruffel S., **Gallois J.L.**, **Moury B.**, **Robaglia C.**, **Palloix A.** and **Caranta C.** (2006) Simultaneous mutations in translation initiation factors eIF4e and eIF(iso)4E are required to prevent pepper veinal mottle virus infection in pepper. *Journal of General Virology*, **87**, 2089–2098.

Rusholme R.L. (2000) The genetic control of resistance to *Turnip mosaic virus* (TuMV) in *Brassica*. Ph.D. thesis, University of East Anglia, Norwich.

Rusholme R.L., **Higgins E.E.**, **Walsh J.A.** and **Lydiate D.J.** (2007) Genetic control of broad-spectrum resistance to turnip mosaic virus in *Brassica rapa* (Chinese cabbage). *Journal of General Virology*, **88**, 3177–3186.

Ruud K.A., **Kuhlow C.**, **Goss D.J.** and **Browning K.S.** (1998) Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. *The Journal of Biological Chemistry*, **273**(17), 10,325–10,330.

Sako N. (1981) Virus diseases of Chinese cabbage in Japan, *in:* **N.S. Talekar** and **T.D. Griggs** (Editors), Chinese Cabbage (The Asian Vegetable Research and Development Centre, Taiwan), pages 129–142.

Sato M., **Nakahara K.S.**, **Yoshii M.**, **Ishikawa M.** and **Uyeda I.** (2005) Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of *Arabidopsis thaliana* by potyviruses. *FEBS Letters*, **579**(5), 1167–1171.

Schaad M.C., **Anderberg R.J.** and **Carrington J.C.** (2000) Strain-Specific Interaction of the Tobacco Etch Virus NIa Protein with the Translation Initiation Factor eIF4E in the Yeast Two-Hybrid System. *Virology*, **273**(2), 300–306.

Schaad M.C., **Lellis A.D.** and **Carrington J.C.** (1997) VPg of Tobacco etch potyvirus is a host genotype-specific determinant for long-distance movement. *Journal of Virology*, **71**(11), 8624–8631.

Shattuck V.I. (1992) The biology, epidemiology, and control of Turnip mosaic virus. *Plant Breeding Reviews*, **14**, 199–238.

Shukla D.D. and **Ward C.W.** (1989) Structure of potyvirus coat proteins and its application in the taxonomy of potyvirus group. *Advances in Virus Research*, **36**, 273–314.

Shukla D.D., **Ward C.W.** and **Brunt A.A.** (1994) The potyviridae (CAB International, Oxon, U.K.).

Stellberger T., **Häuser R.**, **Baiker A.**, **Pothineni V.R.**, **Haas J.** and **Uetz P.** (2010) Improving the yeast two-hybrid system with permutated fusions proteins: the Varicella Zoster Virus interactome. *Proteome Science*, **8**, 8.

Stobbs L.W. and **Shattuck V.I.** (1989) Turnip mosaic virus strains in Southern Ontario, Canada. *Plant Disease*, **73**(33), 208–212.

Stoger E., **Williams S.**, **Christou P.**, **Down R.E.** and **Gatehouse J.A.** (1999) Expression of the insecticidal lectin from snowdrop (*Galanthus nivalis* agglutinin; GNA) in transgenic wheat plants: effects on predation by the grain aphid *Sitobion avenae*. *Molecular Breeding*, **5**, 65–73.

Suh S.K., **Green S.K.** and **Park H.G.** (1995) Genetics of resistance to five strains of turnip mosaic virus in chinese cabbage. *Euphytica*, **81**, 71–77.

Sylvester E.S. (1954) Aphid transmission of nonpersistent plant viruses with special reference to the *Brassica nigra* virus. *Hilgardia*, **23**, 53–98.

Tamura K., **Peterson D.**, **Peterson N.**, **Stecher G.**, **Nei M.** and **Kumar S.** (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, **28**(10), 2731–2739.

Thivierge K., **Cotton S.**, **Dufresne P.J.**, **Mathieu I.**, **Beauchemin C.**, **Ide C.**, **Fortin M.G.** and **Laliberté J.F.** (2008) Eukaryotic elongation factor 1A interacts with turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. *Virology*, **377**, 216–225.

Thompson J.D., **Higgins D.G.** and **Gibson T.J.** (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**(22), 4673–4680.

Thorvaldsdóttir H., **Robinson J.T.** and **Mesirov J.P.** (2012) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, **14**(2), 178–192.

Tomimura K., **Gibbs A.J.**, **Jenner C.E.**, **Walsh J.A.** and **Ohshima K.** (2003) The phylogeny of turnip mosaic virus; comparisons of 38 genomic sequences reveal a eurasian origin and a recent 'emergence' in east asia. *Molecular Ecology*, **12**, 2099–2111.

Tomlinson J.A. (1987) Epidemiology and control of virus diseases of vegetables. *Annals of Applied Biology*, **110**(3), 661–681.

Trapnell C., **Pachter L.** and **Salzberg S.L.** (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**(9), 1105–1111.

Trapnell C., **Williams B.A.**, **Pertea G.**, **Mortazavi A.**, **Kwan G.**, **van Baren M.J.**, **Salzberg S.L.**, **Wold B.J.** and **Pachter L.** (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **28**(5), 511–515.

Truniger V. and **Aranda M.A.** (2009) Recessive resistance to plant viruses. *Advances in Virus Research*, **75**, 119–159.

Truniger V., **Nieto C.**, **González-Ibeas D.** and **Aranda M.** (2008) Mechanism of plant eIF4E-mediated resistance against a Carmovirus (*Tombusviridae*): cap-independent translation of a viral RNA controlled *incis* by an (a)virulence determinant. *The Plant Journal*, **56**(5), 716–727.

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

Verchot J. and **Carrington J.C.** (1995) Evidence that the potyvirus P1 proteinase functions *in trans* as an accessory factor for genome amplification. *Journal of Virology*, **69**(6), 3668–3674.

Vijayapalani P., **Maeshima M.**, **Nagasaki-Takekuchi N.** and **Miller W.A.** (2012) Interaction of the Trans-Frame Potyvirus Protein P3N-PIPO with Host Protein PCaP1 Facilitates Potyvirus Movement. *PLoS Pathogens*, **8**(4), e1002,639.

Wall P.K., **Leebens-Mack J.**, **Chanderbali A.S.**, **Barakat A.**, **Wolcott E.**, **Liang H.**, **Landherr L.**, **Tomsho L.P.**, **Hu Y.**, **Carlson J.E.**, **Ma H.**, **Schuster S.C.**, **Soltis D.E.**, **Soltis P.S.**, **Altman N.** and **dePamphilis C.W.** (2009) Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics*, **10**(1), 347.

Walsh J.A. (1989) Genetic control of immunity to turnip mosaic virus in winter oilseed rape (*Brassica napus* ssp. *oleifera*) and the effect of foreign isolates of the virus. *Annals of Applied Biology*, **115**(1), 89–99.

Walsh J.A. and **Jenner C.E.** (2002) *Turnip mosaic virus* and the quest for durable resistance. *Molecular Plant Pathology*, **3**(5), 289–300.

Walsh J.A., **Rusholme R.L.**, **Hughes S.L.**, **Jenner C.E.**, **Bambridge J.M.**, **Lydiate D.J.** and **Green S.K.** (2002) Different classes of resistance to turnip mosaic virus in *Brassica rapa*. *European Journal of Plant Pathology*, **108**, 15–20.

Walsh J.A., **Sharpe A.G.**, **Jenner C.E.** and **Lydiate D.J.** (1999) Characterisation of resistance to turnip mosaic virus in oilseed rape (*Brassica napus*) and genetic mapping of *TuRB01*. *Theorectical and Applied Genetics*, **99**, 1149–1154.

Walsh J.A. and **Tomlinson J.A.** (1985) Viruses infecting winter oilseed rape (*Brassica napus* ssp. *oleifera*). *Annals of Applied Biology*, **107**(3), 485–495.

Wang A. and **Krishnaswamy S.** (2012) Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Molecular Plant Pathology*, 1.

Wang B.B. and **Brendel V.** (2006) Genomewide comparative analysis of alternative splicing in plants. *Proceedings of the National Academy of Sciences*, **103**(18), 7175–7180.

Wang C.Y., **Liang Y.J.**, **Lin Y.S.**, **Shih H.M.**, **Jou Y.S.** and **Yu W.C.Y.** (2004) YY1AP, A Novel Co-activator of YY1. *Journal of Biological Chemistry*, **279**(17), 17,750–17,755.

Wang X., **Wang H.**, **Wang J.**, **Sun R.**, **Wu J.**, **Liu S.**, **Bai Y.**, **Mun J.H.**, **Bancroft I.**, **Cheng F.**, **Huang S.**, **Li X.**, **Hua W.**, **Wang J.**, **Wang X.**, **Freeling M.**, **Pires J.C.**, **Paterson A.H.**, **Chalhoub B.**, **Wang B.**, **Hayward A.**, **Sharpe A.G.**, **Park B.S.**, **Weisshaar B.**, **Liu B.**, **Li B.**, **Liu B.**, **Tong C.**, **Song C.**, **Duran C.**, **Peng C.**, **Geng C.**, **Koh C.**, **Lin C.**, **Edwards D.**, **Mu D.**, **Shen D.**, **Soumpourou E.**, **Li F.**, **Fraser F.**, **Conant G.**, **Lassalle G.**, **King G.J.**, **Bonnema G.**, **Tang H.**, **Wang H.**, **Belcram H.**, **Zhou H.**, **Hirakawa H.**, **Abe H.**, **Guo H.**, **Wang H.**, **Jin H.**, **Parkin I.A.P.**, **Batley J.**, **Kim J.S.**, **Just J.**, **Li J.**, **Xu J.**, **Deng J.**, **Kim J.A.**, **Li J.**, **Yu J.**, **Meng J.**, **Wang J.**, **Min J.**, **Poulain J.**, **Wang J.**, **Hatakeyama K.**, **Wu K.**, **Wang L.**, **Fang L.**, **Trick M.**, **Links M.G.**, **Zhao M.**, **Jin M.**, **Ramchiary N.**, **Drou N.**, **Berkman P.J.**, **Cai Q.**, **Huang Q.**, **Li R.**, **Tabata S.**, **Cheng S.**, **Zhang S.**, **Zhang S.**, **Huang S.**, **Sato S.**, **Sun S.**, **Kwon S.J.**, **Choi S.R.**, **Lee T.H.**, **Fan W.**, **Zhao X.**, **Tan X.**, **Xu X.**, **Wang Y.**, **Qiu Y.**, **Yin Y.**, **Li Y.**, **Du Y.**, **Liao Y.**, **Lim Y.**, **Narusaka Y.**, **Wang Y.**, **Wang Z.**, **Li Z.**, **Wang Z.**, **Xiong Z.** and **Zhang Z.** (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, **43**(10), 1035–1039.

Ward C.W. and **Shukla D.D.** (1991) Taxonomy of potyviruses: current problems and some solutions. *Intervirology*, **32**, 269–296.

Warwick S.I., **Francis A.** and **Al-Shehbaz I.A.** (2006) Brassicaceae: Species checklist and database on CD-Rom. *Plant Systematics and Evolution*, **259**(2-4), 249–258.

Wei T., **Zhang C.**, **Hong J.**, **Xiong R.**, **Kasschau K.D.**, **Zhou X.**, **Carrington J.C.** and **Wang A.** (2010) Formation of Complexes at Plasmodesmata for Potyvirus Intercellular Movement Is Mediated by the Viral Protein P3N-PIPO. *PLoS Pathogens*, **6**(6), e1000,962.

Whitham S.A., **Dinesh-Kumar S.P.**, **Choi D.**, **Hehi R.**, **Corr C.** and **Baker B.J.** (1994) The Product of the Tobacco Mosaic Virus Resistance Gene N: Similarity to Toll and the Interleukin-1 Receptor. *Cell*, **78**, 1101–1115.

Wittmann S., **Henriette C.**, **Fortin M.G.** and **Laliberté J.F.** (1997) Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of *Arabidopsis thaliana* using the yeast two-hybrid system. *Virology*, **234**, 84–92.

Yoon J.Y., **Green S.K.** and **Opeña R.T.** (1993) Inheritance of resistance to turnip mosaic virus in Chinese cabbage. *Euphytica*, **69**, 103–108.

Publications

The attached pages of this thesis present the following publications in which I have featured as an author during the course of the PhD.

Jenner C.E., Nellist C.F., Barker G.C. and **Walsh J.A.** (2010) *Turnip mosaic virus* (TuMV) is able to use alleles of both *eIF4E* and *eIF(iso)4E* from multiple loci of the diploid *Brassica rapa*. *Molecular Plant-Microbe Interactions*, **23**(11), 1198-1505.

Nellist C.F., **Qian W., Jenner C.E., Moore J.D., Zhang S., Wang X., Briggs W.H., Barker G.C., Sun R.** and **Walsh J.A.** (2014) Multiple copies of eukaryotic translation initiation factors in *Brassica rapa* facilitate redundancy, enabling diversification through variation in splicing and broad-spectrum virus resistance. *The Plant Journal*, **77**(2), 261-268.

e-**X**tra*

Turnip mosaic virus (TuMV) Is Able to Use Alleles of Both *eIF4E* and *eIF(iso)4E* from Multiple Loci of the Diploid *Brassica rapa*

Carol E. Jenner, Charlotte F. Nellist, Guy C. Barker, and John A. Walsh

Warwick HRI, University of Warwick, Warwick, CV35 9EF, U.K.

Three copies of *eIF4E* and three copies of *eIF(iso)4E* have been identified and sequenced from a *Turnip mosaic virus* (TuMV)-susceptible, inbred, diploid *Brassica rapa* line, R-o-18. One of the copies of *eIF4E* lacked exons 2 and 3 and appeared to be a pseudogene. The two other copies of *eIF4E* and two of the three copies of *eIF(iso)4E* were isolated from a bacterial artificial chromosome library of R-o-18. Using an *Arabidopsis* line (Col-0::*dSpm*) with a transposon knock-out of the *eIF(iso)4E* gene which resulted in a change from complete susceptibility to complete resistance to TuMV, complementation experiments were carried out with the two versions of *eIF4E* and the two versions of *eIF(iso)4E*. When transformed into Col-0::*dSpm*, all four *Brassica* transgenes complemented the *Arabidopsis eIF-(iso)4E* knock-out, conferring susceptibility to both mechanical and aphid challenge with TuMV. One of the copies of *eIF4E* did not appear to support viral replication as successfully as the other copy of *eIF4E* or the two copies of *eIF(iso)4E*. The results show that TuMV can use both *eIF4E* and *eIF(iso)4E* from *B. rapa* for replication and, for the first time, that a virus can use *eIF4E* and *eIF(iso)4E* from multiple loci of a single host plant.

A key step in plant mRNA translation involves the binding of the 5′ cap to the 40S ribosomal subunit. This is achieved partly through the interaction of the cap with the eukaryotic initiation complex eIF4F, itself a combination of the cap-binding subunit eIF4E and the larger subunit eIF4G (Browning 1996). Isoforms of each component, eIF(iso)4E and eIF(iso)4G, also exist in plants (Allen et al. 1992). It is becoming clear that most eukaryotic organisms encode and express multiple eIF4E family members, some for general translation and others for specific functions, including control of translation (Rhoads 2009).

Arabidopsis thaliana possesses three genes encoding eIF4E (At4g18040, At1g29550, and At1g29590), one for eIF(iso)4E (At5g35620) and a further gene encoding a similar protein named novel cap-binding protein (nCBP) (At5g18110) (Ruud et al. 1998). The eIF4E and eIF(iso)4E proteins have distinctive functions. Whereas eIF4E is present in all tissues except the root specialization zones, eIF(iso)4E is particularly abundant in floral tissues and young tissue, and the proteins are

thought to have differing roles in plant metabolism and development (Rodriguez et al. 1998). Each is under different regulatory control pathways (Dinkova et al. 2000).

The first evidence that plant viruses interacted directly with the eukaryotic translation initiation complex of plants was provided when it was shown that the virus-encoded genome-linked protein (VPg) of *Turnip mosaic virus* (TuMV; genus *Potyvirus*), which binds to the 5′ end of the viral RNA, bound to *A. thaliana* eIF(iso)4E in yeast two-hybrid binding assays (Wittmann et al. 1997). A recent comparison of TuMV VPg interactions with *A. thaliana* eIF(iso)4E, nCBP, and all three eIF4E proteins in the yeast two-hybrid system revealed that only eIF(iso)4E acted as a functional partner (Gallois et al. 2010). Furthermore, in an enzyme-linked immunosorbent assay (ELISA)-based system, VPg interacted with tagged *A. thaliana* eIF(iso)4E and, to a lesser extent, with tagged *A. thaliana* eIF4E (Léonard et al. 2000).

In contrast to the in vitro results, it is the *eIF(iso)4E* gene alone that is required for successful infection of *A. thaliana* by TuMV (Sato et al. 2005). Insertional mutagenesis of *At.eIF-(iso)4E* using a defective maize transposon tagged with an herbicide resistance gene (*dSpm*) produced a plant line with no detectable mRNA or protein from the gene. This line is able to grow normally and is resistant to potyvirus infection, in particular, TuMV and *Lettuce mosaic virus* (Duprat et al. 2002). Additionally, an EMS-induced mutation of *A. thaliana* named *lsp1* also conferred loss of susceptibility to TuMV and other members of the *Potyviridae*; the mutation was found to be in *eIF(iso)4E* (Lellis et al. 2002). In contrast, TuMV is able to infect *A. thaliana* lines with T-DNA insertions in *eIF4E* (Sato et al. 2005), *eIF4G, eIF(iso)4G1*, or *eIF(iso)4G2* (Nicaise et al. 2007).

Since these original discoveries, there have been many reports of recessive resistance to *Potyvirus* spp. and members of certain other virus groups (*Cucumovirus, Carmovirus*, and *Bymovirus*) resulting from mutations in *eIF4E, eIF(iso)4E*, or the other components of the eukaryotic translation initiation complex, *eIF4G* or *eIF(iso)4G* (Robaglia and Caranta 2006). For example, the alleles *pvr1/2* (*Capsicum* spp.) confer resistance to *Potato virus Y* (Kang et al. 2005), *rym4/5/6* (barley) to *Barley yellow mosaic virus* (Kanyuka et al. 2005), *mo1* (lettuce) to *Lettuce mosaic virus* (Nicaise et al. 2003), *sbm1/wlv/cyv-2* (pea) to various viruses (Andrade et al. 2009; Bruun-Rasmussen et al. 2007; Gao et al. 2004), and *bc-3* (bean) to *Bean common mosaic virus* (Naderpour et al. 2010), and all correspond to *eIF4E*. Pea *sbm2*, acting against *Pea seed-borne mosaic virus*, is linked to *eIF(iso)4E* (Gao et al. 2004), *Capsicum pvr6* is a null allele of pepper *eIF(iso)4E* (Ruffel et al. 2002) and *retr01* in *Brassica rapa* has been suggested to be *eIF(iso)4E*

*The *e*-**X**tra logo stands for "electronic extra" and indicates that a supplementary table is published online and that Figure 3 appears in color online.

(Rusholme et al. 2007). Mutations in *eIF4G* and *eIF(iso)4G* have been identified as the resistance genes *cum1* and *rymv1*, respectively (Albar et al. 2006).

Here, we describe the identification and cloning of multiple copies of *eIF4E* and *eIF(iso)4E* from *B. rapa*. Using a *dSpm* transposon knock-out line of *Arabidopsis* (Duprat et al. 2002) we have been able to show, by complementation, that TuMV can utilize at least two copies of both *eIF4E* and *eIF(iso)4E*. This may help to explain why recessive resistance to TuMV in *Brassica* spp. is less common than recessive resistance to *Potyvirus* spp. in some other plant species.

RESULTS

Identification of *B. rapa eIF4E* and *eIF(iso)4E* genes.

The *B. rapa* line R-o-18 is highly inbred and uniformly susceptible to TuMV. Three copies of *eIF4E* and three copies of *eIF(iso)4E* were identified by polymerase chain reaction (PCR) on this line using primers designed to *Brassica* Genome Sequence Survey (GSS) and expressed sequence tag (EST) sequences. The PCR products were then used to probe a bacterial artificial chromosome (BAC) library of this line (Rana et al. 2004) to obtain the full-length genomic sequences. The structures of the genes were inferred from homology to *A. thaliana* genes and to *B. napus* EST sequences. Each gene consisted of five exons and four introns, with the exception of *BraA.eIF4E.b*, which lacked exons 2 and 3 and appeared to be a pseudogene (Fig. 1). Most of the variation between the *eIF4E* genes was due to size differences of intron 1 whereas the *eIF(iso)4E* genes differed markedly in the length of intron 3. Mapping has shown that *BraA.eIF4E.a*, *BraA.eIF4E.c*, *BraA.eIF(iso)4E.a*, and *BraA.eIF(iso)4E.c* are at different loci (Rusholme et al. 2007) (data not shown).

The sequences of the predicted proteins are shown in Figure 2. BraA.eIF(iso)4E.a and BraA.eIF(iso)4E.c were 91.0% identical to each other and 90.8 and 86.7% identical to At.eIF(iso)4E, respectively. BraA.eIF4e.a and BraA.eIF4E.c had 87.8% identity to each other but only 51.8 and 52.8% identity, respectively, to At.eIF(iso)4E.

B. rapa eIF4E or *eIF(iso)4E* genes complement an *eIF(iso)4E* knock-out in *A. thaliana* plants for TuMV infection.

In order to determine which of the *B. rapa* genes could be used by TuMV during the infection process, copies of the genes were transformed into a Sainsbury Laboratory *Arabidopsis* transposant line (SLAT) possessing a nonfunctional *At.eIF(iso)4E* gene (Col-0::*dSpm*). Because TuMV is unable to infect this line (Duprat et al. 2002), any transformed lines challenged with TuMV that were susceptible would indicate functional complementation by the *Brassica* transgene. *A. thaliana* lines were transformed with *BraA.eIF4E.a*, *BraA.eIF4E.c*, *BraA.eIF(iso)4E.a*, and *BraA.eIF(iso)4E.c*. *BraA.eIF4E.b* was a pseudogene and *BraA.eIF(iso)4E.b* was not found in the BAC library; therefore, they were not included.

Col-0 plants developed severe symptoms and were stunted within 2 weeks of challenge with TuMV CDN 1, whereas Col-0::*dSpm* plants remained symptomless, with no virus detected throughout the period of the test (Fig. 3). Infected plants were prone to secondary fungal infections and frequently rotted within 5 weeks of virus challenge. All mock- and virus-inoculated Col-0::*dSpm* plants remained symptomless and free of virus as determined by ELISA at 3 weeks postinoculation.

PCR tests on two to six plants per transformation confirmed the presence of wild-type *At.eIF(iso)4E* in Col-0 lines and the continued presence of the *dSpm* insertion in *At.eIF(iso)4E* in
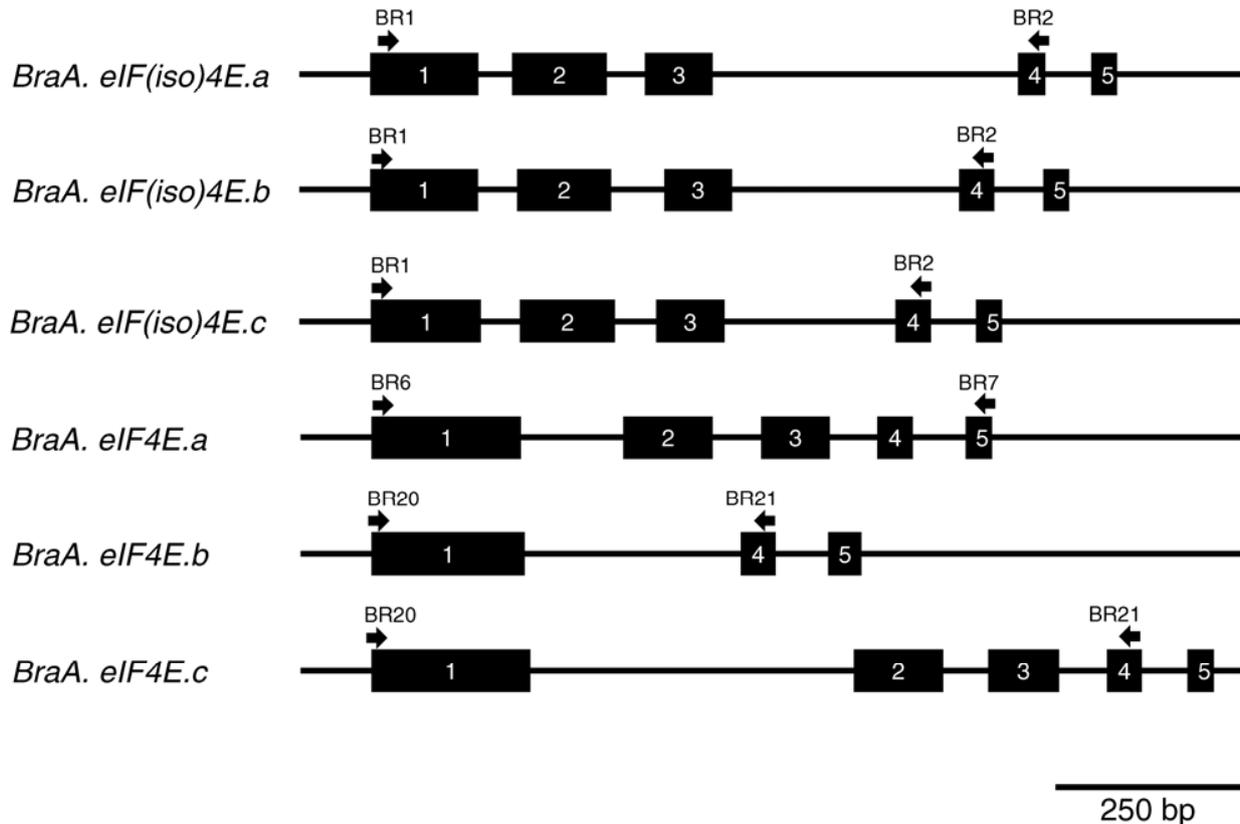

Fig. 1. Structure of *Brassica rapa eIF4E* and *eIF(iso)4E* genes indicating positions of exons, introns, and generic primers used for probe production. Exons are indicated by black boxes and introns are thin lines. Sequences were determined outside the generic primer range using BACs or genome-walking.

all Col-0::*dSpm* plants tested. The T_2 plants segregated for the presence of the transgenes, and the presence of the selection marker (*bar* or *nptII*) was associated with the presence of the appropriate *Brassica* transgene. The absence of the selection marker corresponded with the absence of the *Brassica* transgene. Sequence checks confirmed there were no alterations in the transgenes.

Following TuMV challenge, plant symptoms were assessed visually and virus titers were estimated by ELISA followed by transformation for residual maximum likelihood (REML) analyses. Plants with symptoms were found to have significantly higher virus titers (ELISA optical density [OD] values) than plants with no symptoms ($P < 0.001$) (Table 1). Col-0 T_2 plants with *Brassica* transgenes appeared identical to the parental Col-

A

```
                    10        20        30        40        50        60        70        80
At.eIF(iso)4E     MATDDVNEPLPAAAELPATEAEKQLLPHKLERKWSFWFDNQSLLKKGAAWGASLRKAYTFDTVEDFWGLHETIFQTSKLT  76
BraA.eIF(iso)4E.a MATEDVNEALAAAEVPATETTEKQLPAHKLERKWSFWFDNQSKPKCGAAWGASLRKAYTFDTVCDFWGLHETIFIPSKLT  79
BraA.eIF(iso)4E.c MATEDVNEALAAAEVTAIESTEKQQPPHKLERKWSLWFDNQSKPKCGAAWGVSLRKACTFDTVEDFWGLHETIFVPSRLT  80

                    90        100       110       120       130       140       150       160
At.eIF(iso)4E     ANAEIHLFKAGVEPKWEDPECANGGKWTWVVTANRKEALDKGWLETLMALIGEQFDEADEICGVVASVRPQSKQDKLSLW  156
BraA.eIF(iso)4E.a PNAEIHMFKAGVEPKWEDPECANGGKWTYVVTSNRKPALDKAWLETLMALVGEQFDEADEICGVVASVRPLLKQDKLSLW  157
BraA.eIF(iso)4E.c PNADIHMFKAGVEPKWEDPECANGGKWTYVVTNNRKQALDKAWLETLMALVGEQFDEADEICGVVASVRQLLKQDKLSLW  158

                    170       180       190       200
At.eIF(iso)4E     TRTKSNEAVLMGIGKKWKEILDVTDKITFNNHDDSRRSRFTV                                         198
BraA.eIF(iso)4E.a TRTKSNEAVLMGIGKKWKEILDVTDKITFTNHDDSRRTRFTV                                         199
BraA.eIF(iso)4E.c TRTKSNEAVLMGIGKKWKEILDVTDKISFTNHDDARRSRFSV                                         200
```

B

```
                    10        20        30        40        50        60        70        80
BraA.eIF4E.a      MAVEDTLKPNVATEESNPNSADHPIDRYHEEGDDAEEGLLLLLATVDESSKSAVPESHPLEHSWTLWFDNPSVKSKQTT   74
BraA.eIF4E.c      MAVEDTSKPVVVAEEANPNPTDHPIDRYHEEGDDAEEGEIAGGEGDGDESSKSAVPQSHPLEHSWTFWFDNPSVKLKQAT  80

                    90        100       110       120       130       140       150       160
BraA.eIF4E.a      WGSSLRSVFTFSTVEEFWYNNIRHPSKLANGADLYCFKHNIEPKWEDPICANGGKWTMNFSREKSDKPFLYTLLALIG    154
BraA.eIF4E.c      WGSSLRSVFTFSTVEEFWSLFNNMRGPSKLAGGADFYCKHNIEPKWEDPICANGGKWTMNFPKEKSDKPWLYTLLALIG   160

                    170       180       190       200       210       220       230
BraA.eIF4E.a      EQFDHGDEICGVVVNVRAKQERISIWTKNSSNEAAQVSIGRQWKEFLDYNSSIGFIIHEDAKKLDRGAKSAYTA        228
BraA.eIF4E.c      EQFDHGDEICGAVVNVRGKQERISIWTKNASNEAAQVSIGKQWKEFIDYNNSIGFIIHEDAKKLDRGAKSAYTA        234
```

Fig. 2. Comparison of the *Arabidopsis thaliana* eIF(iso)4E protein sequence with predicted proteins from *Brassica rapa* line R-o-18 *eIF(iso)4E* and *eIF4E* genes. **A,** eIF(iso)4E proteins. Differences from the protein sequence At.eIF(iso)4E are indicated by shading. **B,** eIF4E proteins. Differences from the protein sequence BraA.eIF4E.c are indicated by shading.

Fig. 3. Complementation of *Arabidopsis thaliana* plants with *Brassica rapa* eukaryotic translation initiation factor genes, following challenge by mechanical inoculation with *Turnip mosaic virus* 20 days postinoculation. **A,** Col-0 mock inoculated; **B,** Col-0; **C,** Col-0::*dSpm*; **D,** Col-0 with transgenes; **E,** Col-0::*dSpm* with transgenes; **i,** *BraA.eIF4E.a*; **ii,** *BraA.eIF4E.c*; **iii,** *BraA.eIF(iso)4E.a*; and **iv,** *BraA.eIF(iso)4E.c*. The *dSpm* transposon insertion is in At5g35620, inactivating the gene *At.eIF(iso)4E.*

0 both visually (Fig. 3) and by ELISA (Table 1), indicating no effect of the transgene in this background on virus replication. Col-0::dSpm T$_2$ plants lacking *Brassica* transgenes remained uninfected, like the parental Col-0::dSpm line. T$_2$ Col-0::dSpm plants with *Brassica* transgenes produced symptoms as severe as the Col-0 line, with the exception of Col-0::dSpm transformed with *BraA.eIF4E.a*, where plants had notably milder symptoms. There were significant ($P < 0.001$) differences in the amount of virus detected between lines with symptoms (Table 1, occasion no. 1), although limited numbers of plants were analyzed (both as individual families and by pooling families per transgene). Only Col-0::dSpm families transformed with *BraA.eIF4E.a* had a significantly ($P < 0.05$) lower virus titer than infected Col-0 plants. When larger numbers of plants from one family per transgene were tested (Table 1, occasion no. 2), results confirmed that the mild symptoms seen on Col-0::dSpm plants with the *BraA.eIF4E.a* transgene related to a significantly ($P < 0.05$) lower virus titer than the other infected lines.

Plants challenged with TuMV using aphids were less efficiently infected (all potentially susceptible controls were not infected), and produced milder symptoms than those challenged mechanically. The virus titer as measured by ELISA was also significantly reduced ($P < 0.001$) (Table 2). The lower levels of symptom severity and virus titer achieved by aphid transmission are likely to reflect a lack of TuMV acquisition by aphids or a smaller viral inoculum load.

The symptoms were remarkable in the consistency of visual severity between plants within any one T$_2$ family (derived from a T$_1$ plant with a *Brassica* transgene); plants appeared either stunted to equal degrees or totally healthy. There was also consistency in visual symptom severity between plants of different families containing the same *Brassica* transgene. If the T$_1$ plants had possessed a single copy of the transgene, the T$_2$ families would be expected to segregate 3:1 for the presence of the transgene, with a quarter of plants having two copies of the transgene, half having one copy (heterozygotes), and one quarter having no copies. The consistency of infection levels (visual or ELISA) suggested that homozygous and heterozygous transgenic plants could not be distinguished on the basis of degree of virus susceptibility.

DISCUSSION

It has been shown that very few *Potyvirus* spp. use eIF4E in one host and eIF(iso)4E in another host for protein translation; for example, *Tobacco etch virus* (TEV) uses eIF4E in pepper (Kang et al. 2005) and tomato (Ruffel et al. 2005) but eIF(iso)4E in *A. thaliana* (Lellis et al. 2002). Also, it is very rare that a *Potyvirus* sp. uses both eIF4E and eIF(iso)4E in the same host; for example, *Pepper veinal mottle virus* (PVMV) uses both in pepper (Ruffel et al. 2006). Because TuMV infection of *A. thaliana* requires functional eIF(iso)4E (Duprat et al. 2002; Lellis et al. 2002), we asked whether the same preference would be shown for genes from *B. rapa*, an economically important host. Additionally, because *B. rapa* possesses multiple loci of *eIF4E* and *eIF(iso)4E*, we investigated the ability of TuMV to use the genes at the different loci. A requirement for a particular gene would be an excellent target for the basis of recessive resistance, either natural or engineered via gene silencing (Zhang et al. 2006). Alternatively, it would be impor-

Table 1. Phenotypes of transgenic *Arabidopsis thaliana* plants possessing *Brassica rapa* eukaryotic translation initiation factor genes following mechanical inoculation with *Turnip mosaic virus* (visual symptoms and enzyme-linked immunosorbent assay [ELISA])[a]

| | | No. of plants with symptoms[b] | | | | Mean ELISA optical density (no. of plants)[c] | | | |
| | | Col-0 | | Col-0::dSpm | | Col-0 | | Col-0::dSpm | |
Brassica transgene	T$_2$ family	0	SM	0	SM	0	SM	0	SM
Occasion no. 1									
None	…	0	12	12	0	–	1.77 (3)	–0.03 (12)	–
BraA.eIF4E.a	i	0	12	4	8	–	1.58 (3)	0.00 (2)	0.33 (2)
	ii	0	12	1	11	–	1.65 (3)	0.01 (1)	0.28 (3)
	iii	0	12	4	8	–	1.74 (3)	0.00 (2)	0.40 (3)
	iv	0	12	7	5	–	1.73 (3)	0.02 (2)	0.28 (2)
	Total	0	48	16	32	–	1.68 (12)	0.01*** (7)	0.32* (10)
BraA.eIF4E.c	i	0	12	5	7	–	1.64 (3)	0.02 (1)	1.13 (3)
	ii	0	12	2	10	–	1.57 (3)	0.03 (1)	1.21 (3)
	iii	0	12	2	10	–	1.30 (3)	0.00 (1)	0.75 (3)
	iv	0	12	1	11	–	1.70 (3)	0.03 (1)	0.62 (3)
	Total	0	48	10	38	–	1.55 (12)	0.02*** (4)	0.93 (12)
BraA.eIF(iso)4E.a	i	0	12	2	10	–	1.54 (3)	0.04 (1)	1.53 (3)
	ii	0	12	0	12	–	1.60 (3)	–	0.53 (4)
	iii	0	12	3	9	–	1.31 (3)	0.01 (1)	1.01(3)
	iv	0	12	3	9	–	1.40 (3)	0.01 (1)	0.81 (3)
	Total	0	48	8	40	–	1.50 (12)	0.02*** (3)	0.94 (13)
BraA.eIF(iso)4E.c	i	0	12	7	5	–	1.70 (3)	0.01 (2)	0.25 (2)
	ii	0	12	5	7	–	1.60 (3)	0.02 (2)	1.15 (2)
	iii	–	–	0	12	–	–	–	0.58 (4)
	iv	–	–	1	11	–	–	0.02 (1)	0.92 (3)
	Total	0	24	13	25	1.65 (6)	0.02*** (5)	0.77 (11)
Occasion no. 2									
None	…	0	12	12	0	–	1.14 (12)	0.00*** (12)	–
BraA.eIF4E.a	i	0	12	6	6	–	1.18 (12)	–0.01*** (6)	0.17* (6)
BraA.eIF4E.c	i	0	12	3	9	–	1.17 (12)	–0.01*** (3)	1.23 (9)
BraA.eIF(iso)4E.a	i	0	12	0	12	–	1.20 (12)	–	1.21 (12)
BraA.eIF(iso)4E.c	i	0	12	11	1	–	1.26 (12)	–0.03*** (11)	1.41 (1)

[a] *dSpm* transposon insertion is in At5g35620, inactivating the gene *At.eIF(iso)4E*. Symptoms observed: 0 = no symptoms, SM = systemic mosaic with stunting and leaf deformation, – = not applicable.
[b] Twelve plants inoculated per T$_2$ family. Families have not been selected for transgene homozygosity and are segregating for the presence of the transgene.
[c] ELISA optical density measured at 405 nm after 2 h. Statistically significant differences (by residual maximum likelihood) compared with Col-0 plants with no transgene are indicated by * and *** ($P < 0.05$ and 0.001, respectively).

tant to know whether TuMV is able to use different copies of *eIF4E* or *eIF(iso)4E* or whether it has a requirement for more than one copy, in order to deploy such resistance in *B. rapa*.

Our results are unusual in that TuMV exhibits the ability to use four different translation initiation factor proteins to counteract the absence of At.eIF(iso)4E in the knock-out *A. thaliana* line. This is the first example of a virus being able to use different copies of *eIF4E* and *eIF(iso)4E* from multiple loci of the same plant. The use of simultaneous knock-outs of At.eIF(iso)4G1 and At.eIF(iso)4G2 has been reported to prevent TuMV infection of *A. thaliana*, whereas plants with only one of the two genes inactivated are susceptible (Nicaise et al. 2007). However, the ability to use four isoforms (two each of BraA.eIF4E and BraA.eIF(iso)4E) from one host species is remarkable. A slightly similar situation occurs for *Chilli veinal mottle virus* (ChiVMV), whose VPg interacts (in yeast two-hybrid studies) with both *Capsicum* eIF4E and eIF(iso)4E (Hwang et al. 2009). Mutations or gene silencing of both these genes together (also known as *pvr1/2* and *pvr6*) confers resistance to ChiVMV and PVMV (Ruffel et al. 2006), although it is unclear whether *Capsicum annuum* possesses other copies of either gene.

Of the four translation initiation factor genes, TuMV replication was least supported by *BraA.eIF4E.a* when expressed in *A. thaliana*. Further data would be needed to understand this; for example, whether the underlying cause is due to different levels of transcription or translation (and whether this is also the case in *B. rapa*) or due to protein structure. However, comparison of the protein sequences with the resolved structure of wheat eIF4E (Monzingo et al. 2007) revealed that the four *Brassica* proteins possess all the conserved residues noted as important in cap-binding and for stabilizing the structure of the protein, with the notable exception of a key tryptophan residue lacking in BraA.eIF4E.a (residue 145, replaced by phenylalanine).

It is likely that the viral protein interacting with the *B. rapa* eIF4E/eIF(iso)4E proteins is the VPg. A single amino acid substitution in TuMV VPg that can abolish the interaction with *Arabidopsis* eIF(iso)4E in yeast two-hybrid experiments simultaneously renders the virus unable to infect *B. perviridis* (a synonym for *B. rapa* var. *perviridis*) (Léonard et al. 2000). TuMV VPg has been reported to interact with one version of eIF(iso)4E from *B. rapa* in the yeast two-hybrid system (Hwang et al. 2009), although the authors did not describe different versions of the gene. Co-purification studies from *B. perviridis* have confirmed that TuMV VPg (as a precursor protein with or without the adjacent 6K protein and Pro protease sections) does interact in planta with eIF(iso)4E/eIF4E (Léonard et al. 2004). Infection by TuMV alters the profile of isomers found; Léonard and associates (2004) reported finding eIF(iso)4E in mock-inoculated and TuMV-infected plants but detectable levels of eIF4E in infected plants only. The two proteins appear to have different subcellular locations in *Nico-*

tiana benthamiana, eIF(iso)4E being primarily associated with rough endoplasmic reticulum and eIF4E with free ribosomes (Beauchemin et al. 2007). Because different forms of VPg are found in different cellular locations, Beauchemin and associates (2007) suggested that, in *N. benthamiana* at least, the 6K-VPg-Pro is the most likely form to be involved in viral translation (with eIF(iso)4E) and the VPg-Pro form may be involved in other cellular perturbations. TuMV VPg can bind to, but does not disrupt, the eIF(iso)4E-eIF(iso)4G complex, and it is believed to compete with capped cellular mRNAs for the translation machinery (Miyoshi et al. 2006; Plante et al. 2004). However, it should be noted that other potyviral proteins have associations with eIF4E, such as NIa interacting with eIF4E in yeast two-hybrid tests (Schaad et al. 2000) or CI being involved in eIF4E-based resistance breaking (Abdul-Razzak et al. 2009).

Kang and associates (2007) found that transforming tomato with *eIF4E* from the *pvr1* locus in pepper resulted in dominant resistance to several TEV strains and other *Potyvirus* spp. In our experiments, wild-type Col-0 *A. thaliana* plants transformed with *eIF4E*, or *eIF(iso)4E* from *B. rapa*, were as susceptible to TuMV as untransformed wild-type plants.

Of the known plant virus resistances where heritability has been characterized, recessive resistances are more common for *Potyvirus* spp. than for viruses of other families; 64% of the examples listed by Díaz-Pendón and associates (2004) corresponded to *Potyvirus* spp. In their review of sources of resistance to *Potyvirus* spp., Provvidenti and Hampton (1992) showed that, where heritability had been determined, 60 genes were dominant and 39 were recessive. They listed two examples of recessive resistance, compared with three examples of dominant resistance in *Brassica* spp. A review of resistances to TuMV listed four dominant and three recessive resistances (Shattuck 1992). Despite a large amount of screening over a prolonged period, we have found that recessive resistances to TuMV are relatively scarce in *Brassica* spp.; we have mapped six dominant resistance genes and one recessive gene (Walsh and Jenner 2002) and it took a screen of more than 3,000 lines of the diploid *B. rapa* (Liu et al. 1996) to find the latter. This recessive resistance was effective against a range of TuMV isolates from different parts of the world and representing different serotypes, pathotypes, and genetic groups (Walsh et al. 2002) and, thus far, has not been broken. Two loci controlling the resistance were mapped; the first gene, *retr01*, was recessive and the second, *ConTR01*, was dominant (Rusholme et al. 2007). We also showed that the A genome of *B. napus* (derived from *B. rapa*) appears to have three *eIF4E* loci and three *eIF(iso)4E* loci and that *retr01* appeared to be coincident with one of the *eIF(iso)4E* loci and *ConTR01* appeared to be coincident with one of the loci of *eIF4E*, or possibly one of the other loci of *eIF(iso)4E* (Rusholme et al. 2007). A possible explanation for the relative scarcity of recessive resistances to TuMV in *Brassica* spp. is

Table 2. Comparison of mechanical and aphid inoculation on degree of infection by *Turnip mosaic virus* of transgenic *Arabidopsis thaliana* plants possessing *Brassica rapa* eukaryotic translation initiation factor *eIF(iso)4Ea*[a]

		No. of plants with symptoms[b]				Mean ELISA optical density[c]			
		Col-0		Col-0::dSpm		Col-0		Col-0::dSpm	
Brassica transgene	Inoculation method	0	SM	0	SM	0	SM	0	SM
None	Mechanical	0	5	5	0	–	2.67	0.15***	–
	Aphid	2	3	5	0	0.05***	0.52***	0.04***	–
BraA.eIF(iso)4E.a	Mechanical	0	12	4	8	–	2.72	0.11***	2.58
	Aphid	6	6	7	5	–	0.48***	0.04***	0.41***

[a] *dSpm* transposon insertion is in At5g35620, inactivating the gene *At.eIF(iso)4E*. Symptoms observed: 0 = no symptoms, SM = systemic mosaic with stunting and leaf deformation, and – = not applicable.

[b] Twelve plants inoculated per T_2 family. Families have not been selected for transgene homozygosity and are segregating for the presence of the transgene.

[c] Enzyme-linked immunosorbent assay (ELISA) optical density measured at 405 nm after 1.5 h, all plants were tested. Statistically significant differences (by residual maximum likelihood) compared with Col-0 mechanically-inoculated plants are indicated by *** ($P < 0.001$).

that, for a plant to be resistant, all copies of *eIF4E* and *eIF(iso)4E* would have to be nonfunctional for TuMV. In amphidiploid *Brassica* spp. (*B. napus*, *B. juncea*, and *B. carinata*), recessive resistance based on *eIF4E* or *eIF(iso)4E* is potentially less likely, because there appear to be six loci of each gene present (three of each in the A genome and three of each in the C genome) (Rusholme et al. 2007).

MATERIALS AND METHODS

TuMV.

TuMV CDN 1 is a pathotype 4 isolate (Walsh 1989) able to overcome several dominant *Brassica* resistance genes. It was maintained in *B. juncea* cv. Tendergreen by mechanical inoculation as previously described (Walsh 1989). Aphid transmission was performed using *Myzus persicae* cultures raised on *B. napus* cv. Mikado. Aphids were starved for approximately 2 h, then allowed to feed on infected *B. juncea* leaves for a few minutes before they were transferred onto test plants (five per plant) for 2 days and, finally, killed using Aphox (Syngenta, Cambridge) insecticide.

Plant lines.

B. rapa subsp. *trilocularis* (yellow sarson) line R-o-18 is an inbred line highly susceptible to TuMV (Hughes et al. 2002). A BAC library of genomic DNA of R-o-18 (JBr, using vector pBIBAC2, average insert size 128 kb) was made by I. Bancroft, John Innes Centre (Norwich, U.K.) (Rana et al. 2004).

A. thaliana Col-0 is highly susceptible to TuMV infection. Col-0::*dSpm* is a SLAT line that is homozygous for a transposon (*dSpm*) insertion in exon 2 of At5g35620 (*At.eIF(iso)4E*) (Duprat et al. 2002). The transposon carries resistance to the herbicide glufosinate ammonium. The presence and location of the transposon was verified using the primer pairs dspm1-4E2K01 and dspm2-4E2KO4 as described previously (Duprat et al. 2002). Col-0::*dSpm* is resistant to infection by TuMV CDN 1.

Identification of full-length *eIF* gene sequences from *B. rapa*.

Genomic DNA was extracted from young R-o-18 leaves using a DNeasy plant mini kit (Qiagen, Crawley, U.K.) followed by amplification using the GenomiPhi system (GE Healthcare, Little Chalfont, U.K.), both processes performed according to the manufacturers' recommendations.

Generic primers were designed using GSS and EST sequences with homology to known *eIF4E* and *eIF(iso)4E* genes to encompass the majority of the open reading frame (ORF) (Table 3). The sequences used for *eIF4E* were from *B. oleracea* (BH540903, BZ068573, and BZ449347), *B. rapa* (CX270207,

CX270456, and DY010188), and *B. napus* (CB686150, CD834916, CN733851, CX188254, CX189337, CX194455, CX195737, and CX280257); for *eIF(iso)4E*, sequences were from *B. oleracea* (BZ436642, BZ514129, EH413959, EH414327, EH425132, EH426021, and EH426034), *B. rapa* (CX266566 and CX268848), and *B. napus* (D. Lydiate and E. Higgins, *personal communication*).

PCR was performed on genomic DNA using standard protocols with *Taq* DNA polymerase (Invitrogen, Paisley, U.K.). Primer pair BR1-BR2 produced multiple products of differing sizes from which could be separated *BraA.eIF(iso)4E.a* (1,215 bp), *BraA.eIF(iso)4E.b* (1,123 bp), and *BraA.eIF(iso)4E.c* (995 bp). Primer pairs BR6-BR7 and BR20-BR21 amplified most of the coding region of *BraA.eIF4E.a* (1,158 bp) and *BraA.eIF4E.c* (1,435 bp), respectively. PCR products were cloned into pCR2.1 (Invitrogen) or pMOS*Blue* (GE Healthcare) prior to sequencing at least three colonies with the Big-Dye Terminator system (Applied Biosystems, Warrington, U.K.) with products run on an ABI Prism 3130xl Genetics Analyzer (Applied Biosystems).

The PCR products were labeled with ^{32}P dCTP using the RediPrimeII DNA labeling system (GE Healthcare) and standard protocols. The radiolabeled probes were hybridized to filters printed with 36,864 colonies of the JBr BAC library. Potential bacterial clones of interest were identified, purchased, and subjected to a further PCR using the original primers to verify the genes' presence. Between one and three BAC per gene were thus identified (e.g., *BraA.eIF4E.a* in JBr034F07, *BraA.eIF4E.c* in JBr039L10, *BraA.eIF(iso)4E.a* in JBr043O19, and *BraA.eIF(iso)4E.c* in JBr040N12), although no BAC with *BraA.eIF(iso)4E.b* was found. Two BAC (JBr042K14 and JBr043K20) were found to contain an *eIF4E*-like gene that lacked exons 2 and 3 and had stop codons in the normal reading frame of exons 1 and 4. This defective gene was designated *BraA.eIF4E.b*.

Primers matching the PCR product sequences were designed to extend the known sequence into the BAC DNA. Sequence was determined upstream and downstream of the genes in a stepwise manner for approximately 2 kb in each direction beyond the start and stop codons of the ORF (Supplementary Table 1).

Construction of binary vectors.

Having identified the start and stop codons of the ORF, primers BR57 to BR64 (Table 3) were used to amplify the *Brassica* genes together with likely promoters (200 to 600 bp upstream of the start codon). In the case of *BraA.eIF4E.a*, the upstream primer location was restricted by the need to avoid an upstream ORF discovered in close proximity. Downstream regions were chosen to include likely sites of polyA tail attach-

Table 3. Primers used for the identification and cloning of *Brassica rapa eIF4E* and *eIF(iso)4E* genes

Primer	Sequence (5′–3′)	Genes	Location	Direction
BR1	ATGGCGACAGAGGATGTGAACG	*BraA.eIF(iso)4E.a, b, c*	Exon 1	Forward
BR2	TCTCCTTCCACTTCTTCCCAATAC	*BraA.eIF(iso)4E.a, b, c*	Exon 4	Reverse
BR6	CACTCTCAAGCCTAATGTCGC	*BraA.eIF4E.a*	Exon 1	Forward
BR7	AGAGGTTGTAGGTTTCAGGCAGTG	*BraA.eIF4E.a*	Exon 5	Reverse
BR20	CAATGGCGGTAGAAGACACTT	*BraA.eIF4E.b, c*	Exon 1	Forward
BR21	CATGGATGATGAAACCAATGT	*BraA.eIF4E.b, c*	Exon 4	Reverse
BR57	AAAAAGCAGGCTTTTGGTCTGCAGTTATGTTATTAG	*BraA.eIF4E.a*	Upstream	Forward
BR58	AGAAAGCTGGGTAAAAAGGCTTGCGAGTCA	*BraA.eIF4E.a*	Downstream	Reverse
BR59	AAAAAGCAGGCTTAGGACAAATGATATGGGGAGAGT	*BraA.eIF4E.c*	Upstream	Forward
BR60	AGAAAGCTGGGTAGCTTGGCGACCTTTTGA	*BraA.eIF4E.c*	Downstream	Reverse
BR61	AAAAAGCAGGCTACCCTTAGTATCAATTAATTCAGAAC	*BraA.eIF(iso)4E.a*	Upstream	Forward
BR62	AGAAAGCTGGGTTCCTGGTCGAAATCAATAAGA	*BraA.eIF(iso)4E.a*	Downstream	Reverse
BR63	AAAAAGCAGGCTTTTTTAAGAATGGAGGGAGTAT	*BraA.eIF(iso)4E.c*	Upstream	Forward
BR64	AGAAAGCTGGGTGAAGCGCGGGTCAAAAT	*BraA.eIF(iso)4E.c*	Downstream	Reverse

ment as found in comparable *Brassica* EST. The sizes of the cloned regions were *BraA.eIF4E.a*, 2,144 bp; *BraA.eIF4E.c*, 2,758 bp; *BraA.eIF(iso)4E.a*, 2,076 bp; and *BraA.eIF(iso)4E.c*, 1,818 bp.

The genes were amplified from the BAC using KOD Hotstart DNA polymerase (Merck, Nottingham, U.K.), and *attB* adapters were added by PCR with the same enzyme and Gateway primers attB1 and attB2 (Invitrogen). The products were cloned into the Gateway Entry plasmid pDONR221 using BP clonase (Invitrogen).

Binary vectors pB2GW7 and pK2GW7 (Karimi et al. 2002) contain a 35S constitutive promoter, a 35S terminator, and glufosinate ammonium herbicide resistance (*bar*) or kanamycin resistance (*nptII*), respectively, for in planta selection. The *Brassica* genes were transferred from the Entry vector into the binary vectors using an LR clonase-mediated reaction (Invitrogen). Sequences were verified at all stages of the cloning.

Transformation of *B. rapa* genes into *Arabidopsis thaliana*.

Purified plasmid DNA of each construct was electroporated into *Agrobacterium tumefaciens* GV3101 pMP90 using standard techniques. The presence and stability of the transgenes was verified by PCR.

Arabidopsis thaliana Col-0 and Col-0::*dSpm* lines were grown and transformed with *Agrobacterium tumefaciens* containing the constructs using the floral dip method (Clough and Bent 1998). Constructs in pB2GW7 were used to transform Col-0; pK2GW7 was used to transform Col-0::*dSpm*. Each *Brassica* gene was transformed into each plant line (Table 4).

The primary transformed plants were allowed to set seed and the T_1 seed was harvested. For selection using glufosinate ammonium, seed was grown in trays on F2S seed compost (Levington, Suffolk, U.K.) that was drenched every 3 to 4 days with glufosinate ammonium (Harvest; Bayer Crop Science, Cambridge) at 150 mg/liter for 3 weeks. For selection using kanamycin, seed was grown in deep petri dishes on 0.5× Murashige and Skoog medium (Duchefa Biochemie, Melford Laboratories, Ipswich, U.K.) with kanamycin (Sigma-Aldrich, Gillingham, U.K.) at 50 µg/ml for 3 weeks. Genomic DNA from individual T_1 survivors was obtained as above and PCR was performed to check for the presence of the transgenes (*Brassica* genes and selection markers) prior to potting on the plants in compost without selection to allow the plants to flower and to set seed (T_2).

T_2 plants were grown and genomic DNA was extracted as above. PCR checks were performed using the *Brassica* primers used to generate the inserts for the binary clones with KOD Hotstart DNA polymerase. The products were sequenced to confirm that there were no errors in the final transgenes. The continued presence of *dSpm* in *At.eIF(iso)4E* was also checked in appropriate lines.

TuMV challenge of *A. thaliana* lines transformed with *B. rapa* genes.

For each transformed line, seed of up to four T_2 families (-1 to -4), derived from different T_2 individuals, were sown. Only

Table 4. Designations of lines of *Arabidopsis thaliana* transformed with *Brassica rapa* genes

| *Brassica* gene | A. thaliana | |
	Col-0	Col-0::*dSpm*[a]
BraA.eIF4E.a	TP085	AT25
BraA.eIF4E.c	TP082	AT22
BraA.eIF(iso)4E.a	TP083	AT23
BraA.eIF(iso)4E.c	TP084	AT24

[a] Possesses *dSpm* transposon insertion in At5g35620, inactivating the gene *At.eIF(iso)4E*.

two families were produced from the transformation of Col-0 with *BraA.eIF(iso)4E.c*. Fifteen plants of each family were sown in F2S compost and grown at 20°C with a 9-h day length to maintain the vegetative state. After 33 days, four to six true leaves of 12 plants were mechanically inoculated using TuMV-infected *B. juncea* leaf sap and a cotton bud to hold the inoculum. The remaining plants were inoculated with buffer only.

Symptoms were assessed visually each week for 4 weeks. At 3 weeks postinoculation, leaf samples from two plants were collected for genomic DNA extraction. The presence of *Brassica* transgenes, selection marker genes, and the *dSpm* transposon in *At.eIF(iso)4E* were all checked by PCR. TuMV presence in three other plants was assayed by plate-trapped antigen ELISA using the mouse monoclonal antibody EMA67, essentially as previously described (Jenner et al. 1999), in order to confirm that symptoms were due to viral infection and not transgene presence. Leaf sap was extracted and diluted with 1 ml of water, then diluted one-third before loading on microtiter plates. OD values were measured after 2 h, or earlier if approaching the maximum measureable by the ELISA plate reader. Because infection of all the families containing the transgenes in the first experiment was unexpected, determination of virus titers in every individual plant was technically difficult. Therefore, the experiment was repeated for each of the four transgenes using mechanical inoculation, examining every individual plant of a single family per transgene (family i: TP085, TP082, TP083, TP084, AT25, AT22, AT23, and AT24) in order to determine the virus titer. A further experiment examined one family of lines possessing the *BraA.eIF(iso)4E.a* transgene (family iv: TP083 and AT23) in order to compare mechanical inoculation with aphid inoculation directly.

Due to the nature of the design of the experiments, data were analyzed using REML. REML allows estimation of variance components and treatment effects of a linear model with both fixed and random effects. It can be used, as in this case, to analyze unbalanced data sets. Data were transformed according to $\log_e(100 \times [\text{mean OD} + 0.05])$, then analyzed for each experimental occasion.

LITERATURE CITED

Abdul-Razzak, A., Guiraud, T., Peypelut, M., Walter, J., Houvenaghel, M.-C., Candresse, T., Le Gall, O., and German-Retana, S. 2009. Involvement of the cylindrical inclusion (CI) protein in the overcoming of an eIF4E-mediated resistance against *Lettuce mosaic potyvirus*. Mol. Plant Pathol. 10:109-113.

Albar, L., Bangratz-Reyser, M., Hébrard, E., Ndjiondjop, M.-N., Jones, M., and Ghesquière, A. 2006. Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to *Rice yellow mottle virus*. Plant J. 47:417-426.

Allen, M. L., Metz, A. M., Timmer, R. T., Rhoads, R. E., and Browning, K. S. 1992. Isolation and sequence of the cDNAs encoding the subunits of the isozyme form of wheat-protein synthesis initiation factor-4F. J. Biol. Chem. 267:23232-23236.

Andrade, M., Abe, Y., Nakahara, K. S., and Uyeda, I. 2009. The cyv-2 resistance to *Clover yellow vein virus* in pea is controlled by the eukaryotic initiation factor 4E. J. Gen. Plant Pathol. 75:241-249.

Beauchemin, C., Boutet, N., and Laliberté, J.-F. 2007. Visualization of the interaction between the precursors of VPg, the viral protein linked to the genome of *Turnip mosaic virus*, and the translation eukaryotic initiation factor iso 4E in planta. J. Virol. 81:775-782.

Browning, K. S. 1996. The plant translational apparatus. Plant Mol. Biol. 32:107-144.

Bruun-Rasmussen, M., Møller, I. S., Tulinius, G., Hansen, J. K. R., Lund, O. S., and Johansen, I. E. 2007. The same allele of translation initiation factor 4E mediates resistance against two *Potyvirus* spp. in *Pisum sativum*. Mol. Plant-Microbe Interact. 20:1075-1082.

Clough, S. J., and Bent, A. F. 1998. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16:735-743.

Díaz-Pendón, J. A., Truniger, V., Nieto, C., García-Mas, J., Bendahmane, A., and Aranda, M. A. 2004. Advances in understanding recessive resistance to plant viruses. Mol. Plant Pathol. 5:223-233.

Dinkova, T. D., Aguilar, R., and Sánchez De Jiménez, E. 2000. Expression of maize eukaryotic initiation factor (eIF) iso4E is regulated at the translational level. Biochem. J. 351:825-831.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K. S., and Robaglia, C. 2002. The *Arabidopsis* eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. Plant J. 32:927-934.

Gallois, J.-L., Charron, C., Sánchez, F., Pagny, G., Houvenaghel, M.-C., Moretti, A., Ponz, F., Revers, F., Caranta, C., and German-Retana, S. 2010. Single amino acid changes in the *Turnip mosaic virus* viral genome-linked protein (VPg) confer virulence towards *Arabidopsis thaliana* mutants knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. J. Gen. Virol. 91:288-293.

Gao, Z., Eyers, S., Thomas, C., Ellis, N., and Maule, A. 2004. Identification of markers tightly linked to *sbm* recessive genes for resistance to *Pea seed-borne mosaic virus*. Theor. Appl. Genet. 109:488-494.

Hughes, S. L., Green, S. K., Lydiate, D. J., and Walsh, J. A. 2002. Resistance to *Turnip mosaic virus* in *Brassica rapa* and *B. napus* and the analysis of genetic inheritance in selected lines. Plant Pathol. 51:567-573.

Hwang, J., Li, J., Liu, W. Y., An, S. J., Cho, H., Her, N., Yeam, I., Kim, D., and Kang, B. C. 2009. Double mutations in *eIF4E* and *eIFiso4E* confer recessive resistance to *Chilli veinal mottle virus* in pepper. Mol. Cells 27:329-336.

Jenner, C. E., Keane, G. J., Jones, J. E., and Walsh, J. A. 1999. Serotypic variation in turnip mosaic virus. Plant Pathol. 48:101-108.

Kang, B.-C., Yeam, I., Frantz, J. D., Murphy, J. F., and Jahn, M. M. 2005. The *pvr1* locus in *Capsicum* encodes a translation initiation factor eIF4E that interacts with *Tobacco etch virus* VPg. Plant J. 42:392-405.

Kang, B.-C., Yeam, I., Li, H., Perez, K. W., and Jahn, M. M. 2007. Ectopic expression of a recessive resistance gene generates dominant potyvirus resistance in plants. Plant Biotechnol. J. 5:526-536.

Kanyuka, K., Druka, A., Caldwell, D. G., Tymon, A., Mccallum, N., Waugh, R., and Adams, M. J. 2005. Evidence that the recessive bymovirus resistance locus rym4 in barley corresponds to the eukaryotic translation initiation factor 4E gene. Mol. Plant Pathol. 6:449-458.

Karimi, M., Inzé, D., and Depicker, A. 2002. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci. 7:193-195.

Lellis, A. D., Kasschau, K. D., Whitham, S. A., and Carrington, J. C. 2002. Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus infection. Curr. Biol. 12:1046-1051.

Léonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M. G., and Laliberté, J.-F. 2000. Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. J. Virol. 74:7730-7737.

Léonard, S., Viel, C., Beauchemin, C., Daigneault, N., Fortin, M. G., and Laliberté, J.-F. 2004. Interaction of VPg-Pro of *Turnip mosaic virus* with the translation initiation factor 4E and the poly(A)-binding protein *in planta*. J. Gen. Virol. 85:1055-1063.

Liu, X. P., Lu, W. C., Liu, Y. K., Wei, S. Q., Xu, J. B., Liu, Z. R., Zhang, H. J., Li, J. L., Ke, G. L., Yao, W. Y., Cai, Y. S., Wu, F. Y., Cao, S. C., Li, Y. H., Xie, S. D., Lin, B. X., and Zhang, C. L. 1996. Occurrence and strain differentiation of turnip mosaic potyvirus and sources of resistance in chinese cabbage in China. Acta Hortic. 407:431-440.

Miyoshi, H., Suehiro, N., Tomoo, K., Muto, S., Takahashi, T., Tsukamoto, T., Ohmori, T., and Natsuaki, T. 2006. Binding analyses for the interaction between plant virus genome-linked protein (VPg) and plant translational initiation factors. Biochimie 88:329-340.

Monzingo, A. F., Dhaliwal, S., Dutt-Chaudhuri, A., Lyon, A., Sadow, J. H., Hoffman, D. W., Robertus, J. D., and Browning, K. S. 2007. The struc-

ture of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. Plant Physiol. 143:1504-1518.

Naderpour, M., Lund, O. S., Larsen, R., and Johansen, E. 2010. Potyviral resistance derived from cultivars of *Phaseolus vulgaris* carrying bc-3 is associated with the homozygotic presence of a mutated eIF4E allele. Mol. Plant Pathol. 11: 255-263.

Nicaise, V., Gallois, J.-L., Chafiai, F., Allen, L. M., Schurdi-Levraud, V., Browning, K. S., Candresse, T., Caranta, C., Le Gall, O., and German-Retana, S. 2007. Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in *Arabidopsis thaliana*. FEBS (Fed. Eur. Biochem. Soc.) Lett. 581:1041-1046.

Plante, D., Viel, C., Léonard, S., Tampo, H., Laliberté, J.-F., and Fortin, M. G. 2004. *Turnip mosaic virus* VPg does not disrupt the translation initiation complex but interferes with cap binding. Physiol. Mol. Plant Pathol. 64:219-226.

Provvidenti, R., and Hampton, R. O. 1992. Sources of resistance to viruses in the *Potyviridae*. Arch. Virol. Suppl. 5:189-211.

Rana, D., Van Den Boogaart, T., O'Neill, C. M., Hynes, L., Bent, E., Macpherson, L., Park, J. Y., Lim, Y. P., and Bancroft, I. 2004. Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. Plant J. 40:725-733.

Rhoads, R. E. 2009. eIF4E: New family members, new binding partners, new roles. J. Biol. Chem. 284:16711-16715.

Robaglia, C., and Caranta, C. 2006. Translation initiation factors: A weak link in plant RNA virus infection. Trends Plant Sci. 11:40-45.

Rodriguez, C. M., Freire, M. A., Camilleri, C., and Robaglia, C. 1998. The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. Plant J. 13:465-473.

Ruffel, S., Dussault, M.-H., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C., and Caranta, C. 2002. A natural recessive resistance gene against *Potato virus Y* in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J. 32:1067-1075.

Ruffel, S., Gallois, J. L., Lesage, M. L., and Caranta, C. 2005. The recessive potyvirus resistance gene *pot-1* is the tomato orthologue of the pepper *pvr2-eIF4E* gene. Mol. Genet. Genomics 274:346-353.

Ruffel, S., Gallois, J.-L., Moury, B., Robaglia, C., Palloix, A., and Caranta, C. 2006. Simultaneous mutations in translation initiation factors *eIF4E* and *eIF(iso)4E* are required to prevent *Pepper veinal mottle virus* infection of pepper. J. Gen. Virol. 87:2089-2098.

Rusholme, R. L., Higgins, E. E., Walsh, J. A., and Lydiate, D. J. 2007. Genetic control of broad-spectrum resistance to *Turnip mosaic virus* in *Brassica rapa* (Chinese cabbage). J. Gen. Virol. 88:3177-3186.

Ruud, K. A., Kuhlow, C., Goss, D. J., and Browning, K. S. 1998. Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. J. Biol. Chem. 273:10325-10330.

Sato, M., Nakahara, K., Yoshii, M., Ishikawa, M., and Uyeda, I. 2005. Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of *Arabidopsis thaliana* by potyviruses. FEBS (Fed. Eur. Biochem. Soc.) Lett. 579:1167-1171.

Schaad, M. C., Anderberg, R. J., and Carrington, J. C. 2000. Strain-specific interaction of the tobacco etch virus NIa protein with the translation initiation factor eIF4E in the yeast two-hybrid system. Virology 273:300-306.

Shattuck, V. I. 1992. The biology, epidemiology, and control of *Turnip mosaic virus*. Plant Breed. Rev. 14:199-238.

Walsh, J. A. 1989. Genetic control of immunity to turnip mosaic virus in winter oilseed rape (*Brassica napus* ssp. *oleifera*) and the effect of foreign isolates of the virus. Ann. Appl. Biol. 115:89-99.

Walsh, J. A., and Jenner, C. E. 2002. *Turnip mosaic virus* and the quest for durable resistance. Mol. Plant Pathol. 3:289-300.

Walsh, J. A., Rusholme, R. L., Hughes, S. L., Jenner, C. E., Bambridge, J. M., Lydiate, D. J., and Green, S. K., 2002. Different classes of resistance to turnip mosaic virus in *Brassica rapa*. Eur. J. Plant Pathol. 108:15-20.

Wittmann, S., Chatel, H., Fortin, M. G., and Laliberté, J.-F. 1997. Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso)4E of *Arabidopsis thaliana* using the yeast two-hybrid system. Virology 234:84-92.

Zhang, Y.-Y., Li, H.-X., Ouyang, B., and Ye, Z.-B., 2006. Regulation of eukaryotic initiation factor 4E and its isoform: Implications for antiviral strategy in plants. J. Integr. Plant Biol. 48:1129-1139.

Multiple copies of eukaryotic translation initiation factors in *Brassica rapa* facilitate redundancy, enabling diversification through variation in splicing and broad-spectrum virus resistance

Charlotte F. Nellist[1,†], Wei Qian[2,†], Carol E. Jenner[1], Jonathan D. Moore[1], Shujiang Zhang[2], Xiaowu Wang[2], William H. Briggs[3], Guy C. Barker[1], Rifei Sun[2,*] and John A. Walsh[1,*]

[1]*School of Life Sciences, University of Warwick, Wellesbourne, Warwick CV35 9EF, UK,*
[2]*Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China, and*
[3]*Syngenta Seeds, Westeinde 62, 1601 AA Enkhuizen, the Netherlands*

SUMMARY

Recessive strain-specific resistance to a number of plant viruses in the *Potyvirus* genus has been found to be based on mutations in the eukaryotic translation initiation factor 4E (*eIF4E*) and its isoform, *eIF(iso)4E*. We identified three copies of *eIF(iso)4E* in a number of *Brassica rapa* lines. Here we report broad-spectrum resistance to the potyvirus *Turnip mosaic virus* (TuMV) due to a natural mechanism based on the mis-splicing of the *eIF(iso)4E* allele in some TuMV-resistant *B. rapa* var. *pekinensis* lines. Of the splice variants, the most common results in a stop codon in intron 1 and a much truncated, non-functional protein. The existence of multiple copies has enabled redundancy in the host plant's translational machinery, resulting in diversification and emergence of the resistance. Deployment of the resistance is complicated by the presence of multiple copies of the gene. Our data suggest that in the *B. rapa* subspecies *trilocularis,* TuMV appears to be able to use copies of *eIF(iso)4E* at two loci. Transformation of different copies of *eIF(iso)4E* from a resistant *B. rapa* line into an *eIF(iso)4E* knockout line of *Arabidopsis thaliana* proved misleading because it showed that, when expressed ectopically, TuMV could use multiple copies which was not the case in the resistant *B. rapa* line. The inability of TuMV to access multiple copies of *eIF(iso)4E* in *B. rapa* and the broad spectrum of the resistance suggest it may be durable.

Keywords: *Brassica rapa*, *Turnip mosaic virus*, *eIF(iso)4E*, broad-spectrum virus resistance, mis-splicing, *Arabidopsis thaliana*, Chinese cabbage.

INTRODUCTION

Recruitment of the eukaryotic translation machinery to the 5′ end of mRNA is a crucial regulatory step in initiation of cap-dependent translation. Investigation of recessive resistance to plant viruses identified the involvement of a group of proteins involved in mRNA translation, particularly eukaryotic translation initiation factor 4E (eIF4E), its isoform eIF(iso)4E (Robaglia and Caranta, 2006) and to a lesser extent eIF4G (Le Gall *et al.*, 2011). This was particularly true for recessive resistance to members of the *Potyviridae*, the largest group of plant viruses. The potyvirus 'viral protein genome-linked' (VPg) and *Arabidopsis thaliana* eIF(iso)4E have been shown to interact in yeast two-hybrid binding assays (Wittmann *et al.*, 1997). Studies on pea

eIF4E supported the view that the binding site for the VPg of a potyvirus overlapped with the mRNA cap-binding site (Ashby *et al.*, 2011). Caliciviruses also possess a VPg and some have been shown to utilise eIF4E to translate their genome in mammals (Goodfellow and Roberts, 2008). Mutations in *eIF4E* in a range of plant species have been shown to confer resistance to a range of potyviruses (Robaglia and Caranta, 2006), and knocking out *eIF(iso)4E* in Arabidopsis resulted in resistance to the potyviruses *Turnip mosaic virus* (TuMV) and *Lettuce mosaic virus* (Duprat *et al.*, 2002; Lellis *et al.*, 2002).

The TuMV-resistant RLR22 *Brassica rapa* var. *pekinensis* (Chinese cabbage), derived from an accession identified in

a screen of more than 3000 lines (Liu *et al.*, 1996) has broad-spectrum TuMV resistance (Walsh *et al.*, 2002). Artificial (mechanical) inoculation of RLR22 using leaf sap from TuMV-infected plants resulted in chlorotic spots in inoculated leaves, with no detectable systemic spread of the virus (Rusholme *et al.*, 2007; Figure S1 in Supporting Information). Following natural aphid TuMV challenge, no symptoms were seen and no virus was detected in any leaves (Rusholme *et al.*, 2007), indicating that RLR22 plants would be completely resistant (immune) in the field. Segregation following a cross between RLR22 and the TuMV-susceptible line R-o-18 of the closely related *B. rapa* ssp. *trilocularis* (Roxb.) Hanelt. (yellow sarson; Figure S1), revealed that the resistance was due to a recessive gene, *retr01* (coincident with the copy of *elF(iso)4E* on chromosome A4) that was epistatic to a dominant gene, *ConTR01* (coincident with one of the other copies of *elF(iso)4E*, or one of the copies of *elF4E*, both on chromosome A8; Rusholme *et al.*, 2007). Crosses between a further TuMV-resistant Chinese cabbage line (BP8407) derived from the screened accessions (Liu *et al.*, 1996) and a susceptible Chinese cabbage line, revealed that this resistance was due to a single recessive gene, *retr02*, that mapped to the same locus as *retr01* (Qian *et al.*, 2013).

RESULTS

Three copies of *elF4E* and of *elF(iso)4E* identified in RLR22 *B. rapa*

Three copies of *elF4E* and three copies of *elF(iso)*4E were identified in a genomic library of the *B. rapa* line R-o-18 (Jenner *et al.*, 2010) and a genomic library of RLR22. This was consistent with genomic analyses of *Brassica* diploid species, which indicated that they evolved from genome triplication of an ancestor with a genome similar to *A. thaliana* (Town *et al.*, 2006). Sequencing of the genes and comparison of the RLR22 sequence with a sequence-based genetic map of *B. rapa* (Wang *et al.*, 2011) confirmed the location of *retr01* (*BraA.elF(iso)4E.a*) on chromosome A4 in *B. rapa.* Molecular markers located *retr02* on a scaffold also from chromosome A4 (Qian *et al.*, 2013). The location of the *ConTR01* candidates *BraA.elF4E.c* and *BraA.elF(iso)4E.c* from RLR22 (Rusholme *et al.*, 2007) was confirmed on chromosome A8, again by comparison with the sequence-based genetic map (Wang *et al.*, 2011).

retr01 is mis-spliced in resistant plant lines

To identify the mechanism of broad-spectrum resistance to TuMV in *B. rapa*, several plant lines were studied. *BraA.elF(iso)4E.a* was sequenced from four lines known to have broad-spectrum resistance, or derived from lines with broad-spectrum resistance to TuMV [RLR22 (Walsh *et al.*, 2002; possessing *retr01* and *ConTR01*; Rusholme

et al., 2007), BP058 (Walsh *et al.*, 2002), Jong Bai No. 2 (Hughes *et al.*, 2002)], BP8407 (possessing *retr02*; Qian *et al.*, 2013) and four lines known to be susceptible to all TuMV isolates tested [R-o-18 (Rusholme *et al.*, 2007), Ji Zao Chun (Qian *et al.*, 2013), CK 2 and CK 1 (not susceptible to UK 1 TuMV)]. The genomic sequence of *BraA.elF (iso)4E.a* on chromosome A4 of RLR22 and BP8407 was identical, showing that *retr02* is *retr01*, rather than a different allele. All the TuMV-resistant lines were found to have an extra G (indel), adjacent to the splice site of *BraA.elF(iso)4E.a* exon 1 and intron 1 (at position 201 nucleotides from the ATG), relative to alleles in all susceptible lines.

To study the expression of *BraA.elF(iso)4E.a*, reverse-transcriptase (RT) PCR was carried out on all the lines and the products were sequenced. Plants to be genotyped were challenged with TuMV isolate CDN 1, or TuMV-C4, to verify resistance/susceptibility. All TuMV-susceptible lines (lacking the indel) produced a single product approximately 600 nucleotides (nt) in size (Figure 1a), corresponding to the correctly spliced version of the gene (Figure 1b). All the lines possessing the indel were resistant to TuMV and had a larger major product of approximately 664 nt, plus a smaller minor product of approximately 600 nt (Figure 1a). The larger product retained the extra G and the whole of intron 1 (Figure 1c) and the smaller product was of a similar size to the correctly spliced version. Sequencing of RT-PCR products of *BraA.elF(iso)4E.a* showed that introns 2–4 were correctly spliced in all resistant and susceptible plants. The retention of intron 1 by all lines possessing the indel resulted in the introduction of a premature stop codon in intron 1 at the 234 nt position (Figure 1c). Cloning RT-PCR products from RLR22 revealed a further four less common variants. These included a variant possessing the last 15 nt of intron 1, resulting in a slightly elongated, in-frame mRNA sequence (Figure 1d), one with intron 1 excised along with the last 3 nt of exon 1, resulting in a slightly truncated, in-frame mRNA with a substitution (Figure 1e), one with an extra G at the end of exon 1, resulting in a premature stop codon (Figure S2b), and one that was correctly spliced (Figure 1b). There was a clear association between lines possessing the indel, resulting in mis-splicing of *BraA.elF(iso)4E.a* (*retr01*), and broad-spectrum resistance.

The most common mis-spliced variant of *BraA.elF(iso)4E.a* is non-functional for TuMV

To establish whether the most common splice variant of *BraA.elF(iso)4E.a* (intron 1 retained, Figure 1c) could be functional for TuMV, cDNA of the genes from the TuMV-susceptible line Ji Zao Chun and TuMV-resistant line BP8407 were expressed in *Escherichia coli* Rosetta(DE3) pLysS with a 6× Histidine tag. The deduced sizes of the Ji Zao Chun and BP8407 proteins (minus the 6× Histidine
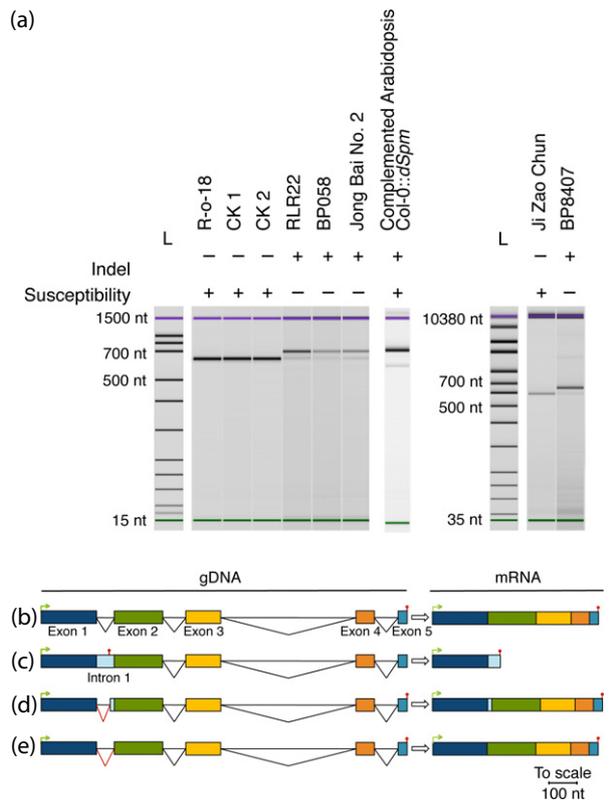
(a)

Figure 1. *BraA.eIF(iso)4E.a* is mis-spliced in resistant *Brassica rapa* and in transgenic *Arabidopsis thaliana*.

(a) Detection of a single transcript in susceptible plants, corresponding to correctly spliced *BraA.eIF(iso)4E.a*, splice variants in resistant *B. rapa* plants and splice variants from an Arabidopsis Col-0::*dSpm* plant complemented with the RLR22 allele of *BraA.eIF(iso)4E.a*, separated using an Agilent 2100 Bioanalyser.

(b) Correctly spliced *BraA.eIF(iso)4E.a* detected in susceptible plants (a).

(c) The most common mis-spliced variant retained the extra G and the whole of intron 1 resulting in a premature stop codon at the 234 nucleotide (nt) position, detected in resistant lines and Arabidopsis Col-0::*dSpm* complemented by the RLR22 allele of *BraA.eIF(iso)4E.a*.

(d) Variant retaining the last 15 nt of intron 1 (in-frame), detected in RLR22 and Col-0::*dSpm* complemented by *BraA.eIF(iso)4E.a* (a).

(e) Variant lacking the last 3 nt of exon 1 (in-frame) and with a substitution, detected in RLR22 and some F_2 plants from the RLR22, CK 1 cross.

(a)

(b)

Figure 2. The most common mis-spliced variant of *BraA.eIF(iso)4E.a* produced a truncated eIF(iso)4E protein that did not interact with *Turnip mosaic virus* (TuMV) 'viral protein genome-linked' (VPg) and two less common splice variants did interact with the VPg of TuMV-C4, in yeast two-hybrid assays.

(a) *Brassica rapa* BP8407 (intron 1 retained) and Ji Zao Chun *BraA.eIF(iso)4E.a* recombinant proteins with a 6× histidine tag expressed in transgenic *Escherichia coli* Rosetta(DE3)pLysS after induction with isopropyl β-D-1-thiogalactopyranoside. BraA.eIF(iso)4E.a from the resistant BP8407 was truncated compared with that from the susceptible Ji Zao Chun and their deduced molecular weights were consistent with the polypeptide sequences predicted from the mRNAs. pET-32a empty vector.

(b) Yeast two-hybrid assay (i) TuMV-C4 VPg + RLR22 BraA.eIF(iso)4E.a (from a splice variant lacking last 3 nucleotide (nt) of exon 1, resulting in loss of an amino acid and a substitution) from a F_2 plant from the RLR22, CK 1 cross, (ii) TuMV-C4 VPg + RLR22 BraA.eIF(iso)4E.a (from splice variant retaining last 15 nt of intron 1, resulting in an extra five amino acids) from *Arabidopsis thaliana* Col-0::*dSpm* transformed with the RLR22 allele of *BraA.eIF(iso)4E.a*, (iii) TuMV-C4 VPg + BP8407 BraA.eIF(iso)4E.a (from splice variant retaining whole of intron 1, resulting in a stop codon in intron 1), (iv) TuMV-C4 VPg + BP8407 BraA.eIF(iso)4E.a (from splice variant retaining an extra G at the end of exon 1, resulting in a stop codon in exon 2), (v) TuMV-C4 VPg + Ji Zao Chun BraA.eIF(iso)4E.a, (vi) TuMV-C4 VPg + Arabidopsis eIF(iso)4E (positive control), (vii) TuMV-C4 VPg + empty prey, (viii) Ji Zao Chun + empty bait.

tag, approximately 18 kDa) appeared to be consistent with the predicted sizes (22.50 and 8.67 kDa, respectively; Figure 2a), confirming that the mis-splicing resulted in a truncated protein. Yeast two-hybrid assays confirmed the physical interaction between the VPg of TuMV-C4 and the BraA.eIF(iso)4E.a protein from the susceptible Ji Zao Chun line and the lack of interaction with truncated BraA.eIF(iso)4E.a proteins [*BraA.eIF(iso)4E.a* mRNA with an extra G at the end of exon 1 (Figure S2b), or intron 1 retained (Figure 1c)] from BP8407 (Figure 2b and Table S1), indicating that these truncated proteins are unlikely to be functional for the virus. Expression of the intron 1-retained construct in the yeast was confirmed by RT-PCR.
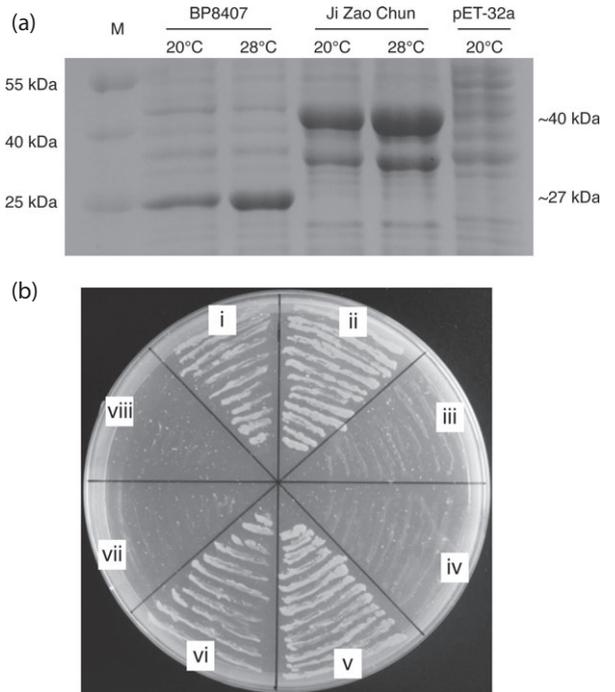
BraA.eIF(iso)4E.c is the only candidate for *ConTR01*

To identify the second gene (*ConTR01*) involved in the TuMV resistance in the progeny of the cross between RLR22 and R-o-18 (Rusholme *et al.*, 2007), following genotyping of B_1S_1 families, a B_1S_2 population homozygous for *retr01* (RLR22 allele of *BraA.eIF(iso)4E.a*), but segregating for *BraA.eIF(iso)4E.c* and homozygous for the RLR22 allele of *BraA.eIF4E.c* (the *ConTR01* candidates; Rusholme *et al.*, 2007), was produced. This family segregated for resistance and susceptibility. At the *BraA.eIF(iso)4E.c* locus, all plants

homozygous for the RLR22 allele were resistant, whereas the heterozygotes and the plants homozygous for the R-o-18 allele segregated for resistance. All plants from a family segregating for *BraA.elF4E.c* (homozygous for the RLR22 alleles of *BraA.elF(iso)4E.a* and *BraA.elF(iso)4E.c*) were resistant to TuMV. These results ruled out *BraA.elF4E.c* as a candidate for *ConTR01*, leaving *BraA.elF(iso)4E.c* as the only candidate. Sequencing *BraA.elF(iso)4E.c* from R-o-18 and RLR22 showed four amino acid differences (L36F, V52A, T80I and Q150P). Attempts to detect an interaction between the R-o-18 or RLR22 alleles of BraA.elF(iso)4E.c and the VPg of TuMV-C4 in yeast two-hybrid experiments were unsuccessful (Table S1). So it is not clear whether any of the amino acid differences affect interaction with the VPg.

Whole transcriptome sequencing of R-o-18 and RLR22 plants revealed qualitative and quantitative differences in *BraA.elF(iso)4E.a* expression between the two lines. It was correctly spliced in R-o-18 (Figure 1b), whereas in RLR22, only the variant retaining intron 1 (Figure 1c) was detected. Expression in RLR22 was significantly lower than in R-o-18. There were no significant differences between the levels of expression of *BraA.elF(iso)4E.c* and *BraA.elF4E.c* between the two lines (Table 1).

In crosses between Chinese cabbage, resistance is inherited as a single recessive gene

As the susceptible parent in the original cross, R-o-18, was a different subspecies from RLR22 (Chinese cabbage), the genetic inheritance of resistance was also investigated in crosses with Chinese cabbage lines. Segregation ratios of genotypes in F_2 generations from crosses between Ji Zao Chun and BP8407 (Qian *et al.*, 2013) and between RLR22 and two Chinese cabbage lines (CK 1 and CK 2) were consistent with those predicted for a single recessive gene ($\chi^2 = 2.15$, 2.89 and 4.40, respectively, all $P > 0.05$). CK 1 and CK 2 crosses demonstrated that most F_2 plants that were homozygous for *retr01* were resistant following artificial (mechanical) inoculation; however, a small number of

plants showed an occasional chlorotic spot on uninoculated leaves. Reverse transcriptase PCR of the latter plants detected mis-spliced variants of *BraA.elF(iso)4E.a*, including one with the last 3 nt of exon 1 missing (Figure 1e), resulting in the loss of one amino acid and a substitution. This mis-spliced variant interacted with the VPg of TuMV-C4 in yeast two-hybrid assays (Figure 2b), indicating that it could have been functional for TuMV.

As the sequence of *BraA.elF(iso)4E.c* from the CK 1, CK 2, BP8407 and Ji Zao Chun lines was identical to that in RLR22, all the susceptible Chinese cabbage parents already had the *ConTR01* candidate, explaining why resistance segregated monogenically in these crosses. This result, along with those for the segregation of resistance in the offspring of the RLR22 R-o-18 cross showed that the presence of three copies of *elF(iso)4E* can lead to different segregation patterns in offspring, depending upon which copies of *elF(iso)4E* from the parental lines can be used and accessed by TuMV. Interestingly, from the CK 1 and CK 2 crosses with RLR22, F_2 plants heterozygous at the *retr01* locus were less susceptible to TuMV than those plants that were homozygous for the allele from the susceptible parent (Figure 3a). Enzyme linked immunosorbent assay (ELISA) confirmed that the heterozygotes [mean A_{405} from ELISA for CK 1 offspring 0.81 ($\pm$0.074) and for CK 2 offspring 0.28 ($\pm$0.04)] accumulated significantly less virus than homozygotes [mean A_{405} for CK 1 offspring 2.09 ($\pm$0.132) and for CK 2 offspring 1.47 ($\pm$0.159), both $P < 0.001$].

Ectopic expression of copies of *BraA.elF4E* and *BraA.elF (iso)4E* from RLR22 *B. rapa* complemented the TuMV-resistant Arabidopsis *elF(iso)4E* knockout line Col-0::dSpm

We investigated the ability of TuMV to use different copies of RLR22 *elF4E* and *elF(iso)4E* by transforming *BraA.elF4E.a*, *BraA.elF4E.c*, *BraA.elF(iso)4E.a* and *BraA.elF(iso)4E.c* into the Arabidopsis Col-0::dSpm line possessing a transposon knock-out of *elF(iso)4E* that conferred resistance to TuMV (Duprat *et al.*, 2002). Reverse transcriptase PCR

Table 1 Expression levels of copies of *elF4E* and *elF(iso)4E* from *Turnip mosaic virus* (TuMV)-susceptible *Brassica rapa* ssp. *trilocularis* R-o-18 and TuMV-resistant *B. rapa* ssp. *pekinensis* RLR22 plants measured in fragments per kilobase of exon per million fragments mapped (FPKM) from transcriptome analysis. The data show that all genes except *BraA.elF(iso)4E.b* were expressed and the only gene where there was a significant difference in expression between the two lines was *BraA.elF(iso)4E.a*

| Plant line | Gene | | | | |
	BraA.elF(iso)4E.a	*BraA.elF(iso)4E.b*	*BraA.elF(iso)4E.c*	*BraA.elF4E.a*	*BraA.elF4E.c*
R-o-18	7.7	0.05	10.5	3.7	13.5
RLR22	2.0	0.07	12.6	2.9	11.8
Q-value	0.0047[a]	1	0.7810	0.8040	0.8573

BraA.elF4E.b is a pseudogene lacking exons 2 and 3 (Jenner *et al.*, 2010).
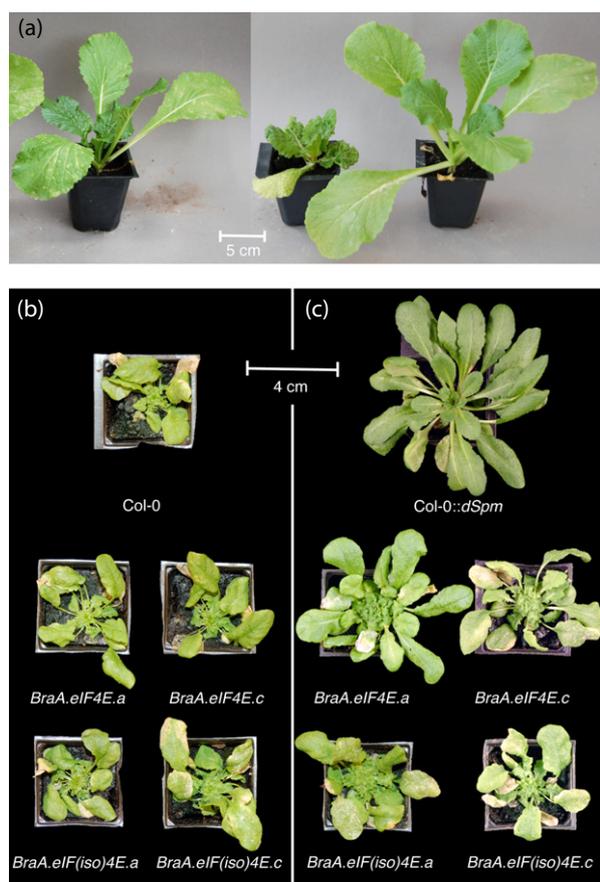[a]Statistically significant ($Q \leq 0.005$).

Figure 3. Susceptibility of *Brassica rapa* plants to *Turnip mosaic virus* (TuMV) CDN 1 and *Arabidopsis thaliana* Col-0::*dSpm* plants complemented by copies of *BraA.eIF(iso)4E* and *BraA.eIF4E* from two RLR22 loci of each gene.
(a) F$_2$ plants derived from a cross between the Chinese cabbage lines RLR22 and CK 1, a plant heterozygous for *BraA.eIF(iso)4E.a* with limited systemic infection (left), a plant homozygous for the CK 1 copy of *BraA.eIF(iso)4E.a* with severe systemic infection (middle) and a plant homozygous for the RLR22 copy of *BraA.eIF(iso)4E.a*, resistant to TuMV (right).
(b) Arabidopsis Col-0 transformed with copies of *eIF4E* and *eIF(iso)4E* from *B. rapa* RLR22, following artificial (mechanical) inoculation with TuMV, 20 days post-inoculation.
(c) Complementation of Arabidopsis Col-0::*dSpm* plants with copies of *eIF4E* and *eIF(iso)4E* from *B. rapa* RLR22, following artificial (mechanical) inoculation with TuMV, 20 days post-inoculation. The *dSpm* transposon insertion is in *At5 g35620*, inactivating *At.eIF(iso)4E*.

confirmed expression of all these genes in the transgenic Col-0 and Col-0::*dSpm* lines. Artificial (mechanical; Figure 3b) and aphid challenge of both transformed Col-0 controls and Col-0::*dSpm* showed that all four copies of *eIF4E*/*eIF(iso)4E* from RLR22 complemented the *eIF(iso)4E* knockout (Figure 3c), indicating that TuMV could use all copies investigated. The ectopic expression was misleading in that it indicated that these genes from R-o-18 (Jenner *et al.*, 2010) and RLR22 were functional for TuMV. Segregation of the phenotypes and genotypes in B$_1$S$_1$ (Rusholme *et al.*, 2007) and B$_1$S$_2$ plants from the cross

between these two lines clearly demonstrated that either these genes were not available or were not functional for the virus in RLR22. The segregation also demonstrated that the two copies of *eIF4E* were either not available or were not functional for the virus in R-o-18. The RT-PCR and transcriptome analyses of transformed Arabidopsis plants did not reveal any correctly spliced copies of RLR22 *BraA.eIF(iso)4E.a*, but did detect eight mis-spliced variants (Figures 1 and S2). Of these, the one possessing the last 15 nt of intron 1 (Figure 1d; in-frame, with an additional five amino acids), could potentially be functional for TuMV. Yeast two-hybrid assays confirmed that this mis-spliced variant interacted with the VPg of TuMV-C4 (Figure 2b and Table S1), and this may explain why the knockout line was complemented by the RLR22 allele of *BraA.eIF(iso)4E.a*.

DISCUSSION

Amino acid sequences of some mis-spliced variants of BraA.eIF(iso)4E.a suggested they could be functional. Comparing the sequences of these with those of correctly spliced BraA.eIF(iso)4E.a and eIF4E and eIF(iso)4E in a number of plant species (Monzingo *et al.*, 2007; German-Retana *et al.*, 2008), indicated that a number of variants had the correct amino acids at important sites (Figure S3). This suggests that some of the mis-spliced variants could be functional for TuMV. The ability of two of these mis-spliced variants to bind the TuMV VPg in yeast two-hybrid assays further supports this possibility. The existence and low frequency of these potentially functional variants, combined with artificial inoculation of large amounts of virus, might explain why unexpected limited infection was seen in some plants. For example, the limited infection of RLR22-inoculated leaves (which did not spread systemically) and the occasional very limited systemic infection of plants homozygous for the RLR22 allele of *BraA.eIF(iso)4E.a* derived from the crosses we made.

Our confirmation of three copies of *eIF(iso)4E* in RLR22 revealed how it was possible for the copy that TuMV would normally use in Chinese cabbage to be non-functional for both plant and virus, without apparently disadvantaging the plant. The durability of the resistance will be dependent upon the virus not mutating to be able to utilise/access other copies *in planta*. It may also be possible that TuMV could evolve to be capable of effective cap-independent translation (Basso *et al.*, 1994). The situation could be further complicated by the variability of the mis-splicing of *eIF(iso)4E* resulting in sufficient quantities of splice variants that could be functional for the virus. The difficulty we experienced in detecting correctly spliced variants in RLR22 and variants with minor amino acid changes indicates that this is highly unlikely. Evidence to date suggests that the VPg protein of potyviruses competes with host plant mRNA cap for eIF4E binding (Gao *et al.*, 2004) and eIF(iso)4E binding (Plante *et al.*, 2004). The fact that it was

possible to induce limited infection in inoculated leaves (which was not able to spread systemically) only by artificially inoculating excessive amounts of TuMV into resistant plants (homozygous for *retr01*), indicates that there was very little eIF(iso)4E protein present that was functional for the virus, and/or the virus was less efficient than mRNA cap in competing for the protein. As plants homozygous for the *ConTR01* candidate that were heterozygous for *retr01* were clearly less susceptible to TuMV than plants homozygous for the *retr01* allele from susceptible plants, the amount of eIF(iso)4E protein in these brassica plants must be limiting for virus replication.

In this study we have demonstrated a mechanism of translation factor-based resistance, which, unlike previous examples that are strain-specific, is broad spectrum. Mis-splicing has also been induced artificially by targeting induced local lesions in genomes (TILLING) of tomato (Piron *et al.*, 2010). This resulted in resistance to the potyviruses *Pepper mottle virus* (PepMoV) and strain-specific resistance to *Potato virus Y* (PVY) through a substitution in the splice site of *eIF4E* (G1485A), causing the deletion of exons 2 and 3 in the mRNA (Piron *et al.*, 2010). At least one PVY isolate was able to overcome the resistance (Piron *et al.*, 2010). It was suggested that the resistance-breaking PVY isolate was able to use a different copy of *eIF4E*. We have shown that TuMV can use at least two copies of *eIF4E* and two of *eIF(iso)4E* from both susceptible (Jenner *et al.*, 2010) and resistant plants, when expressed ectopically in Arabidopsis and that all these copies are expressed in *B. rapa* RLR22. This suggests that copies of *eIF4E* and *eIF(iso)4E* other than *BraA.eIF(iso)4E.a* are inaccessible to TuMV in RLR22, possibly because of where in the cell they are expressed. This, along with the inability of a wide range of virus isolates from around the world, representing different genotypes, pathotypes and serotypes to overcome the resistance in RLR22 (Walsh *et al.*, 2002), indicates that the broad-spectrum resistance mechanism could provide durable potyvirus resistance in a range of plant species. The presence of multiple copies of eukaryotic translation initiation factors in *B. rapa* has facilitated redundancy. The redundancy arising from the mis-splicing has enabled diversification and resulted in the plant being able to evade pathogen infection.

EXPERIMENTAL PROCEDURES

Plant materials and virus isolates

Brassica rapa RLR22, R-o-18 (Rusholme *et al.*, 2007), BP058 (Walsh *et al.*, 2002), Jong Bai No. 2 (Hughes *et al.*, 2002), BP8407, Ji Zao Chun (Qian *et al.*, 2013) and inbred Syngenta lines CK 1 and CK 2 plants were grown in insect-free glasshouses at 18°C. *Arabidopsis thaliana* Col-0::*dSpm* (Duprat *et al.*, 2002) and Col-0 plants were grown in a growth room at 20°C with a 9/15-h day/night light cycle. The TuMV isolates used in this study were CDN 1 (Jenner and Walsh, 1996) and TuMV-C4 (Qian *et al.*, 2013).

Plant inoculation assays

Brassica and Arabidopsis plants were artificially (mechanically) inoculated with TuMV isolates or healthy plant sap, assessed and tested for the presence and quantity of TuMV by ELISA (Rusholme *et al.*, 2007; Jenner *et al.*, 2010). Aphid transmission of TuMV to Arabidopsis was performed using *Myzus persicae* aphids (Jenner *et al.*, 2010).

Construction of a genomic library, identification and cloning of copies of *eIF4E* and *eIF(iso)4E*

A genomic library was prepared from the *B. rapa* line RLR22 with broad-spectrum resistance to TuMV by Warwick Plant Genomic Libraries Limited using the pCC1FOS fosmid vector obtained under licence from Epicentre Technologies (http://www.epibio.com/). Copies of *eIF4E* and *eIF(iso)*4E were identified in the same manner as copies were identified in R-o-18 (Jenner *et al.*, 2010) and sequenced.

Expression of *BraA.eIF(iso)4E.a*

The RNA was extracted from brassica or *A. thaliana* leaves using the Ambion RNAqueous kit (Life Technologies, http://www.life-technologies.com/). Reverse-transcription reactions were carried out with 1 μg of total RNA for 15 min at 70°C using either CN4 antisense primer 5'-AGAAAGCTGGGTTCAGACAGTGAACCTAG TTCTTC-3' (including an attB site, underlined) for R-o-18, or CN5 antisense primer 5'-AGAAAGCTGGGTTCAGACAGTGAACCG AGTTCTTC-3' (including an attB site underlined) for Chinese cabbage lines, except Ji Zao Chun. For PCR, 4 μl of the reverse-transcription reaction was used as a template in 50-μl reactions, with 5 U of Taq polymerase (Invitrogen, http://www.invitrogen.com/) per reaction. The PCR was run for 30 cycles of 30 sec at 95°C, 30 sec at 58°C and 80 sec at 72°C. We used the following primers: CN3 sense 5'-AAAAAGCAGGCTCGATGGCG-ACAGAGGATG-3' (including an attB site underlined) and either CN4 or CN5 for amplifying *BraA.eIF(iso)4E.a* from all *B. rapa* lines except BP8407 and Ji Zao Chun where Bio11535F sense 5'-ATGGCGACAGAGGATGT-3' and Bio11536R antisense 5'-TCAGA-CAGTGAACCGA-3' were used. The RT-PCR products were separated by agarose gel electrophoresis and detected by staining with GelRed (Bioline, http://www.bioline.com/), or GoldView (SBS Genetech Ltd, http://www.sbsbio.com/). Products were also ran on an Agilent 2100 Bioanalyser (Agilent Technologies, http://www.home.agilent.com/) to view the different sized products. Bands were excised from gels and DNA purified using Gel Extraction Kits (Qiagen, http://www.qiagen.com/). The products were cloned for sequencing using the TOPO TA Cloning Kit for Sequencing (Invitrogen).

Protein analysis

BraA.eIF(iso)4E.a was amplified from BP8407 and Ji Zao Chun by first-strand cDNA synthesis with a poly dT primer using a Prime Script® RT-PCR kit (TaKaRa, http://www.takara-bio.com/) according to the manufacturer's instructions. Gene-specific PCR was carried out using Bio11537 sense 5'-GGAATTCCATGGCGACAGAGG ATGTG-3' (*Eco*RI site underlined) and Bio11538 antisense 5'-CCGCTCGAGTCAGACAGTGAACCGAG-3' (*Xho*I site underlined) primers. The amplified fragments were then digested and cloned into pET-32a (Novagen, http://www.novagen.com/) with a 6× histidine tag. The constructs were transformed into *E. coli* Rosetta (DE3)pLysS. Expression of recombinant proteins was induced at 20°C and 28°C for 20 h by the addition of 1 mmol isopropyl β-D-1-

thiogalactopyranoside. Molecular weights were determined using SDS-PAGE (Murphy and Kyle, 1994).

Yeast two-hybrid assays

Protein interactions were tested using the Matchmaker GAL 4 yeast two-hybrid system (Clontech). *BraA.eIF(iso)4E.a* was amplified from BP8407 and Ji Zao Chun RNA by RT-PCR with Bio11538 antisense primer followed by PCR with Bio11537 and Bio11538 primers. The VPg was amplified by RT-PCR from the RNA of brassica plants infected by TuMV-C4, with Bio120214 antisense primer 5′-CCCCGGGTCACTCGTGGTCCACTGGGA-3′ (*Xma*I site underlined), followed by PCR with Bio120213 sense 5′-CCCATATGAT-GGCGAAAGGTAAGAGGC-3′ (*Nde*I site underlined) and Bio120214 antisense primers. The TuMV VPg cDNA was cloned into the pGBKT7 plasmid by the *Nde*I (5′-end) and *Xma*I (3′-end) sites to generate a bait plasmid. *BraA.eIF(iso)4E.a* cDNAs were cloned into the pGADT7 plasmid by the *Eco*RI (5′-end) and *Xho*I (3′-end) sites to generate prey plasmids. Bait and prey constructs were transformed into the yeast strain AH109; the reporter genes were *HIS3* and *ADE2*. Empty vectors pGADT7 and pGBKT7 were used as negative controls along with the manufacturer's positive control.

Genotyping of *B. rapa* plants at *BraA.eIF(iso)4E.a, BraA.eIF(iso)4E.c* and *BraA.eIF4E.c* loci, production of B₁S₂ lines and analysis of virus susceptibility

We extracted DNA from leaves of B_1S_1 *B. rapa* plants using the DNeasy kit (Qiagen). The primers BR2 sense 5′-TCTCCTCCACTTCTTCCCAATAC-3′ and BR14 antisense 5′-TAG-ACAAGGCTTGGCTTGAAACTG-3′ were used to genotype relevant *eIF(iso)4E* copies from RLR22 and R-o-18, giving different sized products for *BraA.eIF(iso)4E.a* (larger for R-o-18) and *BraA.eIF(iso)4E.c* (larger for R-o-18). To genotype *BraA.eIF4E.c*, primers CN55 sense 5′-TCTTTGTTGGTGGGTTAGATTCCG-3′ and CN56 antisense 5′-ATCAACGCAAGCAACTACATCGAG-3′ amplified the R-o-18 allele and primers CN44 sense 5′-TTTCTTGTTGGGTTAAGTGAAG-3′ and CN45 antisense 5′-CAAGCAACTACATGGAAAAAAC-3′ amplified the RLR22 allele. Two plants from B_1S_1 seed (Rusholme *et al.*, 2007), one homozygous for the RLR22 alleles of *BraA.eIF(iso)4E.a* and *BraA.eIF4E.c* but heterozygous for *BraA.eIF(iso)4E.c* and one homozygous for the RLR22 alleles of *BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c* but heterozygous for *BraA.eIF4E.c*, were identified. These were then vernalised, grown on to flower and selfed to produce B_1S_2 seed, segregating for the respective genes. A single nucleotide polymorphism marker designed by Syngenta was used for genotyping F_2 populations from the RLR22 and CK 1 and CK 2 crosses at the *BraA.eIF(iso)4E.a* locus. The ELISA absorbance values of plants homozygous for the R-o-18 allele of *BraA.eIF(iso)4E.c* and heterozygous at this locus were analysed using Student's *t*-test.

Transcriptome analysis

The transcriptomic expression profile of an *A. thaliana* Col-0::*dSpm* plant complemented with *BraA.eIF(iso)4E.a* from RLR22 was determined using total RNA extracted from a young leaf and oligo(dT) selection performed twice using Dynal magnetic beads (Invitrogen). Illumina library preparation was performed using a mRNA-TruSeq sample prep kit version five (Illumina Inc., http://www.illumina.com/), according to the manufacturer's protocol (15018818 revA). The library was sequenced using Illumina's GAIIx sequencing system. Using the Illumina CASAVA pipeline, 70-bp end sequence reads were base-called and scored for read quality.

Sequences of mRNA from total RNA extracted from young leaves from each of three R-o-18 plants and three RLR22 plants were determined by SeqWright using an Illumina HiSeq 2000 system (Illumina Inc.). Ribosomal RNA depletion was performed using the RiboMinus Eukaryote Kit for RNA-Seq (Life Technologies, A10837-08) and libraries constructed using the TruSeq RNA Sample Preparation Kit (Illumina Inc.). The sequencing runs were paired-end, 100-bp reads and were analysed using Illumina CASA-VA version 1.8.

Sequence reads were aligned to the published Arabidopsis (Lamesch *et al.*, 2012) and *B. rapa* (Wang *et al.*, 2011) genome assemblies using Tophat (Trapnell *et al.*, 2009) and Bowtie (Langmead *et al.*, 2009) algorithms, respectively. The cufflinks algorithm (Trapnell *et al.*, 2010) was used to calculate and compare fragments per kilobase of exon per million mapped fragments (FPKM), to estimate relative transcript abundances.

Complementation of *A. thaliana* Col-0::*dSpm*

BraA.eIF(iso)4E.a, BraA.eIF(iso)4E.c, BraA.eIF4E.a and *BraA.eIF4E.c* were amplified from RLR22 genomic DNA, cloned and transformed into the *A. thaliana* line Col-0::*dSpm* possessing a transposon knock-out of *eIF(iso)4E* (Duprat *et al.*, 2002) as described for R-o-18 copies of the genes (Jenner *et al.*, 2010).

RLR22 sequences have been deposited in GenBank. Full-length genomic DNA of: *BraA.eIF(iso)4E.a*, JA722714; *BraA.eIF(iso)4E.c*, JA722768; *BraA.eIF4E.a*, JA722747; *BraA.eIF4E.c*, JA722756. Misspliced variants of *BraA.eIF(iso)4E.a*: with whole of intron 1 retained, cDNA sequence, JA722715; with last 15 nt of intron 1 retained, cDNA sequence, JA722717; with last 3 nt of exon 1 missing, cDNA sequence, JA722719. cDNA sequences of: *BraA.eIF(iso)4E.c*, JA722769; *BraA.eIF4E.a*, JA722748; *BraA.eIF4E.c*, JA722757.

ACCESSION NUMBERS

JA722714, JA722768, JA722747, JA722756, JA722715, JA722717, JA722719, JA722769, JA722748, JA722757.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Plant symptoms following artificial (mechanical) challenge with *Turnip mosaic virus* CDN 1.

Figure S2. Correctly spliced and further mis-spliced variants of *BraA.eIF(iso)4E.a*.

Figure S3. Multiple alignment of amino acid sequences of *eIF(iso)4E*, based on homology with *eIF4E* and *eIF(iso)4E* from other plant species.

Table S1. The interactions between the VPg of TuMV-C4 and proteins from alleles of *BraA.eIF(iso)4E.a* and *BraA.eIF(iso)4E.c* from different *Brassica rapa* lines.

REFERENCES

Ashby, J.A., Stevenson, C.E.M., Jarvis, G.E., Lawson, D.M. and Maule, A.J. (2011) Structure-based mutational analysis of eIF4E in relation to *sbm1* resistance to Pea seed-borne mosaic virus in pea. *PLoS ONE*, **6**, e15873.

Basso, J., Dallaire, P., Charest, P.J., Devantier, Y. and Laliberté, J.-F. (1994) Evidence for an internal ribosome entry site within the 5′ non-translated region of turnip mosaic potyvirus RNA. *J. Gen. Virol.* **75**, 3157–3165.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K.S. and Robaglia, C. (2002) The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *Plant J.* **32**, 927–934.

Gao, Z., Johansen, E., Eyers, S., Thomas, C.L., Noel Ellis, T.H. and Maule, A.J. (2004) The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *Plant J.* **40**, 376–385.

German-Retana, S., Walter, J., Doublet, B. *et al.* (2008) Mutational analysis of plant cap-binding protein eIF4E reveals key amino acids involved in biochemical functions and potyvirus infection. *J. Virol.* **82**, 7601–7612.

Goodfellow, I.G. and Roberts, L.O. (2008) Eukaryotic initiation factor 4E. *Int. J. Biochem. Cell Biol.* **40**, 2675–2680.

Hughes, S.L., Green, S.K., Lydiate, D.J. and Walsh, J.A. (2002) Resistance to *Turnip mosaic virus* in *Brassica rapa* and *B. napus* and the analysis of genetic inheritance in selected lines. *Plant. Pathol.* **51**, 567–573.

Jenner, C.E. and Walsh, J.A. (1996) Pathotypic variation in turnip mosaic virus with special reference to European isolates. *Plant. Pathol.* **45**, 848–856.

Jenner, C.E., Nellist, C.F., Barker, G.C. and Walsh, J.A. (2010) *Turnip mosaic virus* (TuMV) is able to use alleles of both *eIF4E* and *eIF(iso)4E* from multiple loci of the diploid *Brassica rapa*. *Mol. Plant Microbe Interact.* **23**, 1498–1505.

Lamesch, P., Berardini, T.Z., Li, D. *et al.* (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**, D1202–D1210.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.

Le Gall, O., Aranda, M.A. and Caranta, C. (2011) Plant resistance to viruses mediated by translation initiation factors. In *Recent Advances in Plant Virology* (Caranta, C., Aranda, M.A., Tepfer, M. and López-Moya, J.J., eds). Norfolk: Caister Academic Press, pp. 177–194.

Lellis, A.D., Kasschau, K.D., Whitham, S.A. and Carrington, J.C. (2002) Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role of eIF(iso)4E during potyvirus infection. *Curr. Biol.* **12**, 1046–1051.

Liu, X.P., Lu, W.C., Lui, Y.K. *et al.* (1996) Occurrence and strain differentiation of Turnip mosaic potyvirus and sources of resistance in Chinese cabbage in China. *Acta Hort.* **407**, 431–440.

Monzingo, A., Dhaliwal, S., Dutt-Chaudhuri, A., Lyon, A., Sadow, J.H., Hoffman, D.W., Robertus, J.D. and Browning, K.S. (2007) The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. *Plant Physiol.* **143**, 1504–1518.

Murphy, J.F. and Kyle, M.M. (1994) Isolation and viral infection of *Capsicum* leaf protoplasts. *Plant Cell Rep.* **13**, 397–400.

Piron, F., Nicolaï, M., Minoïa, S., Piednoir, E., Moretti, A., Salgues, A., Zamir, D., Caranta, C. and Bendahmane, A. (2010) An induced mutation in tomato eIF4E leads to immunity to two potyviruses. *PLoS ONE*, **5**, e11313.

Plante, D., Viel, C., Léonard, S., Tampo, H., Laliberté, J.-F. and Fortin, M.G. (2004) *Turnip mosaic virus* VPg does not disrupt the translation initiation complex but interferes with cap binding. *Physiol. Mol. Plant Pathol.* **64**, 219–226.

Qian, W., Zhang, S., Zhang, S. *et al.* (2013) Mapping and candidate-gene screening of the novel Turnip mosaic virus resistance gene *retr02* in Chinese cabbage (*Brassica rapa* L.). *Theor. Appl. Genet.* **126**, 179–188.

Robaglia, C. and Caranta, C. (2006) Translation initiation factors: a weak link in plant RNA virus infection. *Trends Plant Sci.* **11**, 40–45.

Rusholme, R.L., Higgins, E.E., Walsh, J.A. and Lydiate, D.J. (2007) Genetic control of broad-spectrum resistance to turnip mosaic virus in *Brassica rapa* (Chinese cabbage). *J. Gen. Virol.* **88**, 3177–3186.

Town, C.D., Cheung, F., Maiti, R. *et al.* (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. *Plant Cell*, **18**, 1348–1359.

Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.

Walsh, J.A., Rusholme, R.L., Hughes, S.L., Jenner, C.E., Bambridge, J.M., Lydiate, D.J. and Green, S.K. (2002) Different classes of resistance to Turnip mosaic virus in *Brassica rapa*. *Eur. J. Plant Pathol.* **108**, 15–20.

Wang, X., Wang, H., Wang, J. *et al.* (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **43**, 1035–1039.

Wittmann, S., Chatel, H., Fortin, M.G. and Laliberté, J.-F. (1997) Interaction of the viral protein genome linked of Turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of *Arabidopsis thaliana* using the yeast two-hybrid system. *Virology*, **234**, 84–92.

Appendix A

GenBank Accession Details

Table A.1 – Details of sequences submitted to GenBank.

Line	Gene	Sequence Details	GenBank Accession	Reference
TuMV	Whole genome	cDNA	AB093610	Tomimura *et al.* (2003)
R-o-18	*BraA.eIF(iso)4E.a*	gDNA	JA722711	Nellist *et al.* (2014)
R-o-18	*BraA.eIF(iso)4E.a*	cDNA	JA722712	Nellist *et al.* (2014)
R-o-18	*BraA.eIF(iso)4E.c*	gDNA	JA722765	Nellist *et al.* (2014)
R-o-18	*BraA.eIF(iso)4E.c*	cDNA	JA722766	Nellist *et al.* (2014)
R-o-18	*BraA.eIF4E.c*	gDNA	JA722753	Nellist *et al.* (2014)
R-o-18	*BraA.eIF4E.c*	cDNA	JA722754	Nellist *et al.* (2014)
RLR22	*BraA.eIF(iso)4E.a*	gDNA	JA722714	Nellist *et al.* (2014)
RLR22	*BraA.eIF(iso)4E.a* (whole of intron 1)[a]	cDNA	JA722715	Nellist *et al.* (2014)
RLR22	*BraA.eIF(iso)4E.a* (last 15 nt of intron 1)[b]	cDNA	JA722717	Nellist *et al.* (2014)
RLR22	*BraA.eIF(iso)4E.a* (last 3 nt of exon 1)[c]	cDNA	JA722719	Nellist *et al.* (2014)

[a] RLR22 *BraA.eIF(iso)4E.a* variant retaining extra G and the whole of intron 1.
[b] RLR22 *BraA.eIF(iso)4E.a* variant retaining extra G and the last 14 nt of intron 1/retaining the last 15 nt of intron 1.
[c] RLR22 *BraA.eIF(iso)4E.a* variant lacking the last 3 nt of exon 1.